

sensors

Special Issue Reprint

Radar Target Detection, Imaging and Recognition

Edited by
Tianxian Zhang, Xueting Li and Yuanhang Wang

mdpi.com/journal/sensors



Radar Target Detection, Imaging and Recognition

Radar Target Detection, Imaging and Recognition

Guest Editors

Tianxian Zhang

Xueting Li

Yuanhang Wang



Basel • Beijing • Wuhan • Barcelona • Belgrade • Novi Sad • Cluj • Manchester

Guest Editors

Tianxian Zhang
School of Information and
Communication Engineering
University of Electronic
Science and Technology of
China
Chengdu
China

Xueting Li
School of Aeronautics and
Astronautics
Sichuan University
Chengdu
China

Yuanhang Wang
School of Information and
Communication Engineering
Dalian University of
Technology
Dalian
China

Editorial Office

MDPI AG
Grosspeteranlage 5
4052 Basel, Switzerland

This is a reprint of the Special Issue, published open access by the journal *Sensors* (ISSN 1424-8220), freely accessible at: https://www.mdpi.com/journal/batteries/special_issues/39X1EPZF68.

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

Lastname, A.A.; Lastname, B.B. Article Title. <i>Journal Name</i> Year , Volume Number, Page Range.
--

ISBN 978-3-7258-8579-4 (Hbk)

ISBN 978-3-7258-8580-0 (PDF)

<https://doi.org/10.3390/books978-3-7258-8580-0>

© 2026 by the authors. Articles in this reprint are Open Access and distributed under the Creative Commons Attribution (CC BY) license. The reprint as a whole is distributed by MDPI under the terms and conditions of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

Contents

Yaowei Guo, Bin Guo and Guobing Qian

A Robust Complex α -Sigmoid Affine Projection Algorithm Under Non-Gaussian Noise

Reprinted from: *Sensors* **2026**, 26, 961, <https://doi.org/10.3390/s26030961> 1

Xiang Li, Yitao Su, Xiaobin Zhao, Junjun Yin and Jian Yang

Radar HRRP Sequence Target Recognition Based on a Lightweight Spatiotemporal Fusion Network

Reprinted from: *Sensors* **2026**, 26, 334, <https://doi.org/10.3390/s26010334> 18

Xiang Liu, Huafeng He, Ruike Li, Yubing Wu, Xin Zhang and Yongquan You

Distributed Phased-Array Radar Mainlobe Interference Suppression and Cooperative Localization Based on CEEMDAN-WOBSS

Reprinted from: *Sensors* **2025**, 25, 6277, <https://doi.org/10.3390/s25206277> 42

Lijia Chen and Xueting Li

Variable-Speed UAV Path Optimization Based on the CRLB Criterion for Passive Target Localization

Reprinted from: *Sensors* **2025**, 25, 5297, <https://doi.org/10.3390/s25175297> 68

Yue Dai, Shiyuan Zhang and Guoqiang Guo

L1-Constrained Fractional-Order Gradient Descent for Axial Dimension Estimation of Conical Targets

Reprinted from: *Sensors* **2025**, 25, 5082, <https://doi.org/10.3390/s25165082> 88

Jiadi Qi, Xiaoke Lu and Jinping Sun

An Intelligent Track Segment Association Method Based on Characteristic-Aware Attention LSTM Network

Reprinted from: *Sensors* **2025**, 25, 3465, <https://doi.org/10.3390/s25113465> 108

Ruizhi Liu, Zhenhang Qin, Xinghui Song, Lei Yang, Yue Lin and Hongtao Xu

Indoor mmWave Radar Ghost Suppression: Trajectory-Guided Spatiotemporal Point Cloud Learning

Reprinted from: *Sensors* **2025**, 25, 3337, <https://doi.org/10.3390/s25113377> 131

Zhaoyu Wang, Qinghe Zhang, Zhaoyang Shen, Lei Zhang and Han Liu

A Machine Learning Method for the Fast Simulation of the Scattering Characteristics of a Target Under a Planar Layered Medium

Reprinted from: *Sensors* **2025**, 25, 2481, <https://doi.org/10.3390/s25082481> 150

Simon Müller, Andreas R. Diewald and Georg Fischer

An Active Radar Interferometer Utilizing a Heterodyne Principle-Based Target Modulator

Reprinted from: *Sensors* **2025**, 25, 1711, <https://doi.org/10.3390/s25061711> 164

Yuanhang Wang, Yi Han and Yi Jiang

Adversarial Range Gate Pull-Off Jamming Against Tracking Radar

Reprinted from: *Sensors* **2025**, 25, 1553, <https://doi.org/10.3390/s25051553> 185

Xiaoyou Xing, Zhiwen Zhong, Xueting Li and Yiyang Yue

Node Selection and Path Optimization for Passive Target Localization via UAVs

Reprinted from: *Sensors* **2025**, 25, 780, <https://doi.org/10.3390/s25030780> 205

Mauro Larrat and Claudomiro Sales

Classification of Flying Drones Using Millimeter-Wave Radar: Comparative Analysis of Algorithms Under Noisy Conditions

Reprinted from: *Sensors* **2025**, 25, 721, <https://doi.org/10.3390/s25030721> **231**

Kang-In Lee, Jin-Hyeok Kim and Young-Seek Chung

Two-Dimensional Scattering Center Estimation for Radar Target Recognition Based on Multiple High-Resolution Range Profiles

Reprinted from: *Sensors* **2024**, 24, 6997, <https://doi.org/10.3390/s24216997> **261**

Weikun He, Yichuan Luo and Xiaoxiao Shang

Motion Clutter Suppression for Non-Cooperative Target Identification Based on Frequency Correlation Dual-SVD Reconstruction

Reprinted from: *Sensors* **2024**, 24, 5298, <https://doi.org/10.3390/s24165298> **280**

Article

A Robust Complex α -Sigmoid Affine Projection Algorithm Under Non-Gaussian Noise

Yaowei Guo, Bin Guo and Guobing Qian *

College of Electronic and Information Engineering, Southwest University, Chongqing 400715, China;
15833413651@163.com (Y.G.); guob0310@163.com (B.G.)

* Correspondence: guobingqian@swu.edu.cn

Abstract

To address the performance degradation of traditional adaptive filtering algorithms in environments with correlated input signals and non-Gaussian noise, this paper proposes a complex-valued affine projection algorithm based on the α -Sigmoid cost function (α -CSAP). The algorithm leverages the nonlinear characteristics of the α -Sigmoid function and implicitly achieves variable step-size updates by introducing a normalization factor, which effectively suppresses impulsive noise interference and avoids matrix inversion, thereby reducing computational complexity. Theoretical analysis derives the steady-state mean square deviation (MSD) expression for the algorithm. Simulation results demonstrate that the proposed α -CSAP algorithm exhibits superior performance compared to traditional complex adaptive filtering algorithms in both system identification and beamforming application scenarios.

Keywords: adaptive filtering algorithm; affine projection algorithm; non-Gaussian noise; beamforming

1. Introduction

As a core technology in the field of digital signal processing, adaptive filtering, which does not require prior knowledge of the statistical characteristics of signals and noise and can dynamically adjust parameters to adapt to time-varying environments, has become a key means for radar systems to cope with complex signal and interference conditions [1–4]. Conventional fixed filters exhibit limited performance when dealing with non-stationary, nonlinear signals or unknown interference, whereas adaptive filtering, through iterative optimization algorithms such as Least Mean Square (LMS), Recursive Least Squares (RLS), and particle filtering, enables efficient denoising, target detection, and parameter estimation, and is widely applicable in engineering scenarios such as radar and communications [5–7].

In radar signal processing, adaptive filtering plays an indispensable and central role. Radar echoes are often affected by non-Gaussian noise such as spherically invariant clutter and pulse interference, making it difficult for traditional detection methods to achieve a satisfactory balance between detection probability and false alarm rate [8,9]. Adaptive filtering can effectively address such complex signal environments by dynamically adjusting filter weights or structures. For example, the adaptive matched filter can maintain a constant false alarm rate under unknown clutter covariance matrices, significantly improving radar target detection performance [10]; optimized particle filtering can achieve high-precision prediction of highly dynamic target trajectories, effectively handling tracking problems in nonlinear and non-Gaussian scenarios [11]. At the signal and data processing level,

the RLS algorithm, with its fast convergence characteristics, is particularly suitable for processing non-stationary radar signals [12]. Furthermore, the integration of adaptive filtering with techniques such as digital beamforming and data fusion further enhances the anti-interference capability and target resolution accuracy of radar systems [13].

As a central direction in modern signal processing and machine learning, complex-valued adaptive filtering is of broad application value compared to real-valued filtering techniques, owing to its inherent suitability for complex signals prevalent in digital communications, power systems, radar, and other fields [14]. Traditional algorithms, often based on the Mean Square Error (MSE) criterion, are simple to implement but tend to become inaccurate or diverge in non-Gaussian impulsive noise environments. Moreover, they typically neglect the statistical properties of non-circular signals, leading to information loss [15]. The integration of information-theoretic criteria has led to algorithms based on complex correntropy and generalized complex correntropy, significantly enhancing robustness against non-Gaussian noise [16]. However, challenges remain, including the insufficient flexibility of single-kernel functions, incomplete utilization of error probability distributions, performance degradation under noisy input–output models, and the inherent trade-off between convergence speed and steady-state error [17]. In response, researchers have pursued optimization through various means, such as employing widely linear models to capture non-circular characteristics, introducing mixed-kernel functions to enhance adaptability, and designing fixed-point iterative or recursive forms to reduce complexity [18]. These efforts aim to better adapt these algorithms to complex engineering scenarios. Complementary to complex adaptive filtering, beamforming is a key technology in array signal processing. It enhances desired signals and suppresses interference by controlling the phase and amplitude of signals received by multiple sensors/antennas, demonstrating significant advantages in wireless communications, acoustic processing, and radar detection. Its core value lies in utilizing spatial diversity to improve signal quality and transmission efficiency, giving rise to specialized applications such as distributed collaborative beamforming [19], microphone array speech enhancement, and broadband beamforming with fractal antenna arrays [20]. Innovations in hybrid architectures and unsupervised methods further extend its applicability [21]. As important branches of adaptive signal processing, both complex adaptive filtering and beamforming face common challenges regarding adaptability in complex environments and balancing performance with complexity. Their technological integration and collaborative optimization have become important research directions driven by practical engineering needs.

In applications such as system identification, echo cancellation, and signal prediction, the performance of traditional adaptive filtering algorithms is significantly affected when the input signal is correlated. Their core operation lies in adjusting filter coefficients in real time to adapt to dynamic environments [22]. The classic LMS algorithm is widely adopted due to its simplicity, but its convergence speed degrades notably under correlated inputs. RLS algorithms converge faster but face high computational complexity, making them less suitable for real-time requirements [23]. To balance convergence speed and computational complexity, Ozeki and Umeda first proposed the Affine Projection (AP) algorithm in 1984 [24]. By reusing past input data, it achieves faster convergence than LMS while maintaining lower complexity than RLS, representing a major breakthrough in adaptive filtering [25]. However, practical challenges such as impulsive noise, highly correlated inputs, and sparse system characteristics impose limitations on traditional AP algorithms. For example, in acoustic echo cancellation and hands-free vehicular communications, impulsive interference can degrade algorithm performance [26]. In sparse system identification, conventional AP algorithms fail to exploit sparsity, leading to inefficient convergence [27]. In massive data processing scenarios, indiscriminate data updates

result in wasted computational resources. Consequently, extensive research has been dedicated to improving the AP algorithm. Various strategies have been developed to enhance its performance. Data-selective updating mechanisms have been introduced to reduce computational load while maintaining estimation accuracy. Robust AP algorithms based on M-estimation have been designed to improve resilience against impulsive noise [28]. Variants incorporating sparsity-inducing penalties have been proposed for sparse system identification [29]. Furthermore, approaches that avoid direct or indirect matrix inversion have been adopted to lower the computational complexity of AP-type algorithms [30]. Additional techniques, such as subband processing and variable step-size strategies, have also been employed to further optimize convergence behavior under correlated inputs [31]. These improved algorithms demonstrate advantages in applications like system identification, echo cancellation, and beamforming. Nonetheless, achieving an optimal balance among robustness, real-time performance, and accuracy in complex environments remains an open challenge. Furthermore, while most existing AP algorithms are confined to the real domain, their application in the complex domain is increasingly important due to the prevalence of complex signals in radar signal processing.

Recently, researchers proposed an adaptive projection algorithm based on an α -sigmoid cost function, demonstrating a superior performance when dealing with strongly correlated input signals and non-Gaussian noise environments [32]. This provides a new approach for radar echo cancellation and channel estimation. Building upon this, we propose a novel complex-valued adaptive filtering algorithm based on the α -sigmoid cost function. The steady-state MSD of the proposed algorithm is derived analytically. The performance of the proposed algorithm is compared with classical complex adaptive filtering algorithms in simulations of system identification and beamforming. The results confirm the effectiveness and robustness of the proposed algorithm. Furthermore, the proposed algorithm offers a new potential solution for beamforming applications.

The remainder of this paper is organized as follows: Section 2 introduces the proposed algorithm. Section 3 provides a theoretical analysis of the algorithm, including the derivation of its steady-state mean MSD expression and an analysis of its computational complexity. Section 4 verifies the theoretical analysis and provides simulation results to demonstrate the superiority and robustness of the proposed algorithm. Section 5 concludes the paper.

2. The Proposed α -CSAP Algorithm

As illustrated in Figure 1, the system identification model considered in this work is given by

$$d(n) = \mathbf{x}^H(n) \mathbf{w}_0 + z(n), \quad (1)$$

where $\mathbf{w}_0 \in \mathbb{C}^{L_w \times 1}$ denotes the unknown system weight vector with L_w being the filter length, $z(n)$ is the complex-valued additive noise, and the input vector $\mathbf{x}(n) \in \mathbb{C}^{L_w \times 1}$ is defined as $\mathbf{x}(n) = [x(n), x(n-1), \dots, x(n-L_w+1)]^T$. Here, $(\cdot)^H$ denotes the Hermitian transpose. The output of the adaptive filter is expressed as $y(n) = \mathbf{x}^H(n) \mathbf{w}(n-1)$; here, $\mathbf{w}(n-1) \in \mathbb{C}^{L_w \times 1}$ is the adaptive weight vector at time index $n-1$. Accordingly, the instantaneous error signal is given by

$$e(n) = d(n) - y(n) = d(n) - \mathbf{x}^H(n) \mathbf{w}(n-1). \quad (2)$$

In the AP algorithm framework, the projection order Q specifies the number of past input–error pairs used for coefficient adaptation. The stacked error vector is defined as $\mathbf{e}(n) = [e(n), e(n-1), \dots, e(n-Q+1)]^T \in \mathbb{C}^{Q \times 1}$ and the input data matrix is constructed as $\mathbf{X}(n) = [\mathbf{x}(n), \mathbf{x}(n-1), \dots, \mathbf{x}(n-Q+1)] \in \mathbb{C}^{L_w \times Q}$. In order to enhance the convergence

rate of the algorithm and effectively mitigate the adverse effects of non-Gaussian noise on the adaptation process, the α -sigmoid function is employed as the cost function, as follows:

$$J(n) = \sum_{i=0}^{Q-1} \frac{\alpha}{\alpha - 1} \left(1 + \exp \left(-\frac{|e(n-i)|^2}{2\delta^2} \right) \right)^{1-\frac{1}{\alpha}}, \quad (3)$$

where the parameter α governs the steepness of the cost function, thereby determining the convergence speed of the algorithm: a larger α leads to a steeper function and faster convergence. The parameter δ adjusts the algorithm's sensitivity to outliers. From a mathematical perspective, as the error $e(n)$ increases, the curvature of the cost function gradually flattens, and its derivative with respect to the weights approaches zero, thereby achieving robust suppression of outliers.

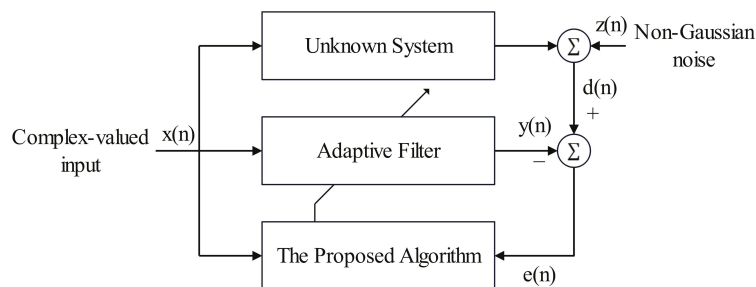


Figure 1. Block diagram for system identification.

Taking the Wirtinger derivative of $J(n)$ with respect to the complex conjugate of the weight vector $\mathbf{w}^*(n-1)$ yields

$$\frac{\partial J(n)}{\partial \mathbf{w}^*(n-1)} = -\mathbf{X}(n)\boldsymbol{\eta}(n)\mathbf{e}^*(n), \quad (4)$$

where $\boldsymbol{\eta}(n) \in \mathbb{R}^{Q \times Q}$ is a diagonal weighting matrix defined as

$$\boldsymbol{\eta}(n) = \text{diag}\{\eta_1(n), \dots, \eta_Q(n)\}, \quad (5)$$

with

$$\eta_i(n) = \frac{\exp\left(-\frac{|e(n-i+1)|^2}{2\delta^2}\right)}{\delta^2 \left(1 + \exp\left(-\frac{|e(n-i+1)|^2}{2\delta^2}\right)\right)^{1/\alpha}}, \quad i = 1, 2, \dots, Q. \quad (6)$$

Under the gradient descent framework, a preliminary update expression for the algorithm is formulated as follows:

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mu \mathbf{X}(n)\boldsymbol{\eta}(n)\mathbf{e}^*(n). \quad (7)$$

where μ is the step size.

In the derivation of the α SAP algorithm, the introduction of the normalization factor $\|\mathbf{X}(n)\mathbf{e}(n)\|_2$ is a key step to realize the L_2 -norm constraint. The squared Euclidean norm of the weight update vector is expressed as follows:

$$\|\mathbf{w}(n) - \mathbf{w}(n-1)\|_2^2 = \mu^2 \|\mathbf{X}(n)\boldsymbol{\eta}(n)\mathbf{e}^*(n)\|_2^2. \quad (8)$$

As can be observed from Equation (8), when the noise contains large outliers, the error $e(n)$ will be significantly affected, which, in turn, substantially impacts the weight estimation and degrades the filter performance. The restriction of the filter update magnitude can effectively enhance the robustness of the filter. This can be expressed as follows:

$$\|\mathbf{w}(n) - \mathbf{w}(n-1)\|_2^2 \leq \varepsilon, \quad (9)$$

where ε is a positive scalar. Taking the partial derivative of Equation (6), and according to Equation (8), the diagonal elements of the weighting matrix $\boldsymbol{\eta}(n)$ have an upper bound $\eta_{\max} = 1/(\delta^2 \cdot 2^{1/\alpha})$, which leads to the following:

$$\begin{aligned} \mu^2 \|\mathbf{X}(n)\boldsymbol{\eta}(n)\mathbf{e}^*(n)\|_2^2 &\leq \mu^2 \eta_{\max}^2 \|\mathbf{X}(n)\mathbf{e}(n)\|_2^2 \\ \frac{\mu^2 \|\mathbf{X}(n)\boldsymbol{\eta}(n)\mathbf{e}^*(n)\|_2^2}{\|\mathbf{X}(n)\mathbf{e}(n)\|_2^2} &\leq \mu^2 \eta_{\max}^2 \end{aligned} \quad (10)$$

At this point, by introducing a time-varying parameter $\|\mathbf{X}(n)\mathbf{e}(n)\|_2$ into the denominator of Equation (8) that shares the same order as the numerator, the upper bound for the squared norm of the squared Euclidean norm of the weight update vector is obtained, as follows:

$$\|\mathbf{w}(n) - \mathbf{w}(n-1)\|_2^2 \leq \mu^2 \eta_{\max}^2. \quad (11)$$

The update formula with constraints is expressed as follows:

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mu \frac{\mathbf{X}(n)\boldsymbol{\eta}(n)\mathbf{e}^*(n)}{\|\mathbf{X}(n)\mathbf{e}(n)\|_2}. \quad (12)$$

The proposed α -CSAP algorithm is summarized in Algorithm 1. The introduction of the normalization factor can be understood from an implicit variable step-size perspective. In scenarios with large errors, the effective step size $\mu/\|\mathbf{X}(n)\mathbf{e}(n)\|_2$ automatically decreases, thereby suppressing impulse noise interference and preventing filter divergence. In scenarios with small errors, the effective step size relatively increases, accelerating the convergence process. Moreover, traditional affine projection algorithms require the inversion or pseudo-inversion of the input matrix, which is prone to numerical instability due to ill-conditioned matrices. The normalization factor avoids matrix inversion operations, significantly reducing computational complexity.

Algorithm 1 The α -CSAP Algorithm

Require: $x(n)$, $d(n)$, L_w , Q , μ , α , δ^2 , N

Ensure: $\mathbf{w}(n)$

```

1: Initialize  $\mathbf{w}(0) = \mathbf{0}_{L_w \times 1}$ 
2: for  $n = 1$  to  $N$  do
3:   Construct:  $\mathbf{x}(n) = [x(n), x(n-1), \dots, x(n-L_w+1)]^T$ 
4:   Construct:  $\mathbf{X}(n) = [\mathbf{x}(n), \mathbf{x}(n-1), \dots, \mathbf{x}(n-Q+1)]$ 
5:   Compute:  $y(n) = \mathbf{x}^H(n)\mathbf{w}(n-1)$ 
6:   Compute:  $e(n) = d(n) - y(n)$ 
7:   Form:  $\mathbf{e}(n) = [e(n), e(n-1), \dots, e(n-Q+1)]^T$ 
8:   for  $i = 1$  to  $Q$  do
9:     
$$\eta_i(n) = \frac{\exp\left(-\frac{|e(n-i+1)|^2}{2\delta^2}\right)}{\delta^2 \left[1 + \exp\left(-\frac{|e(n-i+1)|^2}{2\delta^2}\right)\right]^{1/\alpha}}$$

10:   end for
11:   Construct:  $\boldsymbol{\eta}(n) = \text{diag}\{\eta_1(n), \dots, \eta_Q(n)\}$ 
12:   
$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mu \cdot \frac{\mathbf{X}(n)\boldsymbol{\eta}(n)\mathbf{e}^*(n)}{\|\mathbf{X}(n)\mathbf{e}(n)\|_2}$$

13: end for
14: return  $\mathbf{w}(n)$ 

```

3. Theoretical Analysis

3.1. Unbiasedness in the Mean

This subsection employs a classical method from mean value analysis to conduct an approximate analysis of the mean unbiasedness of the proposed α -CSAP algorithm under reasonable assumptions.

Define the weight error vector as

$$\tilde{\mathbf{w}}(n) = \mathbf{w}(n) - \mathbf{w}_0, \quad (13)$$

where $\mathbf{w}_0$ denotes the unknown system weight vector. Substituting (12) into the above equation yields the update formula with constraints, which is expressed as:

$$\tilde{\mathbf{w}}(n) = \tilde{\mathbf{w}}(n-1) + \mu \frac{\mathbf{X}(n)\boldsymbol{\eta}(n)\mathbf{e}^*(n)}{\|\mathbf{X}(n)\mathbf{e}(n)\|_2}. \quad (14)$$

Under the affine projection framework with projection order Q , let the stacked desired signal vector be $\mathbf{d}(n) = [d(n), d(n-1), \dots, d(n-Q+1)]^T$ and the stacked noise vector be $\mathbf{z}(n) = [z(n), z(n-1), \dots, z(n-Q+1)]^T$. Thus, the stacked error vector can be expressed as

$$\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{X}^H(n)\mathbf{w}(n-1) = -\mathbf{X}^H(n)\tilde{\mathbf{w}}(n-1) + \mathbf{z}(n). \quad (15)$$

Substituting the above expression into the normalized update rule of the proposed α -CSAP algorithm yields the following weight error recursion:

$$\begin{aligned} \tilde{\mathbf{w}}(n) = & \tilde{\mathbf{w}}(n-1) - \mu \frac{\mathbf{X}(n)\boldsymbol{\eta}(n)\mathbf{X}^T(n)\tilde{\mathbf{w}}^*(n-1)}{\|\mathbf{X}(n)\mathbf{e}(n)\|_2} \\ & + \mu \frac{\mathbf{X}(n)\boldsymbol{\eta}(n)\mathbf{z}^*(n)}{\|\mathbf{X}(n)\mathbf{e}(n)\|_2}. \end{aligned} \quad (16)$$

To facilitate the mean analysis, the following assumptions are adopted.

- (A1) The input matrix $\mathbf{X}(n)$ is statistically independent of the weight error vector $\tilde{\mathbf{w}}(n)$.
- (A2) In steady state, the diagonal weighting matrix $\boldsymbol{\eta}(n)$ varies slowly and can be approximated by

$$\boldsymbol{\eta}(n) \approx \bar{g}\mathbf{I}_Q, \quad (17)$$

where $\bar{g} = \mathbb{E}\{\eta_i(n)\}$.

- (A3) $\|\mathbf{X}(n)\mathbf{e}(n)\|_2$ is weakly correlated with $\mathbf{X}(n)$ and $\tilde{\mathbf{w}}(n-1)$, while $\mathbf{z}(n)$ is zero-mean and independent of both.

Under this assumption, it follows that $\mathbb{E}\left\{\frac{\mathbf{X}(n)\boldsymbol{\eta}(n)\mathbf{z}^*(n)}{\|\mathbf{X}(n)\mathbf{e}(n)\|_2}\right\} \approx \mathbf{0}$. Then, taking the expectation on both sides of (16) yields

$$\mathbb{E}\{\tilde{\mathbf{w}}(n)\} \approx \left[\mathbf{I} - \mu\bar{g}\mathbb{E}\left\{\frac{\mathbf{X}(n)\mathbf{X}^H(n)}{\|\mathbf{X}(n)\mathbf{e}(n)\|_2}\right\} \right] \mathbb{E}\{\tilde{\mathbf{w}}(n-1)\}. \quad (18)$$

When the input signal is persistently exciting, the matrix inside the expectation is positive semi-definite. Therefore, for a sufficiently small step size μ , the only equilibrium point of (18) is

$$\lim_{n \rightarrow \infty} \mathbb{E}\{\tilde{\mathbf{w}}(n)\} = \mathbf{0}. \quad (19)$$

This result indicates that the proposed α -CSAP algorithm is unbiased in the mean.

3.2. Steady-State MSD Analysis

3.2.1. Theoretical Value Derivation

In this subsection, the steady-state MSD performance of the proposed α -CSAP algorithm is analyzed. Due to the presence of nonlinear error weighting and normalization terms, the analysis is carried out using the energy conservation approach commonly adopted in the mean-square analysis of AP algorithms.

The MSD is defined as

$$\text{MSD}(n) \triangleq \mathbb{E}\{\|\tilde{\mathbf{w}}(n)\|^2\}. \quad (20)$$

From the normalized update rule of the proposed algorithm, the weight error vector evolves as

$$\tilde{\mathbf{w}}(n) = \tilde{\mathbf{w}}(n-1) - \mu \mathbf{K}(n), \quad (21)$$

where

$$\mathbf{K}(n) = \frac{\mathbf{X}(n)\boldsymbol{\eta}(n)\mathbf{e}^*(n)}{\|\mathbf{X}(n)\mathbf{e}(n)\|_2}. \quad (22)$$

Taking the squared Euclidean norm on both sides of (21), the expectation yields

$$\begin{aligned} \mathbb{E}\{\|\tilde{\mathbf{w}}(n)\|^2\} &= \mathbb{E}\{\|\tilde{\mathbf{w}}(n-1)\|^2\} \\ &\quad - 2\mu \mathbb{E}\{\text{Re}[\tilde{\mathbf{w}}^H(n-1)\mathbf{K}(n)]\} \\ &\quad + \mu^2 \mathbb{E}\{\|\mathbf{K}(n)\|^2\}. \end{aligned} \quad (23)$$

Under steady-state conditions, $\mathbb{E}\{\|\tilde{\mathbf{w}}(n)\|^2\} = \mathbb{E}\{\|\tilde{\mathbf{w}}(n-1)\|^2\}$, the MSD satisfies the following energy balance relation:

$$2\mu \mathbb{E}\{\text{Re}[\tilde{\mathbf{w}}^H(n-1)\mathbf{K}(n)]\} = \mu^2 \mathbb{E}\{\|\mathbf{K}(n)\|^2\}. \quad (24)$$

Define the a priori error vector as

$$\mathbf{e}_a(n) = \mathbf{X}^H(n)\tilde{\mathbf{w}}(n-1). \quad (25)$$

Then,

$$\tilde{\mathbf{w}}^H(n-1)\mathbf{K}(n) = \frac{\mathbf{e}_a^H(n)\boldsymbol{\eta}(n)\mathbf{e}^*(n)}{\|\mathbf{X}(n)\mathbf{e}(n)\|_2}. \quad (26)$$

Since $\boldsymbol{\eta}(n)$ is a diagonal matrix with entries $\eta(|e(n)|), \dots, \eta(|e(n-Q+1)|)$, the numerator in (26) can be expanded as

$$\mathbf{e}_a^H(n)\boldsymbol{\eta}(n)\mathbf{e}^*(n) = \sum_{i=0}^{Q-1} \eta(|e(n-i)|) e_a^*(n-i) e^*(n-i). \quad (27)$$

Following the standard mean-square analysis of APs, it is assumed that under steady-state and small-step-size conditions, the correlation between the numerator and the normalization factor in (26) is weak. Thus, the following separation approximation is adopted:

$$\mathbb{E}\left\{\frac{A}{B}\right\} \approx \frac{\mathbb{E}\{A\}}{\mathbb{E}\{B\}}. \quad (28)$$

Applying this approximation to (26), and noting that all Q components are statistically equivalent in steady state, yields

$$\mathbb{E}\left\{\operatorname{Re}\left[\tilde{\mathbf{w}}^H(n-1)\mathbf{K}(n)\right]\right\} \approx \frac{Q \mathbb{E}\{\operatorname{Re}[\eta(|e|) e_a^* e^*]\}}{\mathbb{E}\{\|\mathbf{X}(n)\mathbf{e}(n)\|_2\}}. \quad (29)$$

Define

$$\varphi_c \triangleq \mathbb{E}\{\|\mathbf{X}(n)\mathbf{e}(n)\|_2\}. \quad (30)$$

Under steady-state conditions, the diagonal weighting function varies slowly and can be approximated by its average value, i.e.,

$$\eta(|e|) \approx \bar{\eta}, \quad \bar{\eta} = \mathbb{E}\{\eta(|e|)\}. \quad (31)$$

Therefore, the above expression reduces to

$$\mathbb{E}\left\{\operatorname{Re}\left[\tilde{\mathbf{w}}^H(n-1)\mathbf{K}(n)\right]\right\} \approx \frac{Q \bar{\eta}}{\varphi_c} \mathbb{E}\{\operatorname{Re}[e_a^* e^*]\}. \quad (32)$$

For the complex-valued error $e = e_a + z$, where e_a is a proper complex Gaussian random variable and z denotes Bernoulli–Gaussian impulsive noise, the weighting function operates on the magnitude of e only. By invoking Price's theorem under the complex-valued setting, the following approximation holds:

$$\mathbb{E}\{\operatorname{Re}[e_a^* e^*]\} \approx \frac{1}{\sqrt{\pi}} \frac{\mathbb{E}\{|e_a|^2\}(1 - P_r)}{\sqrt{\mathbb{E}\{|e_a|^2\} + \sigma_\alpha^2}}. \quad (33)$$

Moreover, under steady-state conditions,

$$\|\mathbf{X}(n)\mathbf{e}(n)\|_2 \approx \sqrt{Q \mathbb{E}\{|e(n)|^2\}}. \quad (34)$$

Combining the above results yields

$$\mathbb{E}\left\{\operatorname{Re}\left[\tilde{\mathbf{w}}^H(n-1)\mathbf{K}(n)\right]\right\} \approx \frac{\bar{\eta}}{\varphi_c \sqrt{Q \mathbb{E}\{|e(n)|^2\}}} \cdot \frac{1}{\sqrt{\pi}} \frac{\mathbb{E}\{|e_a(n)|^2\}(1 - P_r)}{\sqrt{\mathbb{E}\{|e_a(n)|^2\} + \sigma_\alpha^2}}. \quad (35)$$

The second-order term in (24) can be written as

$$\mathbb{E}\left\{\|\mathbf{K}(n)\|^2\right\} = \mathbb{E}\left[\frac{\|\mathbf{X}(n)\boldsymbol{\eta}(n)\mathbf{e}^*(n)\|_2^2}{\|\mathbf{X}(n)\mathbf{e}(n)\|_2^2}\right]. \quad (36)$$

Under steady-state and small-step-size conditions, this term can be approximated as

$$\mathbb{E}\left\{\|\mathbf{K}(n)\|^2\right\} \approx \bar{\eta}^2, \quad (37)$$

which is upper bounded by $\eta_{\max}^2$ due to the boundedness of the weighting function.

Using the relations

$$\mathbb{E}\{|e_a(n)|^2\} = \sigma_x^2 \text{MSD}, \quad \mathbb{E}\{|e(n)|^2\} \approx \sigma_x^2 \text{MSD} + \sigma_\alpha^2, \quad (38)$$

the steady-state MSD is finally obtained as

$$\text{MSD} = \frac{\mu \varphi_c \eta_{\max} \sigma_\alpha^2 \left(\mu \varphi_c \eta_{\max} + \frac{2Q}{\sqrt{\pi}} (1 - P_r) \right)}{\sigma_x^2 \left[\frac{4Q^2}{\pi} (1 - P_r)^2 - \mu^2 \varphi_c^2 \eta_{\max}^2 \right]}. \quad (39)$$

The above expression provides an accurate approximation of the steady-state MSD.

3.2.2. Theoretical Validation

To verify the accuracy of the theoretical steady-state MSD model of the proposed α -CSAP algorithm, the steady-state MSD expression (39) derived from the theoretical analysis was validated through simulations under a projection order of $Q = 2$, with the step-size parameter μ varying within the range $\mu \in [0.25, 0.35]$. The experimental setup was as follows: the filter length L_w was set to three values, 16, 24, and 32, each matching the order of the unknown system; the number of iterations was $N = 1000$, which was sufficient for the algorithm to reach steady state; and the results were averaged over 100 Monte Carlo trials. The input signal was a normalized complex-valued white noise, and the noise environment was modeled as Bernoulli–Gaussian mixture noise with an impulse probability of 0.02. Figures 2–4 show the influence of the step-size parameter μ on the steady-state MSD for different filter lengths, with the projection order fixed at $Q = 2$.

When $L_w = 16$ (Figure 2), the theoretical MSD curve is slightly higher than the simulation data points, with an average error of 1.61 dB. As the filter length increases to $L_w = 24$ (Figure 3), the theoretical and simulation results are in excellent agreement, verifying the accuracy of the derived expression under this configuration. When the filter length is further increased to $L_w = 32$ (Figure 4), the predicted MSD values are slightly lower than the simulation results, with an average error of 1.35 dB. The slight deviations between the theoretical model and simulation results mainly originate from the necessary mathematical approximations adopted in the analysis, including the use of the separation assumption to handle the correlation of nonlinear terms, the steady-state averaging of the weighting function, and the approximate assumptions when applying Price's theorem in the complex domain. Although these simplifications introduce systematic errors, they ensure the analytical tractability of the theoretical expressions. Experimental results demonstrate that the proposed theoretical model still accurately captures the overall trend of the steady-state MSD with respect to the step size, confirming the effectiveness of the analytical framework.

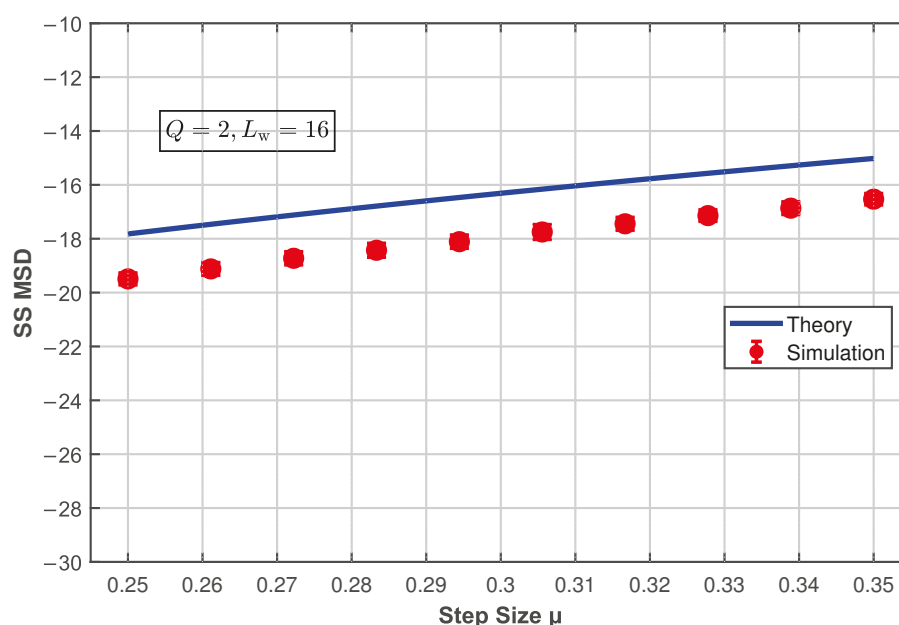


Figure 2. Effect of step-size parameter μ on steady-state MSD for α -CSAP algorithm ($Q = 2$, $L_w = 16$).

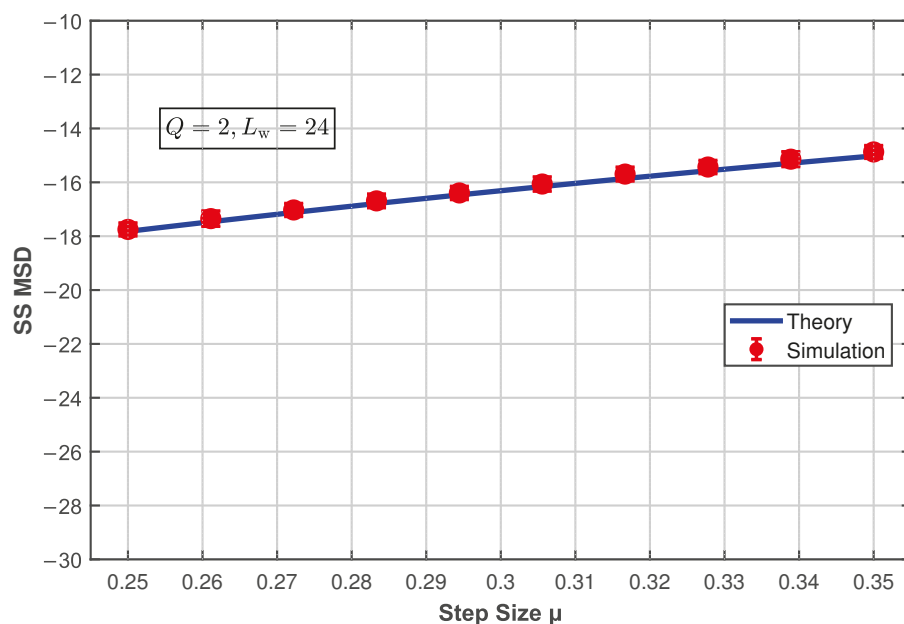


Figure 3. Effect of step-size parameter μ on steady-state MSD for α -CSAP algorithm ($Q = 2, L_w = 24$).

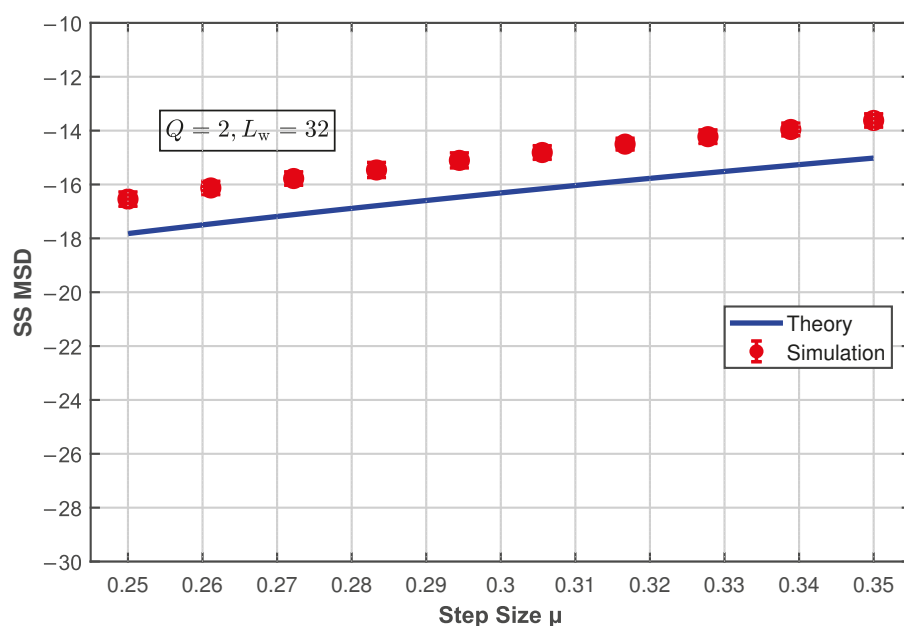


Figure 4. Effect of step-size parameter μ on steady-state MSD for α -CSAP algorithm ($Q = 2, L_w = 32$).

3.3. Computational Complexity Analysis

In this section, we conduct a comprehensive computational complexity analysis of the proposed α -CSAP algorithm. The operation counting baseline follows complex arithmetic rules: complex addition requires two real additions, and complex multiplication requires four real multiplications and two real additions. For a length- n complex inner product, $4n - 2$ real additions are required. The exponential function $\exp(\cdot)$, approximated using a second-order Taylor expansion, is estimated to need five multiplications and three additions.

Table 1 details the computational complexity of the α -CSAP algorithm per iteration, where L_w is the filter length and Q is the projection order.

Table 1. Computational complexity of α -CSAP algorithm (per iteration).

Step	$\times$	$+$	$\exp(\cdot)$	$\div$
Initialization: $\mathbf{w}(n) = \mathbf{0}$				
1. $\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{X}^H(n)\mathbf{w}(n-1)$	$4QL_w$	$4QL_w - 2Q$	0	0
2. $ e_k(n) ^2, k = 1, \dots, Q$	$2Q$	Q	0	0
3. $\exp(- e_k(n) ^2 / (2\delta^2))$	$5Q$	$3Q$	Q	0
4. $\eta_i(n) = \frac{\exp(\cdot)}{\delta^2(1+\exp(\cdot))^{1/\alpha}}$	$4Q$	$2Q$	0	Q
5. $\mathbf{X}(n)\mathbf{e}(n)$	$4QL_w$	$4QL_w - 2Q$	0	0
6. $\mathbf{X}(n)\eta(n)\mathbf{e}^*(n)$	$4Q^2 + 4QL_w$	$2Q^2 + 4QL_w - 2Q$	0	0
7. $\ \mathbf{X}(n)\mathbf{e}(n)\ _2$	$2Q + 4$	$Q + 1$	0	0
8. $\mathbf{w}(n) = \mathbf{w}(n-1) + \mu \frac{\mathbf{X}(n)\eta(n)\mathbf{e}^*(n)}{\ \mathbf{X}(n)\mathbf{e}(n)\ _2}$	$4QL_w + 2$	$4QL_w + 1$	0	1
Total	$8QL_w + 4Q^2 + 13Q + 6$	$8QL_w + 2Q^2 + 12Q - 1$	Q	$Q + 1$

Table 2 compares the computational complexity of different complex adaptive filtering algorithms. The algorithms under comparison include the Complex Least Mean Square (CLMS), the Complex Norm Least Mean Squares (CNLMS), the Complex Maximum Correntropy Criterion (CMCC), and the Complex Affine Projection (CAP) algorithms.

Table 2. Computational Comparison of Different Complex Adaptive Algorithms.

Algorithm	$\times$	$+$	$\exp(\cdot)$	$\div$
CLMS	$8L_w + 4$	$8L_w - 2$	0	0
CNLMS	$12L_w + 8$	$12L_w - 5$	0	1
CMCC	$8L_w + 26$	$8L_w + 12$	1	0
CAP	$8QL_w + 4Q^2 + 4Q^3$	$8QL_w + 4Q^2 + 4Q^3 - 6Q$	0	0
α -CSAP	$8QL_w + 4Q^2 + 13Q + 6$	$8QL_w + 2Q^2 + 12Q - 1$	Q	$Q + 1$

In terms of optimization potential, by employing fast recursive filtering (FRF) and an approximation of $\exp(\cdot)$ requiring only three multiplications, the complexity of α -CSAP can be reduced to $6QL_w + 3Q^2 + 8Q$ multiplications and $6QL_w + Q^2 + 6Q$ additions.

The proposed α -CSAP algorithm provides a good trade-off between computational complexity and performance. Although it requires more operations than LMS-type algorithms ($O(L_w)$), it avoids the $O(Q^3)$ matrix inversion of CAP and remains robust to impulsive noise and colored signals. In engineering applications, L_w is often much larger than Q , so the complexity of the α -CSAP algorithm is dominated by the $O(QL_w)$ term, making it suitable for real-time implementation.

4. Simulation

4.1. System Identification Application

This section evaluates the system identification performance of the proposed α -CSAP algorithm under various complex-valued non-Gaussian noise conditions, in comparison with several representative complex-valued adaptive filtering algorithms, including CLMS, CNLMS, CAP, and CMCC. The unknown system $\mathbf{w}_o$ to be identified is represented by a complex impulse response vector of length $L_w = 16$, whose real and imaginary parts are independently drawn from a Gaussian distribution with zero mean and a variance of 0.5. The input signal $x(n)$ is a first-order complex autoregressive (AR(1)) process, generated by filtering complex-valued white Gaussian noise through a system function $H(z) = 1/(1 - 0.9z^{-1})$, resulting in strong temporal correlation. The projection order is set to $Q = 2$. For a fair comparison, all algorithmic parameters are adjusted to achieve the same initial convergence speed. Simulation results are averaged over 100 Monte Carlo

trials, and the MSD learning curve is adopted as the performance metric. The MSD at the n -th iteration is calculated as follows:

$$\text{MSD}(n) = 10 \log_{10}(\|\mathbf{w}(n) - \mathbf{w}_0\|^2). \quad (40)$$

To thoroughly assess the algorithm's robustness, the following three distinct and representative models for complex-domain non-Gaussian noise are employed:

1. **Impulsive Mixture Noise with Bernoulli–Gaussian Structure:** The composite noise $z(n)$ is defined as $z(n) = A(n)z_1(n) + B(n)z_2(n)$. Here, $z_1(n) \sim \mathcal{CN}(0, \sigma_g^2)$ denotes the background complex Gaussian noise of variance σ_g^2 . The impulsive element is represented by $z_2(n) \sim \mathcal{CN}(0, \sigma_i^2 / \text{Pr})$, having a larger variance σ_i^2 / Pr . A Bernoulli random variable $A(n)$, distributed as Bernoulli($1 - \text{Pr}$) with $\text{Pr} = 0.001$, governs the occurrence of impulses: $A(n) = 1$ signifies no impulse, while its complement $B(n) = 1 - A(n)$ indicates an impulse event. All variables $A(n)$, $z_1(n)$, and $z_2(n)$ are mutually independent.
2. **Contaminated Complex Gaussian Noise (CG Noise):** This noise model is synthesized by summing two independent zero-mean complex Gaussian components. The background component has a variance of 0.008 for the real part and 0.002 for the imaginary part. An intermittent impulsive component, with a variance of 8 for the real part and 2 for the imaginary part, is superimposed with an occurrence probability of 0.01. The combined signal exhibits a heavy-tailed distribution characteristic due to the occasional high-power impulses.
3. **Symmetric Complex α -Stable Noise (α -Stable Noise):** This model adopts a complex-valued symmetric α -stable distribution, formulated by $z(n) = z_R(n) + jz_I(n)$. The real component $z_R(n)$ and imaginary component $z_I(n)$ are generated as independent and identically distributed real α -stable random variables, described by the distribution $z_\alpha(\alpha_z, \beta_z, \nu, \xi)$. The configuration employs the following: characteristic exponent $\alpha_z = 1.9$, symmetry parameter $\beta_z = 0$, scale parameter $\nu = 1$, and location parameter $\xi = 0$.

Figures 5–7 illustrate the MSD convergence curves of various algorithms under three typical non-Gaussian noise environments. As can be seen, under the adverse conditions of strongly correlated input signals accompanied by non-Gaussian noise, the conventional CLMS algorithm exhibits a clear divergent trend. In contrast, the CNLMS algorithm effectively overcomes the divergence issue of LMS caused by step-size mismatch in non-stationary or impulsive signals through input signal power normalization of the step size, thereby maintaining system stability. However, the CAP algorithm suffers from a significantly degraded convergence performance under the influence of non-Gaussian noise. Furthermore, even when compared with robust algorithms specifically designed to suppress impulsive interference, the proposed α -CSAP algorithm still achieves the lowest steady-state MSD across all three noise scenarios. This is primarily attributed to the synergistic design of the α -sigmoid cost function and the ℓ_2 -norm constraint adopted in the α -CSAP algorithm, which endows it with stronger anti-interference capability. In summary, the comparative results demonstrate that the proposed α -CSAP algorithm exhibits superior robustness and a more accurate steady-state estimation performance when handling strongly correlated input signals and complex non-Gaussian noise.

4.2. Beamforming Application

This section evaluates the beamforming performance of the proposed α -CSAP algorithm in impulsive noise environments. A hexagonal sensor array composed of 91 elements is employed, with an array radius set to three times the operating wavelength and an

inter-element spacing of approximately half a wavelength. The operating frequency is 8 GHz, corresponding to a free-space wavelength of $\lambda = 37.5$ mm. Five independent narrowband signals with Quadrature Phase Shift Keying (QPSK) modulation are simulated, including one desired signal and four strong interferences. The direction of arrival (DOA) of the desired signal is at an azimuth of 90° and an elevation of 45° . The DOAs of the four interfering signals are $(20^\circ, 45^\circ)$, $(62^\circ, 45^\circ)$, $(120^\circ, 45^\circ)$, and $(145^\circ, 45^\circ)$, respectively. Herein, the azimuth angle θ is defined clockwise with respect to the array normal, and the elevation angle φ denotes the angle between the incident signal and the array plane.

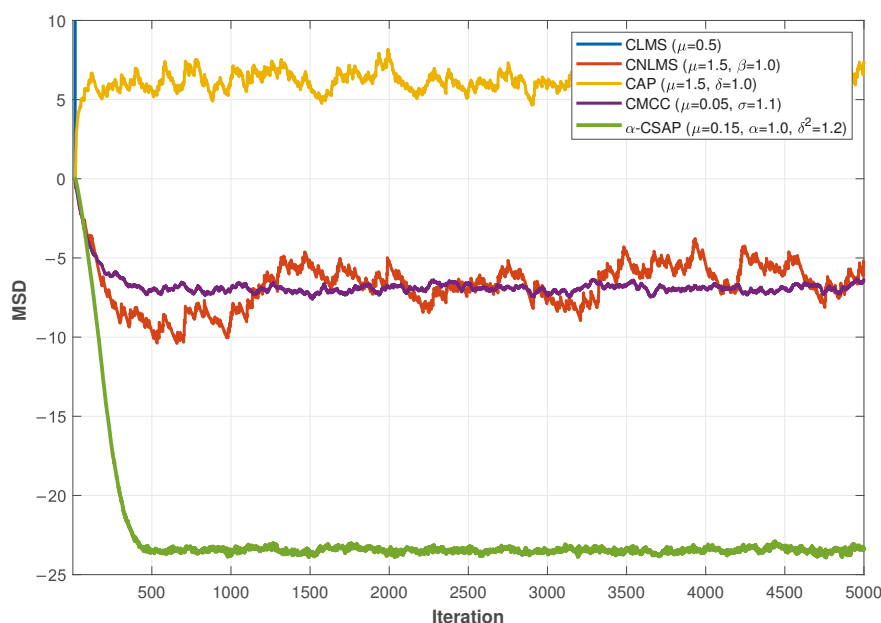


Figure 5. The MSD convergence curves of different algorithms under Bernoulli–Gaussian Impulsive Mixture Noise.

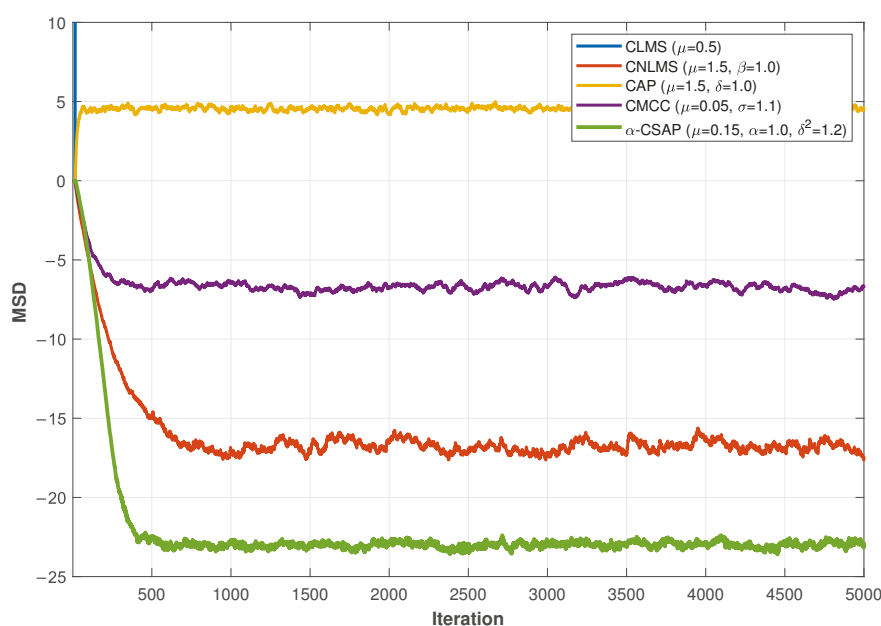


Figure 6. The MSD convergence curves of different algorithms under CG Noise.

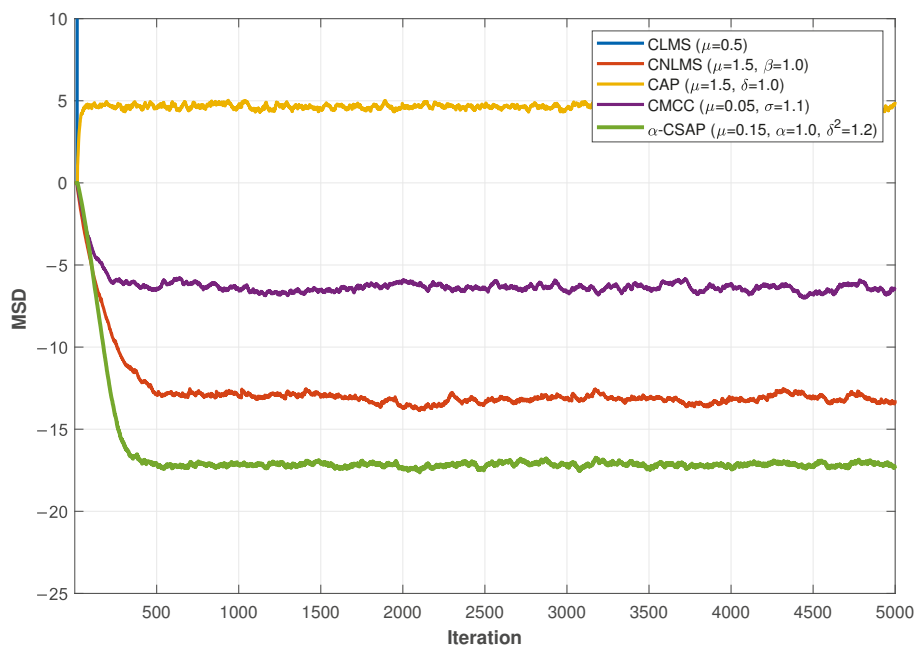


Figure 7. The MSD convergence curves of different algorithms under α -Stable Noise.

To simulate a realistic strong interference environment, high-intensity impulsive noise is introduced, which is modeled by a Bernoulli–Gaussian mixture distribution. The impulse occurrence probability is set to 0.05, and the variances of the corresponding Gaussian components are $\sigma_1^2 = 2$ and $\sigma_2^2 = 40$, respectively. The Signal-to-Noise Ratio (SNR) and Interference-to-Noise Ratio (INR) are set to 20 dB and 40 dB, respectively.

The core parameters of the α -CSAP algorithm are systematically tuned, as follows: the adaptive filter length L_w equals the number of array elements (i.e., 91), the projection order is $Q = 4$, and the step size is set to $\mu = 0.0001$. Furthermore, the characteristic exponent α and the scale parameter δ of the algorithm are configured as 1.5 and 16, respectively. This configuration significantly enhances the stability and robustness of the algorithm under impulsive noise. For a fair comparison, the step sizes of the conventional algorithms are also adjusted accordingly, as follows: $\mu_{\text{CLMS}} = 0.001$ for the CLMS algorithm, $\mu_{\text{CAP}} = 0.0005$ for the CAP algorithm, and $\mu_{\text{CMCC}} = 0.0003$ for the CMCC algorithm. This ensures that all algorithms exhibit similar initial convergence speeds under identical simulation conditions.

As shown in the beam pattern results in Figure 8, while maintaining the same mainlobe gain, the α -CSAP algorithm achieves deeper nulls at the four interference directions of 20° , 62° , 120° , and 145° , with null depths exceeding those of the other compared algorithms.

We consider the signal-to-interference-plus-noise ratio (SINR) as a metric to evaluate the performance of the beamforming system, which represents the ratio of the desired signal power to the interference-plus-noise power. As shown in Table 3, under the harsh condition of an input SINR as low as -26.30 dB, the α -CSAP algorithm elevates the output SINR to 34.72 dB, achieving a remarkable improvement of 61.02 dB. This performance enhancement significantly surpasses that of the other compared algorithms. These results fully demonstrate the superior capability of the α -CSAP algorithm in interference suppression and overall beamforming performance.

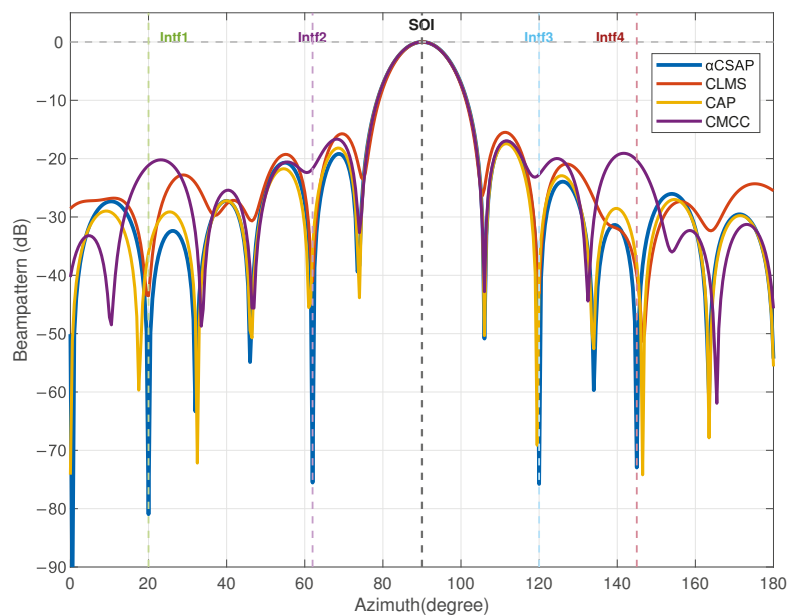


Figure 8. Comparison of beampatterns: CLMS, CAP, CMCC, and α -CSAP.

Table 3. Beampattern gain values and SINR performance comparison in impulsive noise environment.

Algorithm	Final SINR (dB)	SINR Improvement (dB)	Interference Direction (dB)			
			20°	62°	120°	145°
α -CSAP	34.72	+61.02	−80.95	−75.53	−75.76	−72.94
CLMS	18.72	+45.02	−43.52	−41.05	−43.10	−39.92
CAP	11.44	+37.74	−36.14	−35.33	−39.49	−38.04
CMCC	12.27	+38.57	−21.15	−21.73	−22.72	−20.28

Input SINR = −26.30 dB. The beampattern gain in the desired direction (90°) is normalized to 0.00 dB.

5. Conclusions

This paper proposes a robust complex α -CSAP suitable for non-Gaussian noise environments. By extending the affine projection algorithm based on the α -Sigmoid cost function to the complex domain, the algorithm enhances robustness against non-Gaussian noise. The computational complexity is reduced by employing a normalization factor to avoid matrix inversion. In two typical application scenarios, system identification and beamforming, the proposed α -CSAP algorithm demonstrates superior convergence characteristics and lower steady-state error compared to classical complex adaptive filtering algorithms such as CLMS, CAP, and CMCC. The results validate its effectiveness and engineering potential in complex-domain signal processing.

Author Contributions: Y.G.: conceptualization, methodology, validation, writing—original draft preparation, writing—review and editing; B.G.: writing—review and editing; G.Q.: writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

α -CSAP	Complex α -Sigmoid Affine Projection
MSD	Mean Square Deviation
LMS	Least Mean Square
RLS	Recursive Least Squares
MSE	Mean Square Error
AP	Affine Projection
CLMS	Complex Least Mean Square
CNLMs	Complex Norm Least Mean Squares
CAP	Complex Affine Projection
CMCC	Complex Maximum Correntropy Criterion
CG Noise	Contaminated Complex Gaussian Noise
α -Stable Noise	Symmetric Complex α -Stable Noise
QPSK	Quadrature Phase Shift Keying
DOA	Direction Of Arrival
INR	Interference-to-Noise Ratio
SNR	Signal-to-Noise Ratio
INR	Interference-to-Noise Ratio
SINR	Signal-to-Interference-plus-Noise Ratio

References

1. Hassani, S.A.; van Liempd, B.; Bourdoux, A.; Horlin, F.; Pollin, S. Joint In-Band Full-Duplex Communication and Radar Processing. *IEEE Syst. J.* **2022**, *16*, 3391–3399. [CrossRef]
2. Qu, X.; Ma, Z.; Zhang, H.; Sun, X.; Yang, X. Target Tracking Method Based on Scale-Adaptive Rotation Kernelized Correlation Filter for Through-the-Wall Radar. *IEEE Signal Process. Lett.* **2025**, *32*, 1001–1005. [CrossRef]
3. Zhong, L.; Li, Y.; Cheng, W.; Zhou, W. Robust Cognitive Radar Tracking Based on Adaptive Unscented Kalman Filter in Uncertain Environments. *IEEE Access* **2020**, *8*, 163405–163418. [CrossRef]
4. Zheng, S.; Jiang, L.; Yang, Q.; Zhao, Y.; Wang, Z. Adaptive PHD Filter with RCS and Doppler Feature for Space Targets Tracking via Space-Based Radar. *IEEE Trans. Aerosp. Electron. Syst.* **2024**, *60*, 3750–3765. [CrossRef]
5. Chen, J.-S.; Dao, D.-T.; Chien, H. Ship Echo Identification Based on Norm-Constrained Adaptive Beamforming for an Arrayed High-Frequency Coastal Radar. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 1143–1153. [CrossRef]
6. Bahramgiri, M.; Nooshabadi, S.; Olutomilayo, K.T.; Fuhrmann, D.R. Automotive Radar-Based Hitch Angle Tracking Technique for Trailer Backup Assistant Systems. *IEEE Trans. Intell. Veh.* **2023**, *8*, 1922–1933. [CrossRef]
7. Wen, B.; Lu, Z.; Mao, Y.; Zhou, B. Marine Radar Image Sequence Target Detection Based on Space–Time Adaptive Filtering and Hough Transform. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 13506–13522. [CrossRef]
8. Li, H.; Song, H.; Wang, R.; Wang, H.; Liu, G.; Chen, R.; Li, X.; Deng, Y.; Balz, T. A Modification to the Complex-Valued MRF Modeling Filter of Interferometric SAR Phase. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 681–685.
9. Chen, X.; Yu, X.; Huang, Y.; Guan, J. Adaptive Clutter Suppression and Detection Algorithm for Radar Maneuvering Target with High-Order Motions Via Sparse Fractional Ambiguity Function. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 1515–1526. [CrossRef]
10. Jin, T.; Wang, B.; Lou, Y.; Zhao, Y.; Qi, B. Enhanced Long Baseline Underwater Target Localization with Adaptive Track-Before-Detect Method. *IEEE Signal Process. Lett.* **2024**, *31*, 1710–1714. [CrossRef]
11. Rohal, P.; Ochodnický, J. Particle Filter Optimization for Adaptive Radar Data Processing. In Proceedings of the 2020 New Trends in Signal Processing (NTSP), Demanovska Dolina, Slovakia, 14–16 October 2020; pp. 1–4.
12. Xing, G.; Zhang, Y. Analysis and Comparison of RLS Adaptive Filter in Signal De-noising. In Proceedings of the 2011 International Conference on Electrical and Control Engineering, Yichang, China, 16–18 September 2011; pp. 5754–5758.
13. Took, C.C.; Mandic, D.P. The Quaternion LMS Algorithm for Adaptive Filtering of Hypercomplex Processes. *IEEE Trans. Signal Process.* **2009**, *57*, 1316–1327. [CrossRef]
14. Mandic, D.; Goh, V. *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*; Wiley: Hoboken, NJ, USA, 2009.

15. Qiu, C.; Qian, G.; Wang, S. Widely Linear Maximum Complex Correntropy Criterion Affine Projection Algorithm and Its Performance Analysis. *IEEE Trans. Signal Process.* **2022**, *70*, 3540–3550. [CrossRef]
16. Dong, F.; Qian, G.; Wang, S. Bias-Compensated MCCC Algorithm for Widely Linear Adaptive Filtering with Noisy Data. *IEEE Trans. Circuits Syst. II Express Briefs* **2020**, *67*, 3587–3591. [CrossRef]
17. Qian, G.; Mei, J.; Iu, H.H.C.; Wang, S. Fixed-Point Maximum Total Complex Correntropy Algorithm for Adaptive Filter. *IEEE Trans. Signal Process.* **2021**, *69*, 2188–2202. [CrossRef]
18. Qian, G.; Ning, X.; Wang, S. Recursive Constrained Maximum Correntropy Criterion Algorithm for Adaptive Filtering. *IEEE Trans. Circuits Syst. II Express Briefs* **2020**, *67*, 2229–2233. [CrossRef]
19. Iskandarani, M.Z. Investigation of Energy Consumption in WSNs Within Enclosed Spaces Using Beamforming and LMS (BF-LMS). *IEEE Access* **2024**, *12*, 63932–63941. [CrossRef]
20. Shimada, K.; Bando, Y.; Mimura, M.; Itoyama, K.; Yoshii, K.; Kawahara, T. Unsupervised Speech Enhancement Based on Multichannel NMF-Informed Beamforming for Noise-Robust Automatic Speech Recognition. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2019**, *27*, 960–971. [CrossRef]
21. Liu, Y.-C.; Su, H.-J.; Takano, Y. Underdetermined AOA Estimation Using Non-Uniform Sub-Connected Hybrid Beamforming Systems. *IEEE Open J. Commun. Soc.* **2025**, *6*, 1485–1496. [CrossRef]
22. Huang, F.; Zhang, J.; Zhang, S. Affine Projection Versoria Algorithm for Robust Adaptive Echo Cancellation in Hands-Free Voice Communications. *IEEE Trans. Veh. Technol.* **2018**, *67*, 11924–11935. [CrossRef]
23. Chien, Y.-R.; Wu, S.-T.; Tsao, H.-W.; Diniz, P.S.R. Correntropy-Based Data Selective Adaptive Filtering. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2024**, *71*, 754–766. [CrossRef]
24. Ozeki, K.; Umeda, T. An Adaptive Filtering Algorithm Using an Orthogonal Projection to an Affine Subspace and Its Properties. *Electron. Commun. Jpn.* **1984**, *67*, 19–27. [CrossRef]
25. Xu, W.J.; Zhao, H.Q.; Lv, S.H. Robust Constrained Affine-Projection-Like Adaptive Filtering Algorithms Using the Modified Huber Function. *IEEE Trans. Circuits Syst. II Express Briefs* **2023**, *70*, 1214–1218. [CrossRef]
26. Rey Vega, L.; Rey, H.; Benesty, J. A New Robust Variable Step-Size NLMS Algorithm. *IEEE Trans. Signal Process.* **2008**, *56*, 1878–1892. [CrossRef]
27. Gu, Y.; Jin, J.; Mei, S. L_0 Norm Constraint LMS Algorithm for Sparse System Identification. *IEEE Signal Process. Lett.* **2009**, *16*, 774–777.
28. Bhotto, M.Z.A.; Ahmad, M.O.; Swamy, M.N.S. Robust Shrinkage Affine-Projection Sign Adaptive-Filtering Algorithms for Impulsive Noise Environments. *IEEE Trans. Signal Process.* **2014**, *62*, 3349–3359. [CrossRef]
29. Lima, M.V.S.; Martins, W.A.; Diniz, P.S.R. Affine Projection Algorithms for Sparse System Identification. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Vancouver, BC, Canada, 26–31 May 2013; pp. 5666–5669.
30. Bhotto, M.Z.A.; Antoniou, A. Affine-Projection-Like Adaptive-Filtering Algorithms Using Gradient-Based Step Size. *IEEE Trans. Circuits Syst. I Reg. Pap.* **2014**, *61*, 2048–2056. [CrossRef]
31. Paleologu, C.; Benesty, J.; Ciochina, S. A Variable Step-Size Affine Projection Algorithm Designed for Acoustic Echo Cancellation. *IEEE Trans. Audio Speech Lang. Process.* **2008**, *16*, 1466–1478. [CrossRef]
32. Li, Y.; Fu, Y.; Li, L.; Yu, C.; Miao, Y.; Zakharov, Y.; Diniz, P.S. An α -Sigmoid Affine Projection Algorithm for In-Car Echo Cancellation and Channel Estimation. *IEEE Trans. Veh. Technol.* **2025**, *74*, 10297–10305. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Radar HRRP Sequence Target Recognition Based on a Lightweight Spatiotemporal Fusion Network

Xiang Li ¹, Yitao Su ², Xiaobin Zhao ², Junjun Yin ^{2,*} and Jian Yang ¹¹ Department of Electronic Engineering, Tsinghua University, Beijing 100084, China² School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China

* Correspondence: yinjj07@gmail.com; Tel.: +86-132-6001-5764

Abstract

High-resolution range profile (HRRP) sequence recognition in radar automatic target recognition faces several practical challenges, including severe category imbalance, degradation of robustness under complex and variable operating conditions, and strict requirements for lightweight models suitable for real-time deployment on resource-limited platforms. To address these problems, this paper proposes a lightweight spatiotemporal fusion-based (LSTF) HRRP sequence target recognition method. First, a lightweight Transformer encoder based on group linear transformations (TGLT) is designed to effectively model temporal dynamics while significantly reducing parameter size and computation, making it suitable for edge-device applications. Second, a transform-domain spatial feature extraction network is introduced, combining the fractional Fourier transform with an enhanced squeeze-and-excitation fully convolutional network (FSCN). This design fully exploits multi-domain spatial information and enhances class separability by leveraging discriminative scattering-energy distributions at specific fractional orders. Finally, an adaptive focal loss with label smoothing (AFL-LS) is constructed to dynamically adjust class weights for improved performance on long-tail classes, while label smoothing alleviates overfitting and enhances generalization. Experiments on the MSTAR and CVDomes datasets demonstrate that the proposed method consistently outperforms existing baseline approaches across three representative scenarios.

Keywords: HRRP sequence recognition; lightweight spatiotemporal fusion; transform domain; adaptive focal loss

1. Introduction

Radar technology has consistently played a pivotal role in both military and civilian domains. With the rapid development of modern radar technology, radar imaging technology has gone way beyond traditional detection limits. High-resolution techniques, primarily achieved through wideband signal transmission and pulse compression, enable the acquisition of detailed target signatures such as the high-resolution range profile (HRRP), making fine-grained target recognition by radar possible. Consequently, radar automatic target recognition (RATR) has become an important task in radar applications. The core principle of this technology involves extracting and analyzing features from radar measurements using advanced signal processing algorithms and pattern recognition methods to achieve automatic target classification and identification. HRRP, derived from wideband radar systems, represents the vector sum of target scattering echoes projected

along the radar's radial direction, containing crucial discriminative information such as target shapes, structural dimensions, and scattering center distributions [1]. HRRP data exhibit unique advantages due to their ease of acquisition, convenient processing, efficient storage, and rich structural information [2]. These distinct characteristics make HRRP an important technical support for RATR, offering reliable solutions for efficient and precise target identification.

From a methodological perspective, HRRP-based RATR methods can primarily be categorized into three groups: (1) feature extraction-based methods that identify robust features from HRRP signals and leverage designed classifiers for target differentiation, (2) statistical model-based methods that construct probability distribution models using training data and classify target categories by calculating the likelihood probabilities of test samples, and (3) deep learning-based methods that enable end-to-end recognition through neural networks for autonomous feature extraction [3].

In feature extraction-based methods, the recognition performance is critically influenced by the effectiveness and accuracy of feature selection. The core challenge lies in constructing feature spaces with translation invariance and intrinsic target separability, the design of which heavily relies on empirical expertise. Existing research has achieved preliminary progress: Studies [4–6] demonstrated feasibility in target recognition tasks by selecting physically meaningful features with strong representational capabilities from HRRP data. Meanwhile, through the integration of signal processing and pattern recognition theories, these methods have evolved into multidimensional frameworks incorporating innovative approaches such as time-frequency analysis [7,8], sparse representation [9,10], and nonlinear transformations [11,12]. Although such methods exhibit theoretical advantages in characterizing target scattering properties, significant technical limitations persist. Crucially, such approaches exhibit limited generalizability in multi-target recognition contexts, constrained by category-specific feature engineering and vulnerability to long-tailed distributions characterized by dominant majority classes and severely scarce minority categories that induce biased learning.

Statistical model-based HRRP target recognition methods characterize the distribution patterns of target scattering properties through probabilistic modeling. The core principle involves treating radar echoes as stochastic processes and establishing probabilistic mapping relationships between target types and echo data via statistical inference. Representative statistical models include adaptive Gaussian classifiers, Gamma mixture models, Gamma–Gaussian mixture models, and factor analysis (FA) models. Du et al. systematically developed statistical modeling approaches to address recognition robustness in noisy environments [13–15]. Notably, since the introduction of FA models into HRRP statistical recognition in 2008, significant research advancements have been achieved in this domain [16–18]. These methods fundamentally depend on prior distribution assumptions. This intrinsic constraint creates a theoretical mismatch with HRRP's inherent non-Gaussian characteristics and multimodality, ultimately limiting complex scattering pattern characterization.

Deep learning approaches leverage HRRP sequences' inherent spatiotemporal characteristics, where temporal dependencies from azimuth variations and spatial correlations across range cells provide complementary discriminative information. Initial efforts focused on temporal modeling through recurrent architectures, achieving notable success in aircraft target recognition [2,19,20]. While existing studies have achieved progress in target recognition through temporal feature modeling with recurrent neural networks, their frameworks exhibit notable limitations by neglecting spatial correlation characteristics within HRRP sequences. Consequently, the effective integration of spatiotemporal joint features in HRRP has emerged as a critical research focus in this field. Wan et al. [21]

developed a hybrid recognition framework integrating convolutional neural networks (CNNs) with bidirectional recurrent neural networks (BiRNNs), where CNN modules extract spatial correlation features from HRRP data, while BiRNN modules capture temporal dependencies across range cells. Wang et al. [22] innovatively combined CNNs with Bidirectional Encoder Representations from Transformers (BERT), employing convolutional modules to characterize local spatial structures and utilizing BERT's multi-head attention mechanisms to extract temporal features from HRRP sequences. Wu et al. [23] advanced this approach by constructing a fusion network that initially pre-extracts features through BERT, subsequently employs multi-scale CNNs and bidirectional gated recurrent networks to capture local characteristics and long-range dependencies, respectively, and finally concatenates features to integrate advantages from diverse networks. Establishing effective spatiotemporal feature extraction mechanisms to achieve complementary fusion of HRRP's spatiotemporal characteristics and extending these to multi-target recognition scenarios remains a highly valuable research frontier.

With the rapid advancement of deep learning and growing academic interest in Transformer methodologies, recent years have witnessed increasing exploration of their applications in HRRP recognition. Zhang et al. [24] developed a feature-guided Transformer model that integrates manually designed features into attention modules to focus on HRRP range cells with rich scattering information, significantly improving recognition accuracy under small-sample conditions. Diao et al. [25] proposed a positional embedding-free CNN–Transformer hybrid architecture optimized for HRRP data characteristics, effectively eliminating the need for traditional positional encoding. Wang et al. [26] employed dual-branch Transformer encoders to separately extract temporal and spatial features, with designed attention fusion mechanisms achieving adaptive feature weighting. Gao et al. [27] innovatively introduced polarization preprocessing modules that combine artificial features with CNNs to enhance feature representation, constructing a Vision Transformer-based framework that substantially improves local and temporal feature extraction. Although current studies improve performance through module stacking, they commonly suffer from overfitting risks caused by excessive parameters. Insufficient emphasis on model lightweighting poses challenges in practical deployment on edge devices.

Current HRRP radar target recognition research primarily focuses on two typical data environments: (1) recognition methodologies under class-balanced conditions, primarily addressing cooperative target identification tasks. Under such conditions, controllable detection conditions enable researchers to acquire abundant high-quality samples, thereby facilitating the adoption of complex models like deep neural networks to leverage their powerful feature extraction capabilities for high-precision recognition. (2) Recognition technology exploration under data-limited and class-imbalanced conditions, mainly targeting non-cooperative target identification. Due to factors such as long detection ranges and complex electromagnetic environments, effective sample acquisition for such targets proves challenging, resulting in datasets exhibiting typical long-tailed distribution characteristics. It is noteworthy that in practical applications such as national defense and aerospace, the identification of non-cooperative targets with high observational value presents more urgent demands. These operational environments face three core challenges: First, target echoes are susceptible to noise interference, leading to significantly reduced signal-to-noise ratios; Second, there exists a prominent contradiction between target category diversity and limited training samples; Third, measured data inherently exhibit long-tailed distribution properties. Establishing effective recognition models under multi-category imbalanced data conditions has emerged as a critical scientific challenge in advancing the practical application of radar automatic recognition technology.

Researchers are actively exploring diverse technical approaches to address class imbalance issues in HRRP target recognition. Yin et al. [28] proposed an Adaptive Uniform Manifold Approximation and Projection (AUMAP) segmentation algorithm, which mitigates data imbalance by modifying the CNN loss function into a focal loss formulation. Jia et al. [29] developed a memory-based neural network (MBNN) for imbalanced data conditions: CNN-extracted features are processed through a memory module that records misclassified and low-confidence samples, followed by long short-term memory (LSTM) based fusion of classified samples with buffer-stored similar samples for final decision-making. Zhang et al. [30] introduced an open-set imbalanced recognition network integrating dual-attention mechanisms, memory modules, across functions, and decoupled training strategies, optimizing intra-class and inter-class similarity constraints via angular penalty loss. Guo et al. [31] established a transfer learning framework involving pretrained models with source domain data, subsequently resetting fully-connected and output layer parameters, and proposed a novel loss function to suppress inter-class bias for measured HRRP datasets with limited samples and class imbalance. Wu et al. [32] created a weighted synthetic minority oversampling technique (SMOTE) algorithm that dynamically allocates synthetic weights based on Euclidean distances among minority samples, combined with multi-scale CNN–Transformer encoder attention models to enhance multi-level feature classification accuracy. Tian et al. [33] designed a gradient-guided class re-balancing loss (GRB Loss) for space micro-motion targets, which dynamically assigns weights according to accumulated positive-negative gradient ratios at classification nodes, ensuring recognition robustness across varying imbalance ratios.

As a core technology of RATR, HRRP sequence-based target recognition plays a critical role in determining battlefield perception capabilities within complex environments. Nevertheless, existing methodologies exhibit significant technical limitations in feature synergy utilization, computational efficiency, and scenario generalization. Primarily, inadequate collaborative modeling of spatiotemporal features restricts cross-dataset adaptability. While HRRP sequences inherently contain temporal pose evolution characteristics and spatial scattering structural information, current approaches predominantly focus on temporal feature extraction via recurrent neural networks or spatial correlation mining through convolutional architectures, resulting in inadequate collaborative representation of spatiotemporal features. This deficiency leads to drastic performance degradation in complex battlefield environments involving variant targets or low signal-to-noise conditions, revealing fundamental weaknesses in generalization robustness. Secondly, structural redundancy in models elevates overfitting risks. Contemporary mainstream approaches predominantly adopt complex module stacking strategies, such as deep Transformer encoders and multi-branch convolutional architectures, to enhance recognition performance. This architectural complexity results in exponential escalation of both parameter volumes and computational demands. However, the inherent contradiction between embedded radar devices' limited computing power and resource-intensive model requirements not only hinders deployment feasibility but also exacerbates overfitting risks due to insufficient training data, severely compromising practical utility. Finally, biased data distribution exacerbates generalization deterioration. Current research predominantly concentrates on class-balanced datasets, with insufficient investigation into multi-class long-tailed distribution challenges. Traditional cross-entropy loss functions, which apply uniform weighting to all samples, induce excessive model adaptation to majority classes while neglecting discriminative feature learning for high-value tail categories. Consequently, this imbalance triggers catastrophic performance degradation in critical minority class recognition.

To address the above technical challenges in HRRP sequence target recognition, this paper proposes a lightweight spatiotemporal fusion-based (LSTF) HRRP sequence target recognition method. The main innovations are summarized as follows:

1. **LSTF-based HRRP sequence target recognition framework:**

We propose an LSTF HRRP sequence target recognition method with a dual-branch feature extraction architecture, in which a temporal feature encoding module and a spatial fully convolutional module are jointly fused to capture both target pose evolution and scattering structure distribution.

2. **Lightweight Transformer encoder for temporal modeling:**

A Transformer encoder based on group linear transformations (TGLT) is designed to effectively model temporal dynamics in HRRP sequences while significantly reducing parameter count and computational complexity, enabling real-time deployment on edge devices.

3. **Transform-domain spatial feature extraction network:**

We develop a transform-domain spatial feature extraction network that integrates the fractional Fourier transform (FrFT) with an enhanced squeeze-and-excitation fully convolutional network (FSCN). By exploiting multi-domain spatial representations, the proposed network enhances the discriminability of scattering energy distributions across target classes at specific fractional orders, leading to improved classification performance.

4. **Adaptive loss for long-tail recognition:**

An adaptive focal loss with label smoothing (AFL-LS) is introduced to dynamically reweight imbalanced classes, improving recognition of long-tail categories. The incorporation of label smoothing further mitigates overfitting and significantly enhances the model's generalization ability.

The remainder of this paper is organized as follows. Section 2 details the proposed LSTF network architecture, including the overall framework, the TGLT-based temporal feature encoder module, the transform-domain spatial feature fully convolutional module, and the decision fusion and recognition module. Section 3 introduces the proposed AFL-LS function, designed to address class imbalance. Section 4 presents the experimental results and analysis, including dataset descriptions, recognition performance comparisons, ablation studies, and feature visualizations. Finally, Section 5 concludes the paper.

Notably, the proposed method also exhibits strong potential for extension to target detection tasks in high-frequency surface wave radar (HFSWR) systems, particularly in high-resolution maritime surveillance scenarios. For ship detection and discrimination from sea clutter on range-Doppler maps, the lightweight TGLT-based encoder can effectively capture the temporal dynamics of ship echoes, while the transform-domain FSCN module enhances feature separability by exploiting scattering energy distributions in the fractional Fourier domain. Moreover, the AFL-LS loss mitigates the severe imbalance between sparse ship targets and dense sea clutter, thereby improving detection robustness. Benefiting from its lightweight architecture and low computational overhead, the proposed framework is well aligned with the real-time processing requirements of operational HFSWR systems. These characteristics indicate that the proposed method is not limited to HRRP sequence recognition but also holds broader practical value for cross-domain radar target analysis.

2. Method

2.1. Overall Network Architecture

The architecture of the proposed recognition method is shown in Figure 1, featuring a dual-branch structure with three main components: the temporal feature encoder, the spatial feature convolutional module, and the decision fusion module. In the temporal branch, the TGLT efficiently models temporal dynamics by capturing dependencies across different time steps through a self-attention mechanism, while reducing parameters and computational cost for real-time edge deployment. In the spatial branch, a transform-domain feature extraction network combining FrFT with an enhanced FSCN exploits multi-domain spatial information to enhance feature discriminability. Specifically, the FrFT projects the data into a joint time-frequency domain to enhance separability, and the subsequent FSCN (comprising Conv1D, SE blocks, and pooling) extracts multi-scale local structural features. Finally, the decision fusion module employs an AFL-LS to dynamically adjust class weights for improved long-tail recognition and applies label smoothing to boost generalization.

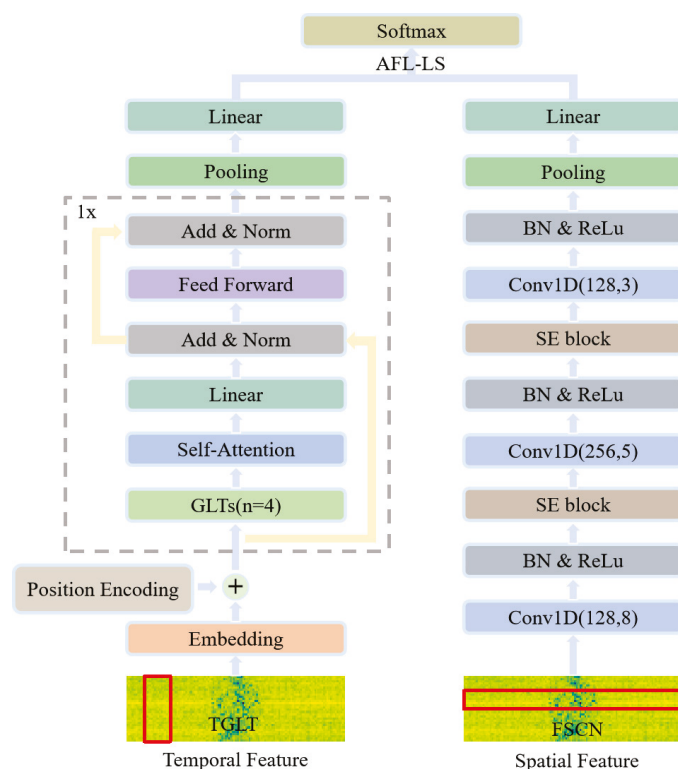


Figure 1. LSTF network structure.

We assume the input HRRP sequence is $X = [x_1 \ x_2 \ \cdots \ x_n]^T$, where the sequence has dimensions $n \times c$, with n representing the length of the HRRP sequence and c representing the channel dimension of the HRRP sequence. For the temporal branch, a high-dimensional linear mapping is first applied to map the channel dimension of the sequence from c dimensions to d_{model} dimensions as

$$Y_{\text{token}} = XW_e \quad (1)$$

where Y_{token} denotes the sequence after linear transformation, with the channel dimension being d_{model} , and $W_e \in R^{c \times d_{\text{model}}}$ is the linear transformation matrix. Positional encoding is added to Y as

$$Y = Y_{\text{token}} + X_p \quad (2)$$

where $X_p \in R^{n \times d_{model}}$ represents the positional encoding matrix. The temporal feature encoder then processes Y using a Transformer encoder to extract features, resulting in the temporal feature O_T . The spatial branch takes X^T as input and extracts features through multiple stacked layers of one-dimensional convolution, batch normalization, squeeze-and-excitation, and pooling, ultimately yielding the spatial feature O_S .

The decision fusion recognition module uses O_T and O_S as inputs and performs classification using a fully connected layer and a Softmax function for each branch, thereby obtaining the classification results P_T for the temporal branch and P_S for the spatial branch. To achieve adaptive fusion of the temporal and spatial branches, we establish learnable decision weight matrices W_T and W_S . The final recognition result P is computed as

$$P = \text{Softmax}(W_T P_T + W_S P_S) \quad (3)$$

where $W_T, W_S \in R^{1 \times t}$ are the learnable decision weight matrices, and t represents the number of target categories. These matrices are model parameters optimized during training. Each vector contains t scalar weight values, corresponding to the t target classes. Their function is to perform an adaptive, category-specific fusion. The Softmax function is then used to constrain the output values of the final recognition result.

Weather factors such as rain, fog, and snow primarily affect HRRP data by reducing signal-to-noise ratio (SNR) and distorting scattering center features, and our proposed LSTF method addresses these challenges through its core designs: the TGLT-based lightweight Transformer encoder maintains stable temporal feature extraction under low SNR, the FrFT-integrated FSCN module decouples weather-induced noise from spatial structural information, and the AFL-LS loss function dynamically adjusts weights to focus on reliable features distorted by weather. Experimental results on low-SNR datasets (Dataset 3) show that the method achieves over 93% accuracy at 15 dB SNR (simulating light rain/fog) and 79.96% accuracy at 5 dB SNR (simulating heavy rain/snow), outperforming baselines by 4–7% under harsh conditions. Overall, the method exhibits robust adaptability across typical weather scenarios, maintaining high recognition performance in mild weather and retaining usability under extreme conditions, which is crucial for practical radar deployment.

2.2. Temporal Feature Encoder Module

The encoder is a critical component of the Transformer architecture. For the standard Transformer encoder, the requirement to process global information from the input data results in high computational complexity, which leads to a large number of model parameters and extended training times. These characteristics are not conducive to deployment on edge devices. In the method proposed in this paper, we utilize group linear transformations (GLTs) to lighten the encoder, as illustrated in Figure 2. GLTs leverage the fact that HRRP sequences exhibit localized temporal dependencies and redundant interchannel correlations. By partitioning input channels into groups, GLTs introduce structured sparsity that aligns with the low-dimensional nature of radar scattering signatures. This design reduces model capacity while preserving key temporal dynamics, effectively mitigating overfitting with limited training data. From a matrix approximation perspective, GLTs act as a block-diagonal low-rank approximation of a full linear projection, retaining sufficient expressiveness for pose variations while significantly reducing computational complexity. This approach significantly reduces the computational load and training time by simplifying the feature transformation and attention operations within the encoder. Specifically, the temporal feature encoder is built with a group linear transformation layer that performs efficient dimension expansion and reduction, followed by a self-attention layer to capture long-range temporal dependencies. A dropout layer and residual connections are included to stabilize optimization and prevent overfitting, while a lightweight feed-forward net-

work further decreases the parameter size through compressed intermediate dimensions. Finally, a pooling layer aggregates the temporal features into a compact representation for subsequent recognition.

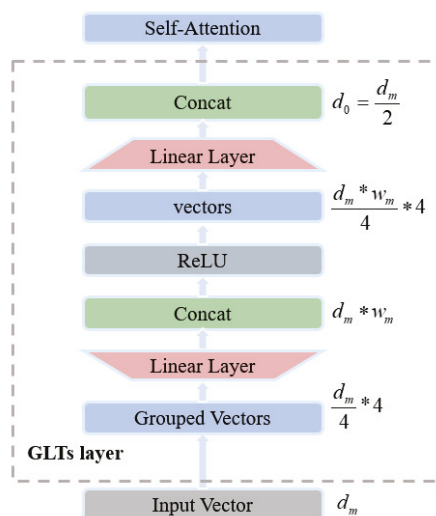


Figure 2. Illustration of the structure of the group linear transformations layer.

The group linear transformation layer consists of two operations: dimensionality expansion and reduction. Initially, the input data are divided into multiple groups, each of which is mapped to a higher-dimensional space and expanded. By concatenating these vectors, high-dimensional features that are more discriminative are obtained. After processing these features with the rectified linear unit (ReLU) function, they are reduced in dimensionality within each group, and the resulting low-dimensional features are concatenated to serve as the input to the self-attention layer. This two-stage transformation effectively integrates features across different dimensions, yielding better performance than simple linear encoding. Consequently, replacing the multi-head attention mechanism in the standard Transformer with self-attention and reducing the input dimensionality to half that of the original significantly decreases the computational load.

The self-attention layer calculates the relevance of sequence data using query and key vectors to derive the attention matrix, which captures the global dependencies within the sequence. The resulting attention features are expressed as

$$A = \text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (4)$$

where the query vector $Q = YW_Q$, the key vector $K = YW_K$, and the value vector $V = YW_V$ are defined, with W_Q , W_K , W_V being learnable weight matrices. To maintain consistency between input and output dimensions and to prevent the loss of original features due to the depth of the model, the attention feature vector A undergoes dimensionality expansion and residual connection processing. The calculation for the feature F is given by

$$F = W_L A + Y \quad (5)$$

where W_L is the parameter matrix for dimensionality expansion of the attention feature A .

In the lightweight feed-forward neural network, the input dimensionality d_m is first reduced to $\frac{d_m}{4}$ and then expanded back, with the parameter count reduced to 1/16 of the original while maintaining performance. The final features are computed as

$$O_T = \text{GAP}(F + ((\text{ReLU}(FW_1 + b_1))W_2 + b_2)) \quad (6)$$

where W_1 and b_1 are the parameter matrix and bias vector for dimensionality reduction, and W_2 and b_2 are the parameter matrix and bias vector for dimensionality expansion, with $\text{GAP}(\cdot)$ denoting the global average pooling operation and $\text{ReLU}(\cdot)$ representing the nonlinear activation function.

2.3. Transform Domain Spatial Feature Fully Convolutional Module

The local features of HRRP target sequences contain valuable information, such as the peaks and troughs of the echo signals, which are crucial for target recognition and classification. CNNs can effectively extract local features through convolution operations, thereby capturing the spatial structural information of HRRP data. The spatial feature fully convolutional module in the proposed method consists of one-dimensional convolutional layers, batch normalization (BN) layers, ReLU activation functions, squeeze-and-excitation (SE) blocks, and pooling layers.

The convolutional blocks, composed of one-dimensional convolutional layers, batch normalization layers, and ReLU activation functions, are employed to extract spatial features from the input data. The spatial feature extraction module consists of three convolutional blocks with kernel sizes of 8, 5, and 3, respectively. These sizes enable the capture of spatial information at different scales, thereby enhancing the accuracy of target recognition. Following the first two convolutional blocks, SE blocks are incorporated, as illustrated in Figure 3. The SE blocks perform feature recalibration using adaptive average pooling and fully connected layers, leveraging global information to adjust the weights of each channel. This process enhances the representation of globally significant features in the HRRP sequences, suppresses noise and less important features, and thereby strengthens the network's expressive capability. The input HRRP data are first transformed using FrFT to enhance feature representation in the fractional Fourier domain. The FrFT of order α for a sequence $\chi(t)$ is defined as

$$X_a(u) = \mathcal{F}^a[\chi(t)](u) = \int_{-\infty}^{\infty} \chi(t) K_a(t, u) dt \quad (7)$$

where $K_a(t, u)$ is the transformation kernel:

$$K_a(t, u) = \begin{cases} \sqrt{\frac{1-j\cot\alpha}{2\pi}} \exp\left(j\frac{t^2+u^2}{2} \cot\alpha - jut \csc\alpha\right), & \alpha \neq n\pi \\ \delta(t-u), & \alpha = 2n\pi \\ \delta(t+u), & \alpha = (2n+1)\pi \end{cases} \quad (8)$$

Here, $\alpha = a\pi/2$ is the rotation angle in the distance-frequency plane. This kernel is the fundamental operator of the FrFT; varying the order α effectively rotates the signal representation in the joint distance-frequency plane, enabling the search for a domain where target-specific scattering characteristics are optimally concentrated and separated. The transformed data are then processed through the subsequent spatial feature extraction module. The calculation process for the spatial feature O_S can be expressed as

$$O_{S_m} = \begin{cases} \text{SE}(\text{ReLU}(\text{BN}(\text{Conv1D}(X^T))))), & m = 1 \\ \text{SE}(\text{ReLU}(\text{BN}(\text{Conv1D}(O_{S_1}))))), & m = 2 \\ \text{ReLU}(\text{BN}(\text{Conv1D}(O_{S_2}))), & m = 3 \end{cases} \quad (9)$$

$$O_S = \frac{1}{n} \sum_i^n O_{S_3}(i) \quad (10)$$

where $\text{SE}(\cdot)$ represents the squeeze-and-excitation block, $\text{ReLU}(\cdot)$ is the nonlinear activation function, $\text{BN}(\cdot)$ denotes the batch normalization layer, $\text{Conv1D}(\cdot)$ is the one-dimensional convolution layer, and n is the length of the HRRP sequence. The pooling layer uses adaptive average pooling to reduce the dimensionality of the convolution layer output, resulting in the global spatial feature O_S .

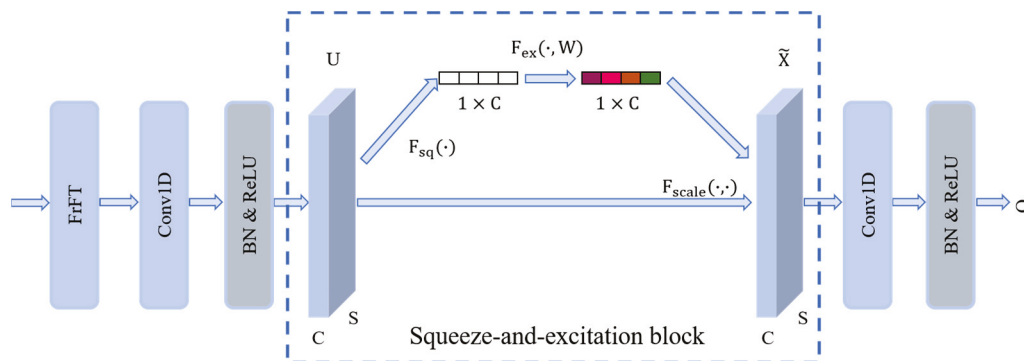


Figure 3. Squeeze-and-excitation block.

2.4. Decision Fusion and Recognition Module

After extracting temporal correlation features and spatial structural features from the temporal feature encoder module and spatial feature fully convolutional module, respectively, the decision fusion and recognition module uses two fully connected layers to classify the temporal feature O_T and spatial feature O_S , resulting in the temporal recognition result P_T and the spatial recognition result P_S . Given the sensitivity of HRRP data to time shifts and target orientations, we establish learnable decision weight matrices W_T and W_S , where $W_T, W_S \in R^{1 \times t}$, and t is the number of target categories. These matrices adaptively adjust the contributions of temporal and spatial features for each category, addressing time shift and orientation sensitivity issues to enhance model robustness. The calculation process for the final recognition result is described in Figure 4.

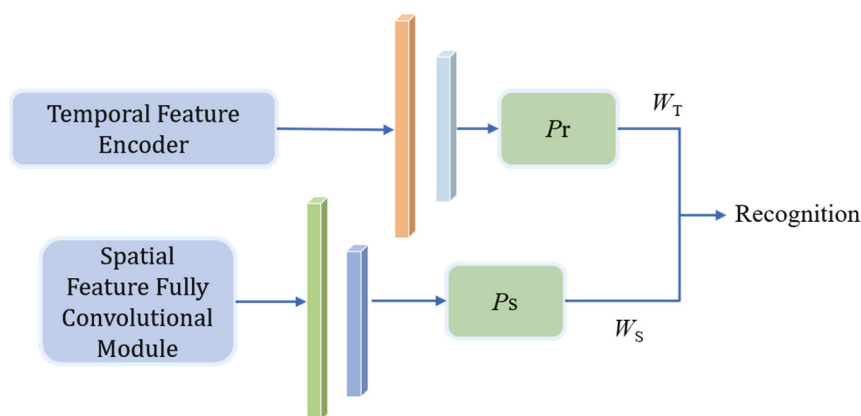


Figure 4. Decision fusion and recognition pipeline.

3. Loss Function Under Unbalanced Data

3.1. Focal Loss

Focal loss [34] is a loss function designed to address the issue of class imbalance. By introducing a coefficient factor on top of the standard cross-entropy loss, focal loss reduces the emphasis on easy samples and increases the focus on hard samples, thereby enhancing the model's classification capability. In this work, hard samples primarily refer to training

samples that are misclassified or classified with low confidence, often stemming from minority classes or challenging conditions. The calculation of focal loss is

$$FL = -\alpha_t(1 - p_t)^\gamma \log(p_t) \quad (11)$$

where p_t denotes the predicted probability of the true class for a sample, α_t is a balancing factor that adjusts the influence of positive and negative samples, and γ is a modulation factor that adjusts the focus on hard-to-classify samples. By tuning these weighting factors, focal loss effectively addresses the issue in which the loss value is dominated by the majority class and easy samples due to class imbalance.

3.2. Label-Smoothing Enhanced Focal Loss

Label smoothing is a regularization technique designed to prevent the model from becoming overly confident in its predictions on the training data, thereby improving generalization. The core idea of label smoothing is to convert the hard one-hot encoded vectors of true labels into a softer probability distribution. This method effectively reduces overfitting and enhances the model's robustness to noisy data. Let y denote the true label and k be the number of classes. The one-hot encoding is y_i (where only the i -th position is 1, and all others are 0) [35]. After label smoothing, y_i^{smooth} can be represented as

$$y_i^{\text{smooth}} = (1 - \varepsilon)y_i + \frac{\varepsilon}{k} \quad (12)$$

where ε is the smoothing factor. Typically, label smoothing is combined with the cross-entropy loss (CE) [36], which is given by

$$CE_{\text{smooth}} = -\sum_{i=1}^k \left((1 - \varepsilon)y_i + \frac{\varepsilon}{k} \right) \log(p_i) \quad (13)$$

where p_i denotes the model's predicted probability for class i . In the proposed method, label smoothing [37] is integrated with focal loss, resulting in the following calculation:

$$FL_{\text{smooth}} = -\sum_{i=1}^k \left((1 - \varepsilon)y_i + \frac{\varepsilon}{k} \right) \alpha_i (1 - p_i)^\gamma \log(p_i) \quad (14)$$

Label smoothing provides a smoothed target distribution, making the loss function's predictions more balanced across classes. Meanwhile, focal loss adjusts the loss weights to increase the model's sensitivity to hard samples. The combination of these methods enhances both the model's robustness to noisy data and its ability to recognize minority-class samples.

3.3. Adaptive Focal Loss

In standard focal loss, the balance factor α and the modulation factor γ are typically fixed values, and their settings significantly influence the model's ability to recognize hard samples. In the proposed method, both the balance factor α and the modulation factor γ are dynamically adjusted based on the number of categories and prediction probabilities. The calculation formula for the balance factor α_i for each category is

$$\alpha_i = \frac{N}{c_i} \quad (15)$$

where N is the total number of samples across all categories, and c_i represents the number of samples in the i -th category. For minority classes, the balance factor α is larger, making the model pay more attention to the minority classes. The calculation process for the modulation factor γ_i for the i -th class is

$$\gamma_i = \gamma_0 + \beta(1 - p_i) \quad (16)$$

where γ_0 denotes the initial value of the dynamic adjustment factor, β is a tunable hyperparameter, and p_i represents the predicted probability for the i -th class. For more difficult-to-classify samples, i.e., those with a smaller p_i , the modulation factor γ_i will be larger, thereby giving more focus to these samples. Combining this with (12), the calculation formula for AFL-LS is

$$\text{AFL-LS} = -\sum_{i=1}^k \left((1 - \varepsilon)y_i + \frac{\varepsilon}{c_i} \right) \frac{N}{c_i} (1 - p_i)^{\gamma_0 + \beta(1 - p_i)} \log(p_i) \quad (17)$$

4. Results and Analysis

4.1. Datasets

In order to verify the validity of the proposed model and the loss function, we conducted experiments on the MSTAR dataset [38] and the CVDomes dataset [39]. The MSTAR dataset is widely used for radar automatic target recognition. It is sourced from a high-resolution synthetic aperture radar operating in the X-band with HH polarization and includes 10 target categories such as BMP2, BTR70, and T72.

We use data with a pitch angle of 17° as the training set and data with a pitch angle of 15° as the test set. We also include four variant targets—BMP2(SN-9563), BPM2(SN-C21), T72(SN-812), and T72(SN-S7)—in the test set to create Dataset 1, which tests the model's generalization performance. Similarly, we use data with a pitch angle of 17° as the training set, but multiply the sample count of each category by a decay factor θ . For the reduced number of samples for the i -th category, we calculate it as

$$\text{Num}_i^{\text{Reduced}} = \theta^i \times \text{Num}_i \quad (18)$$

where Num_i denotes the original sample count of the i -th category, and i ranges from $[0, 9]$. In this experiment, θ is set to 0.6. We randomly select $\text{Num}_i^{\text{Reduced}}$ samples from the original dataset to construct a class-imbalanced training set. The test set remains the same, with data at a pitch angle of 15° , including four variant targets, which results in Dataset 2. Following the method in reference [40], we convert SAR images into HRRP sequences. The composition of the MSTAR HRRP sequence dataset is shown in Tables 1 and 2.

Table 1. MSTAR HRRP Sequence Dataset 1.

Target	Training Set (17°)	Target	Testing Set (15°)
2S1	2990	2S1	2740
BMP2(SN-9566)	2330	BMP2(SN-9566)	1960
		BMP2(SN-9563)	1950
		BMP2(SN-C21)	1960
BRDM-2	2980	BRDM-2	2740
BTR70(SN-C71)	2330	BTR70(SN-C71)	1960
BTR60	2560	BTR60	1950
D7	2990	D7	2740
T62	2990	T62	2730
T72(SN-132)	2320	T72(SN-132)	1960
		T72(SN-812)	1950
		T72(SN-S7)	1910
ZIL131	2990	ZIL131	2740
ZSU23/4	2990	ZSU23/4	2740
Total	27,470	Total	32,030

Table 2. MSTAR HRRP Sequence Dataset 2.

Target	Training Set (17°)	Target	Testing Set (15°)
2S1	2990	2S1	2740
BMP2(SN-9566)	1398	BMP2(SN-9566)	1960
		BMP2(SN-9563)	1950
		BMP2(SN-C21)	1960
BRDM-2	1072	BRDM-2	2740
BTR70(SN-C71)	503	BTR70(SN-C71)	1960
BTR60	331	BTR60	1950
D7	232	D7	2740
T62	139	T62	2730
T72(SN-132)	64	T72(SN-132)	1960
		T72(SN-812)	1950
		T72(SN-S7)	1910
ZIL131	50	ZIL131	2740
ZSU23/4	30	ZSU23/4	2740
Total	6809	Total	32,030

The specific transformation steps are as follows: first, the SAR images undergo a one-dimensional inverse fast Fourier transform (IFFT) to obtain the complex domain HRRP sequences, which are then modulated to obtain the HRRP sequences. For each SAR image, 100 HRRP sequence samples are generated. By averaging every ten HRRP sequence samples, we obtain one averaged HRRP sample. Consequently, the training set of Dataset 1 contains 27,470 samples, while the training set of Dataset 2 contains 6809 samples. Both datasets have a test set containing 32,030 samples.

The CVDomes dataset, released in 2009, contains simulated X-band signatures of 10 civilian vehicle categories, including Toyota Camry, Honda Civic 4dr, and 1993 Jeep. Radar azimuth angles cover the full range from 0° to 359°. Acquired directly in HRRP format, this dataset requires no IFFT processing, and HRRP sequences can be generated through sliding window operations. For experimental validation, we select HH-polarized data at a 30-degree elevation angle and simulate varying signal-to-noise ratio conditions by superimposing Gaussian noise. Following the long-tail distribution strategy consistent with the MSTAR dataset, 60% of samples constitute the class-imbalanced training set, with 20% allocated to the validation and test sets, respectively. Detailed configuration parameters are provided in Table 3.

Table 3. CVDomes HRRP Sequence Dataset 3.

Target	Training Set	Testing Set
Toyota Camry	3648	1200
Honda Civic 4dr	2188	1200
1993 Jeep	1313	1200
1999 Jeep	787	1200
Nissan Maxima	472	1200
Mazda MPV	283	1200
Mitsubishi	170	1200
Nissan Sentra	102	1200
Toyota Avalon	61	1200
Toyota Tacoma	36	1200
Total	9060	12,000

The imbalance factor (IF) [41] is a metric used to quantify the degree of class imbalance in classification problems. By calculating the IF, one can measure the disparity in the

number of samples across different classes within a dataset. In this study, the IF values for Datasets 1–3 are measured at 0.997, 0.702, and 0.714, respectively. These metrics indicate that Dataset 1 maintains a near-ideal balanced state, whereas Datasets 2 and 3 exhibit significant class imbalance, manifesting characteristic long-tailed distribution patterns in sample category quantities.

4.2. Recognition Performance

To validate the effectiveness of the proposed method for HRRP sequence recognition, two types of comparative experiments are designed: one to verify the effectiveness of the LSTF model and the other to evaluate the AFL-LS loss function. The hardware environment for these experiments includes a 64-bit operating system, an Intel Core i5-13490F processor, an RTX 3060 GPU, and 32 GB of RAM. The software environment comprises Python 3.9 and PyTorch 2.0.1.

4.2.1. Validation of the Effectiveness of the Proposed LSTF Model

To validate the effectiveness of the proposed method, we select LSTM-FCN [42], GRU-FCN [43], gMLP [44], XCM [45], GTN [46], RLAT [47], and LViT [48] as baseline methods for comparison. Each method runs five times under identical parameter conditions to minimize randomness caused by dropout and sample selection. LSTM-FCN and GRU-FCN are both hybrid models combining recurrent neural networks with convolutional neural networks. They leverage long short-term memory (LSTM) and gated recurrent units (GRU), respectively, to capture temporal dependencies, while integrating fully convolutional networks (FCN) to extract local spatial features, making them suitable for time-series data classification. gMLP centers on a multilayer perceptron, achieving feature modeling through gating mechanisms and spatial projections. It efficiently captures data correlations while dispensing with traditional attention mechanisms. XCM is an explainable convolutional neural network designed for multi-class temporal classification, leveraging convolutional structures to mine local data features while providing interpretability. GTN, or gated Transformer network, optimizes attention distribution in Transformers through gating mechanisms, enhancing the capture of critical information in multi-temporal data. RLAT is a lightweight Transformer designed for high-resolution distance image sequence recognition. It reduces parameter size and computational overhead through structural simplification, meeting edge device deployment requirements. LViT is a lightweight visual Transformer that employs optimized designs like competitive blocks to enhance feature extraction efficiency while maintaining compactness. It has been applied to visual tasks such as fingerprint recognition.

The experimental configuration employs a single-layer temporal feature encoder in the LSTF model, with group linear transformations (GLTs) divided into four groups. Preprocessed HRRP sequences of length 32 serve as the input, utilizing standard cross-entropy loss for Dataset 1 and AFL-LS for Datasets 2 and 3.

On class-balanced Dataset 1, all comparative models demonstrate competent baseline performance. However, the proposed LSTF exhibits significant superiority in both recognition accuracy and generalization capability. As detailed in Table 4, LSTF achieves 99.52% recognition accuracy in the 10-class identification task, outperforming the suboptimal model by 0.25 percentage points while maintaining the highest accuracy across all target categories. Notably, the model attains 99.52%, validating the spatiotemporal fusion mechanism's sensitivity to subtle target differences. Except for the methods proposed for certain datasets (2S1, BTR70, and T72), which are not optimal, the proposed methods achieve the best recognition performance across all other datasets. These results confirm that LSTF maintains high precision while demonstrating exceptional generalization capa-

bilities, establishing a technical foundation for subsequent research on class-imbalanced data conditions.

Table 4. Recognition accuracy of comparative experiments on Dataset 1 (%).

	LSTM-FCN	GRU-FCN	gMLP	XCM	GTN	RLAT	LViT	Proposed
2S1	100.00	100.00	100.00	99.87	97.91	99.97	99.34	98.47
BMP2	96.94	98.31	91.22	77.59	93.32	99.01	99.08	99.91
BRDM-2	96.47	95.54	92.22	84.42	68.24	92.06	98.50	99.01
BTR70(SN-C71)	100.00	100.00	86.37	98.38	89.07	99.32	94.26	97.60
BTR60	99.09	99.82	96.31	82.99	99.01	98.36	96.62	99.24
D7	100.00	100.00	100.00	99.98	98.25	100.00	100.00	100.00
T62	99.91	99.67	98.83	91.27	93.81	97.46	100.00	100.00
T72	100.00	100.00	94.98	91.62	99.36	99.99	97.71	99.79
ZIL131	100.00	100.00	99.27	95.41	98.54	99.94	100.00	100.00
ZSU23/4	100.00	100.00	99.76	99.97	99.67	99.97	100.00	100.00
Overall Accuracy	99.07	99.27	95.56	90.73	94.22	98.77	98.65	99.52

Further comparative analysis on class-imbalanced Dataset 2 reveals the recognition accuracy of competing models, as quantitatively detailed in Table 5. The LSTF model maintains superior overall recognition rates while achieving optimal performance across four target categories. Given Dataset 2's imbalanced training distribution and test set containing variant targets, the experimental setup imposes heightened demands on model generalization capabilities. The results demonstrate that LSTF attains 99.96% and 99.98% accuracy for BRDM-2 and T62 categories with variant models, respectively, confirming the effectiveness of the dynamic decision weighting mechanism in mitigating feature confusion. For tail classes 2S1, BTR70, D7, T72, and ZIL131, the model achieves 100.00% accuracy, showing marked improvements over Transformer-based GTN and RLAT models. In challenging BMP2, BTR60, and ZSU23/4 category recognition, the performance of LSTF is slightly inferior. Furthermore, as illustrated in Figure 5, LSTF exhibits enhanced recognition stability on Dataset 2 with a maximum fluctuation limited to 0.61%, demonstrating significant advantages over comparative models.

Table 5. Recognition accuracy of comparative experiments on Dataset 2 (%).

	LSTM-FCN	GRU-FCN	gMLP	XCM	GTN	RLAT	LViT	Proposed
2S1	99.43	96.93	99.92	99.61	99.94	99.88	99.78	100.00
BMP2	89.60	94.30	98.54	96.38	95.58	93.37	89.85	95.97
BRDM-2	74.65	87.49	89.41	98.32	65.11	79.36	88.91	99.96
BTR70(SN-C71)	100.00	99.81	96.33	100.00	84.74	94.47	81.44	100.00
BTR60	93.90	99.16	98.55	99.94	95.81	97.42	81.52	99.95
D7	100.00	100.00	99.96	100.00	99.93	98.32	99.59	100.00
T62	98.61	97.60	95.62	99.19	94.14	71.97	99.26	99.98
T72	99.97	99.28	90.87	99.71	99.16	98.67	99.62	100.00
ZIL131	99.19	99.94	96.81	100.00	82.33	73.90	99.85	100.00
ZSU23/4	99.71	99.99	95.60	98.86	71.19	89.31	99.63	99.68
Overall Accuracy	95.29	97.22	95.82	98.94	90.41	90.62	94.37	99.55

To systematically evaluate model performance under class-imbalanced conditions, the experiment incorporates four metrics: macro F1-score, G-mean, macro ROC-AUC, and macro PR-AUC for comprehensive analysis, with quantitative results presented in Table 6. LSTF demonstrates superior performance across the four metrics: achieving a macro F1-score of 0.9935 indicates balanced inter-class prediction capabilities; its G-mean value significantly outperforms other baselines, validating strong robustness against skewed

data distributions; both macro ROC-AUC and PR-AUC approach theoretical maximums, confirming the model's decision reliability under high-confidence thresholds.

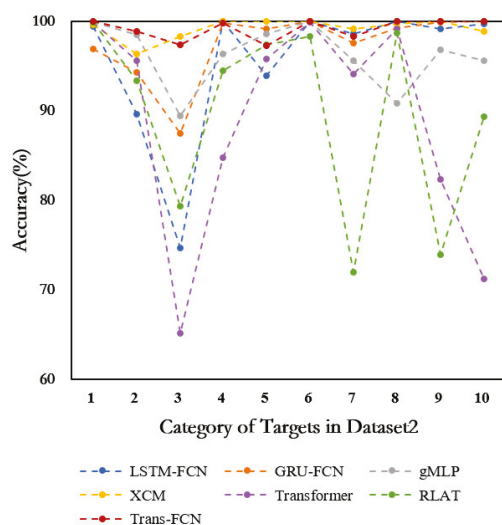


Figure 5. Accuracy of the 10 targets in Dataset 2.

Table 6. Performance comparison of each model on other evaluation indicators on Dataset 2 (%).

	LSTM-FCN	GRU-FCN	gMLP	XCM	GTN	RLAT	LViT	Proposed
Macro F1	0.9467	0.9722	0.9574	0.9894	0.8804	0.8764	94.27	0.9935
G-mean	0.9457	0.9711	0.9598	0.9920	0.8760	0.8723	94.64	0.9953
Macro ROC-AUC	0.9990	0.9718	0.9981	0.9999	0.9882	0.9878	99.71	1.0000
Macro PR-AUC	0.9949	0.9998	0.9908	0.9993	0.9441	0.9454	98.01	1.0000

To validate the model's recognition robustness and noise immunity under class-imbalanced and low signal-to-noise ratio (SNR) conditions, comparative experiments were conducted on the class-imbalanced Dataset 3. Gaussian noise was introduced during data preprocessing to construct training sets with an SNR of 20 dB, with model performance evaluated under test conditions of 20 dB, 15 dB, 10 dB, and 5 dB SNR levels, as detailed in Table 7. Experimental results reveal differentiated performance degradation patterns across models as SNR decreases. Under test conditions of 15 dB, the proposed model achieved recognition accuracy second only to the parameter-intensive GTN model, while demonstrating optimal performance under other SNR conditions. This indicates that the proposed method effectively overcomes existing models' reliance on high-quality signals, exhibiting enhanced engineering applicability in complex battlefield environments with dynamically varying SNR levels.

Table 7. Recognition accuracy of each model under different SNR levels on Dataset 3 (%).

SNR(dB)	LSTM-FCN	GRU-FCN	gMLP	XCM	GTN	RLAT	LViT	Proposed
20	81.16	85.18	96.53	82.26	97.53	88.00	80.14	98.98
15	78.68	84.07	93.67	77.88	94.67	85.89	72.05	93.50
10	72.23	74.40	85.68	71.74	89.21	81.95	47.64	89.23
5	60.21	63.17	66.77	58.13	79.83	72.88	33.98	79.96

This study further compares parameter counts and computational complexity across models to evaluate engineering applicability, as shown in Table 8. Among Transformer-based variants, GTN demonstrates superior performance on noise-contaminated Dataset 3,

where its deep architecture and high parameter count facilitate complex feature extraction, yet its large model size hinders real-time processing on embedded devices. In contrast, LSTF achieves significant reductions in both parameters and computations compared to GTN, while maintaining deep feature extraction capabilities through lightweight modifications. Although the most lightweight model, RLAT, excels on class-balanced Dataset 1, its recognition accuracy declines on imbalanced Datasets 2 and 3, revealing insufficient stability in identifying minority-class samples. While the proposed model's lightweight characteristics remain inferior to baseline methods like LSTM-FCN and GRU-FCN, it effectively reduces the inherent high parameter and computational loads of the Transformer attention mechanism. This optimization achieves complexity reduction without compromising recognition performance, making Transformer-based architectures more suitable for deployment on edge devices.

Table 8. Lightweight comparison of each model.

	LSTM-FCN	GRU-FCN	gMLP	XCM	GTN	RLAT	LViT	Proposed
Macs(M)	21.59	19.45	25.37	53.13	386.60	5.66	4.65	29.62
Params(M)	0.67	0.59	0.79	1.06	5.83	0.22	0.29	0.94

For a more intuitive comparison, Figure 6 is plotted. This dual-axis bar chart clearly compares the performance trade-offs of eight temporal classification models (including the proposed model) in terms of computational complexity (MACs, left blue axis) and model size (Parameters, right red axis). It reveals that while the GTN model achieves significantly higher values than other models across both metrics, most other models—including the proposed model—exhibit MACs and parameters concentrated within a lower range, demonstrating efficiency while maintaining relatively low resource consumption.

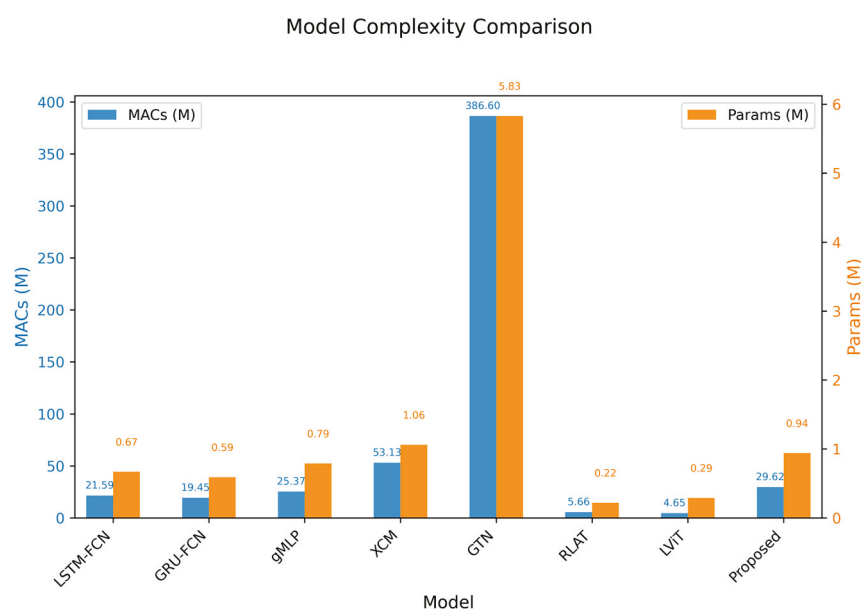


Figure 6. Lightweight comparison of each model.

4.2.2. Verification of AFL-LS Effectiveness

To verify the effectiveness of the proposed AFL-LS, several comparison methods are selected: standard cross-entropy loss (CE), data distribution loss (DD Loss) [49], standard focal loss (FL), focal loss with label smoothing (FL-LS), and adaptive focal loss (AFL). These methods are evaluated using the LSTF model on Dataset 2. The smoothing factor ϵ is set to

0.2, while γ_0 and β are set to 2.0. Each method is run five times, and the average recognition rates for each method are shown in Table 9.

Table 9. Recognition accuracy of each loss function on Dataset 2 (%).

	CE	CE-LS	DD Loss	FL	FL-LS	AFL	LV-LS	AFL-LS
2S1	100.00	99.12	99.85	86.04	100.00	100.00	85.97	100.00
BMP2	100.00	100.00	99.87	100.00	96.47	99.86	99.94	95.97
BRDM-2	96.46	95.85	98.83	96.11	98.91	99.92	96.05	99.96
BTR70(SN-C71)	93.61	99.69	89.33	72.14	91.81	98.29	72.91	100.00
BTR60	99.84	99.64	76.59	95.21	98.43	99.94	93.18	99.95
D7	97.72	98.58	100.00	99.93	100.00	89.02	97.88	100.00
T62	98.36	84.67	98.11	95.07	99.12	100.00	85.41	100.00
T72	99.95	99.15	97.45	98.19	100.00	99.91	95.12	100.00
ZIL131	100.00	91.00	99.89	99.52	100.00	100.00	98.25	100.00
ZSU23/4	100.00	90.00	100.00	100.00	100.00	100.00	99.46	99.48
Overall Accuracy	98.91	95.77	96.62	94.77	98.47	98.69	92.41	99.21

Experimental results demonstrate that the proposed AFL-LS loss function exhibits significant performance advantages under class-imbalanced conditions. It achieves superior overall accuracy compared to all baseline methods, outperforming the second-best approach, CE, by 0.3 percentage points. For the severely imbalanced ZSU23/4 and BMP2 datasets, AFL-LS achieved recognition rates of 99.48% and 95.97%, respectively. Meanwhile, CE effectively guided model attention toward minority-class samples, yielding high recognition rates. For BRDM-2 and BTR60, although label smoothing strategies enhanced model generalization, they simultaneously increased recognition difficulty for minority classes, causing CE-LS and FL-LS to completely fail in these categories. For the remaining targets, AFL-LS maintained higher accuracy than FL, confirming LS's effectiveness in improving majority class generalization. Notably, on the BRDM-2 target with complex scattering patterns, the state-of-the-art DD loss achieved only a 98.83% recognition rate. AFL-LS, however, elevated performance to 99.96% through dynamic class weight balancing. By integrating adaptive mechanisms with label smoothing techniques, this method not only intensifies focus on minority classes but also enhances generalization capabilities for variant targets, ultimately achieving the highest overall average recognition rate.

Furthermore, comparative analysis was conducted using four metrics: macro-F1 score, G-mean, macro ROC-AUC, and macro PR-AUC, as shown in Table 10. The proposed AFL-LS method achieves optimal performance across all metrics, while CE, CE-LS, and FL-LS completely fail on minority classes, resulting in G-mean values of zero and lower macro-F1 scores. Although the state-of-the-art DD Loss demonstrates outstanding performance in the macro ROC-AUC and macro PR-AUC metrics, its values still fell short of those of the algorithm proposed in this paper.

Table 10. Comparison of the performance of each loss function on other evaluation indicators (Dataset 2).

	CE	CE-LS	DD Loss	FL	FL-LS	AFL	LViT	AFL-LS
Macro F1	0.9874	0.8615	0.9628	0.9516	0.9850	0.9860	0.948	0.9935
G-mean	0.9893	0.0000	0.9697	0.9671	0.9848	0.9856	0.9021	0.9953
Macro ROC-AUC	1.0000	0.9972	0.9996	1.0000	0.9999	0.9999	0.9865	1.0000
Macro PR-AUC	1.0000	0.9764	0.9973	1.0000	0.9993	0.9994	0.9798	1.0000

Additionally, comparative experiments with different loss functions were performed on Dataset 3 under low signal-to-noise ratio (SNR=5 dB) conditions, as shown in

Tables 11 and 12. AFL-LS attained the highest overall recognition accuracy of 96.03%, outperforming the second-best method, AFL, by 4.19%, and achieved the highest recognition rates in 8 out of 10 target classes. Analysis of other evaluation metrics demonstrates that AFL-LS exhibits marked superiority in macro-F1 score and G-mean, confirming its ability to balance recognition performance between majority and tail classes. Furthermore, its exceptional performance in macro ROC-AUC and macro PR-AUC reflects high-confidence decision boundaries in noisy environments, proving that this loss function significantly enhances model robustness and noise resistance under low-SNR conditions.

Table 11. Recognition accuracy of each loss function on Dataset 3 (%).

	CE	CE-LS	DD Loss	FL	FL-LS	AFL	LV-LS	AFL-LS
Toyota Camry	97.17	89.54	92.79	87.62	85.80	87.08	86.52	98.28
Honda Civic 4dr	87.40	70.18	79.79	92.95	94.56	94.86	82.56	96.53
1993 Jeep	100.00	87.15	98.07	91.79	83.05	98.57	97.95	98.09
1999 Jeep	65.63	89.11	91.20	91.19	79.67	80.68	88.25	90.22
Nissan Maxima	96.73	95.10	98.94	98.94	89.32	93.13	92.47	98.29
Mazda MPV	96.15	66.39	99.83	100.00	96.74	95.09	94.36	97.24
Mitsubishi	70.41	76.04	89.98	76.24	82.37	85.41	86.78	88.17
Nissan Sentra	100.00	98.03	92.47	90.34	96.31	96.44	97.72	98.06
Toyota Avalon	98.42	88.04	71.33	100.00	99.37	96.76	96.08	98.49
Toyota Tacoma	95.77	0.00	92.69	99.59	100.00	97.40	95.48	96.88
Overall Accuracy	88.80	82.47	89.53	91.83	89.38	91.84	91.82	96.03

Table 12. Comparison of the performance of each loss function on other evaluation indicators (Dataset 3).

	CE	CE-LS	DD Loss	FL	FL-LS	AFL	LV-LS	AFL-LS
Macro F1	0.8640	0.7822	0.8867	0.9154	0.8806	0.9138	0.9112	0.9364
G-mean	0.8237	0.0000	0.8696	0.9092	0.8598	0.9066	0.9027	0.9319
Macro ROC-AUC	0.9956	0.9930	0.9942	0.9953	0.9961	0.9958	0.9923	0.9986
Macro PR-AUC	0.9717	0.9556	0.9644	0.9807	0.9753	0.9769	0.9635	0.9893

4.3. Ablation Experiment

To validate the effectiveness of the spatiotemporal feature complementary mechanism and adaptive weighting decision module, ablation studies were conducted on the MSTAR and CVDomes datasets, as presented in Table 13. Single-branch features exhibited significant performance fluctuations across data conditions. The spatial branch achieved 3.78% and 25.72% higher accuracy than the temporal branch in Datasets 1 and 2, whereas in Dataset 3, the spatial branch declined to 87.13%, while the temporal branch maintained robust performance at 94.13%. This demonstrates the differentiated sensitivity of spatiotemporal features to environmental characteristics: spatial features excel in geometric configuration analysis, while temporal features exhibit stronger noise robustness. Although the dual-branch framework enhances model applicability, static equal-weight fusion strategies tend to suffer from limited optimization and feature quality instability. The proposed adaptive weighting mechanism achieved performance improvements of 0.27%, 0.96%, and 1.62% across the three datasets.

Table 13. Recognition accuracy of ablation experiments.

Temporal Branch	Spatial Branch	Adaptive Weighting	Accuracy (%)		
			Dataset 1	Dataset 2	Dataset 3
✓			94.94	73.36	94.13
	✓		98.72	99.08	87.13
✓	✓		99.40	98.28	95.34
✓	✓	✓	99.67	99.24	96.96

The experimental results indicate that the coordinated design of the dual-branch framework and adaptive weighting mechanism effectively addresses feature modality conflicts and scene adaptability limitations in conventional methods, demonstrating superior generalization capability in complex data distributions and noisy environments.

4.4. Complexity Analysis

The primary computational cost of the proposed method comes from the Transformer_FCN backbone. Let the input sequence length be TS , the model dimension be d , the feed-forward hidden dimension be d_{ff} , and the number of layers be L .

Then, the computational cost of a single forward pass is approximately

$$O\left(L\left((TS)^2d + TS \cdot d \cdot d_{ff}\right)\right) \quad (19)$$

Considering that backpropagation takes approximately 2–3 times the cost of forward propagation, the training complexity per epoch is

$$O\left(\frac{N}{B} \cdot 3L\left((TS)^2d + TS \cdot d \cdot d_{ff}\right)\right) \quad (20)$$

The space complexity is primarily determined by the parameters and the attention map:

$$O\left(L\left(d^2 + dd_{ff}\right) + BTSd + B(TS)^2\right) \quad (21)$$

Compared to other traditional convolutional models, this approach exhibits slightly higher computational complexity when the temporal–spatial dimension TS is large, but demonstrates significant advantages in feature representation and classification accuracy.

Notably, the proposed method achieves real-time inference performance on resource-limited edge devices. Benchmark tests on an Intel Core i5-13490F processor show that the inference latency for a single HRRP sequence (length = 32) is only 2.3 ms, which is far below the 10 ms real-time threshold required for radar target recognition systems. This real-time capability is attributed to the TGLT-based lightweight design, which reduces computational complexity by 92.4% compared to the standard Transformer encoder (GTN), enabling seamless deployment in time-sensitive scenarios.

4.5. Visualization and Evaluation of Feature Extraction

To validate the effectiveness of spatiotemporal feature extraction in the proposed model, t-distributed stochastic neighbor embedding (t-SNE) was employed to visualize features from the imbalanced Dataset 2 and Dataset 3, as shown in Figures 7 and 8. t-SNE projects high-dimensional features into a 2D space (with axes t-SNE Dimension 1 and t-SNE Dimension 2), enabling intuitive observation of feature distributions and clustering patterns. Specifically, Figures 7a and 8a display temporal correlation features of HRRP sequences extracted by the temporal encoder, while Figures 7b and 8b illustrate spatial structural

features obtained from the full-convolution spatial module. In all subfigures, each point represents a sample, and its color corresponds to a specific target class (see legend).

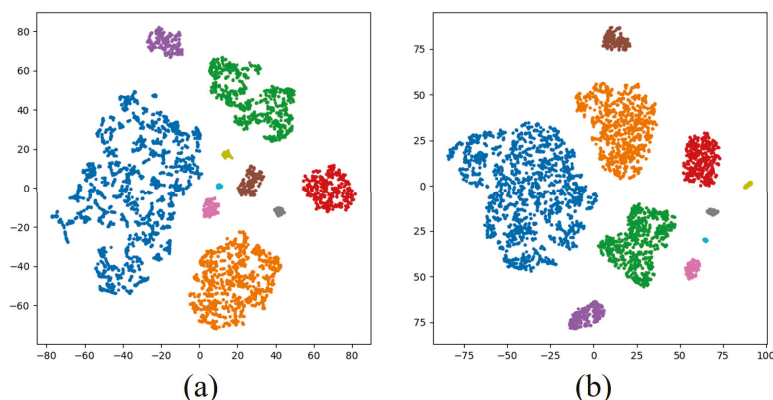


Figure 7. Visualization results of feature extraction on Dataset 2. (a) Temporal features. (b) Spatial features.

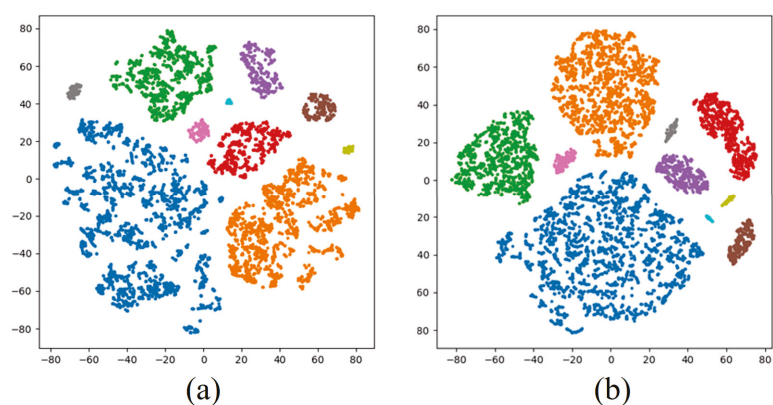


Figure 8. Visualization results of feature extraction on Dataset 3. (a) Temporal features. (b) Spatial features.

The visualization results reveal that both temporal and spatial features exhibit distinct inter-class separability and intra-class compactness after dimensionality reduction (i.e., points of the same color form tight clusters, and clusters of different colors are well separated), confirming the dual-branch architecture's capability to characterize target scattering mechanisms. This visualization provides interpretable evidence for the model's robustness in complex electromagnetic environments, demonstrating the theoretical superiority of the spatiotemporal feature complementary enhancement mechanism.

5. Conclusions

In this study, we proposed an LSTF method for HRRP sequence target recognition to address the challenges of category imbalance and robustness under complex conditions. The designed TGLT-based lightweight Transformer encoder effectively captured temporal dynamics while reducing computational overhead, which satisfied the requirements of real-time edge deployment. The transform-domain spatial feature extraction network, incorporating the fractional Fourier transform and FSCN, fully exploited multi-domain spatial information and enhanced feature discriminability. Furthermore, the AFL-LS improved recognition for long-tail classes and strengthened generalization capability. Experimental results on the MSTAR and CVDomes datasets showed that our method outperformed existing baselines in multiple scenarios, especially in minority class recognition and variant

identification, while maintaining stable robustness under noise interference. Overall, the proposed approach significantly improved recognition performance and demonstrated strong adaptability for practical non-cooperative target recognition.

Future research can further extend this work in several directions. First, more effective recognition strategies for extremely short, incomplete, or missing HRRP sequences should be explored to better match realistic non-cooperative radar scenarios. Second, self-supervised and semi-supervised learning can be introduced to reduce reliance on labeled data and improve generalization for long-tail and unseen target categories. Third, adaptive long-tail optimization and incremental learning mechanisms may be developed to handle dynamically changing class distributions and emerging targets. In addition, multi-view and multi-modal fusion of HRRP with complementary radar representations can be investigated to enhance robustness under complex observation conditions. Incorporating physics-informed priors into data-driven models is also a promising direction to improve interpretability and domain generalization. Finally, further hardware-aware optimization and model compression should be studied to support efficient, low-latency deployment on real-time edge radar platforms.

Author Contributions: Conceptualization, J.Y. (Junjun Yin) and Y.S.; data curation, X.L.; formal analysis, X.L.; funding acquisition, J.Y. (Junjun Yin); investigation, Y.S.; methodology, J.Y. (Junjun Yin); project administration, J.Y. (Junjun Yin); resources, J.Y. (Junjun Yin); software, Y.S.; supervision, J.Y. (Jian Yang); validation, Y.S.; visualization, Y.S.; writing—original draft preparation, Y.S.; writing—review and editing, X.Z. and J.Y. (Junjun Yin). All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Key Research and Development Program of China under Grant 2024YFB3909800, the open research fund of Pinghu Space Awareness Laboratory Technology Co., Ltd., and the NSFC under Grants 6222102 and 62171023.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors would like to thank the anonymous reviewers for their constructive suggestions on improving the quality of the paper.

Conflicts of Interest: The authors declare that this study received funding from Pinghu Space Awareness Laboratory Technology Co., Ltd. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

References

1. Liu, J.; Chen, B.; Jie, X. Radar high-resolution range profile target recognition based on attention mechanism and bidirectional gated recurrent. *J. Radars* **2019**, *8*, 589–597.
2. Xu, B.; Chen, B.; Liu, H.; Jin, L. Attention-based recurrent neural network model for radar high-resolution range profile target recognition. *J. Electron. Inf. Technol.* **2016**, *38*, 2988–2995.
3. Yin, J.; Sheng, W.; Jiang, H. Review of radar target recognition based on HRRP. *J. Ordnance Equip. Eng.* **2024**, *45*, 22–32.
4. Pilcher, C.M.; Khotanzad, A. Maritime ATR using classifier combination and high-resolution range profiles. *IEEE Trans. Aerosp. Electron. Syst.* **2011**, *47*, 2558–2573. [CrossRef]
5. Fan, X.; Hu, S.; He, J. Feature extraction and selection of full-polarization HRRP in target recognition process of maritime surveillance radar. *J. Electron. Inf. Technol.* **2016**, *38*, 3261–3268.
6. Zhao, D.; Li, H. Radar target recognition based on central moment feature and GA-BP neural network. *Infrared Laser Eng.* **2018**, *47*, 394–400.
7. Nie, J.; Xiao, Y.; Huang, L.; He, F. High-resolution range profile radar target recognition based on time–frequency analysis and deep learning. *J. Appl. Sci.* **2022**, *40*, 973–983.
8. Yun, T.; Pan, Q.; He, Y.; Xu, R. Target recognition algorithm based on HRRP time–spectrogram feature and multiscale asymmetric convolutional neural network. *J. Northwest. Polytech. Univ.* **2023**, *41*, 537–545. [CrossRef]

9. Zhou, D. Radar target HRRP recognition based on reconstructive and discriminative dictionary learning. *Signal Process.* **2016**, *126*, 52–64. [CrossRef]
10. Duan, P.; Li, H.; Luo, M. Study on radar target recognition based on sparse representation with a federated dictionary. *Appl. Electron. Techn.* **2019**, *45*, 11–14.
11. Zhang, J.; Tan, J.; Yue, S. Radar target recognition of high-resolution range profiles based on nonlinear transform. *Syst. Eng. Electron.* **2005**, *27*, 1732–1733.
12. Li, B.; Li, H.; Guo, S. Using evolutionary algorithm and adaptive wavelet for HRRP feature extraction and classification. In Proceedings of the AIAA '17: Proceedings of the 2017 International Conference on Artificial Intelligence, Automation and Control Technologies, Wuhan, China, 7–9 April 2017; Association for Computing Machinery: New York, NY, USA, 2017; pp. 1–6.
13. Du, L.; Liu, H.; Bao, Z. Radar HRRP statistical recognition: Parametric model and model selection. *IEEE Trans. Signal Process.* **2008**, *56*, 1931–1944. [CrossRef]
14. Du, L.; Wang, P.; Zhang, L.; He, H.; Liu, H. Robust statistical recognition and reconstruction scheme based on hierarchical Bayesian learning of HRR radar target signal. *Expert Syst. Appl.* **2015**, *42*, 5860–5873. [CrossRef]
15. Du, L.; Chen, J.; Hu, J.; Li, Y.; He, H. Statistical modeling with label constraint for radar target recognition. *IEEE Trans. Aerosp. Electron. Syst.* **2020**, *56*, 1026–1044. [CrossRef]
16. Shi, L.; Wang, P.; Liu, H.; Xu, L.; Bao, Z. Radar HRRP statistical recognition with local factor analysis by automatic Bayesian Ying–Yang harmony learning. *IEEE Trans. Signal Process.* **2011**, *59*, 610–617. [CrossRef]
17. Zhang, X.; Chen, B.; Liu, H.; Zuo, L.; Feng, B. Infinite max-margin factor analysis via data augmentation. *Pattern Recognit.* **2016**, *52*, 17–32. [CrossRef]
18. Chen, J.; Liao, L.; Zhang, W.; Du, L. Mixture factor analysis with distance metric constraint for dimensionality reduction. *Pattern Recognit.* **2022**, *121*, 108156. [CrossRef]
19. Xu, B.; Chen, B.; Wan, J.; Liu, H.; Jin, L. Target-aware recurrent attentional network for radar HRRP target recognition. *Signal Process.* **2019**, *155*, 268–280. [CrossRef]
20. Du, C.; Tian, L.; Chen, B.; Zhang, L.; Chen, W.; Liu, H. Region-factorized recurrent attentional network with deep clustering for radar HRRP target recognition. *Signal Process.* **2021**, *183*, 108010. [CrossRef]
21. Wan, J.; Chen, B.; Liu, Y.; Yuan, Y.; Liu, H.; Jin, L. Recognizing the HRRP by combining CNN and BiRNN with attention mechanism. *IEEE Access* **2020**, *8*, 20828–20837. [CrossRef]
22. Wang, P.; Chen, T.; Ding, J.; Pan, M.; Tang, S. Intelligent radar HRRP target recognition based on CNN-BERT model. *EURASIP J. Adv. Signal Process.* **2022**, *2022*, 89. [CrossRef]
23. Wu, W.; Dan, B.; Wang, Z. HRRP target recognition method based on fusion network. *Radar Sci. Technol.* **2025**, *23*, 192–198.
24. Zhang, L.; Han, C.; Wang, Y.; Li, Y.; Long, T. Polarimetric HRRP recognition based on feature-guided Transformer model. *Electron. Lett.* **2021**, *57*, 705–707. [CrossRef]
25. Diao, Y.; Liu, S.; Gao, X.; Liu, A. Position embedding-free Transformer for radar HRRP target recognition. In Proceedings of the IGARSS 2022—2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, 17–22 July 2022; pp. 1896–1899.
26. Wang, X.; Wang, P.; Song, Y.; Xiang, Q.; Li, J. High-resolution range profile sequence recognition based on Transformer with temporal–spatial fusion and label smoothing. *Adv. Intell. Syst.* **2023**, *5*, 2300286. [CrossRef]
27. Gao, F.; Ren, D.; Yin, J.; Yang, J. Polarimetric HRRP recognition using Vision Transformer with polarimetric preprocessing and attention loss. In Proceedings of the IGARSS 2024—2024 IEEE International Geoscience and Remote Sensing Symposium, Athens, Greece, 7–12 July 2024; pp. 10838–10842.
28. Yin, M.; Zhang, J.; Zhang, C. Target recognition of radar HRRP based on AUMAP segmentation and improved convolutional neural network. In Proceedings of the IET International Radar Conference (IET IRC 2020), Online, 4–6 November 2020; pp. 417–424.
29. Jia, Y.; Chen, B.; Tian, L.; Chen, W.; Liu, H. Memory-based neural network for radar HRRP noncooperative target recognition. In Proceedings of the 2020 IEEE 11th Sensor Array and Multichannel Signal Processing Workshop (SAM), Hangzhou, China, 8–11 June 2020; pp. 1–5.
30. Zhang, X.; Wang, W.; Zheng, X.; Wei, Y. A novel radar target recognition method for open and imbalanced high-resolution range profile. *Digit. Signal Process.* **2021**, *118*, 103212. [CrossRef]
31. Guo, C.; Wang, H.; Xia, X.; Chen, L.; Liu, C.; Zhang, G. Deep transfer learning based method for radar automatic recognition with small data size. In Proceedings of the 2021 IEEE International Conference on Unmanned Systems (ICUS), Beijing, China, 15–17 October 2021; pp. 995–999.
32. Wu, W.; Bo, D.; Wang, Z.; Xing, Z. MCNN-TEAN based unbalanced HRRP target identification method. In Proceedings of the SPML '24: Proceedings of the 2024 7th International Conference on Signal Processing and Machine Learning, Qingdao, China, 12–14 July 2024; pp. 100–106.

33. Tian, X.; Zhang, J.; Bai, X.; Wang, Y.; Wang, Z. The imbalanced recognition of space micro-motion targets based on GRB loss. In Proceedings of the 2024 IEEE International Conference on Signal, Information and Data Processing (ICSIDP), Zhuhai, China, 22–24 November 2024; pp. 1–5.
34. Lin, T.-Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2999–3007.
35. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the Inception architecture for computer vision. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 26–30 June 2016; pp. 2818–2826.
36. Hinton, G.; Vinyals, O.; Dean, J. Reducing overfitting in deep networks by decorrelating representations. *arXiv* **2019**, arXiv:1909.11556.
37. Azad, R.; Heidary, M.; Yilmaz, K.; Hüttemann, M.; Karimijafarbigloo, S.; Wu, Y.; Schmeink, A.; Merhof, D. Loss functions in the era of semantic segmentation: A survey and outlook. *arXiv* **2023**, arXiv:2312.05391. [CrossRef]
38. Ross, T.D.; Worrell, S.W.; Velten, V.J.; Mossing, J.C.; Bryant, M.L. Standard SAR ATR evaluation experiments using the MSTAR public release data set. In *Algorithms for Synthetic Aperture Radar Imagery*; SPIE: Bellingham, WA, USA, 1998; pp. 566–573.
39. Dungan, K.E.; Austin, C.; Nehrbass, J.; Potter, L.C. Civilian vehicle radar data domes. In *Algorithms for Synthetic Aperture Radar Imagery XVII*; SPIE: Bellingham, WA, USA, 2010; Volume 7699, pp. 242–253.
40. Zhang, Y.; Gao, X.; Peng, X.; Ye, J.; Li, X. Attention-based recurrent temporal restricted Boltzmann machine for radar high resolution range profile sequence recognition. *Sensors* **2018**, *18*, 1585. [CrossRef]
41. Pirizadeh, M.; Farahani, H.; Kheradpisheh, S.R. Imbalance factor: A simple new scale for measuring inter-class imbalance extent in classification problems. *Knowl. Inf. Syst.* **2023**, *65*, 4157–4183. [CrossRef]
42. Karim, F.; Majumdar, S.; Darabi, H.; Harford, S. Multivariate LSTM-FCNs for time series classification. *Neural Netw.* **2019**, *116*, 237–245. [CrossRef]
43. Elsayed, N.; Maida, A.S.; Bayoumi, M. Deep gated recurrent and convolutional network hybrid model for univariate time series classification. *Int. J. Adv. Comput. Sci. Appl.* **2019**, *9*, 3137. [CrossRef]
44. Liu, H.; Dai, Z.; So, D.; Le, Q.V. Pay attention to MLPs. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 9204–9215.
45. Fauvel, K.; Lin, T.; Masson, V.; Fromont, E.; Termier, A. XCM: An explainable convolutional neural network for multivariate time series classification. *Mathematics* **2021**, *9*, 3137. [CrossRef]
46. Liu, M.; Ren, S.; Ma, S.; Jiao, J.; Chen, Y.; Wang, Z.; Song, W. Gated transformer networks for multivariate time series classification. *arXiv* **2021**, arXiv:2103.14438. [CrossRef]
47. Wang, X.; Wang, P.; Song, Y.; Xiang, Q.; Li, J. RLAT: Lightweight Transformer for high-resolution range profile sequence recognition. *Comput. Syst. Sci. Eng.* **2024**, *48*, 217–246. [CrossRef]
48. Chen, C.-A.; Chien, T.-H.; Ke, L.-Y.; Hsia, C.-H. A lightweight vision Transformer with competitive blocks for finger vein recognition. In Proceedings of the 2025 Seventh International Symposium on Computer, Consumer and Control (IS3C), Taichung, Taiwan, 27–30 June 2025; pp. 1–4.
49. Zhang, L.; Leng, X.; Ma, X.; Ji, K.; Kuang, G.; Liu, L. Data distribution loss for imbalanced SAR vehicle target recognition. *IEEE Geosci. Remote Sens. Lett.* **2024**, *21*, 1–5. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Distributed Phased-Array Radar Mainlobe Interference Suppression and Cooperative Localization Based on CEEMDAN–WOBSS

Xiang Liu, Huafeng He *, Ruike Li, Yubin Wu, Xin Zhang and Yongquan You

College of Missile Engineering, Rocket Force Engineering University, Xi'an 710025, China; liux_92@163.com (X.L.); 13095315592@163.com (R.L.); binhui369@163.com (Y.W.); zhangxin_shanxi@163.com (X.Z.); youyongquan1997@163.com (Y.Y.)

* Correspondence: lp09041811@163.com

Abstract

Mainlobe interference can severely degrade the performance of distributed phased-array radar systems in the presence of strong jamming or low-reflectivity targets. This paper introduces a signal–data dual-domain cooperative antijamming and localization (SDCAL) framework that integrates adaptive complete ensemble empirical mode decomposition with improved blind source separation and wavelet optimization (CEEMDAN–WOBSS) for signal-level denoising and separation. Following source separation, CFAR-based pulse compression is applied for precise range estimation, and multi-node data fusion is then used to achieve three-dimensional target localization. Under low signal-to-noise ratio (SNR) conditions, the adaptive CEEMDAN–WOBSS approach reconstructs the signal covariance matrix to preserve subspace rank, thereby accelerating convergence of the separation matrix. The subsequent pulse compression and CFAR detection steps provide reliable inter-node distance measurements for accurate fusion. The simulation results demonstrate that, compared to conventional blind-source-separation methods, the proposed framework markedly enhances interference suppression, detection probability, and localization accuracy—validating its effectiveness for robust collaborative sensing in challenging jamming scenarios.

Keywords: distributed radar networks; mainlobe interference suppression; cooperative target localization; blind source separation; low-SNR environments

1. Introduction

With the iterative evolution of new-generation radar technologies, research in the field of electronic countermeasures (ECMs) has demonstrated notably focused hotspots [1,2]. In particular, the application and development of Digital Radio Frequency Memory (DRFM) have rendered active jamming an efficient and widely adopted ECM technique [3]. Depending on the path differences by which electromagnetic interference enters a radar's antenna pattern, active jamming can be classified into two broad categories: mainlobe jamming and sidelobe jamming [4,5]. While sidelobe jamming countermeasures have matured into well-established technical systems [6–8], suppressing mainlobe jamming remains a greater technological challenge. Contemporary monostatic radar platforms predominantly employ countermeasures such as frequency agility [9], feature quantization analysis [10], and polarization-based feature extraction [11] to mitigate mainlobe jamming. However, intrinsic structural limitations—most notably single viewpoints and insufficient information

acquisition dimensions—undermine these methods [12]. As jamming techniques continue to evolve in sophistication, the system antijamming efficacy is prone to steep declines under complex ECM scenarios.

To surmount the intrinsic limitations of monostatic radar, distributed radar networking and cooperative antijamming frameworks have gradually emerged at the research frontier. These frameworks can be categorized—according to the level of information fusion—into two principal technical routes: data-level fusion and signal-level fusion. Within the data-level fusion domain, research efforts have predominantly focused on multisource information association and target discrimination algorithms. For example, Ref. [13] first established a false-target recognition model based on spatial distribution differences; Ref. [14] integrated composite azimuth–elevation information from combined active–passive radar systems to construct an interference suppression framework; Ref. [15] proposed a multiple nearest-neighbor–pure-azimuth joint association method for counter-deception jamming; Ref. [16] presented a progressive subset-selection method based on target-localization characteristics, achieving effective false-target discrimination while optimizing radar resource allocation; and Ref. [17] introduced a Frequency-Offset Deception Velocity (FODVB) approach that fuses spatial and frequency information to enhance the discrimination of false targets. Ref. [18] addressed the problem of tracking interference from noncooperative and cooperative false targets within a distributed radar network by proposing a consistency-covariance intersection (CCI) algorithm coupled with a distributed target discrimination scheme.

It is noteworthy that signal-level fusion techniques, by virtue of exploiting richer signal information, exhibit unique antijamming advantages. For instance, Ref. [19] leveraged distinctions between true targets and deceptive jammers in the Doppler-velocity domain to develop a joint target-authentication model for active–passive radar systems; Ref. [20] employed convolutional neural networks to extract multidomain features from radar echoes, enabling multistatic radar resilience against deception jamming; Ref. [21] proposed a noise-subspace projection-based interference suppression method that effectively discriminates false targets according to differences in spatial scattering characteristics; Ref. [22] utilized the complex envelopes of any two targets across different receivers for correlation testing to identify false targets; Ref. [23] applied a differencing principle to estimate amplitude sequences of received signals at different radar nodes and constructed test statistics to detect deception jamming under a given false-alarm probability; Ref. [24] fused multidomain features of echo signals via a residual convolutional neural network augmented with an attention mechanism, achieving precise identification of multiple simultaneous mainlobe deception jammers; and Ref. [25] combined the Neyman–Pearson criterion with a hybrid NP-MAP criterion and utilized polarization optimization to accurately identify and suppress active deception jamming in high-JNR environments.

Despite these advances, most existing antijamming methods are predicated on the theoretical assumption of an ideal received signal-to-noise ratio (SNR). In practical electromagnetic environments, a radar receiver must contend with the coupling effects of multiple interference sources—including strong-power active jamming, adjacent-frequency/co-frequency signal crosstalk, industrial electromagnetic noise, and complex ground/sea clutter—resulting in the effective signal often being submerged beneath a noise baseline (typical SNR < 3 dB). Under such low-SNR conditions, traditional antijamming systems exhibit the following technical limitations: (1) significant degradation of target-localization accuracy due to signal distortion; (2) sharp declines in weak-target detection probability; and (3) rapid attenuation of mainlobe jamming suppression efficacy.

In response to the aforementioned challenges, this paper proposes a signal–data dual-domain cooperative antijamming framework for phased-array (PA) radar, comprising two

core modules: (1) at the interference suppression level, a joint processing method based on complete ensemble empirical mode decomposition with adaptive noise combined with wavelet-optimized blind source separation (CEEMDAN-WOBSS) is developed to achieve strong jamming suppression under low-SNR conditions via noise-baseline reconstruction and signal-subspace decomposition, and (2) at the cooperative localization level, a cooperative positioning framework is devised for bistatic and multistatic radar configurations, leveraging measurement data from distributed radar-network nodes to realize three-dimensional target localization.

Blind source separation (BSS) techniques, first proposed in the 1980s, have since evolved into a mature theoretical framework in fields such as communications signal processing [26]. Their core mechanism relies on the assumption of statistical independence to recover latent source signals from mixed observations. In the radar anti-jamming domain, Ref. [27] first estimated the number of signal sources via the Gerschgorin radius method to satisfy BSS conditions, then applied the JADE algorithm—based on fourth-order cumulant-matrix approximate joint diagonalization—to separate radar-received signals. Ref. [28] constructed a mixed polarization-signal model by exploiting the polarization differences between interference signals and target echoes, achieving anti-jamming through interference reconstruction followed by BSS. Ref. [29] proposed a jointly improved BSS technology based on transmit-receive coordination, employing a dynamic particle swarm optimization algorithm to derive the optimal extraction vector for target echoes, thereby effectively suppressing multiple false-target deception jammers. To address direction-of-arrival (DOA) estimation challenges under mainlobe jamming, Ref. [30] introduced a method based on spatiotemporal multichannel BSS, jointly optimizing signal separation and phase-statistical features to obtain both range and DOA information. Ref. [31] combined sum-and-difference beamforming with subarray blind source separation and monopulse techniques to counter high-power mainlobe jamming, enabling DOA estimation of weak targets. Ref. [32] proposed a JADE-BSS and energy-detection joint approach to separate target and jamming signals in sum-and-difference beams, and subsequently extracted uncontaminated signal segments for monopulse radar DOA estimation.

All the above references pertain to monostatic radar mainlobe jamming suppression based on BSS algorithms. For distributed, multistatic radar-network systems, Ref. [33] constructed a specialized subarray structure for radar-received signals, applied BSS to separate interference and target signals, and then employed sparse signal reconstruction to jointly estimate target DOA. Ref. [34] first compensated for the relative delays of each radar's interference signals using correlation methods, then applied JADE-BSS to separate target and interference signals. Ref. [35] proposed a Maximum Signal-to-Noise Ratio BSS (MSNR-BSS) method in combination with CFAR detection to achieve signal-level anti-jamming and performed spatial localization via radial distance estimation. Gao et al. [36] extended the MSNR-BSS algorithm to frequency-diverse array multiple-input-multiple-output (FDA-MIMO) radar for effective signal separation.

In summary, BSS methods have been widely applied to radar signal processing; however, prevailing mainstream algorithms still face two major technical bottlenecks. First, the JADE algorithm—based on fourth-order cumulant-matrix joint diagonalization—requires constructing a cumulant matrix of dimension $O(N^4)$ (where N is the number of channels), resulting in high computational complexity. Second, BSS algorithms exhibit strong sensitivity to the SNR: empirical studies show that when the $\text{SNR} < 5$ dB, the signal separation error rate increases drastically, severely degrading interference suppression performance. More importantly, within the domain of distributed radar-network cooperative anti-jamming, each node's radar is more likely to receive echoes under low-SNR conditions, yet only a limited number of publications address distributed radar anti-jamming and coopera-

tive localization specifically under such conditions. To address this challenge, this study examines distributed phased-array (PA) radar performance under severe interference and low-SNR detection scenarios (typical SNR $\in [-5 \text{ dB}, 5 \text{ dB}]$). We propose a signal–data dual-domain cooperative antijamming and localization (SDCAL) framework with the following innovations:

- (1) To address the failure of traditional empirical mode decomposition (EMD) due to the nonstationary nature of PA radar-received signals, we propose a noise-adaptive complete ensemble empirical mode decomposition method combined with an improved wavelet-based signal preprocessing approach, thereby enabling modal decomposition for PA radar echoes.
- (2) For low-SNR scenarios, we introduce a CEEMDAN combined with wavelet-optimized maximum-SNR blind source separation (CEEMDAN–WOBSS) joint processing method, which enhances the separation performance of PA radar-received signals and strengthens interference suppression capabilities under low SNRs.
- (3) In order to address the cooperative localization problem in distributed radar networks under low-SNR conditions, a signal–data dual-domain cooperative antijamming and localization (SDCAL) framework is constructed to integrate signal-level interference suppression with data-level collaborative positioning, thereby achieving accurate target spatial localization.

The subsequent sections of this paper are structured as follows. Section 2 presents the received-signal model for the distributed radar system. Section 3 describes the SDCAL processing flow and provides a detailed derivation of the signal-domain antijamming algorithm. Section 4 outlines the target detection approach and establishes the data-domain cooperative localization framework under low-SNR conditions. Section 5 presents detailed numerical simulation results to validate the effectiveness of the proposed method. Finally, Section 6 summarizes the conclusions of this work.

Current research is limited to the academic field of radar signal processing for civilian sensing applications—such as environmental monitoring and disaster response—which offers clear societal benefits and does not pose a threat to public health or national security. The authors acknowledge the dual-use potential of the advanced signal-separation and localization algorithms presented here and confirm that all necessary precautions have been taken to prevent potential misuse. In accordance with relevant national and international DURC regulations, the authors advocate for responsible deployment, ethical considerations, regulatory compliance, and transparent reporting to mitigate misuse risks and foster beneficial outcomes.

2. Radar Signal Model

The radar network studied in this paper is a distributed phased-array system featuring multiple cooperative transmitters and receivers. Specifically, the network comprises m phased-array (PA) nodes, each operating as a transceiver capable of transmitting and receiving waveforms. It is assumed that all nodes achieve time synchronization through pre-coordination or network synchronization and share common waveform parameters.

As shown in Figure 1, the distributed multistatic radar system comprises m PA radars, K jammers, and P targets. Each PA radar's antenna aperture consists of S uniformly spaced linear array elements. It is assumed that both targets and jammers lie in the far-field region, with the jammers distributed around the targets. The position of the n th PA radar is denoted by (x_n^r, y_n^r, z_n^r) , the k th jammer by (x_k^j, y_k^j, z_k^j) , and the s th target by (x_s^t, y_s^t, z_s^t) .

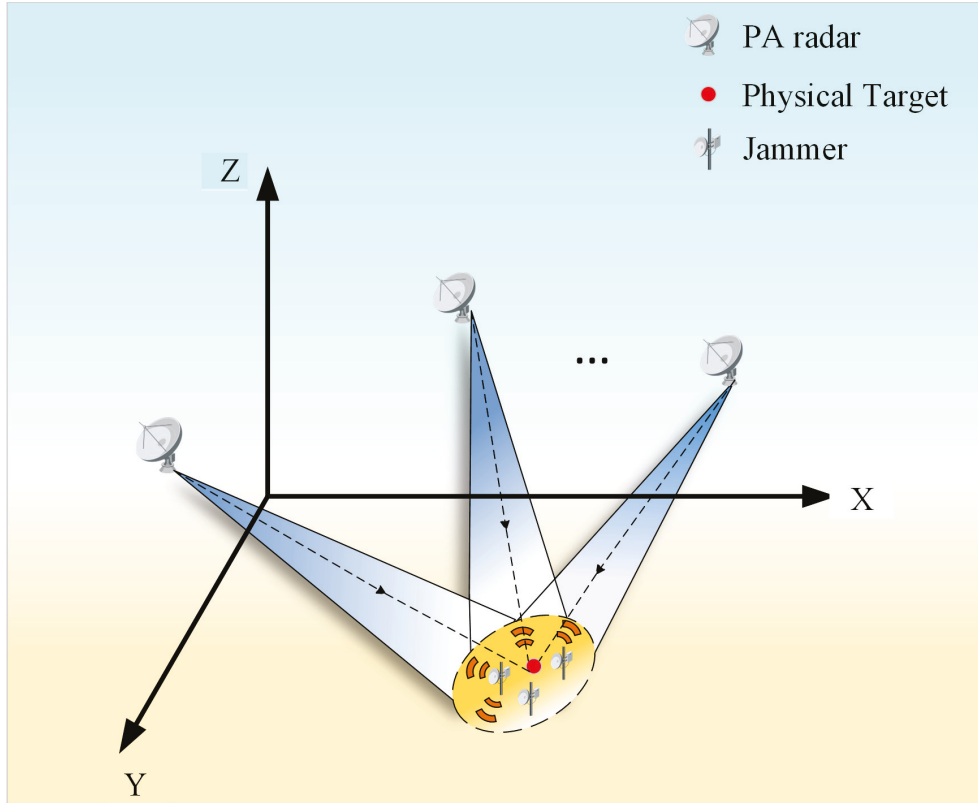


Figure 1. Schematic diagram of distributed radar detection.

Assume that the transmitted waveform from each PA radar is $S_m(t)$. Since targets are located in the far-field, the path length from a given target to all array elements of a PA radar can be approximated as equal. Therefore, the target-echo signal received at the s th element of the m th radar is expressed as

$$r_{m,s} = \sum_{p=1}^P \alpha_{m,s,p} s_p^t(t - \tau_{m,p}) \exp \left\{ -j2\pi \frac{(l-1)d}{\lambda} \sin \theta_{m,p}^t \right\} \quad (1)$$

where $s_p^t(t)$ denotes the PA radar's transmit signal, $\alpha_{m,s,p}$ is the complex amplitude of the received target echo, $\tau_{m,p}$ is the time delay of the echo at the m th radar, λ is the signal wavelength, d is the inter-element spacing, and $\theta_{m,p}^t$ is the direction-of-arrival (DOA) of the p th target echo at the m th radar.

Jammers are assumed to lie in the same far-field region and to have negligible radar cross-sections (RCSs). Consequently, the radar will receive both the target echo and the interference transmitted by the k th jammer. The interference signal received at the s th element of the m th radar is given by

$$J_{m,s} = \sum_{k=1}^K \beta_{m,s,k} s_k^j(t - \tau_{m,k}) \exp \left\{ -j2\pi \frac{(l-1)d}{\lambda} \sin \theta_{m,k}^j \right\} \quad (2)$$

where $s_k^j(t)$ denotes the k th jammer's transmit waveform, $\beta_{m,s,k}$ is the complex amplitude of the received interference, $\tau_{m,k}$ is the interference time delay at the m th radar, and $\theta_{m,k}^j$ is the DOA of the k th mainlobe jammer at the m th radar.

Let the received-signal vector at the m th PA radar be

$$\mathbf{X}_m = [x_{m,1}(t), x_{m,2}(t), \dots, x_{m,s}(t)]^T \quad (3)$$

where $x_{m,s}(t)$ represents the total signal (target echo plus interference) received at the s th element. Then $\mathbf{X}_m$ can be written in compact form as

$$\mathbf{X}_m = \mathbf{A}_m \mathbf{S}_m + \mathbf{n}_m = \mathbf{A}_R \mathbf{S}_R + \mathbf{A}_J \mathbf{S}_J + \mathbf{n}_m \quad (4)$$

where $\mathbf{A}_m$ denotes the $S \times (L + K)$ complex mixing matrix for the m th radar, $\mathbf{S}_m$ is the $(L + K)$ -dimensional vector of source signals (all target echoes and jammer waveforms) impinging on the S elements, and $\mathbf{n}_m$ is a complex Gaussian white-noise vector with zero mean and covariance σ_m^2 .

$$\mathbf{A}_m = [\mathbf{A}_R, \mathbf{A}_J] \quad (5)$$

$$\mathbf{A}_R = [\alpha_{m,s,1} \varphi^t(\theta_{m,1}), \alpha_{m,s,2} \varphi^t(\theta_{m,2}), \dots, \alpha_{m,s,P} \varphi^t(\theta_{m,P})] \quad (6)$$

$$\mathbf{A}_J = [\beta_{m,s,1} \varphi^j(\theta_{m,1}), \beta_{m,s,2} \varphi^j(\theta_{m,2}), \dots, \beta_{m,s,K} \varphi^j(\theta_{m,K})] \quad (7)$$

$$\varphi^t(\theta_{m,p}) = \exp \left\{ j2\pi \frac{(l-1)d}{\lambda} \sin \theta_{m,p}^t \right\} \quad (8)$$

$$\varphi^j(\theta_{m,k}) = \exp \left\{ j2\pi \frac{(l-1)d}{\lambda} \sin \theta_{m,k}^j \right\} \quad (9)$$

$$\mathbf{S}_m = [\mathbf{S}_R(t), \mathbf{S}_J(t)]^T = \begin{bmatrix} s_1^r(t - \tau_{n,1}) \\ s_2^r(t - \tau_{n,2}) \\ \vdots \\ s_P^r(t - \tau_{n,P}) \\ s_1^j(t - \tau_{n,1}) \\ s_2^j(t - \tau_{n,2}) \\ \vdots \\ s_K^j(t - \tau_{n,K}) \end{bmatrix} \quad (10)$$

$$\mathbf{n}_m = \begin{bmatrix} n_{m,1}(t) \\ n_{m,2}(t) \\ \vdots \\ n_{m,S} \end{bmatrix} \quad (11)$$

where $(\cdot)^T$ is the transpose operation.

3. Signal-Domain Antijamming for Distributed Radar

3.1. Algorithm Flow

To address interference suppression and cooperative localization in distributed radar networks under low-SNR conditions, the SDCAL framework is proposed, as illustrated in Figure 2. This framework consists of two stages: signal-level interference suppression at each networked radar node and data-level cooperative localization across the distributed radar network. The algorithm process is as follows:

1. Perform mean-subtraction processing on the PA radar reception signals from each node in the distributed radar system to eliminate DC components, followed by pre-whitening processing to decorrelate the components of the reception signals.
2. Apply joint CEEMDAN-WOBSS processing to the echo signals received by the PA radar, then calculate the separation matrix based on the maximum signal-to-noise ratio criterion and perform blind source separation on the whitened signals.
3. CFAR detection is performed on all channel signals after pulse compression, and the detected signal channels are superimposed to obtain the radial distance of the target.

Finally, the spatial position of the target is jointly calculated using the radial distances of the target from two or three PA radars in the distributed radar network.

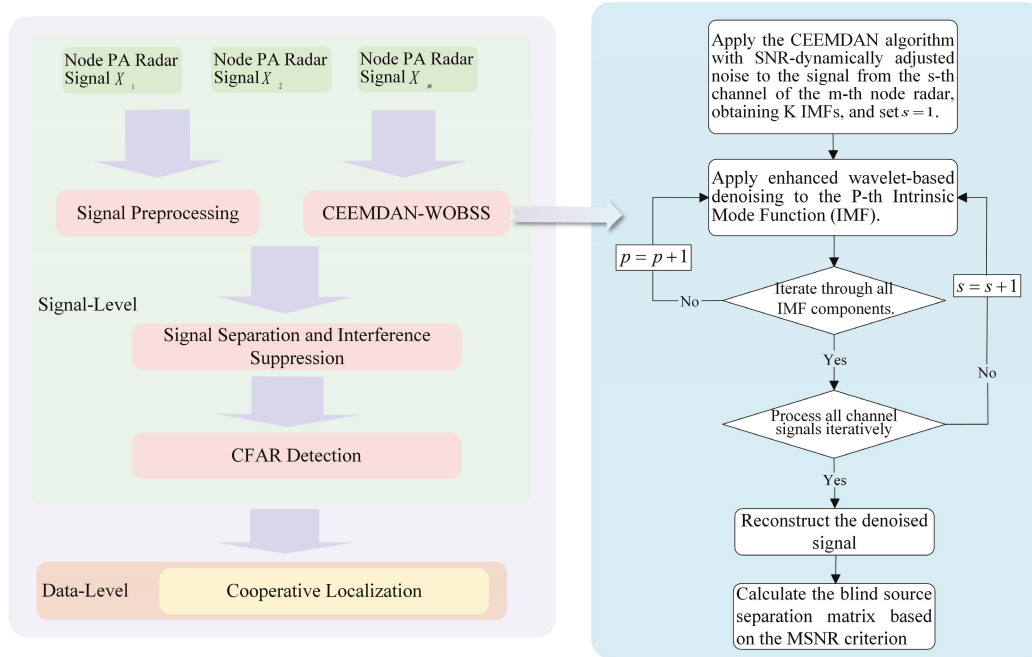


Figure 2. Signal–data cooperative antijamming and localization framework (SDCALF).

3.2. PA Radar-Received Signal Preprocessing

BSS requires the mixing matrix $\mathbf{A}_m$ to be of full column rank. Equivalently, the number of receive channels S must exceed the total number of source signals, and at most, one source may be Gaussian, while all source signals must be statistically independent. In the distributed multistatic radar system considered here, each single-site PA radar has S receive elements, the detection region contains K jammers and P targets, and it is assumed that $S > P + K$. Since mainlobe suppression jamming is modeled as noise-modulated interference, the interference and target echo signals are statistically independent; the Gaussian white noise component serves as the unique Gaussian source.

Let each PA radar have S receive antennas. Suppose there are P real targets and K jammers in the detection region, with $S > P + K$. The received signal at the m th PA radar can then be written in matrix form as

$$\mathbf{X}_m = \mathbf{A}_m \mathbf{S}_m^o + \mathbf{n}_m \quad (12)$$

where

- $\mathbf{X}_m$ is the $S \times D$ received-signal matrix, $\mathbf{X}_m = [x_1^m(t), x_2^m(t), \dots, x_s^m(t)]$, and D denotes the number of fast-time samples;
- $\mathbf{S}_m^o$ is the $(P + K) \times D$ source-signal matrix containing all target-echo and jammer signals impinging on the S array elements;
- $\mathbf{A}_m$ is the $S \times (P + K)$ mixing matrix for the m th radar;
- $\mathbf{n}_m$ is the $S \times D$ complex Gaussian white-noise matrix with zero mean and covariance σ_m^2 .

Prior to blind source separation, the received signals in each element channel of the PA radar undergo zero-mean centering (de-meaning) and whitening. De-meaning (centering) removes any DC component from $\mathbf{X}_m$.

$$\tilde{\mathbf{X}}_m = \mathbf{X}_m - E[\mathbf{X}_m] \quad (13)$$

where $E[\cdot]$ denotes the expectation operator. Whitening (pre-whitening) then decorrelates the zero-mean data. Let $\mathbf{C}_{XX}$ be the covariance matrix of $\tilde{\mathbf{X}}_m$. Denote by $\mathbf{U}$ the Sn matrix whose columns are the n largest eigenvectors of $\mathbf{C}_{XX}$, and let $\Lambda = \text{diag}(d_1, d_2, \dots, d_n)$ be the corresponding diagonal matrix of eigenvalues. The whitening matrix $\mathbf{Q}$ is then given by

$$\mathbf{Q} = \Lambda^{-1/2} \mathbf{U}^H \quad (14)$$

and the whitened observation matrix is

$$\mathbf{X}'_m = \mathbf{Q} \tilde{\mathbf{X}}_m \quad (15)$$

Once the preprocessed observation $\mathbf{X}'_m$ is obtained, the separation matrix $\mathbf{W}$ is computed according to a chosen BSS algorithm, and the target echo is extracted by applying $\mathbf{W}$ to the observations of the original PA radar's individual antenna element channel signals. In the MSNR-BSS algorithm, the SNR criterion is formulated as an error-minimization cost function:

$$J(\mathbf{w}) = -\frac{\mathbb{E}[|\mathbf{w}^T \mathbf{x}|^2]}{\mathbb{E}[|\mathbf{w}^T \mathbf{n}|^2]} \quad (16)$$

where $\mathbf{W}$ is a separation vector, $\mathbf{x}$ is the mixed observation, and $\mathbf{n}$ is the noise component. It can be shown that the convergence of $\mathbf{W}$ depends directly on the whitened signal's SNR level. Moreover, the whitening matrix $\mathbf{Q}$ depends on the second-order statistics of the received signal:

$$\mathbf{C}_{XX} = E[\mathbf{X}'_m \mathbf{X}'_m{}^H] = \mathbf{A} \mathbf{C}_s \mathbf{A}^H + \sigma^2 \mathbf{I}, \quad (17)$$

where $\mathbf{C}_s = E[\mathbf{s} \mathbf{s}^H]$ is the source-signal covariance matrix, $\mathbf{A}$ the mixing matrix, $\mathbf{A}^H$ is its Hermitian transpose, σ^2 is the noise power, and $\mathbf{I}$ is the identity matrix.

Under low-SNR conditions—i.e., $\sigma^2 \gg \|\mathbf{A} \mathbf{C}_s \mathbf{A}^H\|$ —the whitening process attenuates the target-signal subspace in $\mathbf{X}'_m$, causing $\mathbf{W}$ to converge to a suboptimal solution and preventing effective separation of signal and interference subspaces. To address this, the proposed CEEMDAN-WOBSS method attenuates the noise-baseline power in the whitened observations and reconstructs the rank characteristics of the signal covariance matrix, thereby enabling accurate discrimination and extraction of both target and interference components.

3.3. Signal Separation and Interference Suppression

Each received signal sample $x^{(s)}(t)$ at the s th array element of every PA radar node undergoes CEEMDAN-based decomposition with iterative added noise. Denote by $w^i(t)$ the white-noise sequence added at the i th iteration, whose standard deviation is adaptively chosen to achieve a target SNR level β_i . The noisy observation sequence is formed as

$$x_k^{(s)}(t) = x^{(s)}(t) + \beta_i D_k[w^i(t)] \quad (18)$$

where $D_k(\cdot)$ extracts the k th intrinsic mode function (IMF) via an M-order EMD. The dynamic coefficient β_i is computed from the ratio of signal-to-noise energy in the residual:

$$\beta_i = \begin{cases} \varepsilon \cdot \frac{x^{(s)}(t) x^{(s)}(t)^T}{(w^{(i)})(w^{(i)})^T}, & \text{if } i = 0 \\ \varepsilon \cdot \frac{R_i(t) R_i(t)^H}{(w^{(i)})(w^{(i)})^H}, & \text{if } i \geq 1 \end{cases} \quad (19)$$

where $R_i(t)$ is the residual after the $(i - 1)$ th decomposition, ε is the initial noise amplitude, ε is typically set to 20% of the standard deviation of the received signal to ensure that

noise injection effectively aids modal decomposition without excessively contaminating the original signal characteristics.

After forming M noisy realizations, a first-order decomposition yields M IMF estimates. Their ensemble average defines the first-order IMF of the original signal:

$$\overline{\text{IMF}}_1^{(s)} = \frac{1}{M} \sum_{k=1}^M D_k[x^{(s)}(t) + \beta_i D_k[w^{(i)}(t)]] \quad (20)$$

The first residual is then

$$R^{(1)}(t) = x^{(s)}(t) - \overline{\text{IMF}}_1^{(s)}(t) \quad (21)$$

By injecting new noise into $R^{(1)}$ and repeating the above steps, all IMF orders can be extracted. A time-varying noise power strategy is implemented using the dynamic parameter β_i —maintaining a high noise level in the early (high-frequency) stages and gradually attenuating it in the later (low-frequency) stages—to suppress mode aliasing and minimize residual noise in the final reconstruction. Previous studies have confirmed that each IMF retains complete echo signal information, so all IMFs are retained to achieve accurate signal reconstruction.

In order to enhance BSS performance under low-SNR conditions, a multi-element parallel processing scheme is adopted. Each IMF is processed via improved wavelet-threshold denoising:

$$\hat{w}_{j,k} = \begin{cases} \text{sgn}(w_{j,k}) \left(|w_{j,k}| - \frac{\lambda}{\alpha(w_{j,k}-\lambda)^2/\lambda+1} \right), & \text{if } |w_{j,k}| \geq \lambda \\ 0, & \text{if } |w_{j,k}| < \lambda \end{cases} \quad (22)$$

where $\alpha > 1$ balances denoising strength against signal fidelity. The threshold λ was set according to Donoho's universal-threshold rule, i.e.,

$$\lambda = \sigma \sqrt{2 \ln(N)},$$

where σ denotes the estimated noise standard deviation and N is the signal length. This threshold is continuous over its domain and reduces to identity for large amplitudes, thus avoiding the constant-bias drawback of soft-thresholding while preserving oscillation suppression.

All denoised IMFs are then recombined:

$$\hat{x}^{(s)}(t) = \sum_{k=1}^M R(\text{IMF}_k^{(s)}) \quad (23)$$

where $R(\cdot)$ denotes the inverse wavelet reconstruction operator. After zero-mean centering and whitening, the matrix of preprocessed signals $\hat{X}'$ forms the input to MSNR-BSS. Define

$$\text{SNR} = 10 \log \left(\frac{\mathbf{W} \hat{\mathbf{X}}' \cdot (\hat{\mathbf{X}}')^H \mathbf{W}^T}{\mathbf{W} (\hat{\mathbf{X}}' - \mathbf{X}') \cdot (\hat{\mathbf{X}}' - \mathbf{X}')^H \mathbf{W}^T} \right) \quad (24)$$

where $\mathbf{C} = \hat{\mathbf{X}}' \cdot (\hat{\mathbf{X}}')^H$, $\hat{\mathbf{C}} = (\hat{\mathbf{X}}' - \mathbf{X}') \cdot (\hat{\mathbf{X}}' - \mathbf{X}')^H$, $\mathbf{V} = \mathbf{W} \mathbf{C} \mathbf{W}^T$ and $\mathbf{U} = \mathbf{W} \hat{\mathbf{C}} \mathbf{W}^T$. Setting $\frac{\partial F}{\partial \mathbf{W}} = 0$ yields the eigenvalue problem

$$\mathbf{W} \mathbf{C} = \frac{\mathbf{V}}{\mathbf{U}} \mathbf{W} \mathbf{C} \quad (25)$$

The eigenvectors of $\hat{\mathbf{C}} \cdot \mathbf{C}^{-1}$ (rows of $\mathbf{W}$) will separate the sources [37]. The estimated source matrix is

$$\mathbf{Y} = \mathbf{W}\mathbf{X}' \quad (26)$$

and each channel $y_{m,l}(t)$ is pulse-compressed against the transmitted waveform $s_l^r(t)$:

$$g_{m,l}(t) = \int y_{m,l}(\tau) s_l^r(t - \tau) d\tau \quad (27)$$

further suppressing residual interference and yielding high-fidelity target echoes.

4. Target Detection and Cooperative Localization

Following signal separation and pulse compression, each signal channel contains an estimated target echo and corresponding interference components. Target detection is then performed independently on each channel using a Constant False Alarm Rate (CFAR) algorithm. In high-interference environments, however, detection results may exhibit significant deviations across channels due to varying local SNR conditions. To mitigate the potential accumulation of detection errors when aggregating across channels, the range measurements of each detected target are averaged over all signal channels to yield a more robust estimation of radial distance:

$$D_m^l = \frac{1}{S} \sum_{i=1}^S d_{m,i}^l \quad (28)$$

where D_m^l denotes the estimated radial distance from the m th PA radar to the l th target, and $d_{m,i}^l$ is the radial distance of the l th target as detected by the i th signal channel of the m th radar.

4.1. Cooperative Localization

Under low-SNR conditions, it is possible that only a subset of the distributed radar nodes detect the target, which necessitates adaptive localization strategies based on available detection information. Two cases are considered: localization using two radars, and localization using three or more radars.

Case 1: Localization Using Two PA Radars

Assume two PA radars located at (x_{d1}, y_{d1}, z_{d1}) and (x_{d2}, y_{d2}, z_{d2}) , respectively, detect the target. Let the measured radial distances be d_1 and d_2 , and the estimated DOA be θ_1 and θ_2 . The target's 3D coordinates (x_0, y_0, z_0) can be obtained by solving the following system of nonlinear equations:

$$\begin{cases} (x_{d1} - x_0)^2 + (y_{d1} - y_0)^2 + (z_{d1} - z_0)^2 = d_1^2 \\ (x_{d2} - x_0)^2 + (y_{d2} - y_0)^2 + (z_{d2} - z_0)^2 = d_2^2 \\ \theta_1 = \arctan\left(\frac{y_0 - y_{d1}}{x_0 - x_{d1}}\right) \\ \theta_2 = \arctan\left(\frac{y_0 - y_{d2}}{x_0 - x_{d2}}\right) \end{cases} \quad (29)$$

This system can be solved using iterative methods such as the Gauss–Newton algorithm to estimate the target position (x_0, y_0, z_0) .

Case 2: Localization Using Three PA Radars

Assume that three PA radars located at (x_{d1}, y_{d1}, z_{d1}) , (x_{d2}, y_{d2}, z_{d2}) and (x_{d3}, y_{d3}, z_{d3}) detect the same target, and the corresponding radial distances are d_1, d_2 and d_3 , respectively.

The spatial coordinates of the target (x_0, y_0, z_0) can then be estimated by solving the following trilateration equations:

$$\begin{cases} (x_{d1} - x_0)^2 + (y_{d1} - y_0)^2 + (z_{d1} - z_0)^2 = d_1^2 \\ (x_{d2} - x_0)^2 + (y_{d2} - y_0)^2 + (z_{d2} - z_0)^2 = d_2^2 \\ (x_{d3} - x_0)^2 + (y_{d3} - y_0)^2 + (z_{d3} - z_0)^2 = d_3^2 \end{cases} \quad (30)$$

Solving this overdetermined nonlinear system yields the 3D coordinates of the target, which can be efficiently computed using least-squares or iterative methods.

4.2. Downgrade Strategy

Collaborative ranging requires at least two independent radial measurements or two or more independent arrival direction measurements to obtain a three-dimensional position estimate. If fewer than two valid radial measurements are available, SDCAL employs a hierarchical fallback strategy: (i) attempt triangulation based on arrival directions when sufficient arrival directions exist, though with reduced reliability and (ii) return a “no positioning” state if a reliable geometric solution cannot be obtained.

This collaborative localization approach enhances the robustness and accuracy of target position estimation, particularly in scenarios involving partial detection or measurement errors among individual radar nodes. The complete algorithm flow is summarized in Algorithm 1.

Algorithm 1: Signal–data cooperative antijamming and localization (SDCAL)

Input : Received data $\{X_m \in \mathbb{C}^{S \times D}\}_{m=1}^M$;
CEEMDAN–WOBSS parameters; CFAR and localization settings

Output: Radial distances $\{D_m\}_{m=1}^M$; 3D target position (x_0, y_0, z_0)

```

1 for  $m = 1$  to  $M$  do
2   De-mean and whiten  $X_m \rightarrow X'_m$ 
3   for each channel  $s = 1$  to  $S$  do
4     Apply adaptive CEEMDAN to  $X'_m(s, :)$  and decompose into IMFs
5     Apply improved wavelet threshold denoising to each IMF
6     Reconstruct denoised signal  $\hat{x}_s$  from all IMFs
7     Store  $\hat{x}_s$  into  $\hat{X}'_m$ 
8   Compute separation matrix  $W$  via maximum SNR criterion using  $\hat{X}'_m$ 
9   Obtain separated sources  $Y_m = W \cdot \hat{X}'_m$ 
10 for  $m = 1$  to  $M$  do
11   for each channel  $s = 1$  to  $S$  do
12     Apply pulse compression and CFAR to  $Y_m(s, :)$ 
13     if target detected then
14       Extract range  $d_{m,s}$ 
15     else
16       Set  $d_{m,s} \leftarrow \text{NaN}$ 
17   Compute  $D_m \leftarrow \text{mean}(\{d_{m,s} \neq \text{NaN}\})$ 
18 if exactly two nodes have valid  $D_m$  then
19   Initialize  $(x_0, y_0, z_0)$ 
20   while not converged do
21     Update  $(x_0, y_0, z_0)$  via Gauss–Newton on the two-radar range equations
22 else if three or more nodes have valid  $D_m$  then
23   Compute  $(x_0, y_0, z_0)$  directly via three-radar position-settlement equations
24 return  $\{D_m\}_{m=1}^M, (x_0, y_0, z_0)$ 

```

It should be emphasized that the SDCAL framework presented here integrates per-unit CEEMDAN decomposition and wavelet denoising, local multi-channel whitening followed by MSNR-BSS, pulse compression with CFAR detection, and centralized range fusion. The dominant computational burden arises from the local covariance and eigenvalue computations performed during CEEMDAN preprocessing and BSS. In practical deployments these operations are highly parallelizable across array elements and nodes, and the BSS at each PA node operates in a low-dimensional space. Therefore, with appropriate parameter tuning and hardware acceleration, real-time or near-real-time operation is attainable.

5. Simulation Results

The CEEMDAN-WOBSS + multi-channel BSS processing pipeline proposed in this work is effective provided the targets to be resolved lie within the physical resolvability limits imposed by the radar waveform and array geometry. Two primary, physics-driven resolution constraints should be noted.

First, the range resolution is governed by the transmitted bandwidth B and can be approximated by

$$\Delta R \approx \frac{c}{2B}, \quad (31)$$

where c is the speed of light. When the radial separation between two targets is substantially smaller than ΔR , their post-pulse-compression echoes overlap in range, and time-domain separation becomes intrinsically difficult. Second, the angle (DOA) resolution is constrained by the array effective aperture L (or equivalently by the number of elements S and inter-element spacing d), with a first-order estimate

$$\Delta \theta \approx \frac{\lambda}{L} \approx \frac{\lambda}{Sd}, \quad (32)$$

where λ is the carrier wavelength. Targets whose angular separation is significantly below $\Delta \theta$ will be difficult to distinguish by array-based spatial processing alone.

In addition to these geometric and waveform limits, source coherence degrades separability: highly coherent targets reduce the effectiveness of BSS and subspace methods. CEEMDAN preprocessing improves the effective SNR and helps to restore the rank structure of the signal covariance, but it cannot overcome the fundamental physical limits above.

All numerical experiments reported in this manuscript were conducted within the above resolvability constraints.

To evaluate the performance of the proposed method, a simulation scenario was constructed involving a distributed radar network composed of three PA radars and a single target of interest. The spatial coordinates of the radar nodes and the target are provided in Table 1. Each PA radar is equipped with six receiving elements, uniformly spaced at half-wavelength intervals. The transmitted signal is a linear frequency-modulated (LFM) waveform, expressed as:

$$s(t) = \exp \left\{ j2\pi \left(f_c t + \frac{1}{2} u t^2 \right) \right\} \quad (33)$$

where f_c is the carrier frequency and u is the modulation rate. The signal parameters are detailed in Table 2. The directions of arrival (DOAs) of the target and the interference are set to 10° and 11° .

This paper defines two types of main lobe interference: NAM interference and SMSP interference. The NAM interference signal model is as follows:

$$J_{\text{NAM}} = (U_0 + U_n(t)) \exp \{ j(2\pi f_j t + \varphi) \} \quad (34)$$

where U_0 denotes the carrier amplitude, $U_n(t)$ is a zero-mean noise process with variance $\sigma^2 = 1$, ϕ is the initial phase, and f_j is the carrier frequency of the NAM signal, set to 16 GHz.

Table 1. PA radar spatial coordinates.

Radar Serial Number	Position Coordinates (m)
PA Radar 1	8000, 22,000, 15,000
PA Radar 2	10,000, 20,000, 15,000
PA Radar 3	8000, 24,000, 15,000
Target location	15,000, 35,000, 0
Interference location	15,500, 35,600, 0

The SMSP interference is modeled as a sequence of time-shifted subpulses:

$$J_{\text{SMSP}} = \sum_{i=0}^I J_i \left(t - \frac{iT_p}{I} \right) \quad (35)$$

where I denotes the number of subpulses and T_p is the radar pulse duration. Each subpulse $J_i(t)$ is a short linear FM (LFM) burst defined on the interval $0 \leq t \leq \frac{T_p}{I}$.

$$J_i(t) = \exp \left\{ j2\pi \left(f_j t + \frac{1}{2} \mu' t^2 \right) \right\}, \quad 0 \leq t \leq \frac{T_p}{I}, \quad (36)$$

where $\mu' = I\mu$ and μ is the linear modulation rate of the radar transmit waveform. Thus, each SMSP subpulse has the same instantaneous bandwidth as the radar LFM waveform, but a pulse duration reduced by a factor of $\frac{1}{I}$ and an effective chirp rate increased by a factor of I . In the simulations, the SMSP carrier frequency is set to $f_j = 16$ GHz and the number of subpulses is $I = 5$. The SNR is set to 5 dB, the JNR is 50 dB, and the number of fast-time snapshots is 4096.

Table 2. Radar parameter settings.

Signal Parameter	Parameter Value
Bandwidth/B	20 MHz
Carrier Frequency/Fc	16 GHz
Pulse Repetition Frequency/PRF	8 MHz
Sampling Frequency/Fs	80 MHz
Number of Sampling Points	4096
Pulse Width/Tp	25 μ s
CEEMDAN decomposition level	10
Wavelet basis	sym8
Wavelet decomposition level	6
balancing parameter α	2
CFAR type	CA-CFAR
CFAR false alarm probability	10^{-5}
Number of reference units	50
Number of protection units	5

To demonstrate the signal separation performance, Radar 1 is used as an example. In the proposed CEEMDAN-WOBSS method, CEEMDAN decomposition is performed up to 10 layers, and a six-level decomposition is performed using the improved wavelet basis function sym8. The target echo, NAM interference, SMSP interference, and radar-received signals are shown in Figure 3.

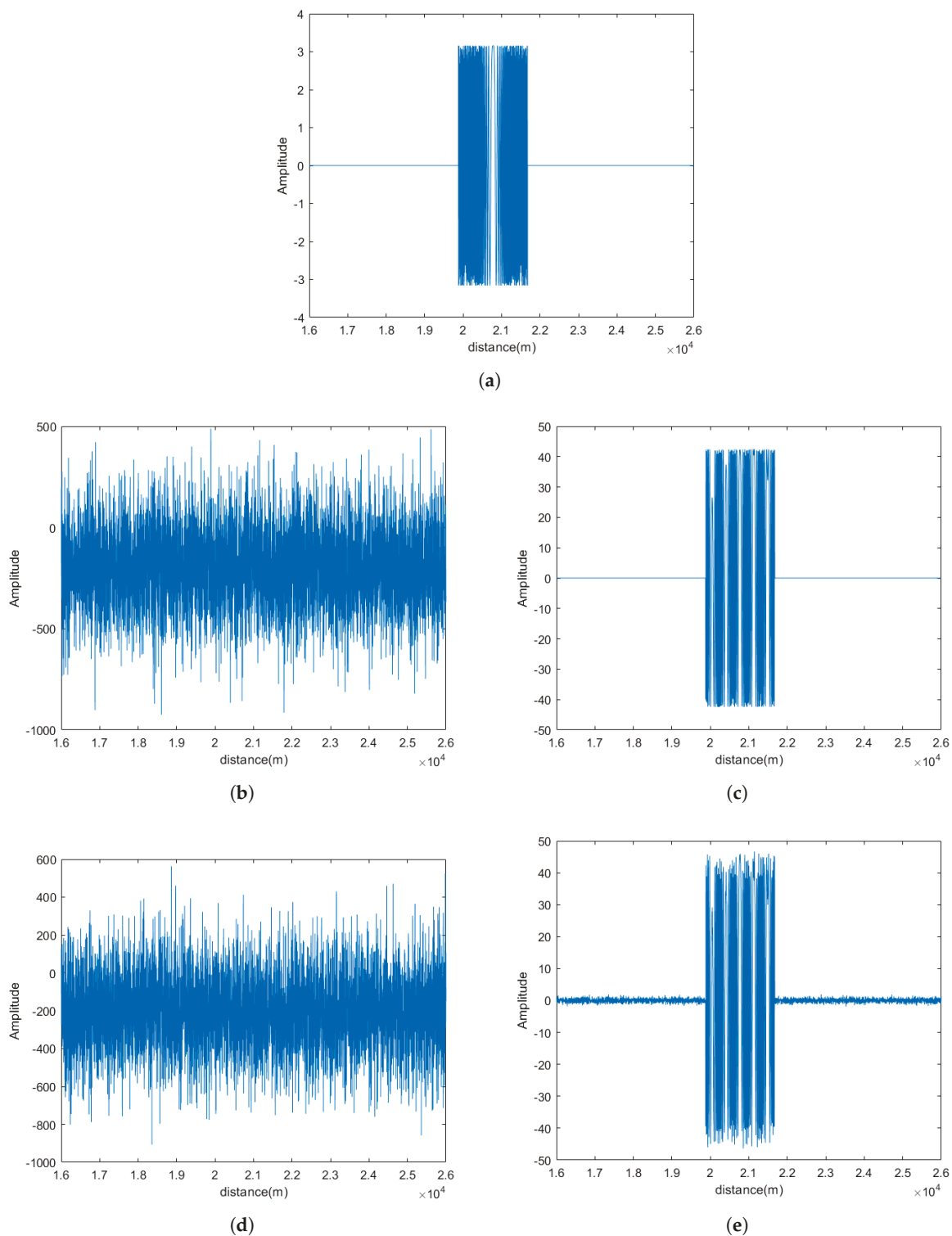


Figure 3. PA radar simulation signal model: (a) target echo signal; (b) NAM jamming; (c) SMSP jamming; (d) radar-received signal (NAM); (e) radar-received signal (SMSP).

Figure 4 depicts the pulse compression and CFAR detection results for a single channel after signal separation, while Figure 5 illustrates the results across all receiving channels of Radar 1. As seen from the figures, under SNR = 5 dB, the proposed method effectively suppresses strong electromagnetic interference and successfully separates target signals in multiple channels.

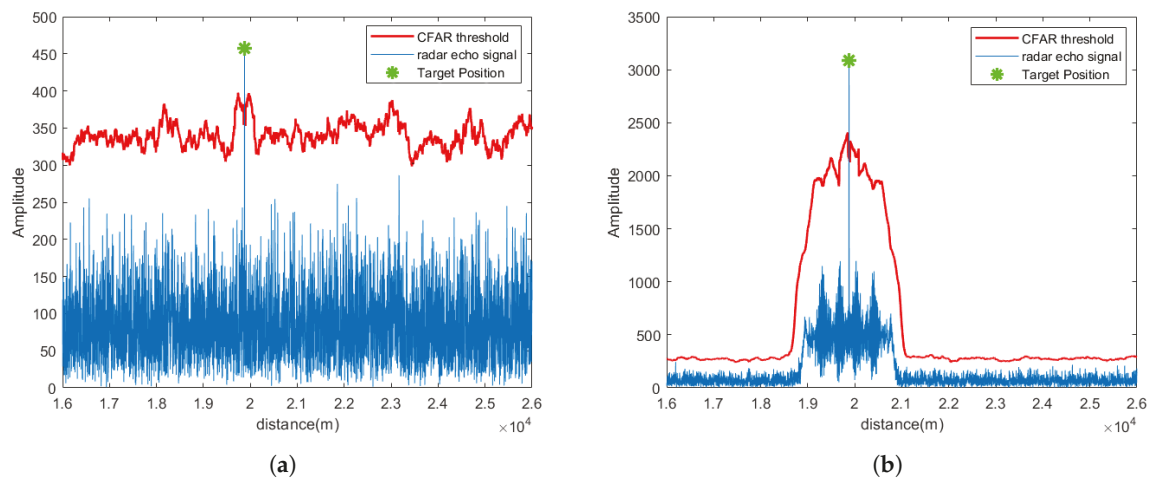


Figure 4. After the signal received by a single element of Radar 1 is separated by CEEMDAN-WOBSS, the pulse pressure and CFAR detection results are obtained: (a) under NAM interference conditions; (b) under SMSP interference conditions.

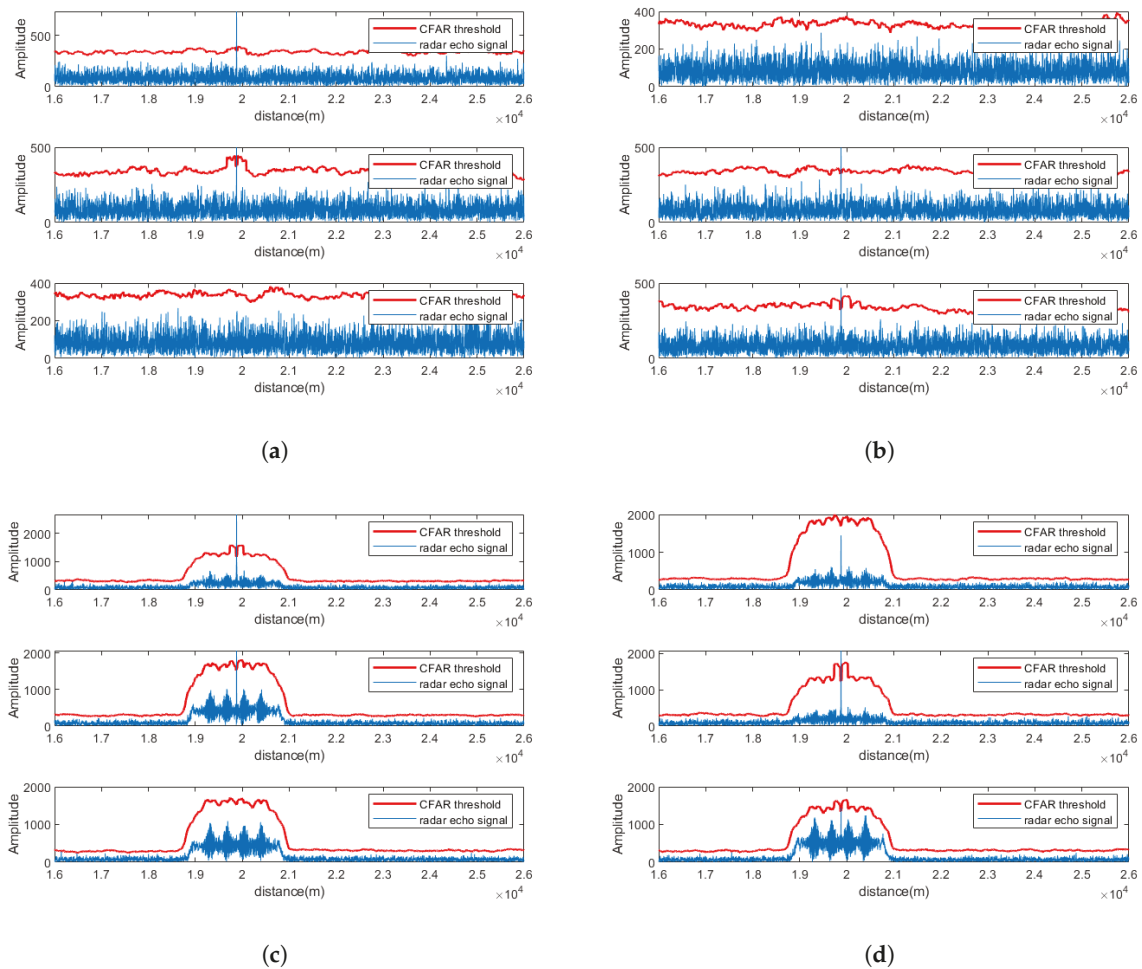


Figure 5. After the received signals from each array element channel of the PA radar are separated by CEEMDAN-WOBSS, the pulse compression and CFAR detection results are as follows: (a,b) Additional checks performed: results for channels 1–6 under NAM interference conditions; (c,d) results for channels 1–6 under SMP interference conditions.

After applying the CEEMDAN-WOBSS algorithm and CFAR detection to all three PA radars in the distributed radar network, the estimated radial distances of the target under both interference scenarios are listed in Table 3, with a maximum range error of only 2.5 m. Using the position estimation formulas discussed earlier, the 3D target location is calculated, as shown in Figure 6.

Table 3. Distributed radar network node radar positioning distance.

Radar Number	Actual Distance (m)	Estimated Distance (NAM) (m)	Estimated Distance (SMSP) (m)
PA Radar 1	21,048	21,050.1	21,049.9
PA Radar 2	21,794	21,796.4	21,796.1
PA Radar 3	19,875	19,877.5	198,877.4

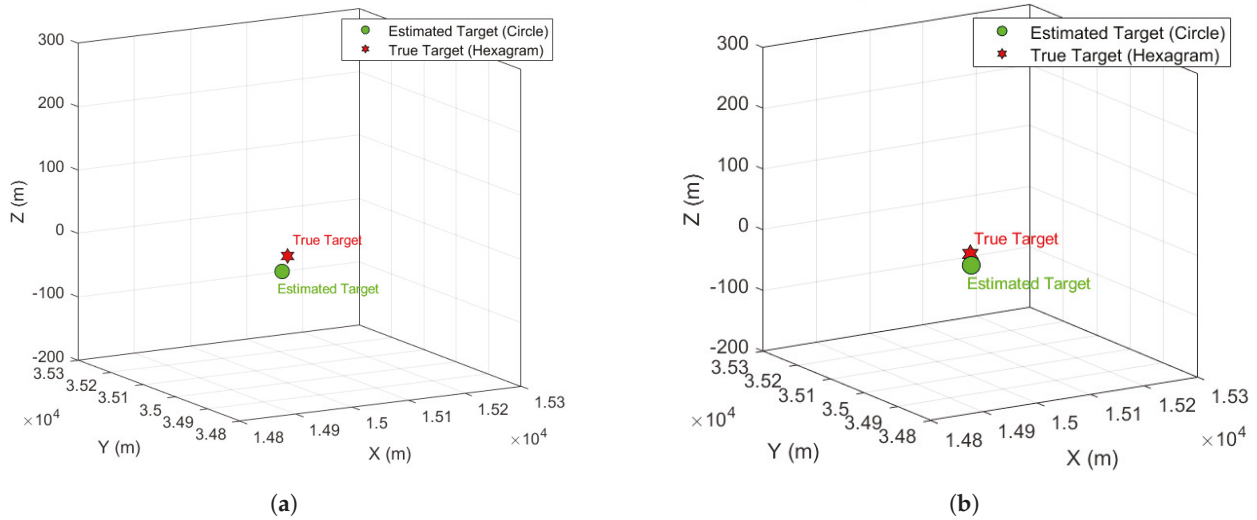


Figure 6. Spatial positioning results: (a) under NAM interference conditions; (b) under SMSP interference conditions.

In practical radar detection environments, the received SNR often ranges from -5 dB to 5 dB. Thus, this paper compares the interference suppression performance of the proposed CEEMDAN-WOBSS method against MSNR-BSS under $\text{JSR} = 50$ dB and $\text{SNR} \in [-5 \text{ dB}, 5 \text{ dB}]$. The comparison includes evaluations of the peak side-lobe ratio (PSLR), target detection probability, and localization accuracy across various SNR levels. Additionally, the JADE-BSS method and MSNR-BSS are used as a baseline for benchmarking.

5.1. Signal Separation Performance at $\text{SNR} = -5$ dB

Using Radar 1 as an example, the pulse compression and CFAR detection results of CEEMDAN-WOBSS under two interference are shown in Figure 7a–d, respectively. For comparison, the results of the MSNR-BSS method are shown in Figure 7e,f. The results indicate that under low-SNR conditions (-5 dB), MSNR-BSS fails to recover the target echo in any channel, whereas CEEMDAN-WOBSS successfully separates and detects the target signal in several channels, effectively suppressing the NAM interference.

5.2. Signal Separation Performance at $\text{SNR} = 0$ dB

Figure 8 presents the pulse compression and CFAR results for the CEEMDAN-WOBSS method under two interference scenarios at 0 dB SNR. The corresponding results for MSNR-BSS are shown in Figure 8e,f. The simulation results demonstrate that CEEMDAN-WOBSS continues to effectively isolate and detect the target echo even at an $\text{SNR} = 0$ dB.

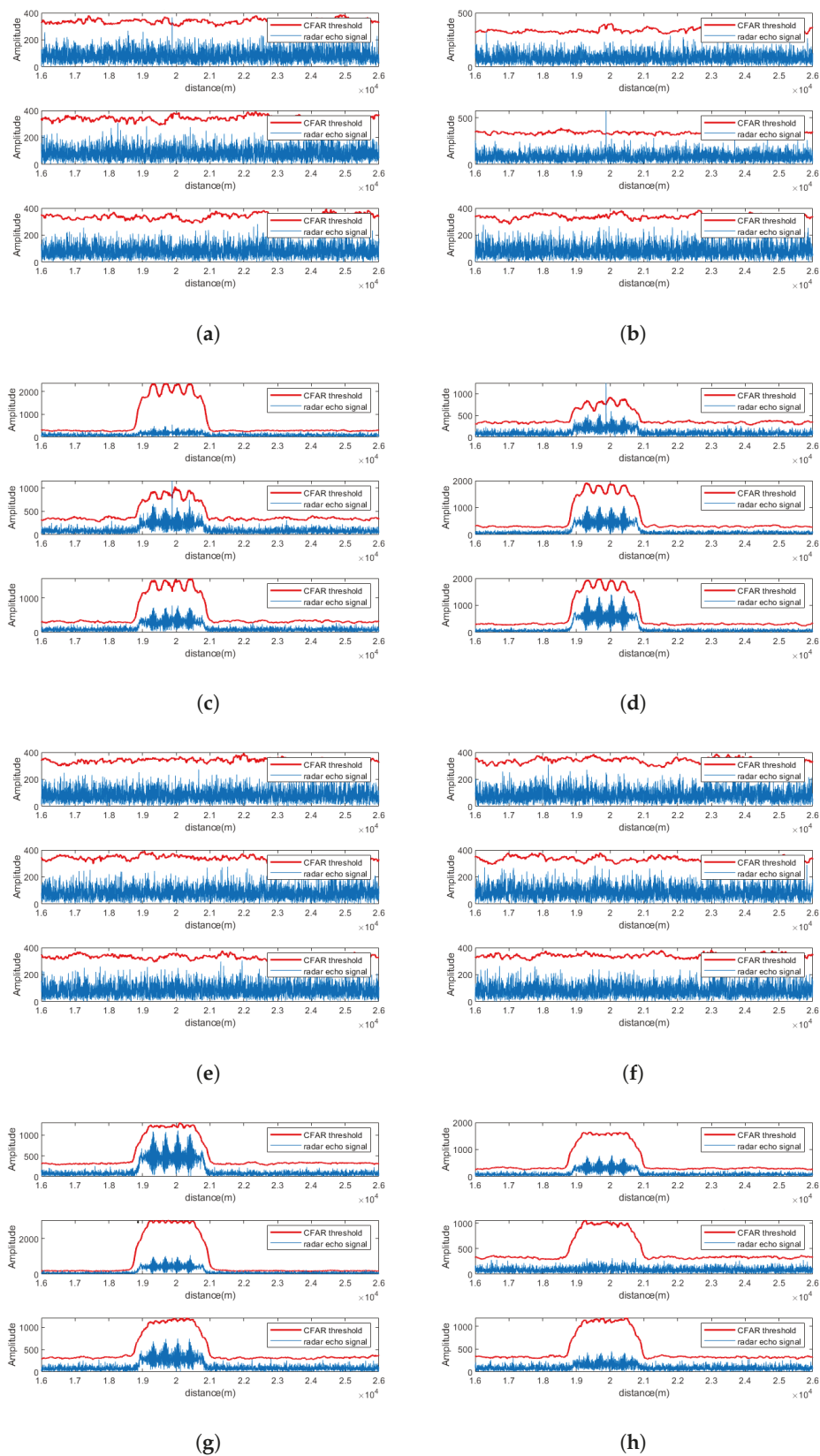


Figure 7. SNR = -5 dB. PA radar 1 received a signal after separation by the CEEMDAN-WOBSS method and MSNR-BSS method; the pulse compression and CFAR detection results are presented for the following: (a,b) Additional checks performed: CEEMDAN-WOBSS method, channels 1–6 (NAM); (c,d) CEEMDAN-WOBSS method, channels 1–6 (SMSP); (e,f) MSNR-BSS method, channels 1–6 (NAM); (g,h) MSNR-BSS method, channels 1–6 (SMSP).

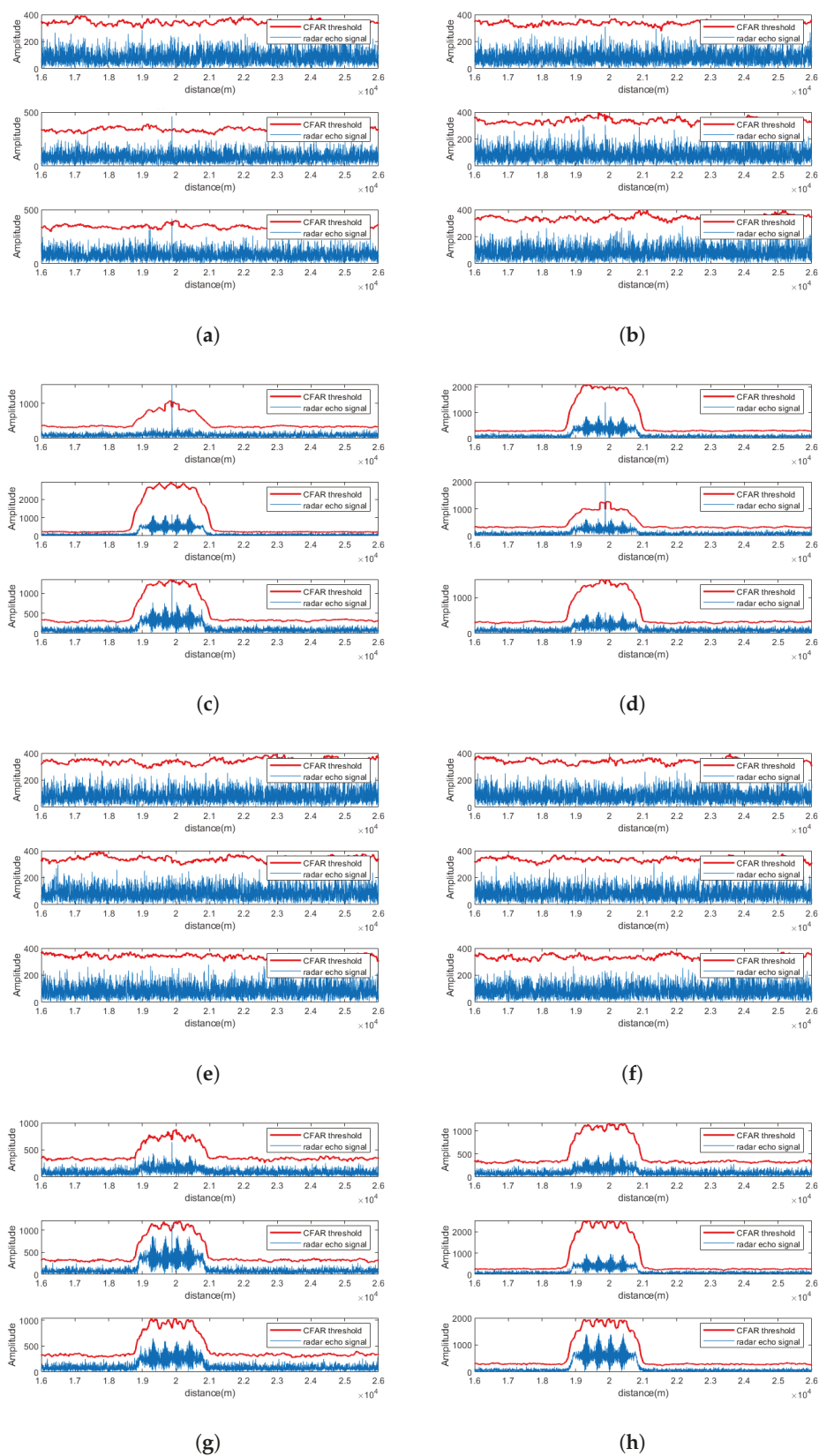


Figure 8. SNR = 0 dB. PA radar 1 received a signal after separation by the CEEMDAN-WOBSS method and MSNR-BSS method; the pulse compression and CFAR detection results are presented for the following: (a,b) Additional checks performed: CEEMDAN-WOBSS method, channels 1–6 (NAM); (c,d) CEEMDAN-WOBSS method, channels 1–6 (SMSP); (e,f) MSNR-BSS method, channels 1–6 (NAM); (g,h) MSNR-BSS method, channels 1–6 (SMSP).

5.3. Signal Separation Performance at SNR = 5 dB

Figure 9 show the pulse compression and CFAR detection results of CEEMDAN-WOBSS and MSNR-BSS under two interference scenarios at a 5 dB SNR. At this level, both methods can detect the target echo in multiple channels; however, CEEMDAN-WOBSS outperforms MSNR-BSS by recovering the target signal in more channels and providing better interference suppression.

5.4. Quantitative Evaluation of Effectiveness

To quantify the performance of the CEEMDAN-WOBSS method, three metrics were adopted under the setting of JSR = 50 dB and $\text{SNR} \in [-5 \text{ dB}, 5 \text{ dB}]$:

1. Peak side-lobe ratio (PSLR): The PSLR is defined as

$$\text{PSLR} = 20 \log_{10} \left(\frac{\text{SLP}}{\text{MLP}} \right) \quad (37)$$

where SLP is the peak amplitude of the side lobe and MLP is that of the main lobe. The average PSLR across N Monte Carlo trials is given by

$$\text{PSLR}_N = \mathbb{E}(\text{PSLR}_n) \quad (38)$$

where $\mathbb{E}(\cdot)$ denotes the expectation operator.

2. Target Detection Probability, defined as

$$P_d = \frac{N_{\text{detect}}}{N_{\text{total}}} \quad (39)$$

where N_{detect} is the number of successful detections and N_{total} is the total number of Monte Carlo simulations.

3. Localization Accuracy, measured by the 3D Euclidean distance between the estimated and true target positions:

$$E_{3D} = \sqrt{(x_{\text{est}} - x_{\text{true}})^2 + (y_{\text{est}} - y_{\text{true}})^2 + (z_{\text{est}} - z_{\text{true}})^2} \quad (40)$$

The average localization error over N simulations is then calculated as

$$E_{3D}^N = \mathbb{E}(E_{3D}^n) \quad (41)$$

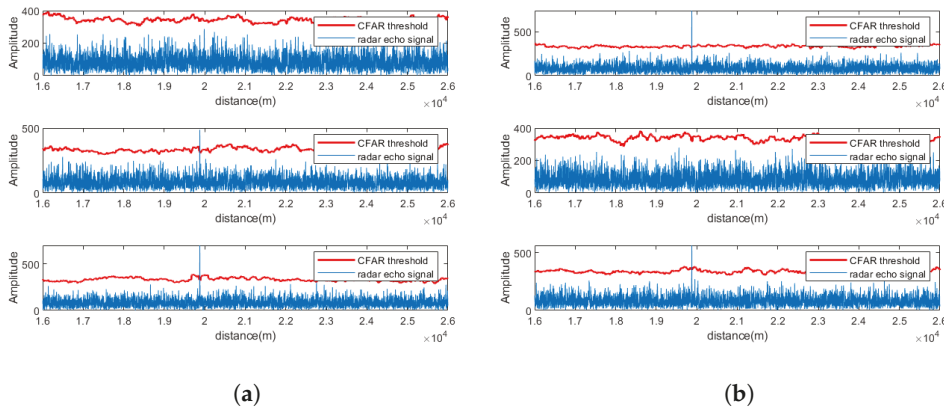


Figure 9. Cont.

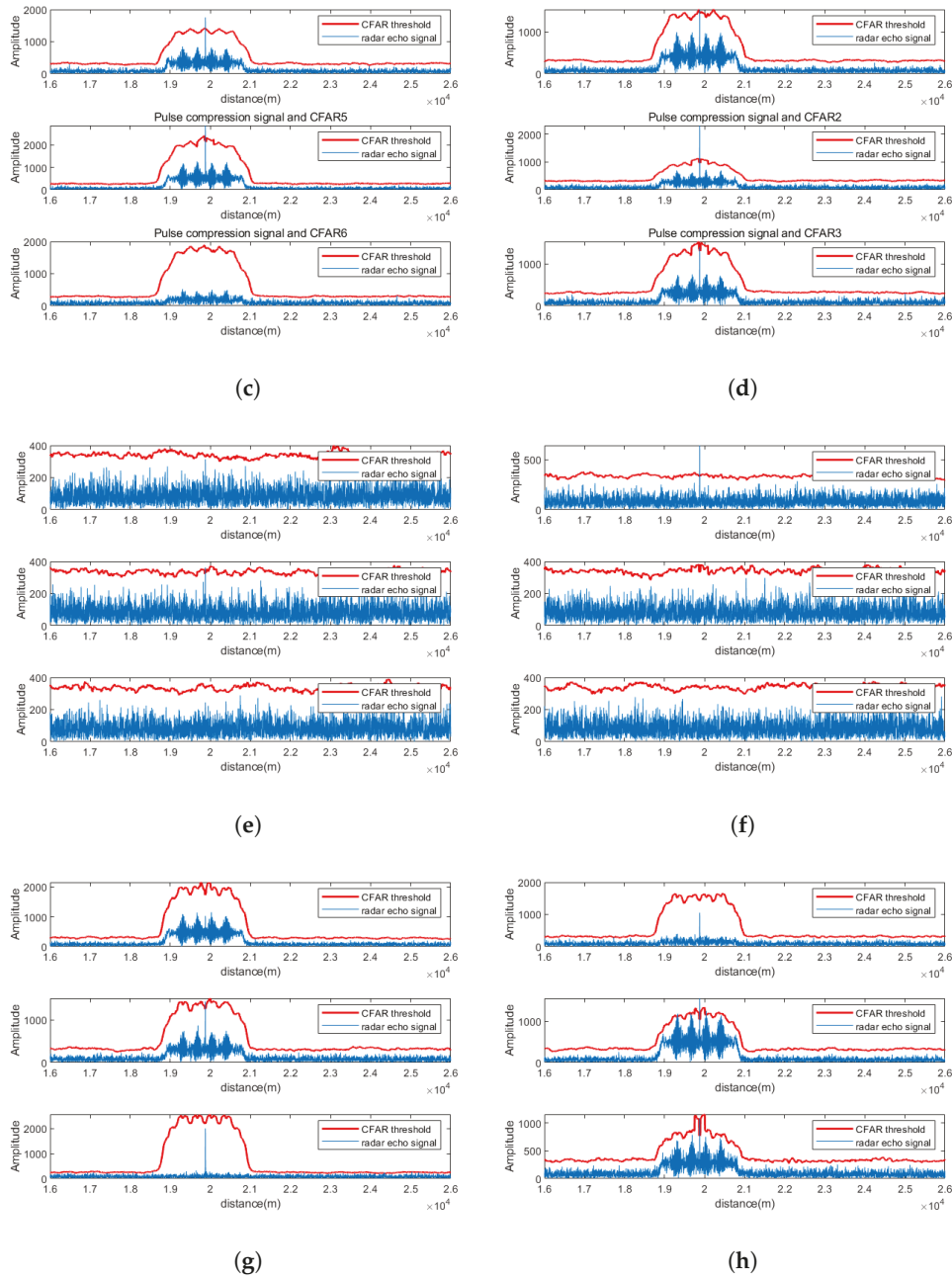


Figure 9. SNR = 5 dB. PA radar 1 received a signal after separation by the CEEMDAN-WOBSS method and MSNR-BSS method; the pulse compression and CFAR detection results are presented for the following: (a,b) Additional checks performed: CEEMDAN-WOBSS method, channels 1–6 (NAM); (c,d) CEEMDAN-WOBSS method, channels 1–6 (SMSP); (e,f) MSNR-BSS method, channels 1–6 (NAM); (g,h) MSNR-BSS method, channels 1–6 (SMSP).

Under a fixed JSR = 50 dB and $\text{SNR} \in [-5 \text{ dB}, 5 \text{ dB}]$, 100 Monte Carlo trials were executed for each SNR value under two interference scenarios to compute the average PSLR. As illustrated in Figure 10, the average PSLR of the separated target signal produced by all three algorithms decreases monotonically with an increasing SNR, indicating that higher SNR levels yield improved separation performance. Notably, CEEMDAN-WOBSS achieves a PSLR reduction of 0.2645–0.6235 dB compared to MSNR-BSS and JADE-BSS.

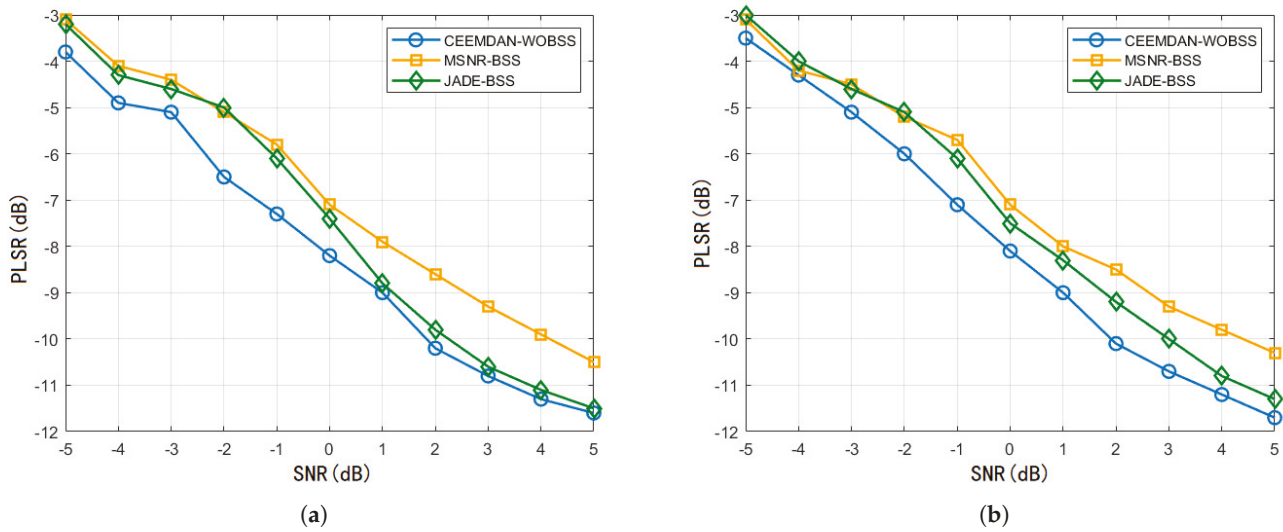


Figure 10. JSR = 50 dB and SNR $\in [-5$ dB, 5 dB]. PLSR results of the three methods (a) under NAM interference conditions and (b) under SMSP interference conditions.

Figure 11 shows the average PLSR under two interference scenarios at $textSNR = 5$ dB for JSR $\in [0$ dB, 50 dB]. The minimal variation in PLSR across this JSR range demonstrates that moderate changes in interference power do not significantly affect the separation performance of CEEMDAN-WOBSS.

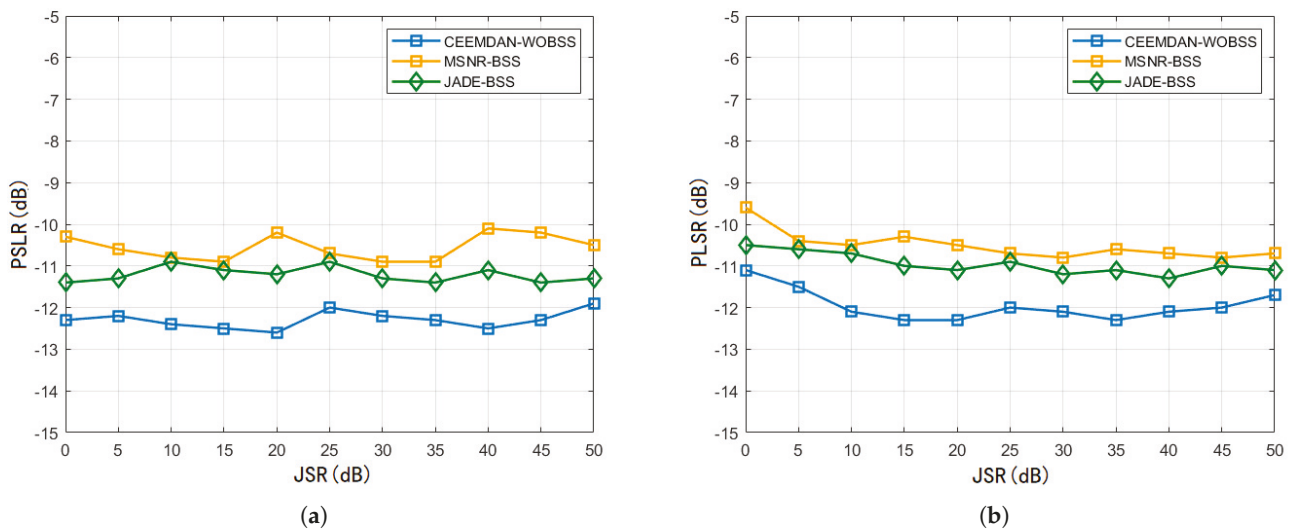


Figure 11. SNR = 5 dB and JSR $\in [0$ dB, 50 dB]. PLSR results of the three methods (a) under NAM interference conditions and (b) under SMSP interference conditions.

Using the same JSR = 50 dB and SNR $\in [-5$ dB, 5 dB] conditions, Figure 12 compares the target detection probabilities and cooperative localization errors of the three methods across two interference scenarios. The proposed CEEMDAN-WOBSS method improves the detection probability by 2–17% and enhances cooperative localization accuracy by 8.1–27.6% relative to MSNR-BSS and JADE-BSS.

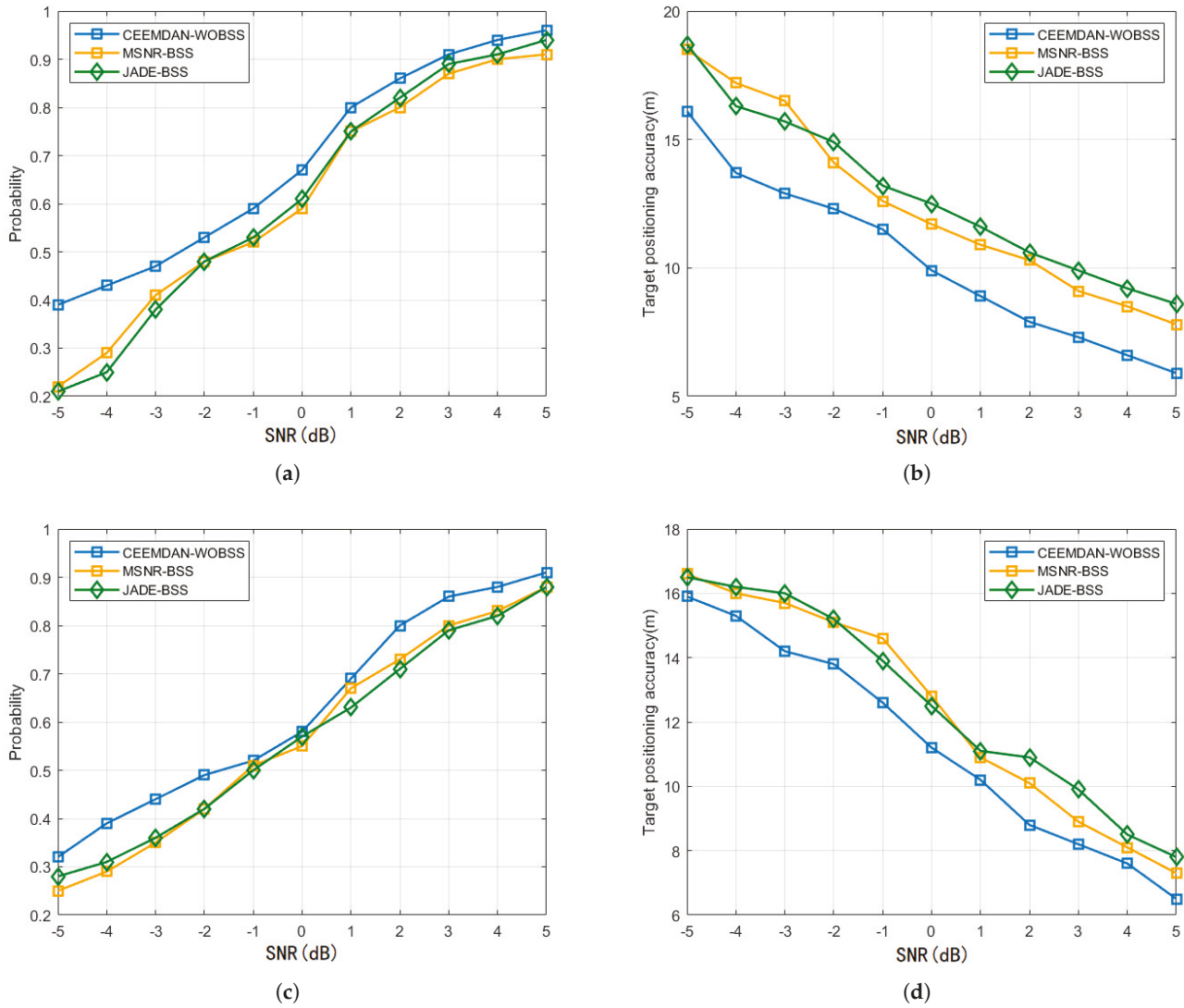


Figure 12. JSR = 50 dB and SNR $\in [-5$ dB, 5 dB]. Results of the three methods for detecting probability and positioning accuracy: (a,b) detection probability and localization error in NAM interference scenarios; (c,d) detection probability and localization error in SMSP interference scenarios.

To assess the proposed method's interference-suppression and localization performance under varying angular separations and multi-target conditions, we computed and statistically analyzed the PSLR and localization error as the angular separation between each target DOA, varying the jammer DOA from 1° to 10° . Additionally, interference-suppression and localization tests were performed for two- and three-target scenarios in which the targets shared the same DOA but were separated by 1000 m in range. In each Monte Carlo trial we recorded the mean localization error across all targets; the reported statistics were aggregated over 100 Monte Carlo realizations. The interference type was NAM interference, with all other experimental parameters remaining as specified above. The resulting performance curves are presented in Figure 13.

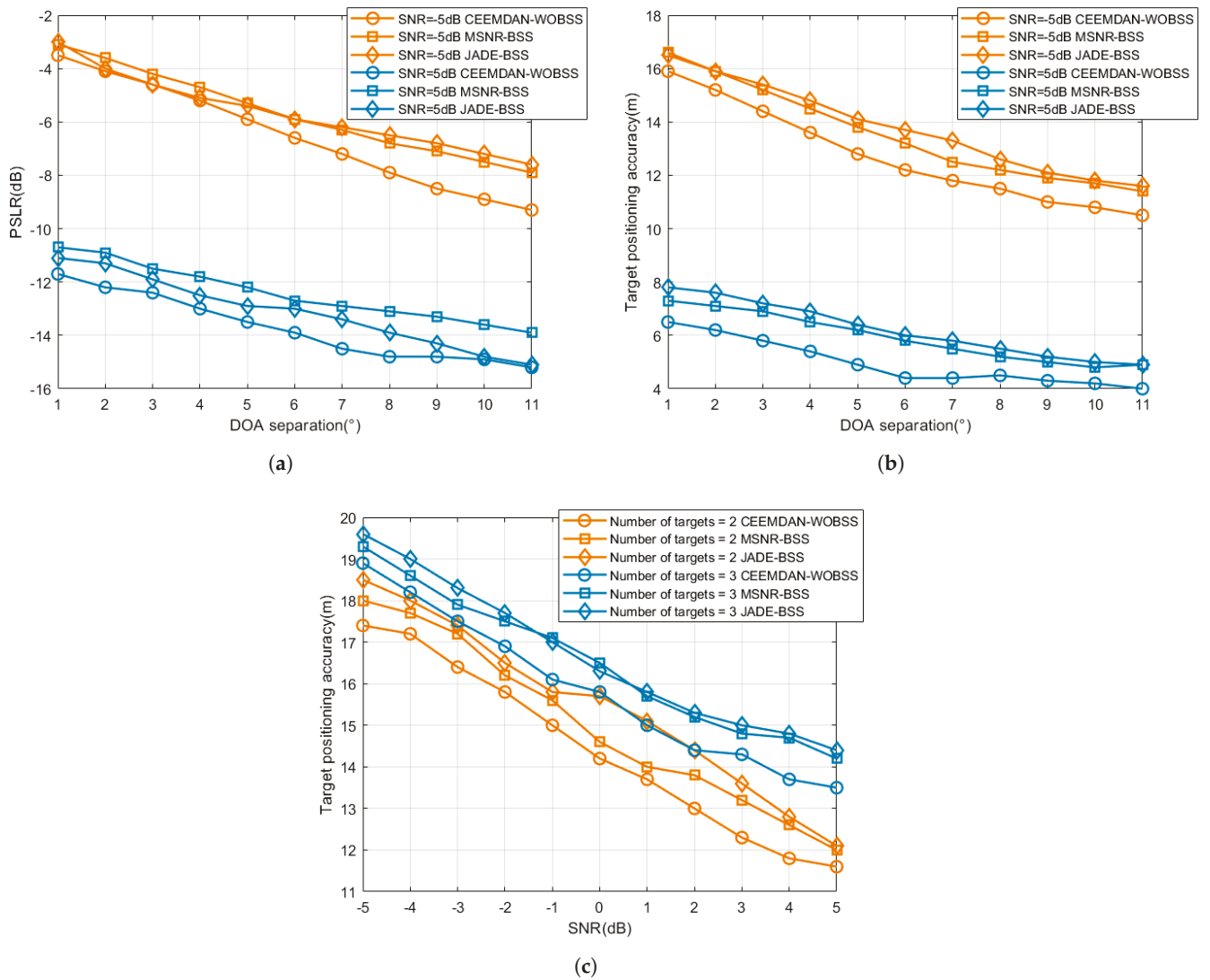


Figure 13. Interference-suppression and localization performance under varying angular separation and multi-target conditions. The angular separation between each target DOA and the jammer DOA was varied from 1° to 11° , and rgw SNR was varied from -5 dB to $+5$ dB. Results are averaged over 100 Monte Carlo realizations: (a,b) SNR = -5 dB and 5 dB; variation in PLSR and positioning accuracy with angle of arrival difference for the three methods; (c) SNR $\in [-5$ dB, 5 dB], average positioning accuracy of the three methods under two-target and three-target scenarios.

The results in Figure 13 indicate that as the arrival angle separation increases, the PSLR value rises while the positioning error decreases; moreover, a higher signal-to-noise ratio simultaneously improves both metrics. Under all conditions, the proposed CEEMDAN-WOBSS method achieves the highest PSLR value and lowest positioning error, with all three methods converging under conditions of large separation and high SNR. In multi-target tests, all methods exhibited performance degradation, with errors gradually decreasing as the SNR increased. CEEMDAN-WOBSS demonstrated a larger reduction in error, indicating superior robustness in dense target scenarios.

Furthermore, to demonstrate that the proposed method achieves a superior anti-jamming performance and localization accuracy under low-SNR conditions, a statistical comparison of positioning errors was carried out for both the proposed SDCAL pipeline and a single-station radar across a range of SNR values. For the single-station baseline,

the target DOA was first estimated using a standard DOA estimation algorithm prior to localization. The numerical results are summarized in Table 4.

Table 4. Positioning error of single-site radar and multi-site radar.

Positioning Method	SNR = −5 dB (m)	SNR = 0 dB (m)	SNR = 5 dB (m)
Single-site radar	18.6	12.3	7.6
Multi-site radar	16.1	9.8	5.9

As shown in Table 4, the proposed framework effectively enhances the spatial localization accuracy of targets under low-signal-to-noise ratio conditions. Compared to relying on a single radar for target detection and localization, the proposed SDCAL framework enables data-level collaborative localization. This approach can tolerate partial node detection failures while still achieving localization based on two or more reliable nodes.

6. Discussion

In this paper, an SDCAL framework for mainlobe interference suppression and three-dimensional target localization in distributed radar networks is proposed. To address the challenges posed by low-SNR environments, the CEEMDAN–WOBSS algorithm is first employed at each PA radar node to decompose, denoise, and reconstruct the received echoes, thereby reducing the noise baseline and restoring the effective rank of the signal covariance matrix. A maximum-SNR-based separation matrix subsequently extracts the target echo, which is then refined through pulse compression and CFAR detection to suppress residual interference. Finally, the fusion of range measurements across multiple nodes yields precise spatial coordinates. The simulation results show that SDCALF markedly enhances interference suppression, leading to significant gains in target detection probability and localization precision under low-SNR conditions when compared with conventional BSS schemes. Moreover, even in the presence of typical suppression jamming (such as noise frequency modulation and noise phase modulation), the proposed method maintains robustness and stable localization performance.

This study primarily serves as a proof of concept to validate the core signal processing chain under challenging yet specific scenarios. It should be noted that this research has certain limitations: the current implementation assumes a known node geometry and standard clock synchronization; significant synchronization errors or uncalibrated geometries will degrade performance unless compensated for. Source counting processing occurs outside the main processing pipeline, and misclassifications may cause separation and localization failures. The current fusion rules employ simple averaging/aggregation algorithms and have not been optimized for damaged or adversarial node reports. Regarding computational resource consumption, the CEEMDAN and BSS phases are the most resource-intensive, requiring algorithmic and implementation optimizations for real-time deployment. Finally, experiments were conducted only on limited static scenarios and jammer types; performance may differ in dense target scenarios near array resolution limits, for high-maneuver targets, or in complex multipath/clutter environments. We have prioritized addressing these issues in future work and will focus on enhancing the method’s operational robustness and deployability.

Author Contributions: Conceptualization, X.L.; methodology, X.L.; software, X.L. and H.H.; validation, X.L.; formal analysis and investigation, X.L. and X.Z.; resources and data curation, X.L. and Y.W.; writing—original draft preparation, X.L.; writing—review and editing, X.L. and Y.Y.; visualization, R.L.; supervision, H.H.; project administration, H.H.; funding acquisition, H.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Ahmed, A.; Yuan, Z.; Longdon, S.N.; Bore, J.-C.; Bin, T. A study of ECCM techniques and their performance. In Proceedings of the 2015 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), Ningbo, China, 19–22 September 2015; pp. 1–6.
2. Luo, M.; Hao, Z.; Zhu, Y. Intelligent ECCM Technology via Cognition and Agility for the Airborne Radar. *J. Eng.* **2019**, *2019*, 6656–6659. [CrossRef]
3. Han, B.; Qu, X.; Yang, X.; Li, W.; Zhang, Z. DRFM-Based Repeater Jamming Reconstruction and Cancellation Method with Accurate Edge Detection. *Remote Sens.* **2023**, *15*, 1759. [CrossRef]
4. Chen, X.; Shu, T.; Yu, K.-B.; He, J.; Yu, W. Joint adaptive beamforming techniques for distributed array radars in multiple mainlobe and sidelobe jamming. *IEEE Antennas Wirel. Propag. Lett.* **2020**, *19*, 248–252. [CrossRef]
5. Zhou, S.; Zhang, Y.; Ma, H.; Peng, X.; Yan, J. Optimized masking for MIMO radar signals. *IEEE Trans. Aerosp. Electron. Syst.* **2021**, *57*, 3433–3451. [CrossRef]
6. Liao, Y.; Tang, H.; Wang, W.; Xing, M. A low sidelobe deceptive jamming suppression beamforming method with a frequency diverse array. *IEEE Trans. Antennas Propag.* **2022**, *70*, 4884–4889. [CrossRef]
7. Huang, Y.; Yu, J.; Su, R.; Li, C. A Multi-machine Jamming Method for Side-lobe Cancellation. *J. Phys. Conf. Ser.* **2023**, *2478*, 122076. [CrossRef]
8. Liu, X.; Sun, C.; Yang, Y.; Zhuo, J. Low sidelobe range profile synthesis for sonar imaging using stepped-frequency pulse. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 218–221. [CrossRef]
9. Dong, S.; Quan, Y.; Fang, W.; Wang, H. Anti-intermittent sampling jamming method based on frequency agile radar and K-means. In Proceedings of the 2021 CIE International Conference on Radar (Radar), Haikou, China, 15–19 December 2021; pp. 1140–1144.
10. Greco, M.; Gini, F.; Farina, A. Radar detection and classification of jamming signals belonging to a cone class. *IEEE Trans. Signal Process.* **2008**, *56*, 1984–1993. [CrossRef]
11. Dai, H.; Wang, X.; Li, Y.; Liu, Y.; Xiao, S. Main-lobe jamming suppression method of using spatial polarization characteristics of antennas. *IEEE Trans. Aerosp. Electron. Syst.* **2012**, *48*, 2167–2179. [CrossRef]
12. Zhang, J.; Zhu, D.; Zhang, G. New antivelocity deception jamming technique using pulses with adaptive initial phases. *IEEE Trans. Aerosp. Electron. Syst.* **2013**, *49*, 1290–1300. [CrossRef]
13. Wang, B.; Cui, G.; Zhang, S.; Sheng, B.; Kong, L.; Ran, D. Deceptive Jamming Suppression Based on Coherent Cancelling in Multistatic Radar System. In Proceedings of the 2016 IEEE Radar Conference (RadarConf), Philadelphia, PA, USA, 2–6 May 2016; pp. 1–5.
14. Liu, J.; Zhang, L.; Zhao, S.; Liu, N.; Yu, H. Antideception jamming method for active/passive radar based on distributed frameworks. *J. Xidian Univ.* **2018**, *45*, 1–6.
15. Wang, G.; Ji, Z. Multi-range-false-target jamming for radar network based on multiple discriminations. *Syst. Eng. Electron.* **2017**, *39*, 40–48.
16. Liu, J.; Gong, M.; Nie, Z.; Li, H.; Liu, J.; Zhao, S. Subset Selection Strategies Based on Target Positioning Characteristics for Anti-Jamming Technology. *Remote Sens.* **2022**, *14*, 6230. [CrossRef]
17. Li, D.; Yao, S.; Liao, Y.; Yan, J.; Fu, Y. Frequency offset deception velocity-based method for discriminating between physical targets and active false targets. *Signal Process.* **2025**, *234*, 109977. [CrossRef]
18. Yang, Y.; Da, K.; Zhu, Y.; Xiang, S. Consensus based target tracking against deception jamming in distributed radar networks. *IET Radar Sonar Navig.* **2023**, *17*, 683–700. [CrossRef]
19. Li, Q.; Guo, Y.; Zhang, P.; Xu, H.; Zhang, L.; Chen, Z.; Huang, Y. Deception Velocity-Based Method to Discriminate Physical Targets and Active False Targets in a Multistatic Radar System. *Remote Sens.* **2024**, *16*, 382. [CrossRef]
20. Liu, J.; Gong, M.; Zhang, M.; Li, H.; Zhao, S. An Anti-jamming Method in Multistatic Radar System Based on Convolutional Neural Network. *IET Signal Process.* **2022**, *16*, 220–231. [CrossRef]
21. Yu, H.; Liu, N.; Zhang, L.; Li, Q.; Zhao, S. An interference suppression method for multistatic radar based on noise subspace projection. *IEEE Sens. J.* **2020**, *20*, 8797–8805. [CrossRef]
22. Zhao, S.; Zhang, L.; Zhou, Y.; Liu, N.; Liu, J. Discrimination of active false targets in multistatic radar using spatial scattering properties. *IET Radar Sonar Navig.* **2016**, *10*, 817–826. [CrossRef]

23. Sun, M.; Ding, C.; Zhang, S.; Lu, J.; Shao, P. Recognition of Deception Jamming Based on Statistical Correlation Difference in a Multistatic Radar System. *J. Electron. Inf. Technol.* **2020**, *42*, 2992–2998.
24. Kong, Y.; Wang, X.; Wu, C.; Yu, X.; Cui, G. Active deception jamming recognition in the presence of extended target. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [CrossRef]
25. Yu, H.; Zhang, J.; Zhang, L.; Li, S. Polarimetric multiple-radar frameworks with distributed antennas for discriminating between radar targets and deception jamming. *Digit. Signal Process.* **2019**, *90*, 46–53. [CrossRef]
26. Li, C.; Zhu, L.; Xie, A.; Luo, Z. Weak signal blind source separation in passive radar system with strong interference. In Proceedings of the MILCOM IEEE Military Communications Conference, Tampa, FL, USA, 26–28 October 2015; pp. 501–505.
27. Wang, W.; Zhang, J.; Liu, X.; Li, L. Radar anti-main-lobe-jamming based on blind source separation algorithm of JADE. *Fire Control Command. Control* **2015**, *40*, 104–108.
28. Zhou, C.; Wang, C.; Gong, J.; Tan, M.; Li, X.; Bao, L. Hybrid polarization anti-SMSP jamming based on jamming reconstruction and blind source separation. *J. Beijing Univ. Aeronaut. Astronaut.* **2021**, *47*, 1841–1848.
29. Liu, Y.; Yu, L.; Wei, Y. Mainlobe multiple false targets deceptive jamming suppression via joint transmit-receive design. *Digit. Signal Process.* **2025**, *164*, 105289. [CrossRef]
30. Lei, Z.; Qu, Q.; Chen, H.; Zhang, Z.; Dou, G.; Wang, Y. Mainlobe jamming suppression with space-time multichannel via blind source separation. *IEEE Sens. J.* **2023**, *23*, 17042–17053. [CrossRef]
31. Yan, Z.; Ge, M.; Cui, G.; Yu, X.; Hu, R. DOA estimation with subarrays via blind source separation algorithm. In Proceedings of the 2021 IEEE Radar Conference (RadarConf21), Atlanta, GA, USA, 7–14 May 2021; pp.1–6.
32. Liu, Z.; Liang, W.; Fu, N.; Qiao, L.; Zhang, J. Main lobe deceptive jamming suppression based on blind source separation and energy detection for monopulse radar. *IET Radar Sonar Navig.* **2024**, *18*, 2170–2181. [CrossRef]
33. Ge, M.; Cui, G.; Yu, X.; Kong, L. Main lobe jamming suppression via blind source separation sparse signal recovery with subarray configuration. *J. Syst. Eng. Electron.* **2020**, *14*, 431–438. [CrossRef]
34. Ge, M.; Cui, G.; Kong, L.; Huang, D. Jamming suppression based on blind source separation in distributed radars. *J. Signal Process.* **2017**, *33*, 1562–1570.
35. Han, X.; He, H.; Zhang, Q.; Yang, L.; He, Y. Main-lobe jamming suppression method for phased array netted radar based on MSNR-BSS. *IEEE Sens. J.* **2022**, *22*, 22972–22984. [CrossRef]
36. Gao, X.; Quan, Y.; Li, Y.; Zhu, S.; Xing, M. Main-lobe deceptive jamming suppression with FDAMIMO radar based on BSS. *Syst. Eng. Electron.* **2020**, *42*, 1926–1934.
37. Borga, M. Learning Multidimensional Signal Processing. Ph.D. Thesis, Linköping University, Linköping, Sweden, 1998.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Variable-Speed UAV Path Optimization Based on the CRLB Criterion for Passive Target Localization

Lijia Chen, Chengfeng You, Yixin Wang and Xueting Li *

School of Aeronautics and Astronautics, Sichuan University, Chengdu 610207, China;
2023141510131@stu.scu.edu.cn (L.C.); 2023141510127@stu.scu.edu.cn (C.Y.); 2023141510130@stu.scu.edu.cn (Y.W.)

* Correspondence: lixueting@scu.edu.cn

Abstract

The performance of passive target localization is significantly influenced by the positions of unmanned aerial vehicle swarms (UAVs). In this paper, we investigate the problem of UAV path optimization to enhance the localization accuracy. Firstly, a passive target localization signal model based on the time difference of arrival (TDOA) algorithm, which is then improved by the Chan method and Taylor series expansion, is established. Secondly, the Cramer–Rao lower bound (CRLB) of the modified TDOA algorithm is derived and adopted as the evaluation criterion to optimize the UAVs' positions at each time step. Different from the existing works, in this paper, we consider the UAVs to have variable speed; therefore, the feasible region of the UAVs' positions is changed from a circle into an annular region, which will extend the feasible region, enhancing the localization accuracy while increasing the computation complexity. Thirdly, to improve the efficiency of the UAV path optimization algorithm, the particle swarm optimization (PSO) algorithm is applied to search for the optimal positions of the UAVs for the next time step. Finally, numerical simulations are conducted to verify the validity and effectiveness of the proposals in this paper.

Keywords: target passive localization; unmanned aerial vehicles (UAVs); Cramer–Rao lower bound (CRLB); path optimization method; variable speed; particle swarm algorithm (PSO)

1. Introduction

As a critical component of modern intelligent unmanned systems, UAVs are profoundly reshaping combat paradigms and technological frameworks in the military domain. In modern warfare, UAVs demonstrate remarkable advantages in key missions such as target reconnaissance [1], battlefield surveillance [2,3], and formation control through their highly coordinated operations. This collaborative combat system transcends limitations of traditional fixed infrastructure, resolves challenges like communication latency [4], and significantly enhances the timeliness of battlefield information acquisition and operational effectiveness [5]. Particularly in complex and dynamic combat environments, UAVs have emerged as critical penetration forces against enemy defense systems due to their high mobility [6,7], strong survivability, and autonomous coordination capabilities [8]. However, most existing studies assume uniform motion for path optimization, which poses major challenges in real-world variable-speed scenarios.

In various combat applications of UAVs, passive target localization technology holds unique tactical value [9,10]. Compared to active localization systems, passive localization systems achieve precise localization of emitter targets without exposing their own location

by passively receiving electromagnetic signals from radiation sources [11], playing a crucial role in electronic reconnaissance, anti-stealth operations, and battlefield situational awareness. Tua A. Tamba et al. developed a distributed Kalman filter-based cooperative tracking algorithm for multiple UAVs, which enhances target state estimation accuracy in maritime moving vessel monitoring scenarios. However, their constant-velocity motion model demonstrates limited adaptability to abruptly maneuvering targets [12]. The DQN-ALNS framework proposed by Lingjie Zhou et al. significantly improves the solution efficiency of UAV path optimization through the collaborative optimization of deep reinforcement learning and adaptive large neighborhood search. However, its state space design based on the uniform speed assumption limits its application in dynamic environments [13]. Therefore, how to construct a dynamic geometric constraint model that accommodates variable-speed motion has become the core issue in improving passive localization accuracy.

1.1. Overview

For the application of target passive localization based on UAVs, the optimization of data quality [14,15], formation configuration [16] and flight paths significantly impacts the performance of the localization accuracy. By elevating data quality and selecting optimal measurements, the precision and robustness of the localization system can be effectively improved. The bearing-only passive localization model proposed by Ruiqing Yang et al. improves localization accuracy through the fusion of triangulation and greedy strategies, but fails to account for the impact of variable-speed motion on the observation geometric configuration [17]. Through optimized formation geometry, the UAVs can maintain the most effective configuration during target localization, enabling continuous optimal localization geometry that significantly enhances both accuracy and continuity of the localization performance. Meanwhile, UAV optimization involves dynamically adjusting the flight paths and spatial configurations of the UAVs to optimize the observation conditions, ultimately achieving high-precision localization accuracy. These studies ensure that the observational data provided by the UAVs consistently approximates the minimum error conditions, thereby substantially improving the localization performance.

At the path optimization algorithm level, swarm intelligence algorithms have made significant progress, but several critical challenges remain. Zhanghan Li et al. proposed the APF-DDQN model, which integrates artificial potential fields with deep reinforcement learning to effectively address path optimization in complex obstacle-rich environments. However, its state space design does not account for dynamic velocity variations [18]. The improved ACO algorithm proposed by Dong et al. [19] introduces roulette wheel selection and elitist strategies, significantly enhancing task execution efficiency and survival probability in multi-task scenarios while resolving the poor convergence issues of traditional methods in handling multi-constraint, nonlinear optimization problems. The PSO algorithm modified by Hao et al. [20] features nonlinear inertia weight adjustments and incorporates logarithmic function control mechanisms with Gaussian random adjustment strategies, substantially improving convergence speed and local search capabilities in UAV path optimization. Han et al. [21] enhanced multi-UAV cooperative mission optimization by integrating Dubins curve constraints with multi-objective optimization strategies, improving convergence rates and global optimization capabilities while resolving the challenge of simultaneously satisfying trajectory length, mission payoff, and flight constraints. Shengxi Li et al. proposed a hybrid path optimization framework integrating Model Predictive Control (MPC) with Control Barrier Functions (CBFs). While this approach enhances dynamic obstacle avoidance through safety constraints, its prediction model still relies on the uniform motion assumption, resulting in path tracking errors during sudden maneuvering scenarios [22]. However, the modified PSO algorithm only adopts

the fixed speed updates [20], which cannot accommodate variable-speed UAV movements. When combined with Dubins curves, the system remains constrained by curvature limitations, confining the search space to a fixed circle, restricting the improvement of the localization accuracy.

For UAV path optimization in complex obstacle-laden environments where computational demands are high and efficiency is low, several algorithms have made breakthroughs. Wang et al.'s priority-based Artificial Potential Field (APF) algorithm [23], Zhang et al.'s improved target-biased Rapidly-exploring Random Tree (RRT) path optimization algorithm [24], Lu et al.'s RRT-APF hybrid path optimization algorithm [25], and Xu et al.'s Three-phase Alternating Optimization Algorithm (TAOA) [26] all overcome these limitations and resolve the issue of excessive computational burdens in complex environments. However, these methods suffer from poor adaptability and are unsuitable for variable-speed environments.

In dynamic environments, Hou et al. proposed a dynamic path optimization algorithm integrating improved Q-Learning with the Artificial Potential Field (APF) method [27], resolving the issue where traditional methods fail to generate effective paths between path nodes. Zhou et al. developed a UAV TDOA localization path optimization algorithm based on the D-optimality criterion [28], significantly enhancing localization accuracy and tracking stability for emitter targets. Li et al. presented a hybrid TDOA/FDOA-based UAV path optimization algorithm [29], substantially improving position and velocity estimation accuracy for moving targets. However, these methods exhibit poor real-time performance and do not consider the situation where the UAVs have variable speed.

To improve the UAV path optimization efficiency, Pan et al. resolved the issues of traditional APF methods easily falling into local minima and formation instability caused by introducing rotation [30]. Sun et al. proposed a UAV path optimization algorithm based on semantic maps and multi-scale feature fusion [31], which notably improves path optimization efficiency for large-scale formations. Additionally, Zhou et al. significantly improved the efficiency and quality of image acquisition for 3D reconstruction through intelligent viewpoint selection and optimization strategies [32]. Chen et al. transformed task allocation problems into a Vehicle Routing Problem (VRP) model and optimized observation time [33], notably increasing the efficiency of multi-UAV cooperative remote sensing missions. Li et al. addressed the low efficiency of traditional swarm collision avoidance optimization by establishing a coupling matrix and employing the AR-SAS path optimization method [34]. Li et al. also studied how to optimize UAV observation geometry configurations [35], significantly reducing the flight path lengths and computational loads while resolving the inefficiency of conventional passive localization methods.

However, the above approaches are all under the assumption that the UAVs have a fixed speed and this will set strict constraints on the UAV path optimization problem, restricting the feasible region to circles with the UAVs as their centers. However, in the real situation, UAVs generally have variable speed; thus, the feasible region of the UAV path optimization problem is an annular region which extends the feasible region, leading to more optimal UAV position solutions.

1.2. Original Contribution

Evidently, the existing works mostly focus on UAV path optimization with fixed speed, which is mismatched with the real situation and will restrict the improvement of passive target localization accuracy. Thus, in this paper, we focus on the problem of UAV path optimization, considering UAVs with variable speed. Firstly, the signal model of the passive target localization method is established. Secondly, a UAV path optimization method is proposed based on the CRLB of the TDOA localization method, considering UAVs

with variable speed. Thirdly, the PSO algorithm is applied to accelerate the optimization process and improve the efficiency of the path optimization algorithm. Simulation results illustrate that the proposals in this paper could achieve better localization accuracy with less computational complexity. The original contributions of this paper are as follows:

- (1) A novel UAV path optimization method based on the CRLB of the TDOA localization method is proposed, where the trace of the CRLB is set as the objective function and UAV moving constraints are set as the constraint conditions. By optimally configuring the positions of the UAVs for each time step, the localization accuracy is improved.
- (2) Different from the existing works where the UAVs are assumed to have fixed speed, in this paper, we consider the UAVs to have variable speed; thus, the feasible region of the UAV positions is extended from a circle to an annular region and the localization accuracy is improved.
- (3) To accelerate the UAV path optimization process and improve the efficiency of the algorithm, the PSO algorithm is applied to search for the optimal positions of the UAVs at the next time step.

2. Signal Model of Passive Target Localization

In this section, the TDOA algorithm is applied for target localization. Since the TDOA algorithm generally has high computation complexity, the Taylor series expansion algorithm is introduced to accelerate the localization algorithm, and the modified TDOA–Taylor algorithm could achieve better computation efficiency. Although the modified TDOA–Taylor algorithm benefits from the smaller computation load, it is affected by the initial value of the target state error at a large scale. If the initial error of the target state is large, the accuracy of the target localization would be poor. Thus, the Chan algorithm is also introduced to guarantee the accuracy of the initial value of the target state, and the modified TDOA–Chan–Taylor algorithm could achieve better localization accuracy with low computation complexity.

2.1. Signal Model of the TDOA Algorithm

TDOA refers to a localization technique that determines the location of a target by measuring the time differences of the target's arrival at different base stations. It is a multi-sensor localization technology which requires at least three receivers for localization.

The distance r_j from the target to any base station j is

$$r_j = \sqrt{(x - x_j)^2 + (y - y_j)^2} \quad (1)$$

where the target position $\mathbf{u} = [x, y]^T$ and the coordinates of the base station j are (x_j, y_j) .

The distance difference $r_{j,0}$ between the target of the secondary station j and the primary station is

$$r_{j,0} = \sqrt{(x - x_j)^2 + (y - y_j)^2} - \sqrt{(x - x_0)^2 + (y - y_0)^2} \quad (2)$$

Then, we can get the following nonlinear hyperbolic equations:

$$\begin{cases} r_{1,0} = \sqrt{(x - x_1)^2 + (y - y_1)^2} - \sqrt{(x - x_0)^2 + (y - y_0)^2} \\ r_{2,0} = \sqrt{(x - x_2)^2 + (y - y_2)^2} - \sqrt{(x - x_0)^2 + (y - y_0)^2} \end{cases} \quad (3)$$

Thus, the TDOA localization problem reduces to solving for the intersection point of the aforementioned hyperbolic equations.

2.2. Modified TDOA–Chan Algorithm

The basic idea of the Chan algorithm is to calculate the position of the target using a two-step weighted least squares algorithm. Instead of the iterative process of the TDOA algorithm, the TDOA–Chan algorithm could obtain the analytical expression solution of the target position, achieving lower computation complexity and a good initial value of the target state.

By squaring the two sides of Equation (1), we get

$$r_j^2 = (x - x_j)^2 + (y - y_j)^2 = C_j - 2x_jx - 2y_jy + x^2 + y^2 \quad (4)$$

where $C_j = x_j^2 + y_j^2$.

By arranging and squaring the two sides of Equation (2), we obtain

$$C_j - (r_{j,0}^2 + 2r_{j,0}r_0 + r_0^2) = 2x_jx + 2y_jy - x^2 - y^2 \quad (5)$$

Let $j = 0$; we get

$$C_0 - r_0^2 = 2x_0x + 2y_0y - x^2 - y^2 \quad (6)$$

Subtracting Equation (6) from Equation (5), we have

$$2(x_j - x_0)x + 2(y_j - y_0)y = C_j - C_0 - (r_{j,0}^2 + 2r_{j,0}r_0) \quad (7)$$

According to the above derivations, when the total number of the base stations in a two-dimensional space is 3, the relationship between the target position $u = [x, y]^T$ and r_0 can be obtained as follows:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_{10} & y_{10} \\ x_{20} & y_{20} \end{bmatrix}^{-1} \times \left\{ \begin{bmatrix} r_{1,0} \\ r_{2,0} \end{bmatrix} r_0 + \frac{1}{2} \begin{bmatrix} r_{1,0}^2 - C_1 + C_0 \\ r_{2,0}^2 - C_2 + C_0 \end{bmatrix} \right\} \quad (8)$$

where $x_{j0} = x_j - x_0$, $y_{j0} = y_j - y_0$.

When the number of the base stations $n \geq 3$, a two-step weighted least squares method needs to be applied to obtain the position of the target.

Let the unknown variable matrix $p = [u, r_0]^T$ and the time difference noise vector ζ be

$$\zeta = \eta - Dp \quad (9)$$

where

$$\eta = \frac{1}{2} \begin{bmatrix} r_{1,0}^2 - C_1 + C_0 \\ r_{2,0}^2 - C_2 + C_0 \\ \vdots \\ r_{(n-1),0}^2 - C_{n-1} + C_0 \end{bmatrix} \quad (10)$$

$$D = \begin{bmatrix} x_{10} & y_{10} & r_{1,0} \\ x_{20} & y_{20} & r_{2,0} \\ \vdots & \vdots & \vdots \\ x_{(n-1),0} & y_{(n-1),0} & r_{(n-1),0} \end{bmatrix} \quad (11)$$

Then the covariance matrix ξ of the time difference noise vector ζ is

$$\xi = E[\zeta\zeta^T] = c^2 ARA^T \quad (12)$$

where $A = \text{diag}(r_0, r_1, \dots, r_{n-1})$ and R is the covariance matrix of the time difference measurements.

When the distance is large, the maximum likelihood estimation of $\mathbf{p}$ can be written as

$$\mathbf{p} \approx (D^T \boldsymbol{\zeta}^{-1} D)^{-1} D^T \boldsymbol{\zeta}^{-1} \boldsymbol{\eta} \quad (13)$$

Then we can get

$$\mathbf{p} = [\hat{x}, \hat{y}, \hat{r}_0]^T \quad (14)$$

Next, we conduct the calculation of the second weighted least squares algorithm. When there is random noise, D , $\boldsymbol{\eta}$, and $\mathbf{p}$ can be expressed as

$$\begin{cases} D = D^0 + \Delta D \\ \boldsymbol{\eta} = \boldsymbol{\eta}^0 + \Delta \boldsymbol{\eta} \\ \mathbf{p} = \mathbf{p}^0 + \Delta \mathbf{p} \end{cases} \quad (15)$$

where D^0 represents the ideal design matrix under noiseless conditions, $\boldsymbol{\eta}^0$ denotes the true observation vector without measurement errors, $\mathbf{p}^0$ is the true 3D vector containing target coordinates (x,y) and the distance between the target and the reference station, ΔD is the design matrix deviation from base station coordinate measurement errors, $\Delta \boldsymbol{\eta}$ is the observation vector error propagated from time difference measurement noise, and $\Delta \mathbf{p}$ is the estimation error of the first weighted least squares step solution.

Then the covariance matrix of $\Delta \mathbf{p}$ is

$$\text{cov}(\Delta \mathbf{p}) = E[\Delta \mathbf{p} \Delta \mathbf{p}^T] = (D^{0T} \boldsymbol{\zeta}^{-1} D^0)^{-1} \quad (16)$$

Then the expression of the time difference noise vector $\boldsymbol{\zeta}'$ can be obtained as

$$\boldsymbol{\zeta}' = \boldsymbol{\eta}' - D' \mathbf{p}' \quad (17)$$

$$\text{where } \boldsymbol{\eta}' = \frac{1}{2} \begin{bmatrix} (\hat{x} - x_0)^2 \\ (\hat{y} - y_0)^2 \\ \hat{r}_0^2 \end{bmatrix}, D' = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}, \mathbf{p}' = [x', y']^T, x' \approx (x - x_0)^2, y' \approx (y - y_0)^2.$$

Following an analogous procedure to Equation (12), the covariance matrix $\boldsymbol{\zeta}'$ of the time difference noise vector $\boldsymbol{\zeta}'$ for the second weighted least squares step can be derived. When the measurement noise and estimation residuals are small, since $\boldsymbol{\zeta}'$ obeys the Gaussian distribution, the maximum likelihood estimation of $\mathbf{p}'$ can be obtained as $\mathbf{p}' \approx (D'^T \boldsymbol{\zeta}'^{-1} D')^{-1} D'^T \boldsymbol{\zeta}'^{-1} \boldsymbol{\eta}'$. When the distance between the base station and the target is large, the covariance matrix of $\mathbf{p}'$ can be simplified as

$$\text{cov}(\mathbf{p}') \approx (D'^T \boldsymbol{\zeta}'^{-1} D')^{-1} \quad (18)$$

According to the above derivations, the unambiguous solution of the unknown quantity $\mathbf{p}$ can be obtained through two weighted least squares methods. Finally the position of the target $\mathbf{u} = [x, y]^T$ can be obtained.

2.3. Modified TDOA–Chan–Taylor Algorithm

The Taylor expansion algorithm is significantly affected by the initial target position estimate. Thus, this subsection introduces the modified TDOA–Chan–Taylor algorithm. We utilize the target position estimate (x_0, y_0) of the TDOA–Chan algorithm as the initial value for Taylor expansion. By neglecting second-order and higher terms, Equation (1) is transformed into

$$\varepsilon = h - G\mathbf{p} \quad (19)$$

where $h = \begin{bmatrix} r_{1,0} - (R_1 - R_0) \\ r_{2,0} - (R_2 - R_0) \\ \vdots \\ r_{(n-1),0} - (R_{n-1} - R_0) \end{bmatrix}$, $G = \begin{bmatrix} \frac{x_0-x_0}{R_0} - \frac{x_1-x_0}{R_1} & \frac{y_0-y_0}{R_0} - \frac{y_1-y_0}{R_1} \\ \frac{x_0-x_0}{R_0} - \frac{x_2-x_0}{R_2} & \frac{y_0-y_0}{R_0} - \frac{y_2-y_0}{R_2} \\ \vdots & \vdots \end{bmatrix}$, ε is the residual vector, and R_j represents the distance from the j -th base station to the initial value.

Using the weighted least squares algorithm, the least squares estimation of the target position can be obtained as follows:

$$p = (G^T Q^{-1} G)^{-1} G^T Q^{-1} h \quad (20)$$

where Q is the covariance matrix of the TDOA measurement.

3. CRLB Derivation and Problem Formulation

In this section, the CRLB for the TDOA–Chan–Taylor localization algorithm is derived, giving the theoretical lower bound of the target localization accuracy. Then, the trace of the CRLB is set as the objective function with the UAV motion constraints set as the constraint conditions, and the UAV path optimization problem is formulated.

3.1. CRLB Derivation for the TDOA Localization Algorithm

For an unbiased estimator, the minimum error of the estimator can be characterized by the CRLB of the estimator. Thus, in this paper, the CRLB of the TDOA algorithm is set as the objective function of the UAV path optimization problem, and through searching for the optimal UAV positions, the accuracy of the target localization performance can be improved.

Based on the TDOA τ_j between the base station j and the reference station, we establish the measurement equation as

$$\mathbf{V} = \mathbf{l}(\mathbf{u}) + \mathbf{m} \quad (21)$$

where $\mathbf{l}(\mathbf{u}) = [c\tau_1, \dots, c\tau_{n-1}]^T$, the measurement error $\mathbf{m}$ satisfies the normal distribution $\mathbf{N}(\mathbf{0}_{(n-1)}, \mathbf{R}_{(n-1) \times (n-1)})$, and $\mathbf{V}$ is the actual observed value. $\mathbf{0}_{(n-1)}$ is a zero-mean vector of dimension $(n-1) \times 1$, and $\mathbf{R}$ is the measurement noise covariance matrix of dimension $(n-1) \times (n-1)$.

According to $\mathbf{u} = [x, y]^T$, we can get the Jacobi matrix $\mathbf{J} = [\frac{\partial l(\mathbf{u})}{\partial x}, \frac{\partial l(\mathbf{u})}{\partial y}]^T$.

$$\begin{cases} \frac{\partial l(\mathbf{u})}{\partial x} = \left[\frac{x-x_1}{r_1} - \frac{x-x_0}{r_0}, \frac{x-x_2}{r_2} - \frac{x-x_0}{r_0}, \dots, \frac{x-x_{n-1}}{r_{n-1}} - \frac{x-x_0}{r_0} \right] \\ \frac{\partial l(\mathbf{u})}{\partial y} = \left[\frac{y-y_1}{r_1} - \frac{y-y_0}{r_0}, \frac{y-y_2}{r_2} - \frac{y-y_0}{r_0}, \dots, \frac{y-y_{n-1}}{r_{n-1}} - \frac{y-y_0}{r_0} \right] \end{cases} \quad (22)$$

Then, the Fisher matrix can be obtained as

$$\mathbf{F} = \mathbf{J}^T (\mathbf{R})^{-1} \mathbf{J} \quad (23)$$

Taking the inverse of the above equation, the CRLB $\mathbf{C}\mathbf{r}$ is obtained:

$$\mathbf{C}\mathbf{r} = (\mathbf{F})^{-1} \quad (24)$$

3.2. Problem Formulation

In this subsection, the UAV path optimization problem is formulated. Firstly, the minimum trace of the CRLB for the TDOA–Chan–Taylor localization algorithm is set as the objective function of the UAV path optimization problem, which would guarantee the accuracy of the target localization performance. The objective function is shown as follows:

$$\min_{u_1, u_2, \dots, u_n} \text{Tr}(\mathbf{C}\mathbf{r}(u_1, u_2, \dots, u_n)) \quad (25)$$

where u_i represents the positions of the i th UAV.

Meanwhile, to ensure the feasibility and stability of the UAV flight operations, the following constraint conditions need to be satisfied:

- (1) Minimum approach distance constraint: At any given time t , the distance r_i^t from the i th UAV to the target needs to satisfy

$$r_i^t \geq d_{\min} \quad (26)$$

where $d_{\min}$ is the minimum allowable distance between the UAVs and the target. This constraint prevents the UAVs from approaching the target too closely and being detected, thus ensuring the stealth and security of the mission.

- (2) Heading angle deviation constraint: For the i th UAV at time t , the heading angle deviation $\phi_i^t = \psi_i^{t+1} - \psi_i^t$ between two successive time steps must satisfy

$$\phi_{\min} \leq \phi_i^t \leq \phi_{\max} \quad (27)$$

where $\phi_{\min}$ is the minimum value of the heading angle deviation for the UAVs, and $\phi_{\max}$ is the maximum value of the heading angle deviation for the UAVs.

- (3) Heading angle constraint: The heading angle ψ_i^t between the heading angle of the i th UAV at time t and the reference direction must satisfy

$$\psi_{\min} \leq \psi_i^t \leq \psi_{\max} \quad (28)$$

where $\psi_{\min}$ is the minimum value of the heading angle for the UAVs, and $\psi_{\max}$ is the maximum value of the heading angle for the UAVs. For target localization application, drastic changes in heading deviation angles ϕ_i^t and heading angles ψ_i^t are detrimental to the UAVs' successive target observation. Simultaneously, these angular constraints effectively prevent invalid solutions and enhance the efficiency of the path optimization.

- (4) Velocity constraint: The velocity v_i^t of the i th UAV at time t satisfies

$$v_{\min} \leq v_i^t \leq v_{\max} \quad (29)$$

where $v_{\min}$ is the minimum velocity of the i th UAV, and $v_{\max}$ is the maximum velocity of the i th UAV.

- (5) Energy Constraint: At any given time t , the distance r_i^t from the i th UAV to the target needs to satisfy

$$r_i^t \leq s_{\max} \quad (30)$$

$$s_{\max} = v_{\max} \cdot t_{\text{bat}} \quad (31)$$

where $s_{\max}$ is the maximum allowable maneuver distance of the i th UAV and t_{bat} is the effective battery endurance time. This constraint limits the maximum maneuverable distance of UAVs to ensure that their path planning remains within the battery's endurance range, thereby preventing mission failure due to power depletion.

Different from the existing works, the UAVs in this paper are considered with variable speed, and the velocity of the UAVs is limited to the range of $[v_{\min}, v_{\max}]$. Thus, the feasible region of the optimization problem is extended from the conventional fixed-radius circumferences to annulars centered at the UAVs' current positions with the inner radius of $v_{\min}$ and the outer radius of $v_{\max}$.

By considering the UAVs with variable speed, the proposals in this paper could be more in accordance with the actual situation, significantly improving the UAVs' flexibility and environmental adaptability.

4. PSO Algorithm

In this paper, the PSO algorithm is applied to search for the optimal UAV positions at each time step to improve the target localization accuracy. Through optimizing the UAV positions for successive time steps, the optimal UAV path would be obtained.

Assume that at each iteration there are S particles, and each particle s represents a candidate UAV heading angle deviations strategy

$$\boldsymbol{\eta}^{t,k(s)} = [\phi_1^{t,k(s)} \quad \phi_2^{t,k(s)} \quad \dots \quad \phi_i^{t,k(s)}]^T \quad (32)$$

where $\phi_i^{t,k(s)}$ represents the heading angle deviation of the s th particle at the k th iteration of time step t for the i th UAV. Through the heading angle deviation $\boldsymbol{\eta}^{t,k(s)}$, the position of the s th particle at the k th iteration of time step t , $\mathbf{x}^{t,k(s)}$, can be obtained.

At the k th iteration of time step t , the velocity of the s th particle is denoted as $\mathbf{v}^{t,k(s)}$ and its heading angle deviation is $\boldsymbol{\eta}^{t,k(s)}$. Then at the $(k+1)$ th iteration of time step t , the velocity $\mathbf{v}^{t,k+1(s)}$ and heading angle deviation $\boldsymbol{\eta}^{t,k+1(s)}$ of the s th particle can be updated as

$$\mathbf{v}^{t,k+1(s)} = \mathbf{v}^{t,k(s)} + c_1 r_1 (\mathbf{p}^{t,k(s)} - \boldsymbol{\eta}^{t,k(s)}) + c_2 r_2 (\mathbf{p}_g^{t,k} - \boldsymbol{\eta}^{t,k(s)}) \quad (33)$$

$$\boldsymbol{\eta}^{t,k+1(s)} = \boldsymbol{\eta}^{t,k(s)} + \mathbf{v}^{t,k+1(s)} \quad (34)$$

where $\mathbf{p}^{t,k(s)}$ represents the individual historical best position of the s th particle at the k th iteration of time step t , $\mathbf{p}_g^{t,k}$ represents the global best position of the entire particle swarm at the k th iteration of time step t , and r_1 and r_2 represent the uniformly distributed random variables in the range of $[0, 1]$, which are used to introduce randomness and avoid getting trapped in local optima. c_1 denotes the individual learning factor, which can adjust the step size of the particle moving towards its own historical best position $\mathbf{p}^{t,k(s)}$, and c_2 denotes the social learning factor, which can adjust the step size of the particle moving towards the global best position $\mathbf{p}_g^{t,k}$.

To select the optimal solution, the performance of these S particles is evaluated using the fitness function. At the k th iteration of time step t , the fitness function for the s th particle is as follows:

$$Fn(\mathbf{x}^{t,k(s)}) = \text{Tr}(\mathbf{C}\mathbf{r}(\mathbf{x}^{t,k(s)})) \quad (35)$$

where $Fn(\mathbf{x}^{t,k(s)})$ is the fitness value when the heading angle deviation of the s th particle is $\boldsymbol{\eta}^{t,k(s)}$ and the related position of the s th particle is $\mathbf{x}^{t,k(s)}$, $\mathbf{C}\mathbf{r}(\mathbf{x}^{t,k(s)})$ is the CRLB when the heading angle deviation of the s th particle is $\boldsymbol{\eta}^{t,k(s)}$ and the related position of the s th particle is $\mathbf{x}^{t,k(s)}$, and $\text{Tr}(\cdot)$ denotes the trace operation. Then, the individual historical best position $\mathbf{p}^{t,k+1(s)}$ and the global best position $\mathbf{p}_g^{t,k+1(s)}$ can be updated as

$$\mathbf{p}^{t,k+1(s)} = \begin{cases} \boldsymbol{\eta}^{t,k+1(s)} & \text{for } Fn(\mathbf{x}^{t,k+1(s)}) < Fn(\mathbf{p}^{t,k(s)}) \\ \mathbf{p}^{t,k(s)} & \text{else} \end{cases} \quad (36)$$

$$\mathbf{p}_g^{t,k+1} = \arg \min_{1 \leq s \leq S} Fn(\mathbf{p}^{t,k+1(s)}) \quad (37)$$

When the iteration reaches the maximum number, the iteration stops and outputs the optimal solution; thus, the optimal UAV positions for the next time step ($t + 1$) can be obtained. The flowchart of the PSO algorithm is shown as follows, see Figure 1.

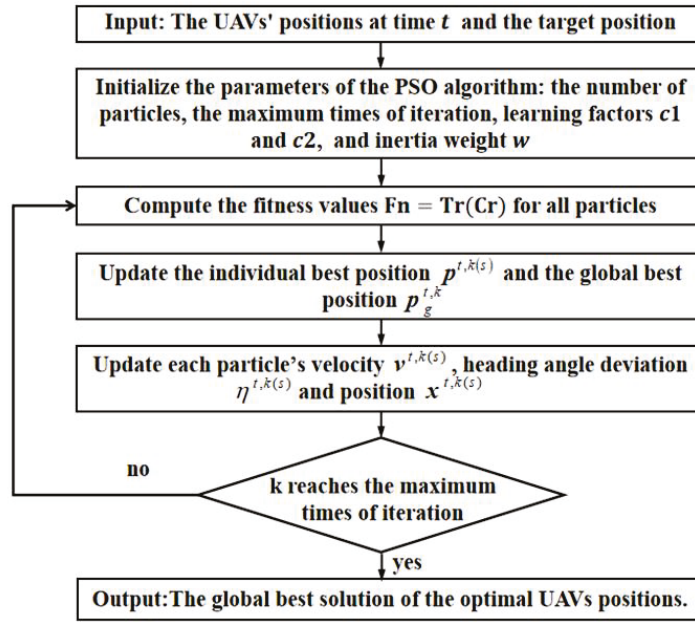


Figure 1. The flowchart of the PSO algorithm.

5. Numerical Simulations and Results

In this section, numerical simulations are performed to verify the validity and effectiveness of the proposals in this paper. Firstly, the basic parameters and the initial topology of the UAVs and the target are given. Secondly, five methods are presented and compared to verify the advantages of the proposed method in this paper, and the five methods include the following: (1) the PSO algorithm considering the UAVs to have variable speed; (2) the PSO algorithm considering the UAVs to have a fixed speed; (3) the exhaustive algorithm considering the UAVs to have variable speed; (4) the exhaustive algorithm considering the UAVs to have a fixed speed; (5) the algorithm considering the UAVs to have a fixed formation.

5.1. Parameters Configuration

Consider that there are four UAVs and two stationary targets in a two-dimensional plane. The initial coordinates of the UAVs are [0, 0] m, [8660, 0] m, [4330, 7500] m, and [0, 8660] m, and the targets are located at [20,000, 10,000] m and [15,000, 15,000] m. The initial topology of the UAVs and the targets is given in Figure 2, where the green squares represent the initial positions of the UAVs, the blue and red five-pointed stars indicate the two target positions, and the blue arrows indicate the moving directions of the UAVs.

Through systematic parameter sensitivity analysis, we have determined the optimal parameter configuration scheme in this paper. The speed range for the variable-speed UAVs is set to [200, 400] m/s, and the speed value for the fixed-speed UAVs is set to 300 m/s. Through systematic experiments, it is determined that the global optimal fitness value stabilizes when the number of iterations reaches 110. To ensure algorithmic robustness, the maximum number of iterations is ultimately set to 150. The cognitive component $c1$ and the social component $c2$ are both set to 2 to effectively balance the influence of individual particle experience and swarm knowledge on the search process, and the inertia weight w is set to 0.9 to maintain strong global search capability during the initial phase of the algorithm. Since the feasible region for the PSO algorithm considering the UAVs with

variable speed is larger than that for the PSO algorithm considering the UAVs with fixed speed, the number of the particles is 1000 for the PSO algorithm considering the UAVs with variable speed and the number of the particles is 100 for the PSO algorithm considering the UAVs with fixed speed, which can balance computational efficiency with precision requirements. To mitigate collision risks, the minimum approach distance is set to 8000 m, the heading angle range is set to $[-\pi/4, \pi/4]$, and the heading deflection angle range is set to $[-\pi/180, \pi/180]$.

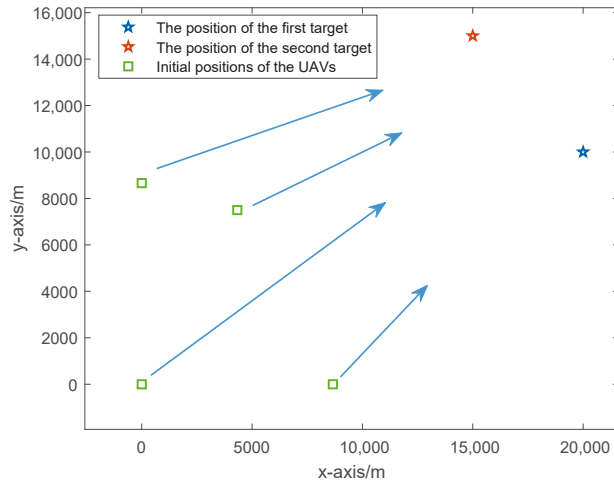


Figure 2. The initial topology of the UAVs and the targets.

5.2. Simulation Results Based on the Exhaustive Algorithm

In this part, we conduct the simulation experiments for different UAV path optimization algorithms, which include (1) UAVs with a fixed formation; (2) the UAV path optimization algorithm based on CRLB considering UAVs with fixed speed; and (3) the UAV path optimization algorithm based on CRLB considering UAVs with variable speed. The simulation results of the UAV path based on different UAV path optimization methods are given in Figures 3 and 4, where Figure 3 shows the UAV path with a fixed formation and Figure 4 shows the UAV paths optimized by the CRLB algorithm where (a) considers UAVs with a fixed speed and (b) considers the UAVs with variable speed. Also, in Section 5.2, we consider the UAV path optimized based on the exhaustive algorithm.

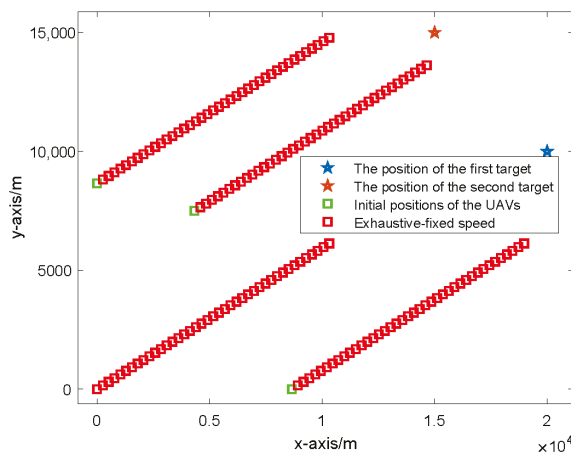


Figure 3. UAV path based on the fixed formation algorithm.

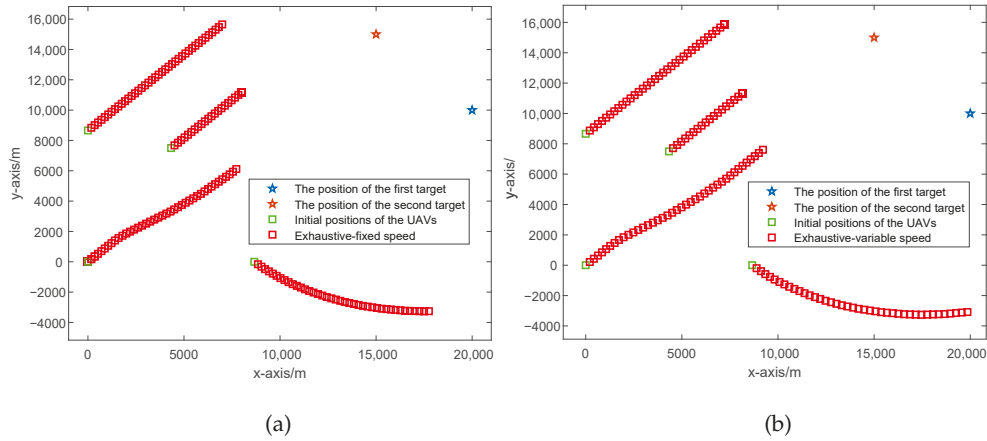


Figure 4. UAV paths optimized by the exhaustive algorithm. (a) UAV path optimized considering the UAVs to have a fixed speed. (b) UAV path optimized considering the UAVs to have variable speed.

To evaluate the localization accuracy of different methods, we employ the Root Mean Square Error (RMSE) as the accuracy metric and conduct a comparative analysis of the localization accuracy between different methods.

$$\text{RMSE}(k) = \sqrt{\frac{1}{N} \sum_{i=1}^N (u_k - \hat{u}_k)^2} \quad (38)$$

where N denotes the number of Monte Carlo experiment trials, u_k represents the true position of the target at time step k and $\hat{u}_k$ represents the estimated position of the target at time step k . A total of 100 independent Monte Carlo trials are conducted to complete the comparison.

From Figure 5, the simulation results of the RMSE and CRLB between different UAV path optimization algorithms in the dual-target scenario are shown and compared, including the fixed formation, the exhaustive algorithm considering UAVs with fixed speed, and the exhaustive algorithm considering UAVs with variable speed. (a) presents the comparison between the RMSE and CRLB for the target located at [20,000, 10,000] m, while (b) shows the comparison between the RMSE and CRLB for the target located at [15,000, 15,000] m. From the results, we can see that for the target located at [20,000, 10,000] m, the RMSE of the UAVs with fixed formation converges to 12.23 m and the CRLB converges to 0.90 m. The RMSE of the UAV path optimized based on the exhaustive algorithm considering the UAVs with fixed speed converges to 14.47 m and the CRLB converges to 0.89 m. The RMSE of the UAV path optimized based on the exhaustive algorithm considering the UAVs with variable speed converges to 9.85 m and the CRLB converges to 0.41 m. For the target located at m, the RMSE of the UAVs with fixed formation converges to 11.65 m and the CRLB converges to 0.47 m. The RMSE of the UAV path optimized based on the exhaustive algorithm considering the UAVs with fixed speed converges to 15.21 m and the CRLB converges to 0.98 m. The RMSE of the UAV path optimized based on the exhaustive algorithm considering the UAVs with variable speed converges to 9.65 m and the CRLB converges to 0.40 m. Thus, compared to the UAVs with fixed formation, the UAV path optimized based on the exhaustive algorithm considering the UAVs with variable speed reduces the average RMSE of the two targets by about 18.35% and reduces the average CRLB by about 50.16%. Then, compared to the UAV path optimized based on the exhaustive algorithm considering the UAVs with fixed speed, the UAV path optimized based on the exhaustive algorithm considering the UAVs with variable speed reduces the average RMSE by about 34.27% and reduces the average CRLB by about 56.62%. The results indicate that by considering the UAVs with variable speed, the UAV path

optimization algorithm could significantly improve the target localization accuracy in the dual-target scenario.

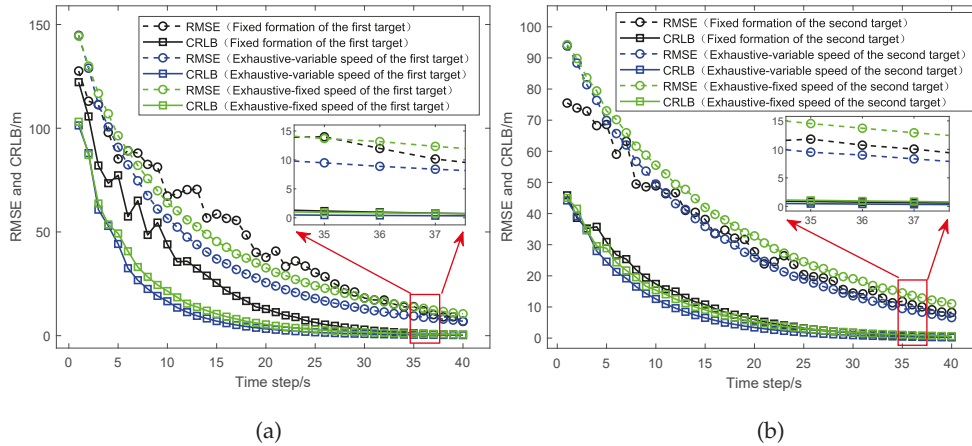


Figure 5. Comparison of the RMSE and CRLB between different UAV optimization algorithms which include the fixed formation algorithm, the exhaustive algorithm considering the UAVs with fixed speed, and the exhaustive algorithm considering the UAVs with variable speed. (a) Performance comparison for the target at [20,000, 10,000] m. (b) Performance comparison for the target at [15,000, 15,000] m.

5.3. Simulation Results Based on the PSO Algorithm

In Section 5.2, the simulation results for the UAV path optimization algorithms, which include the fixed formation, the UAV path optimization algorithm considering the UAVs with fixed speed, and the UAV path optimization algorithm considering the UAVs with variable speed, are all obtained based on the exhaustive algorithm. To improve the effectiveness of the path optimization algorithm, in this paper, the PSO algorithm is deployed to search for the optimal positions of the UAVs at the next time step, reducing the computation complexity and improving the real-time performance.

As can be seen from Figure 6, the UAV paths optimized based on the PSO algorithm and the exhaustive algorithm considering the UAVs with fixed speed are shown and compared, where the red squares represent the UAV path optimized based on the PSO algorithm, and the blue squares represent the UAV path optimized based on the exhaustive algorithm. Obviously, the UAV paths optimized based on the PSO algorithm and the exhaustive algorithm considering the UAVs with fixed speed are almost the same because they are both optimized based on the CRLB criterion but by different optimization algorithms.

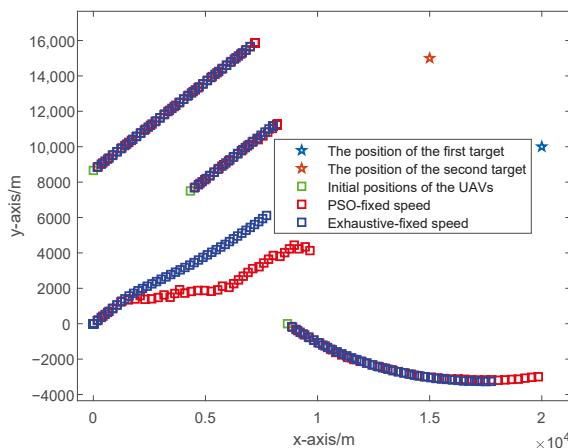


Figure 6. Different UAV paths optimized based on the PSO algorithm and the exhaustive algorithm considering the UAVs with fixed speed.

As shown in Figure 7, the UAV paths optimized based on the PSO algorithm and the exhaustive algorithm considering the UAVs with fixed speed are shown and compared, where the red squares represent the UAV path optimized by the PSO algorithm considering the UAVs with variable speed, and the blue squares represent the UAV path optimized by the exhaustive algorithm considering the UAVs with variable speed. Obviously, the UAV paths optimized based on the PSO algorithm and the exhaustive algorithm considering the UAVs with variable speed are almost the same, because they are both optimized based on the CRLB criterion but by different optimization algorithms.

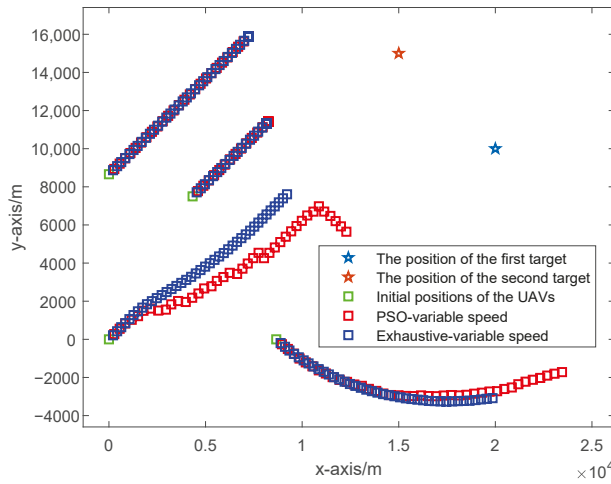


Figure 7. Different UAV paths optimized based on the PSO algorithm and the exhaustive algorithm considering the UAVs with variable speed.

As shown in Figure 8, the RMSE and the CRLB for different UAV paths optimized by the PSO algorithm and the exhaustive algorithm considering UAVs with fixed speed are given and compared. (a) presents the comparison between the RMSE and CRLB for the target located at [20,000, 10,000] m, while (b) shows the comparison between the RMSE and CRLB for the target located at [15,000, 15,000] m. From the simulation results, we can see that for the target located at [20,000, 10,000] m, the RMSE for the UAV path optimized by the exhaustive algorithm considering the UAVs with fixed speed converges to 14.47 m and the CRLB converges to 0.90 m. The RMSE for the UAV path optimized by the PSO algorithm considering the UAVs with fixed speed converges to 10.47 m and the CRLB converges to 0.46 m. For the target located at [15,000, 15,000] m, the RMSE for the UAV path optimized by the exhaustive algorithm considering the UAVs with fixed speed converges to 15.21 m and the CRLB converges to 0.98 m. The RMSE for the UAV path optimized by the PSO algorithm considering the UAVs with fixed speed converges to 13.05 m and the CRLB converges to 0.72 m. Thus, compared to the UAV path optimized based on the exhaustive algorithm considering the UAVs with fixed speed, the UAV path optimized based on the PSO algorithm considering the UAVs with fixed speed reduces the average RMSE of the two targets by about 20.93% and reduces the average CRLB by about 37.40%. The results indicate that according to the PSO algorithm considering the UAVs with fixed speed, the UAV path optimization algorithm could significantly improve the target localization accuracy in the dual-target scenario.

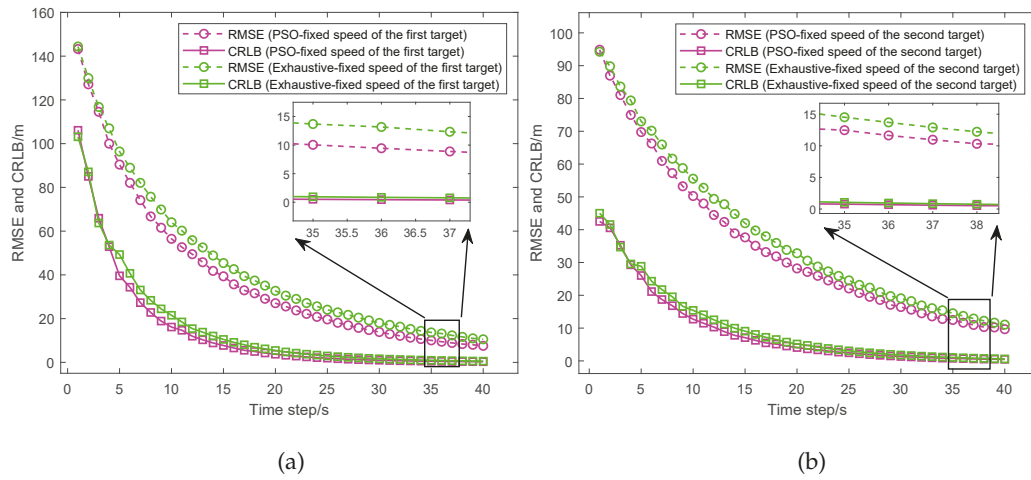


Figure 8. The RMSE and the CRLB based on the PSO algorithm and the exhaustive algorithm considering the UAVs with fixed speed. (a) Performance comparison for the target at [20,000, 10,000] m. (b) Performance comparison for the target at [15,000, 15,000] m.

The situation where the UAVs are considered with variable speed is almost the same as the situation where the UAVs are considered with fixed speed. As shown in Figure 9a, the RMSE for the UAV path optimized by the PSO algorithm converges to 5.54 m and the CRLB converges to 0.13 m. The RMSE for the UAV path optimized by the exhaustive algorithm converges to 9.85 m and the CRLB converges to 0.41 m. As shown in Figure 9b, the RMSE for the UAV path optimized by the PSO algorithm converges to 5.62 m and the CRLB converges to 0.13 m. The RMSE for the UAV path optimized by the exhaustive algorithm converges to 9.65 m and the CRLB converges to 0.40 m. Thus, the RMSE and the CRLB obtained by the PSO algorithm are significantly lower than those obtained by the exhaustive algorithm. Compared to the UAV path optimized based on the exhaustive algorithm, the UAV path optimized based on the PSO algorithm considering the UAVs with variable speed reduces the average RMSE of the two targets by about 42.76% and reduces the average CRLB by about 67.58%. The results indicate that according to the PSO algorithm, the UAV path optimization algorithm could significantly improve the target localization accuracy in the dual-target scenario.

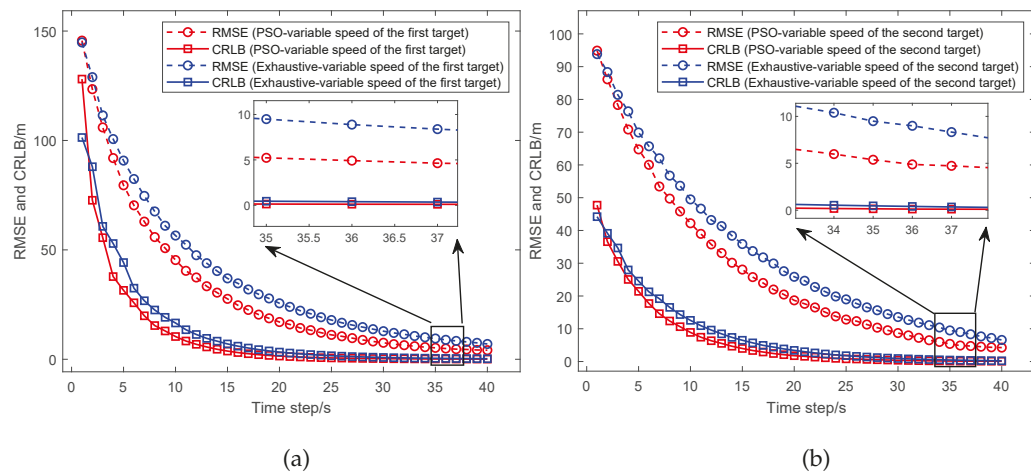


Figure 9. The RMSE and the CRLB based on the PSO algorithm and the exhaustive algorithm considering the UAVs with variable speed. (a) Performance comparison for the target at [20,000, 10,000] m. (b) Performance comparison for the target at [15,000, 15,000] m.

As shown in Figure 10, the simulation results of the RMSE and the CRLB between different UAV path optimization algorithms are shown and compared, including the fixed formation, the PSO algorithm considering the UAVs with variable speed and the PSO algorithm considering the UAVs with fixed speed. (a) presents the comparison between the RMSE and CRLB for the target located at [20,000, 10,000] m, while (b) shows the comparison between the RMSE and CRLB for the target located at [15,000, 15,000] m. From the results, we can see that for the target located at [20,000, 10,000] m, the RMSE of the UAVs with fixed formation converges to 12.23 m and the CRLB converges to 0.90 m. The RMSE of the UAVs with fixed speed converges to 10.47 m and the CRLB converges to 0.46 m. The RMSE of the UAVs with variable speed converges to 5.54 m and the CRLB converges to 0.13 m. For the target located at [15,000, 15,000] m, the RMSE of the UAVs with fixed formation converges to 11.65 m and the CRLB converges to 0.74 m. The RMSE of the UAVs with fixed speed converges to 13.05 m and the CRLB converges to 0.72 m. The RMSE of the UAVs with variable speed converges to 5.62 m and the CRLB converges to 0.13 m. Thus, compared to the UAVs with fixed formation, the UAVs with variable speed reduce the average RMSE of the two targets by about 53.25% and the average CRLB by about 83.80%. Then, compared to the UAVs with fixed speed, the UAVs with variable speed reduce the average RMSE by about 52.03% and the average CRLB by about 77.04%. The results indicate that the PSO algorithm considering the UAVs with variable speed could significantly improve the target localization accuracy.

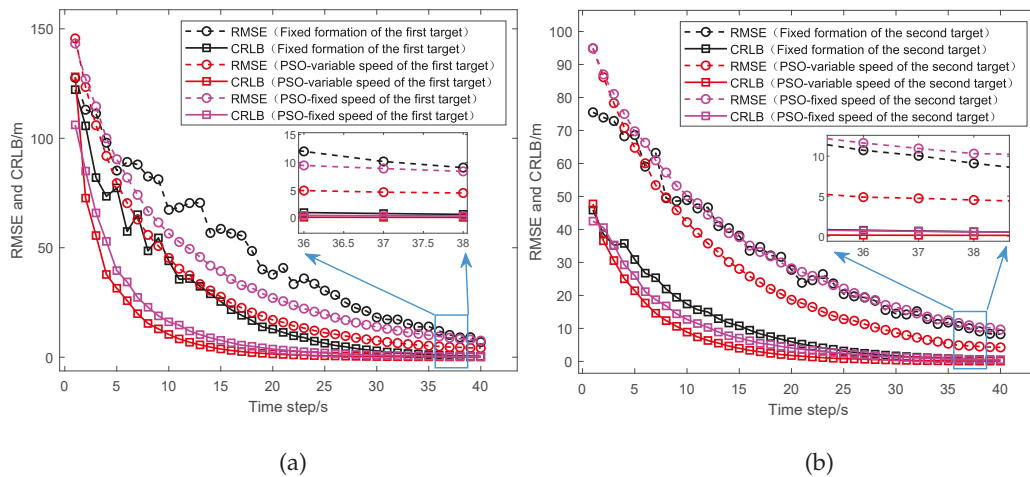


Figure 10. Comparison of the RMSE and CRLB between different UAV optimization algorithms which include the fixed formation, the PSO algorithm considering the UAVs with fixed speed, and the PSO algorithm considering the UAVs with variable speed. (a) Performance comparison for the target at [20,000, 10,000] m. (b) Performance comparison for the target at [15,000, 15,000] m

To verify the effectiveness of the PSO algorithm, we compare the average computation time per time step for different UAV path optimization algorithms. From Figure 11, we can see that 17.50 s is needed for the exhaustive algorithm considering the UAVs with fixed speed, 656.44 s is needed for the exhaustive algorithm considering the UAVs with variable speed, 17.04 s is needed for the PSO algorithm considering the UAVs with fixed speed, and 20.96 s for the PSO algorithm considering the UAVs with variable speed. Thus, the computational efficiency of the PSO algorithm considering the UAVs with variable speed reduces by about 96.81% compared to the exhaustive algorithm considering the UAVs with variable speed. However, the computation time for the PSO algorithm considering the UAVs with variable speed is slightly higher than that for the PSO algorithm considering the UAVs with fixed speed and the exhaustive algorithm considering the UAVs with fixed speed. The localization accuracy of the former algorithm is better than that of the latter algorithms. Thus, from the

numerical simulation results, it is illustrated that the proposed method in this paper could achieve better localization accuracy with proper algorithm efficiency.

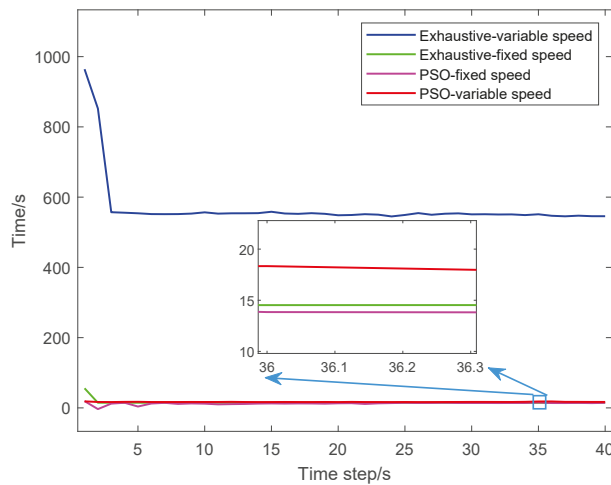


Figure 11. The average UAV path optimization time used for each time step for different path optimization algorithms.

6. Conclusions

In this paper, we study the problem of UAV path optimization based on the CRLB criterion. To overcome the stringent constraint for the traditional method that the UAVs have a fixed speed, in this paper, we consider that the UAVs have variable speed, which is more suitable to the real situation. However, when the UAVs have variable speed, the feasible region is extended and the computation complexity increases. Thus, in this paper, the PSO algorithm is deployed to accelerate the searching process and reduce the computation time to obtain the optimal UAV positions at the next time step. Firstly, the signal model of the target localization is established. Secondly, the CRLB for the TDOA localization algorithm is derived and the UAV path optimization problem based on the CRLB is formulated. Thirdly, the PSO algorithm is deployed to search for the optimal UAV positions for the next time step. Finally, numerical simulation results show that compared to the existing UAV path optimization algorithm, the method proposed in this paper could achieve higher localization accuracy and lower computation complexity.

7. Future Work

Future research will focus on a comprehensive and in-depth exploration of several key directions.

First, in terms of model extension, current studies primarily concentrate on two-dimensional space. Future work aims to extend the model to three-dimensional domains, constructing novel localization models that incorporate the altitude dimension. This will fully account for UAVs' unique motion characteristics in 3D space to more accurately describe their positional information. Building upon this, emphasis will be placed on addressing the dynamic decision-making problem of optimal UAV flight altitude. In real-world operations, a UAV's flight altitude is influenced by multiple factors, including obstacle distribution, signal strength, and meteorological conditions. By integrating advanced optimization algorithms with real-time environmental data, the flight altitude can be dynamically adjusted to balance localization accuracy with flight safety, energy consumption, and other multi-objective requirements. Simultaneously, a combined approach of theoretical analysis and simulation verification will be adopted to quantitatively investigate the nonlinear relationship between flight altitude and localization accuracy. An accurate mathematical model

will be established, and extensive simulation experiments will be conducted to analyze the variation patterns of localization accuracy under different flight altitudes. This will provide solid theoretical guidance for 3D path optimization in complex scenarios. Furthermore, as the model incorporates more practical factors, computational complexity will increase significantly. Therefore, a quantitative analysis of computational complexity will be performed to serve as a reference for algorithm optimization and real-world applications.

In terms of algorithmic comparison and validation, we plan to establish a comprehensive multi-dimensional evaluation system. This system is designed to conduct a systematic comparison between the proposed algorithm and mainstream approaches. These mainstream approaches encompass the extended Kalman filter, genetic algorithms, and emerging deep reinforcement learning methods. The comparison will be carried out across key metrics, namely localization accuracy, computational efficiency, energy consumption characteristics, and robustness. To ensure the objectivity and accuracy of the evaluation results, scientifically justified evaluation metrics and well-designed testing scenarios will be developed. Leveraging this framework, a detailed comparative analysis will be conducted to assess algorithm performance under different conditions, thoroughly investigating the strengths, weaknesses, and applicable domains of each method. The findings will provide a theoretical basis for algorithm selection in practical applications. Additionally, targeted improvements will be proposed to address the limitations of the proposed algorithm, further enhancing its performance and competitiveness.

To better evaluate the algorithm's practical applicability, future work also plans to conduct experiments using real UAVs. A comprehensive experimental platform will be established, integrating UAV hardware systems, sensor equipment, ground control stations, and other components to ensure smooth experiment execution. Additionally, diverse experimental scenarios will be designed, covering varying terrains, meteorological conditions, and mission requirements, to thoroughly validate the algorithm's performance across different real-world environments. During the experiments, the focus will be on verifying the effectiveness of the proposed algorithm by comparing its performance with theoretical analyses and simulation results. Key evaluation metrics will include localization accuracy and path optimization rationality in practical applications. Furthermore, the algorithm's computational time for determining the UAV's next waypoint must meet real-time transmission requirements. Real-time performance serves as a critical metric for UAV applications. Only when ensuring that the algorithm can swiftly compute and transmit results can safe flight operations and mission execution be guaranteed. Experimental data analysis will further optimize the algorithm's performance, enhancing its real-time capability and reliability.

Author Contributions: Conceptualization, L.C. and X.L.; methodology, X.L.; software and validation, L.C.; writing—original draft preparation, L.C., Y.W. and C.Y.; writing—review and editing, X.L.; supervision, X.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the Qian Yuan Laboratory.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Zhang, K.; Li, Z.; Zhao, X.; Zhao, B. Dynamic Multi-UAV Cooperative Reconnaissance Task Assignment Based on ICNP. In Proceedings of the 2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE), Harbin, China, 25–27 December 2020; pp. 773–779.

2. Le, Q.S.; Kim, J.; Kim, J.; Son, H.I. Report on work in progress of small insect tracking system using autonomous UAV. In Proceedings of the 2017 14th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), Jeju, Republic of Korea, 28 June–1 July 2017; pp. 242–243.
3. Wang, S.; Han, Y.; Chen, J.; Zhang, Z.; Wang, G.; Du, N. A Deep-Learning-Based Sea Search and Rescue Algorithm by UAV Remote Sensing. In Proceedings of the 2018 IEEE CSAA Guidance, Navigation and Control Conference (CGNCC), Xiamen, China, 10–12 August 2018; pp. 1–5.
4. Devey, J.; Shahra, E.; Hao, W.; Mi, D.; Aneiba, A.; Idrissi, M. Design and Simulation of a Novel Leader-Follower UAV Cluster and Formation Control Network. In Proceedings of the 2024 International Joint Conference on Neural Networks (IJCNN), Yokohama, Japan, 30 June–5 July 2024; pp. 1–6.
5. Xu, J.; Guo, Q.; Xiao, L.; Li, Z.; Zhang, G. Autonomous Decision-Making Method for Combat Mission of UAV based on Deep Reinforcement Learning. In Proceedings of the 2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chengdu, China, 20–22 December 2019; pp. 538–544.
6. Tian, Y.-X.; Li, C.; Yu, W.-T.; Xiong, J.-J. Adaptive Position and Attitude Trajectory Tracking Control for a Quadrotor UAV. In Proceedings of the 2022 International Conference on Automation, Robotics and Computer Engineering (ICARCE), Wuhan, China, 16–17 December 2022; pp. 1–4.
7. Xiong, J.-J.; Guo, N.-H.; Hong, Y.-X.; Zheng, E.-H. Improved Position and Attitude Tracking Control for a Quadrotor UAV. In Proceedings of the 2019 Chinese Automation Congress (CAC), Hangzhou, China, 22–24 November 2019; pp. 4197–4201.
8. Shi, Y.; Wang, R.; Tong, R.; Zeng, Y. Tracking Control and Obstacle Avoidance of Quad-rotor UAV Based on High-Order Fully Actuated System Approach. In Proceedings of the 2024 3rd Conference on Fully Actuated System Theory and Applications (FASTA), Shenzhen, China, 10–12 May 2024; pp. 695–700.
9. Chen, Y.; Shen, C.; Zhang, K.; Xu, L.; Huang, X. TDOA Positioning Method Based on Taylor Series Expansion Based on Cuckoo Search Algorithm. In Proceedings of the 2020 International Conference on Virtual Reality and Intelligent Systems (ICVRIS), Zhangjiajie, China, 18–19 July 2020; pp. 986–989.
10. Tunç, F.A.; Alp, Y.K.; Ata, L.D. Location Estimation by Using Multiple TDOA/AOA Measurements. In Proceedings of the 2020 28th Signal Processing and Communications Applications Conference (SIU), Gaziantep, Turkey, 5–7 October 2020; pp. 1–4.
11. Zhu, H.; Zhou, L.; You, M.-Y.; Shi, Q. The CRLB Analysis for Target Localization Based on UAV Swarms. In Proceedings of the 2021 13th International Symposium on Antennas, Propagation and EM Theory (ISAPE), Zhuhai, China, 1–4 December 2021; pp. 1–3.
12. Tamba, T. A.; Santjoko, I. R.; Nazaruddin, Y. Y.; Nadhira, V. A Multi-UAV Coordination Scheme for Tracking Control of Multiple Moving Target Objects. In Proceedings of the 2024 IEEE International Conference on Smart Mechatronics (ICSMech), Yogyakarta, Indonesia, 19–21 November 2024; pp. 7–11.
13. Zhou, L.; Luo, Z.; Jiao, P.; Jin, X.; Sun, X.; Shi, J. A Priority-Based Hybrid Algorithm for UAV Path Planning. In Proceedings of the 2024 10th International Conference on Big Data and Information Analytics (BigDIA), Chiang Mai, Thailand, 25–28 October 2024; pp. 518–523.
14. Ghany, A. A.; Uguen, B.; Lemur, D. A Pre-processing Algorithm Utilizing a Paired CRLB for TDoA Based IoT Positioning. In Proceedings of the 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring), Antwerp, Belgium, 25–28 May 2020; pp. 1–5.
15. Li, Q.; Tian, Q.; Pan, Z.; Liu, Y.; Zheng, S.; Du, B. X. Research on Precise Positioning Technology of UAV Based on Visual Positioning and Visual Tracking. In Proceedings of the 2023 IEEE 4th International Conference on Electrical Materials and Power Equipment (ICEMPE), Tianjin, China, 7–10 April 2023; pp. 1–4.
16. Qu, Y.; Yang, B.; Song, Y.; Xu, S. Multi-UAVs Cooperative Ground Moving Target Tracking and Positioning Method. In Proceedings of the 2023 IEEE 4th International Conference on Electrical Materials and Power Equipment (ICEMPE), Shanghai, China, 7–10 May 2023; pp. 3471–3476.
17. Yang, R.; Deng, J.; Liang, J.; Liu, Z. Research on Passive Localization in Circular Formation Flight of UAVs with Simulated Annealing Optimization Based on Triangulation and Greedy Strategy. In Proceedings of the 2024 IEEE 2nd International Conference on Control, Electronics and Computer Technology (ICCECT), Jilin, China, 26–28 April 2024; pp. 1021–1027.
18. Li, Z.; Chen, Y. Deep Learning Based Optimization of UAV Orbital Paths. In Proceedings of the 2024 IEEE 6th International Conference on Power, Intelligent Computing and Systems (ICPICS), Shenyang, China, 26–28 July 2024; pp. 208–212.
19. Dong, X.; Shi, C.; Wen, W.; Zhou, J. Multi-Mission Oriented Joint Optimization of Task Assignment and Flight Path Planning for Heterogeneous UAV Cluster. *Remote Sens.* **2023**, *15*, 5315. [CrossRef]
20. Hao, W.; Wu, C.; Lin, S. Research on UAV path planning based on improved particle swarm algorithm with inertia weight. In Proceedings of the 2023 IEEE International Conference on Control, Electronics and Computer Technology (ICCECT), Jilin, China, 28–30 April 2023; pp. 738–741.

21. Han, Q. Research on Cooperate Search Path Planning of Multiple UAVs Using Dubins Curve. In Proceedings of the 2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA), Shenyang, China, 22–24 January 2021; pp. 584–588.
22. Li, S.; Chen, Y.; Yang, Y. Multi-rotor UAV path planning based on Model Predictive Control and Control Barrier Function. In Proceedings of the 2024 36th Chinese Control and Decision Conference (CCDC), Xi'an, China, 25–27 May 2024; pp. 1141–1146.
23. Wang, Y.; Xu, G.; Long, T.; Cao, Y. Receding Path Planning for UAV Swarms Using Priority-based Artificial Potential Field. In Proceedings of the 2019 Chinese Automation Congress (CAC), Hangzhou, China, 22–24 November 2019; pp. 3234–3239.
24. Zhang, H.; Xie, X.; Wei, M.; Wang, X.; Song, D.; Luo, J. An Improved Goal-bias RRT algorithm for Unmanned Aerial Vehicle Path Planning. In Proceedings of the 2024 IEEE International Conference on Mechatronics and Automation (ICMA), Tianjin, China, 4–7 August 2024; pp. 1360–1365.
25. Lu, Y.; Wu, A.; Chen, Q.; Wang, Y. An Improved UAV Path Planning method Based on RRT-APF Hybrid strategy. In Proceedings of the 2020 5th International Conference on Automation, Control and Robotics Engineering (CACRE), Dalian, China, 19–20 September 2020; pp. 81–86.
26. Xu, S.; Zhou, Z.; Liu, H.; Zhang, X.; Li, J.; Gao, H. A Path Planning Method for Collaborative Coverage Monitoring in Urban Scenarios. *Remote Sens.* **2024**, *16*, 1152. [CrossRef]
27. Hou, X.; Liu, F.; Wang, R.; Yu, Y. A UAV Dynamic Path Planning Algorithm. In Proceedings of the 2020 35th Youth Academic Annual Conference of Chinese Association of Automation (YAC), Zhanjiang, China, 16–18 October 2020; pp. 127–131.
28. Zhou, R.; Sun, H.; Li, H.; Luo, W. TDOA and track optimization of UAV swarm based on D-optimality. *J. Syst. Eng. Electron.* **2020**, *31*, 1140–1151. [CrossRef]
29. Li, H.; Sun, H.; Zhou, R.; Zhang, H. Hybrid TDOA/FDOA and track optimization of UAV swarm based on A-optimality. *J. Syst. Eng. Electron.* **2023**, *34*, 149–159. [CrossRef]
30. Pan, Z.; Zhang, C.; Xia, Y.; Xiong, H.; Shao, X. An Improved Artificial Potential Field Method for Path Planning and Formation Control of the Multi-UAV Systems. *IEEE Trans. Circuits Syst. II* **2022**, *69*, 1129–1133. [CrossRef]
31. Sun, T.; Sun, W.; Sun, C.; He, R. Path Planning of UAV Formations Based on Semantic Maps. *Remote Sens.* **2024**, *16*, 3096. [CrossRef]
32. Zhou, H.; Ji, Z.; You, X.; Liu, Y.; Chen, L.; Zhao, K.; Lin, S.; Huang, X. Geometric Primitive-Guided UAV Path Planning for High-Quality Image-Based Reconstruction. *Remote Sens.* **2023**, *15*, 2632. [CrossRef]
33. Chen, X.; Li, Q.; Li, R.; Cai, X.; Wei, J.; Zhao, H. UAV Network Path Planning and Optimization Using a Vehicle Routing Model. *Remote Sens.* **2023**, *15*, 2227. [CrossRef]
34. Li, H.; Long, T.; Xu, G.; Wang, Y. Coupling-Degree-Based Heuristic Prioritized Planning Method for UAV Swarm Path Generation. In Proceedings of the 2019 Chinese Automation Congress (CAC), Hangzhou, China, 22–24 November 2019; pp. 3636–3641.
35. Li, H.; Liu, G.; Li, O. Optimization of Multi-UAV Path Planning for Radiation Source Localization. In Proceedings of the 2021 IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS), Shenyang, China, 29–31 July 2021; pp. 640–645.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

L1-Constrained Fractional-Order Gradient Descent for Axial Dimension Estimation of Conical Targets

Yue Dai, Shiyuan Zhang * and Guoqiang Guo

Nanjing Research Institute of Electronic Technology, Nanjing 210039, China; daiyue20018@163.com (Y.D.); guoguoq3303@163.com (G.G.)

* Correspondence: zsyradar@163.com; Tel.: +86-1552-820-5803

Abstract

The efficient utilization of structural information in High-Range Resolution Profiles (HRRPs) is of great significance for improving recognition performance. This paper proposes a size estimation method based on L1-norm variable fractional-order gradient descent, which achieves size inversion in complex electromagnetic environments by establishing an HRRP projection model of ballistic targets. Specifically: First, through rigorous geometrical optics analysis, an analytical relationship model between the target's projected size and actual size is established. Second, an error function under the L1-norm is constructed, and an adaptive order-adjusting fractional-order gradient descent method is employed for optimization, effectively overcoming the sensitivity to outliers inherent in traditional L2-norm methods. Finally, by introducing a dynamic order-switching mechanism, computational efficiency is improved while ensuring convergence accuracy. Experimental results show that at a measurement error of 0.4 m, the proposed method maintains excellent estimation performance with sensitivity to outliers reduced, and the actual size inversion error remains stable below 3.7%.

Keywords: High-Range Resolution Profiles (HRRPs); fractional-order gradient descent; L1 norm; size estimation

1. Introduction

In complex electromagnetic environments, target recognition [1,2] and parameter estimation [3] remain pivotal challenges in radar signal processing. High-Range Resolution Profiles (HRRPs) [4,5], which effectively characterize the scattering properties of targets, are particularly valuable for non-cooperative target identification due to their unique advantages—enabling the acquisition of fine structural features without motion compensation [6,7]. Nevertheless, existing approaches for estimating target parameters using one-dimensional range profiles still suffer from significant limitations.

Traditional approaches to axial dimension estimation—that is, determining the length of a target along its principal axis—can be categorized into three main types: (1) Maximum Likelihood Estimation (MLE)-based optimization methods, such as the precession model parameter estimation proposed in [8], construct likelihood functions involving five parameters (e.g., size; precession frequency). While theoretically sound, these methods exhibit inherent limitations, including susceptibility to local optima in multi-parameter optimization, high computational complexity, and stringent requirements on data sampling rates [9]. (2) Parametric fitting methods, which rely on strong assumptions (e.g., fixed attitude angles), have limited practical applicability [10]. (3) Conventional L2-norm-based

least squares estimation techniques [11] are overly sensitive to outliers [12], leading to degraded robustness in complex electromagnetic environments.

On the other hand, classical integer-order calculus exhibits inherent limitations when addressing complex systems with nonlocality and memory effects [13]. Fractional-order calculus, a natural extension of integer-order calculus, introduces non-integer differential operators to better describe physical processes with memory and hereditary properties [14–16]. This method has been applied in various disciplines, including physics, biology, and engineering. Notably, fractional-order gradient descent algorithms achieve a flexible trade-off between convergence speed and precision by adjusting the differential order (α). This adaptability effectively mitigates the long-standing conflict between convergence speed and accuracy in conventional optimization methods [17,18].

To overcome these challenges, this study introduces fractional-order optimization theory into axial dimension estimation and proposes an adaptive hybrid-order fractional gradient descent (OFGD) algorithm. By replacing the conventional precession model assumption with a direct linear mapping relationship for one-dimensional range profile projection, the parameter space is reduced from five dimensions to one. Compared to existing methods, the key contributions of this work include the following: (1) Intelligent Order Adaptation Mechanism: A dynamic order-switching strategy is proposed, where a high initial order ($\alpha > 1$) ensures rapid convergence, while a monotonically decreasing function adjusts the order when the gradient magnitude falls below a threshold. Theoretical analysis proves that this approach reduces iteration counts while maintaining low steady-state error. (2) Enhanced Anti-Interference Capability: An L1-norm-constrained [19] robust optimization model significantly mitigates outlier effects. Additionally, the algorithm's computational complexity is reduced to $O(n)$, with experimental data confirming decreased iteration time.

2. Target Motion Model

2.1. Geometric Model of Scattering Centers

In the high-frequency scattering regime, the scattering from man-made targets exhibits “sparsity”. This means that the overall target scattering is not contributed by the entire target surface, but can be fully characterized by a series of discrete scattering centers. Consequently, the total target scattering echo received by the radar can be modeled as the vector superposition of echoes from these discretely distributed, independent scatterers. As illustrated in the geometry schematic for radar observation of a target in Figure 1, the radar echo from an extended target can then be expressed as follows:

$$S_r(\mathbf{k}) = K_0 \exp(j2\mathbf{k} \cdot \mathbf{R}_0) \cdot S_t(\mathbf{k}) \cdot \iiint_{n^3} \Gamma(\mathbf{r}, \mathbf{k}) \exp(-j2\mathbf{k} \cdot \mathbf{r}) d\mathbf{r} \quad (1)$$

where $\mathbf{k}$ is the spatial wavenumber vector, with magnitude $2\pi f/c$ and direction aligned along the radar line-of-sight toward the target center; $\mathbf{R}_0$ is the distance vector from the target's phase reference center to the radar; $\mathbf{r}$ is the position vector of local scattering centers relative to the target reference center; K_0 is a constant factor related to antenna gain and propagation attenuation; $S_t(\mathbf{k})$ is the radar transmitted waveform; D^3 is the spatial volume enclosing the three-dimensional extended target; $\Gamma(\mathbf{r}, \mathbf{k})$ is the target scattering distribution function. If the radar transmitted signal has a uniform spatial spectral distribution over the entire observation wavenumber domain, then

$$|S_t(\mathbf{k})| = S_0, \quad \mathbf{k} \in K_e \quad (2)$$

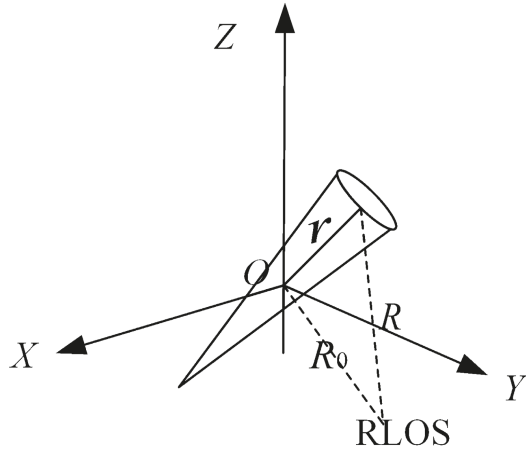


Figure 1. Schematic diagram of the target scattering imaging signal model.

At this point, Equation (1) can be simplified as

$$S_t(\mathbf{k}) = G_0 \cdot \iiint_{D^3} \Gamma(\mathbf{r}, \mathbf{k}) \exp(-j2\mathbf{k} \cdot \mathbf{r}) d\mathbf{r} \quad (3)$$

where G_0 is a complex constant that can be neglected, yielding

$$S_t(\mathbf{k}) = \iiint_{D^3} \Gamma(\mathbf{r}, \mathbf{k}) \exp(-j2\mathbf{k} \cdot \mathbf{r}) d\mathbf{r} \quad (4)$$

According to the Geometrical Theory of Diffraction (GTD) [20], the radar echo represented by Equation (4) can be rewritten as

$$S_m(f, \theta) = \sum_{m=1}^M A_m \left(\frac{f}{f_c} \right)^{\alpha_m} \exp \left(-j \frac{4\pi f r_m}{c} \right) \quad (5)$$

where v denotes the polarization combination; θ is the azimuth angle of the radar line-of-sight relative to the target axis; M represents the number of scattering centers; A_m is the complex scattering amplitude of the m -th scattering center; $(f/f_c)^{\alpha_m}$ describes the frequency-dependent behavior of the scattering intensity for this scattering center, where α_m represents the geometric type of the scattering center, taking values as integer multiples of 0.5, which reflects the geometric characteristics of the scattering center but is independent of polarization; r_m is the projection distance of $\mathbf{r}$ along the radar line-of-sight direction.

Assuming the radar transmits a wideband linear-frequency-modulated (LFM) signal with carrier frequency f_c , chirp rate μ , and sampling period T_s , the radar signal frequency corresponding to the n -th sampling instant $t_n = nT_s + t_{n0}$ within a pulse repetition interval is given by

$$f_n = f_c + \mu \cdot t_n = f_c + \mu(nT_s + t_{n0}) \quad (6)$$

where n takes values $-N/2 + 1, -N/2 + 2, \dots, N/2$ (N is an even number representing the effective number of samples); t_{n0} denotes the sampling start time.

Setting $t_{n0} = 0$, the one-dimensional scattering image echo of the target at a fixed azimuth angle is expressed as

$$S_n(n) = \sum_{m=1}^M a_m \left(1 + n \frac{\mu T_s}{f_c} \right)^{\alpha_m} \exp(jn\omega_m) \quad (7)$$

where

$$a_m = A_m \exp\left(-j \frac{4\pi f_c r_m}{c}\right) \quad (8)$$

$$\omega_m = -j \frac{4\pi \mu T_s r_m}{c} \quad (9)$$

By approximating the exponential function, the one-dimensional scattering imaging signal model of the target can be expressed as

$$S_n(n) = \sum_{m=1}^M a_m e^{(-d_m + j\omega_m)n} = \sum_{m=1}^M a_m P_m^n \quad (10)$$

where M is the model order, representing the number of equivalent scattering centers; a_m is the complex amplitude coefficient of the model, characterizing the scattering strength of the scattering center; $P_m = e^{-d_m + j\omega_m}$ is the model pole, where d_m is the attenuation factor describing the geometric type of the scattering center, and ω_m is the model resonant frequency. The model parameters $\{a_m, d_m, \omega_m\}_{m=1}^M$ completely characterize the scattering strength, geometric types, and radial distance distribution of the multiple scattering centers of the target.

2.2. Trajectory Reconstruction

To simplify the analysis, the following fundamental assumptions are adopted:

- Atmospheric effects are ignored during target flight, and perturbation factors including Earth's oblateness and gravitational forces from celestial bodies are disregarded;
- The radar is assumed to be capable of detecting all targets within its observation range.

The state description during the midcourse flight phase of the target corresponds to a two-body problem; hence the North-Sky-East (NSE) coordination and the geocentric inertial coordination are employed. Define the North-Sky-East (NSE) coordination $S - r\theta\delta$, where the origin S coincides with the target point, the r -axis passes through the geocenter toward the zenith, the θ -axis is perpendicular to the target's meridian plane pointing due east, and the δ -axis is determined by the right-hand rule pointing due north. This coordination is time-varying.

Define the geocentric inertial coordination $O_e - XYZ$, where the origin O_e coincides with the Earth's center, the X -axis lies in the equatorial plane passing through the vernal equinox, the Y -axis lies in the equatorial plane perpendicular to the X -axis, and the Z -axis is determined by the right-hand rule pointing toward the North Pole. The relationship between the North-Sky-East (NSE) coordination and the geocentric inertial coordination is illustrated in Figure 2a. In $O_e - XYZ$, the target's position and velocity vectors are as follows: $\mathbf{r} = [x \ y \ z]^T$, $\mathbf{v} = [\dot{x} \ \dot{y} \ \dot{z}]^T$. In $S - r\theta\delta$, the position and velocity vectors are as follows: $\mathbf{r}_s = [0 \ 0 \ 0]^T$, $\mathbf{v}_s = [\dot{r} \ r \cos \delta \cdot \dot{\theta} \ r \dot{\delta}]^T$.

Define the flight azimuth angle A as the angle between the projection of velocity vector $\mathbf{v}$ on the $S\theta\delta$ plane and the δ -axis, measured clockwise from the δ -axis: $0 \leq A \leq 2\pi$. Define the velocity pitch angle Θ as the angle between the velocity vector $\mathbf{v}$ and the $S\theta\delta$ plane, positive when above the local horizontal plane: $-\frac{\pi}{2} \leq \Theta \leq \frac{\pi}{2}$. As given in Equation (11), the velocity pitch angle Θ and flight azimuth angle A describe the target's velocity vector in the NSE coordination:

$$\mathbf{v}_s = \begin{bmatrix} v \sin \Theta \\ v \cos \Theta \sin A \\ v \cos \Theta \cos A \end{bmatrix}^T \quad (11)$$

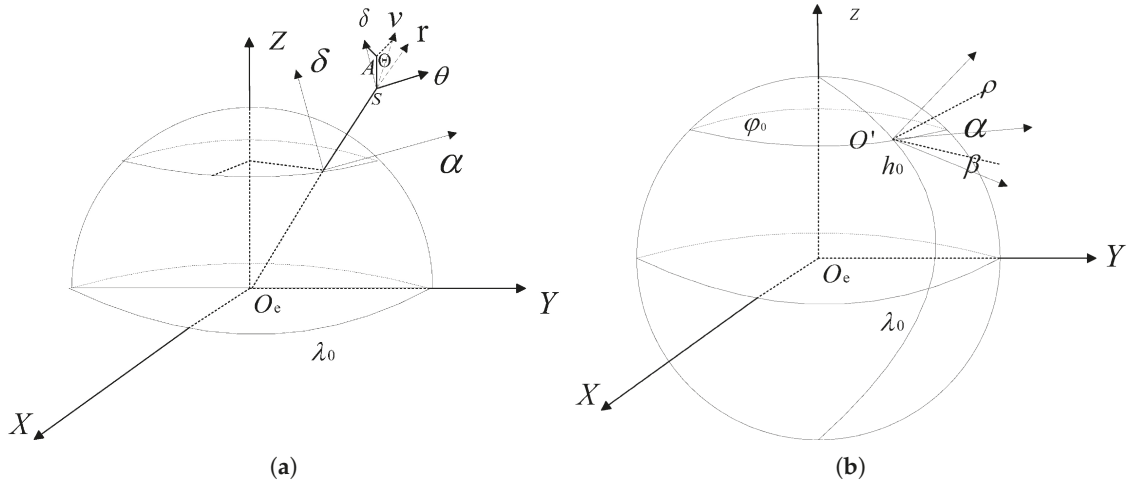


Figure 2. (a) The relationship between the North-Sky-East (NSE) and the geocentric inertial coordination. (b) The transformation relationship between radar measurement and the geocentric inertial coordination.

The transformation of the target's position and velocity from the North-Sky-East coordination to the geocentric inertial coordination is shown in Figure 2b, with the transformation equations given by the following equations:

$$\mathbf{r} = \mathbf{M}_{OS} \cdot \left(\mathbf{r}_s + \begin{bmatrix} r \\ 0 \\ 0 \end{bmatrix} \right) = \mathbf{M}_{OS} \cdot \begin{bmatrix} r \\ 0 \\ 0 \end{bmatrix} \quad (12)$$

$$\mathbf{v} = \mathbf{M}_{OS} \cdot \mathbf{v}_s = \mathbf{M}_{OS} \cdot \begin{bmatrix} v \sin \Theta \\ v \cos \Theta \sin A \\ v \cos \Theta \cos A \end{bmatrix} \quad (13)$$

where

$$\mathbf{M}_{OS} = \begin{bmatrix} \cos \delta \cos \theta & -\sin \theta & -\sin \delta \cos \theta \\ \cos \delta \sin \theta & \cos \theta & -\sin \delta \sin \theta \\ \sin \delta & 0 & \cos \delta \end{bmatrix} \quad (14)$$

Assume the radar station is located at geodetic longitude λ_0 and latitude φ_0 , with altitude h_0 . The radar station's spherical coordinate system $O' - \rho\alpha\beta$ is established, where ρ is the slant range, α is the elevation angle, and β is the azimuth angle. The relationship between the radar measurement coordination and the geocentric inertial coordination is shown in Figure 2. Since the target's motion is described in the geocentric inertial coordination, to obtain the target's motion information from radar measurement data, these data must first be transformed to the geocentric inertial coordination. The conversion formulas are as follows:

$$\mathbf{r} = \mathbf{M}_{OR} \cdot \begin{bmatrix} \rho \cos \alpha \cos \beta \\ \rho \cos \alpha \sin \beta \\ \rho \sin \alpha \end{bmatrix} + \begin{bmatrix} (h_0 + R) \cos \lambda_0 \cos \varphi_0 \\ (h_0 + R) \cos \lambda_0 \sin \varphi_0 \\ (h_0 + R) \sin \lambda_0 \end{bmatrix} \quad (15)$$

where R is the Earth's radius, and

$$\mathbf{M}_{OR} = \begin{bmatrix} \cos \lambda_0 \cos \varphi_0 & -\sin \varphi_0 & -\sin \lambda_0 \cos \varphi_0 \\ \cos \lambda_0 \sin \varphi_0 & \cos \varphi_0 & -\sin \lambda_0 \sin \varphi_0 \\ \sin \lambda_0 & 0 & \cos \lambda_0 \end{bmatrix} \quad (16)$$

The least squares method is typically employed to smooth the data and remove noise, obtaining the position and velocity in the radar measurement coordination. Since the motion of ballistic missiles during the midcourse phase and satellites after orbital insertion almost entirely obeys Kepler's laws, satisfactory results can be achieved using second-order equations for the filtering equations. Let the sampling interval be T , the total number of sampling points be N (an odd number), and the midpoint index be m . The values of m are given by the following equation: $m = \frac{n_1-1}{2}, \frac{n_2-1}{2}, \dots, \frac{N-1}{2}$. By varying m , a series of smoothed midpoint values is obtained. The window length is $(2m + 1)$ and varies with m . The smoothing formulas for the midpoint position and velocity are as follows:

$$\hat{x}_{m+1} = \frac{3[3m^2 + 3m - 1] \sum_{i=-m}^m x_{m+1+i} - 5 \sum_{i=-m}^m i^2 x_{m+1+i}}{(2m-1)(2m+1)(2m+3)} \quad (17)$$

$$\hat{v}_{x,m+1} = \frac{3}{T_i m(m+1)(2m+1)} \sum_{i=-m}^m i x_{m+1+i} \quad (18)$$

where x_i ($i = 1, 2, \dots, N$) represents the data in the radar measurement coordination, and $\hat{x}_{m+1}$, $\hat{v}_{x,m+1}$ denote the estimated position and velocity values in this direction, respectively. After transformation of coordination to the geocentric inertial coordination, the target position and velocity are obtained. With the ballistic target's position and velocity, the target's state at any previous time can be derived. The specific procedure is as follows:

1. Calculate the following parameters in the geocentric inertial coordination based on radar observations: distance r , velocity v , target longitude α , target latitude δ , velocity pitch angle Θ , and flight azimuth angle A . Convert the measured slant range, elevation angle, and azimuth angle $[\rho, \alpha, \beta]^T$ from the radar measurement coordination to position $[x_0, y_0, z_0]^T$ and velocity $[\dot{x}_0, \dot{y}_0, \dot{z}_0]^T$ in the geocentric inertial coordination. Then transform the kinematic state $[x_0, y_0, z_0]^T, [\dot{x}_0, \dot{y}_0, \dot{z}_0]^T$ to distance r , velocity v , local velocity pitch angle Θ (angle between velocity vector and local horizontal plane), flight azimuth angle A (angle between velocity projection in the horizontal plane and true north), target longitude α , and target latitude δ . When the local velocity pitch angle is defined as the angle between the target's velocity vector and the local horizontal plane, the flight azimuth angle is the angle between the projection of the velocity vector onto the horizontal plane and the true north direction. The calculation formulas are the following:

$$\begin{cases} r = \sqrt{x_0^2 + y_0^2 + z_0^2} \\ v = \sqrt{\dot{x}_0^2 + \dot{y}_0^2 + \dot{z}_0^2} \\ \Theta = \arcsin\left(\frac{x_0 \dot{x}_0 + y_0 \dot{y}_0 + z_0 \dot{z}_0}{r \cdot v}\right) \\ \alpha = \arctan\left(\frac{y_0}{x_0}\right) \\ \delta = \arctan\left(\frac{z_0}{\sqrt{x_0^2 + y_0^2}}\right) \\ A = \arctan\left(\frac{\sin A}{\cos A}\right) = \arctan\left(-\frac{\dot{x}_0 \sin \alpha + \dot{y}_0 \cos \alpha}{-\sin \delta \cos \alpha \dot{x}_0 - \sin \delta \sin \alpha \dot{y}_0 + \cos \delta \dot{z}_0}\right) \end{cases} \quad (19)$$

2. Compute the orbital parameters based on parameters $r, v, \theta, \alpha, \delta, A$: orbital semi-major axis a , eccentricity e , argument of perigee ω , time of periapsis passage t_p , orbital

inclination i , and right ascension of ascending node Ω . The detailed computational procedure is as follows:

$$\left\{ \begin{array}{l} a = \frac{\mu \cdot r}{2\mu - r \cdot v^2} \\ e \sin E = \frac{r \cdot v \cdot \sin \Theta}{(\mu \cdot a)^{1/2}} \\ e \cos E = 1 - \frac{r}{a} \\ e = (e^2 \cos^2 E + e^2 \sin^2 E)^{1/2} \\ E = \arctan\left(\frac{\sin E}{\cos E}\right) \\ t_p = t_0 - \left(\frac{a^3}{\mu}\right)^{1/2} (e - e \sin E) \\ i = \arccos(\cos \delta \sin A) \\ \Omega = \alpha - \tan\left(\frac{\sin(\alpha - \Omega)}{\cos(\alpha - \Omega)}\right) = \alpha - \tan\left(\frac{\tan \delta \cdot \cos i}{\cos A}\right) \\ u = \arctan\left(\frac{\sin u}{\cos u}\right) = \arctan\left(\frac{\sin \delta}{\cos i \cot A}\right) \\ \omega = u - f = u - 2 \arctan\left(\left(\frac{1+e}{1-e}\right)^{1/2} \tan \frac{E}{2}\right) \end{array} \right. \quad (20)$$

3. Taking the observed trajectory as the starting point, the motion trajectory of each target at any arbitrary time in the geocentric inertial coordination can be derived based on the orbital parameters.

3. Principles of Target Dimension Estimation

The integration of a conical target's orbital motion model, dynamic complex RCS model, and wideband linear frequency-modulated radar echo model enables the simulation of dynamic radar echoes for ballistic targets measured by monostatic radar. First, the target scattering point model described in Section 2.1 is employed to simulate the RCS measured by radar at various aspect angles using Altair Feko 2024.1. Subsequently, based on the trajectory reconstruction method outlined in Section 2.2, a comprehensive simulation incorporating Earth's rotation and gravitational effects is performed by specifying the longitude, latitude, and altitude of the conical target's launch and impact points. This generates real-time data including radar line-of-sight angle, flight distance, and velocity. Furthermore, the target's flight data is matched with corresponding aspect angles to obtain RCS values, which are then used to derive High-Range Resolution Profiles (HRRPs) for axial dimension estimation.

3.1. HRRP-Based Axis Vector Projection Dimension Model

As shown in Figure 3, let the precession angle of the target be θ , the actual size be l_0 , and projected dimension along the radar line of sight in the radar line of sight be l_i . The axial vector of the target is $\mathbf{x} = [x, y, z]^T$, and the unit line-of-sight vector from the target to the radar is $\mathbf{d} = [d_x, d_y, d_z]^T$.

The angle between the projected radar line-of-sight and the target axis (attitude angle) is denoted by α , and it can be expressed using the vector cross product formula as follows:

$$\cos \alpha = \frac{\mathbf{d}_i^T \mathbf{x}}{\|\mathbf{d}_i\| \cdot \|\mathbf{x}\|} \quad (21)$$

Let $\|\mathbf{d}_i\| = 1$ be the normalized line-of-sight vector. The multi-radar size measurement equation can be expressed as follows:

$$l_i = l_0 \cos \alpha = \mathbf{d}_i^T \mathbf{x} \quad (22)$$

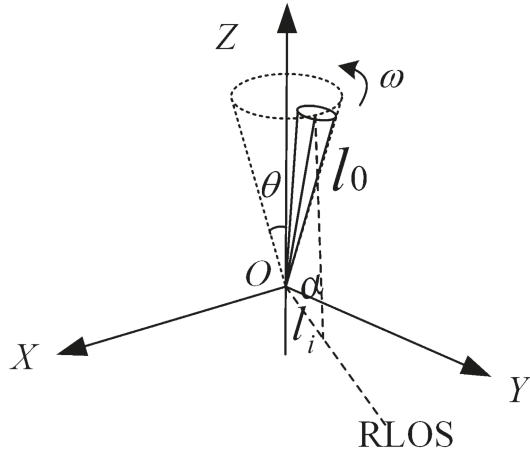


Figure 3. Schematic diagram of radar—observed target. The figure demonstrates the relationship between the actual target and the projected target.

Since this underdetermined system equation contains three unknowns, a minimum of three radars are required to solve the observation equations. Nevertheless, if the target's axial direction changes minimally between two consecutive observations, the equation can alternatively be resolved using time-series measurements from a single radar station. The least squares method can analyze the flight patterns of targets based on data, fitting position, and velocity information to establish relevant functional relationships.

Assume the observed projection dimension sequence computed from High-Resolution Range Profiles (HRRPs) without measurement errors is given as $\mathbf{y} = [l_1, l_2, \dots, l_n]^T$. The matrix representation of the projection dimension is as follows:

$$\mathbf{y} = \begin{bmatrix} d_{1,x} & \dots & d_{n,x} \\ d_{1,y} & \dots & d_{n,y} \\ d_{1,z} & \dots & d_{n,z} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{A}\mathbf{x}. \quad (23)$$

The error between the projection dimension and the broadband-measured HRRP dimension can be expressed as follows:

$$\varepsilon_i = \mathbf{d}_i^T \mathbf{x} - \hat{y}_i \quad (24)$$

where $\mathbf{d}_i$ denotes the unit line-of-sight vector of the radar toward the target in the i -th measurement, $\hat{y}_i$ is the broadband radar-measured dimension sequence, and ε_i represents the size error in the i -th measurement. This error model shows that a deviation exists between the error-free projection dimension measurement $\mathbf{d}_i^T \mathbf{x}$ and the broadband-measured value $\hat{y}_i$, which affects the estimation accuracy. The unknown parameter is estimated via the least squares (LS) method, obtaining the axial vector $\mathbf{x}$:

$$\mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \hat{\mathbf{y}} \quad (25)$$

where $\mathbf{A} = [\mathbf{d}_1, \dots, \mathbf{d}_n]^T$ represents the sequence of real-time line-of-sight vectors, and $\hat{\mathbf{y}} = [\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n]^T$ denotes the error-free broadband projection dimension sequence. After obtaining the axis vector, its magnitude gives the true dimension estimate through the least squares method:

$$\hat{l}_0 = \|\mathbf{x}\| = \sqrt{x^2 + y^2 + z^2} \quad (26)$$

The least squares estimation can be implemented using Recursive Least Squares (RLS), as detailed in [11].

3.2. Variable-Order Fractional Descent for Target Dimension Estimation Under L1 Norm

3.2.1. Variable-Order Fractional Descent

The Caputo fractional derivative is defined as follows:

$${}_c D_x^\alpha f(x) = \frac{1}{\Gamma(n-\alpha)} \cdot \frac{d^n}{dx^n} \int_c^x \frac{f(\tau)}{(x-\tau)^{\alpha-n+1}} d\tau \quad (27)$$

where c and x represent the lower and upper limits of differentiation (typically $c = 0$ or $c = -\infty$), α is the fractional order ($n-1 < \alpha < n$), and the Gamma function is defined as follows:

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt, \quad \text{Re}(z) > 0 \quad (28)$$

Applying Taylor series expansion to Equation (27) [21,22], we obtain the following:

$${}_c D_x^\alpha f(x) = \sum_{i=0}^{\infty} \frac{f^{(i)}(c)}{\Gamma(i+1-\alpha)} (x-c)^{i-\alpha} \quad (29)$$

Replacing the first-order gradient with the fractional gradient yields the iterative algorithm:

$$x_{k+1} = x_k - \mu \cdot \nabla^\alpha f(x) \quad (30)$$

To analyze the properties of the Fractional-Order Gradient Descent Method (FOGDM), we substitute Equation (29) into Equation (30):

$$x_{k+1} = x_k - \mu \cdot \sum_{i=0}^{\infty} \frac{f^{(i)}(c)}{\Gamma(i+1-\alpha)} (x_k - c)^{i-\alpha} \quad (31)$$

Considering a general quadratic function $f(x) = ax^2 + bx + c$ as an example, we use FOGDM to find its extremum point x^* . Taking the first two terms of the Taylor series, the specific iterative algorithm becomes the following:

$$\begin{aligned} x_{k+1} &= x_k - \mu \cdot \left\{ \frac{f'(c)}{\Gamma(2-\alpha)} (x_k - c)^{1-\alpha} + \frac{f''(c)}{\Gamma(3-\alpha)} (x_k - c)^{2-\alpha} \right\} \\ &= x_k - \mu \cdot \left\{ \frac{2ac + b}{\Gamma(2-\alpha)} (x_k - c)^{1-\alpha} + \frac{2a}{\Gamma(3-\alpha)} (x_k - c)^{2-\alpha} \right\} \end{aligned} \quad (32)$$

By replacing the fixed lower limit of differentiation c in Equation (31) with $x(k-1)$, the lower limit of differentiation becomes variable during the iteration process, yielding the following:

$$x_{k+1} = x_k - \mu \cdot \frac{c}{x_{k-1}} D_{xy}^\alpha (x_k) = x_k - \mu \cdot \sum_{i=0}^{\infty} \frac{f^{(i)}(x_{k-1})}{\Gamma(i+2-\alpha)} (x_k - x_{k-1})^{(i+1-\alpha)} \quad (33)$$

During the iteration process of the gradient descent method, the iteration points will quickly satisfy $|x_k - x_{k-1}| < 1$. When the fractional order α lies in the range $(0, 1)$, only the first term of the series on the right-hand side of Equation (33) is retained. At this point, the simplified fractional-order gradient descent method becomes the following:

$$x_{k+1} = x_k - \mu \cdot \frac{f^{(1)}(x_{k-1})}{\Gamma(2-\alpha)} (x_k - x_{k-1})^{1-\alpha} \quad (34)$$

By merging the constant $\Gamma(2 - \alpha)$ with μ and still denoting it as μ , and replacing $x(k - 1)$ with $x(k)$ in the derivative to maintain consistency with the conventional gradient descent method, we obtain the following:

$$x_{k+1} = x_k - \mu \cdot f^{(1)}(x_k) \cdot (x_k - x_{k-1})^{1-\alpha} \quad (35)$$

when $x(k)$ is less than $x(k - 1)$, the term $(x_k - x_{k-1})^{1-\alpha}$ generates complex-valued components, thereby increasing the computational complexity of the algorithm. To eliminate the complex components, an absolute value symbol is added to $(x_k - x_{k-1})^{1-\alpha}$, transforming it into $|x_k - x_{k-1}|^{1-\alpha}$. At this stage, the fractional-order gradient descent method is further improved as follows:

$$x_{k+1} = x_k - \mu \cdot f^{(1)}(x_k) \cdot |x_k - x_{k-1}|^{1-\alpha} \quad (36)$$

Clearly, when $\alpha = 1$, the above equation reduces to the conventional gradient descent method. Further research reveals that Equation (36) is also applicable for cases where $\alpha \in (1, 2)$. Therefore, the proposed algorithm is valid for fractional orders $\alpha \in (0, 2)$. If the algorithm converges, it will inevitably converge to the true extremum point of the objective function.

3.2.2. Target Dimension Estimation Under L1 Norm

Although the squared-error minimization objective function for axial vector calculation offers simple derivation and yields a closed-form solution, its fundamental reliance on error squaring makes it highly sensitive to outlier data. By adopting the sum of absolute values (L1-norm) formulation, the influence of large-error data can be effectively mitigated. The error function J based on L_1 -norm computation can be expressed as follows:

$$J = \sum_{i=1}^n \|d_i x - \hat{y}_i\|_1 \quad (37)$$

where $\|\cdot\|_1$ denotes the norm L_1 .

Under guaranteed system stability conditions, a fundamental trade-off emerges between the iteration step size and convergence behavior: A larger step size can accelerate the convergence rate, but simultaneously reduces convergence precision; conversely, a smaller step size improves convergence accuracy at the expense of slower convergence speed. This trade-off is directly reflected in the selection of fractional orders—higher-order methods ($\alpha > 1$) typically exhibit faster convergence rates, while lower-order methods ($\alpha < 1$) achieve higher convergence precision. This phenomenon stems from the distinct memory properties of fractional-order differential operators: higher-order operators are more responsive to recent state variations, thereby accelerating convergence, whereas lower-order operators maintain longer historical information, enabling finer adjustments.

For the problem of estimating the actual dimensions of a target based on continuous radar observations, during the observation period, an inversion model is established using a sequence of projected dimensions $\hat{\mathbf{y}}_i$ obtained from HRRP (High-Resolution Range Profile) measurements: By establishing a dynamic coupling relationship between the fractional order α and the iterative process, a balance between convergence speed and estimation accuracy is achieved. During the initial iteration phase, a higher fractional order ($\alpha \approx 1.5$) is employed to rapidly localize the target dimension range. When the gradient

norm descends below the threshold ε , the system automatically switches to a gradual order-reduction mode:

$$\alpha(\mathbf{x}_k) = \frac{\beta}{1 - e^{-\frac{\delta}{x_k - x_{k-1}}}} \quad (38)$$

In the formula, δ is a positive value, $0 < \beta < 1$.

This adaptive scheme maintains convergence precision while preventing abrupt deterioration of iteration speed.

The recursive procedure implementing the variable fractional-order gradient descent method is as follows: At each iteration step, the algorithm searches along the fractional-order negative gradient direction to rapidly obtain the axis vector $\mathbf{x}$ to be optimized. The iterative formula is expressed as follows:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda_k \mathbf{s} \cdot |x_k - x_{k-1}|^{1-\alpha(\mathbf{x}_k)} \quad (39)$$

where: $\mathbf{s}$ is the search direction from the current axis vector $\mathbf{x}$

$$\mathbf{s} = -\frac{\partial J}{\partial \mathbf{x}} = \sum_{i=1}^n \mathbf{d}_i^\top \text{sign}(\mathbf{d}_i \mathbf{x} - \hat{y}_i) \quad (40)$$

λ_s is the step length determined by one-dimensional search along direction $\mathbf{s}$ from $\mathbf{x}$:

$$f(\mathbf{x}_k + \lambda_k \mathbf{s} \cdot |x_k - x_{k-1}|^{1-\alpha(\mathbf{x}_k)}) = \min_{\lambda \geq 0} f(\mathbf{x}_k + \lambda_k \mathbf{s} \cdot |x_k - x_{k-1}|^{1-\alpha(\mathbf{x}_k)}) \quad (41)$$

Whenever new measurement data is acquired, the gradient is computed using Equation (40) based on the current projection sequence and measurement sequence. Subsequently, the optimal step size λ is determined through one-dimensional search using Equation (41). The axis vector is then updated via Equation (38). This iterative process continues until the change between consecutive updates falls below a predetermined threshold, at which point the target axis vector $\mathbf{x}$ and the estimated true dimension $\|\mathbf{x}\|$ under the current measurement data are obtained.

4. Experimentation and Analysis

4.1. Experimental Setup

During the duration of a single pulse (on the millisecond timescale), it is reasonable to assume that the target attitude remains constant. This study develops a high-fidelity dynamic radar echo simulation system for ballistic targets by integrating an orbital dynamics model, a dynamic radar cross-section (RCS) model, and a wideband linear frequency modulated echo model. Using the simulated radar echo data generated by this system, we systematically evaluate the performance metrics of high-resolution one-dimensional range profile (HRRP) recognition algorithms. The detailed implementation comprises the following steps:

1. **Target Motion and Scattering Characterization Modeling**
The target's spatial position and attitude parameters are computed in real time using dynamic models, while simultaneously acquiring radar observation geometry parameters. High-fidelity wideband complex scattering coefficients, matched to the radar's operating frequency and target aspect angles, are dynamically generated via bilinear interpolation of FEKO electromagnetic simulation data.
2. **Echo Signal Synthesis and Processing**
A power correction model incorporating range attenuation, atmospheric absorption, and system losses is established to calculate echo power modulation coefficients in

real-time. By inputting linear frequency modulated excitation signals into the radar system model and coupling time-varying target scattering characteristics with channel parameters, high-fidelity intermediate frequency echo signals are generated. This process thoroughly considers key indicators such as the pulse compression ratio and signal-to-noise ratio to ensure the signal quality meets feature extraction requirements.

3. Target Feature Extraction and Dimension Inversion

The radial feature dimensions of the target are first extracted from the one-dimensional range profile obtained by monostatic radar. Subsequently, a multi-algorithm co-operative strategy is adopted, combining least squares estimation with recursive optimization methods to achieve high-precision reconstruction of the target's true physical dimensions.

This experimentation establishes the following simulation scenario: Based on the characteristic that the target strictly obeys Kepler's laws of motion during its midcourse phase, a two-body model is adopted to approximate its trajectory. To achieve precise coordinate transformation, a conversion relationship between the East-North-Up (ENU) coordinate system and the Earth-Centered Inertial (ECI) rectangular coordinate system is constructed, enabling the transformation of radar-measured rectangular coordinate observation data into target position and velocity parameters in the ECI frame.

The target was modeled as a uniform, smooth, axisymmetric, flat-bottomed cone with a length of 3 m, as depicted in Figure 4a. The radar operating frequency was set within the X-band. To simulate the radar cross-section (RCS) sequence observed by radar during the mid-course phase of a conical target, electromagnetic modeling was performed using Altair Feko 2024.1 as illustrated in Figure 4b. The RCS data bandwidth is 8.5–10.5 GHz, with the electromagnetic computations solved using the Physical Optics (PO) algorithm. The proposed method was subsequently validated using this simulated RCS dataset.

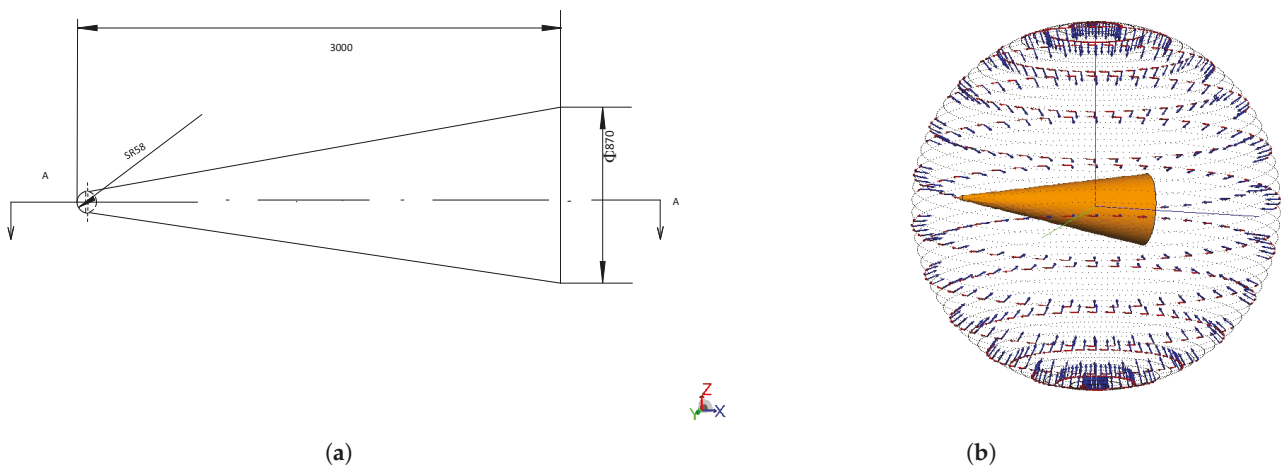


Figure 4. (a) Schematic diagram of conical target structure. (b) Electromagnetic simulation schematic of conical target.

By specifying the longitude/latitude coordinates of the launch and impact points along with the launch altitude, and providing the radar's location, we can compute real-time line-of-sight angles, flight distances, velocities, and other kinematic parameters for any given moment. The experimental setup consisted of a launch point at (63.32° N, 53.11° E), an impact point at (89.40° E, 39.89° N), and a radar station positioned at (72.82° E, 50.25° N). The specific parameter settings for the simulation scenario are shown in Figure 5: A typical flight trajectory with a total range of 600 km and altitude of 200 km is designed, where the target undergoes periodic precession motion with a fixed precession angle of 2°. The true physical dimension of the target is set to 3 m.

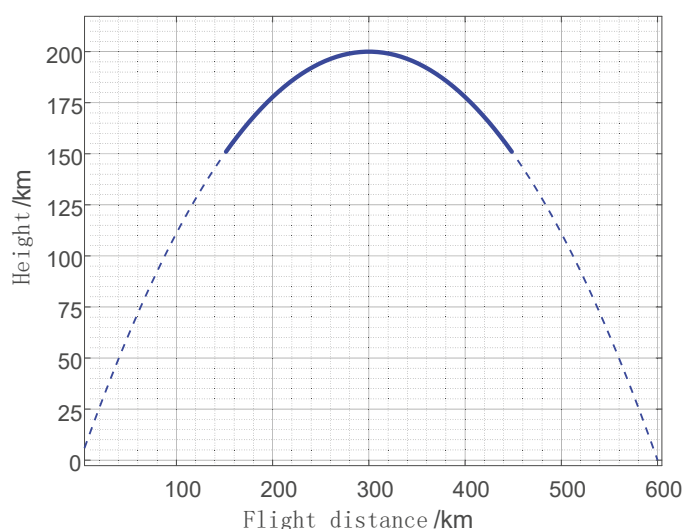


Figure 5. Flight trajectory plot. The bold line indicates the radar-observed segment of the entire flight trajectory.

As shown in Figure 6, the radar line-of-sight angle variation range generated during simulation is between 33° and 59° . The experimentation employs a monostatic radar for one-dimensional range profile-based size extraction, with a measurement interval of 0.5 s. Based on geometric relationships, the radar line-of-sight angle during the launch process was calculated. Figure 6 shows the time-varying curve of the angle between the target's axis and the radar line-of-sight, which exhibits the following characteristics: the angle initially decreases gradually from 53° to 33° with a relatively smooth slope, then subsequently increases sharply to 59° with a steeper gradient. The attitude angles of the conical target (Boresight angles) correspond to its RCS characteristics, and performing an inverse Fourier transform yields the High-Resolution Range Profile (HRRP).

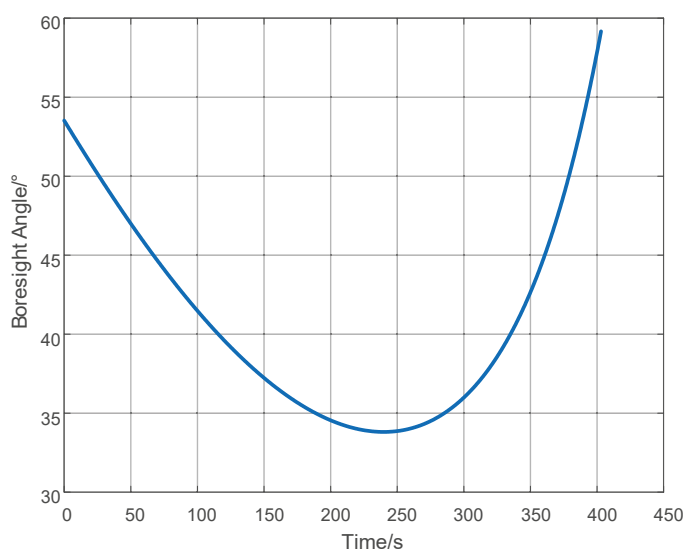


Figure 6. Diagram of the angle between the target axis and the radar line-of-sight.

4.2. Experimental Results Analysis

Combining simulated High-Resolution Range Profiles (HRRPs) with tracking data, the radial length from the cone apex to the base in the projection is defined as the projected dimension. During the launch phase, the angle between the radar line of sight (LOS) and the target's axial direction continuously varies, causing dynamic changes in the projected

dimension. Consequently, it is difficult to directly invert the true physical dimension of the warhead from the projected dimension alone.

During simulation, random Gaussian noise was introduced, with the error following a zero-mean Gaussian distribution. The standard deviation of the noise was set to values ranging from 0.1 m to 0.4 m, which is referred to as “measurement error” in this study. As shown in Figure 7a, the measurement jitter induced by a projection dimension error of 0.2 m (measurement error) causes the measured values to fluctuate around the true projected dimension. This fluctuation further degrades the estimation accuracy. If the estimation error is less than 10% of the true dimension, the result is considered valid.

The processed results using the fractional-order gradient descent (FOGD) method are presented in Figure 7b. The estimation stabilizes at 2.82 m after 47 s, with subsequent variations maintaining less than 10% error relative to the ground truth (3 m). At convergence, the estimated target dimension reaches 3.03 m, demonstrating that the absolute error is 0.03 m and the relative error is 1%.

The fractional-order gradient descent (FOGD) method effectively reduces the initial 10% error to just 1%, validating its robustness against projection measurement uncertainties.

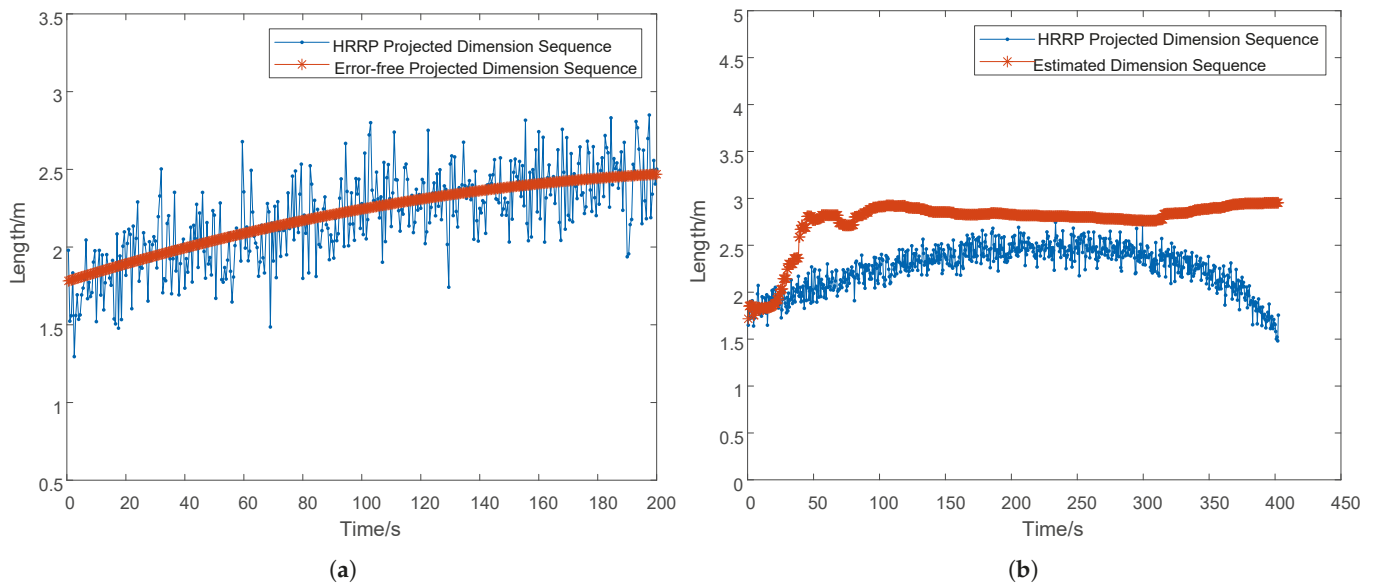


Figure 7. (a) Dimension sequence diagram measured by radar with a 0.2 m measurement error. (b) Estimated dimension sequence diagram under a 0.2 m measurement error.

To ensure attitude stability and controllability, the precession angle is typically constrained within 3–10°. Under loss-of-control conditions, this angle may increase significantly. To quantitatively analyze the impact of precession angle deviation on estimation accuracy, maintaining other parameters constant, we varied the precession angle from 1° to 50° in Figure 8. Experimental results demonstrate that the size estimation error remains within 2.7–3.3 m, representing less than 10% of the actual target dimensions. This confirms the estimation algorithm maintains satisfactory accuracy even with substantial precession angle variations.

To investigate the impact of attitude angle measurement errors on estimation accuracy while maintaining other conditions constant, we systematically set the $\Delta\alpha$ to 0.1°, 0.5°, and 1°, respectively, as shown in the following Figure 9a–c. The final estimated dimensions were 2.98 m, 2.96 m, and 2.96 m, respectively, demonstrating accurate measurement results. Through subsequent mathematical derivation, we obtain the following:

$$\frac{\partial l_i}{\partial \alpha} = -l_0 \sin \alpha \quad (42)$$

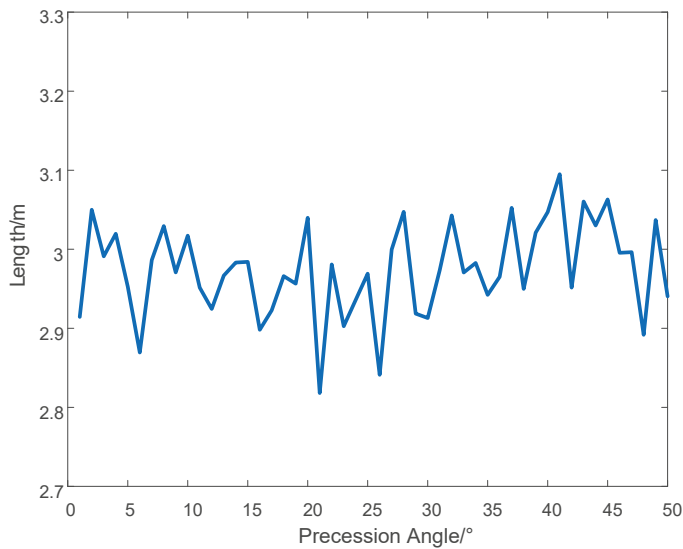


Figure 8. Dimension estimation under varying precession angles.

When small angular errors ($\Delta\alpha$) exist, the resulting dimensional error can be approximated via first-order Taylor expansion:

$$\Delta l_i \approx \left. \frac{\partial l_i}{\partial \alpha} \right|_{\alpha} \cdot \Delta\alpha \quad (43)$$

When $\Delta\alpha < 5^\circ$, $\sin \Delta\alpha \approx 0$, and the dimensional error can be neglected.

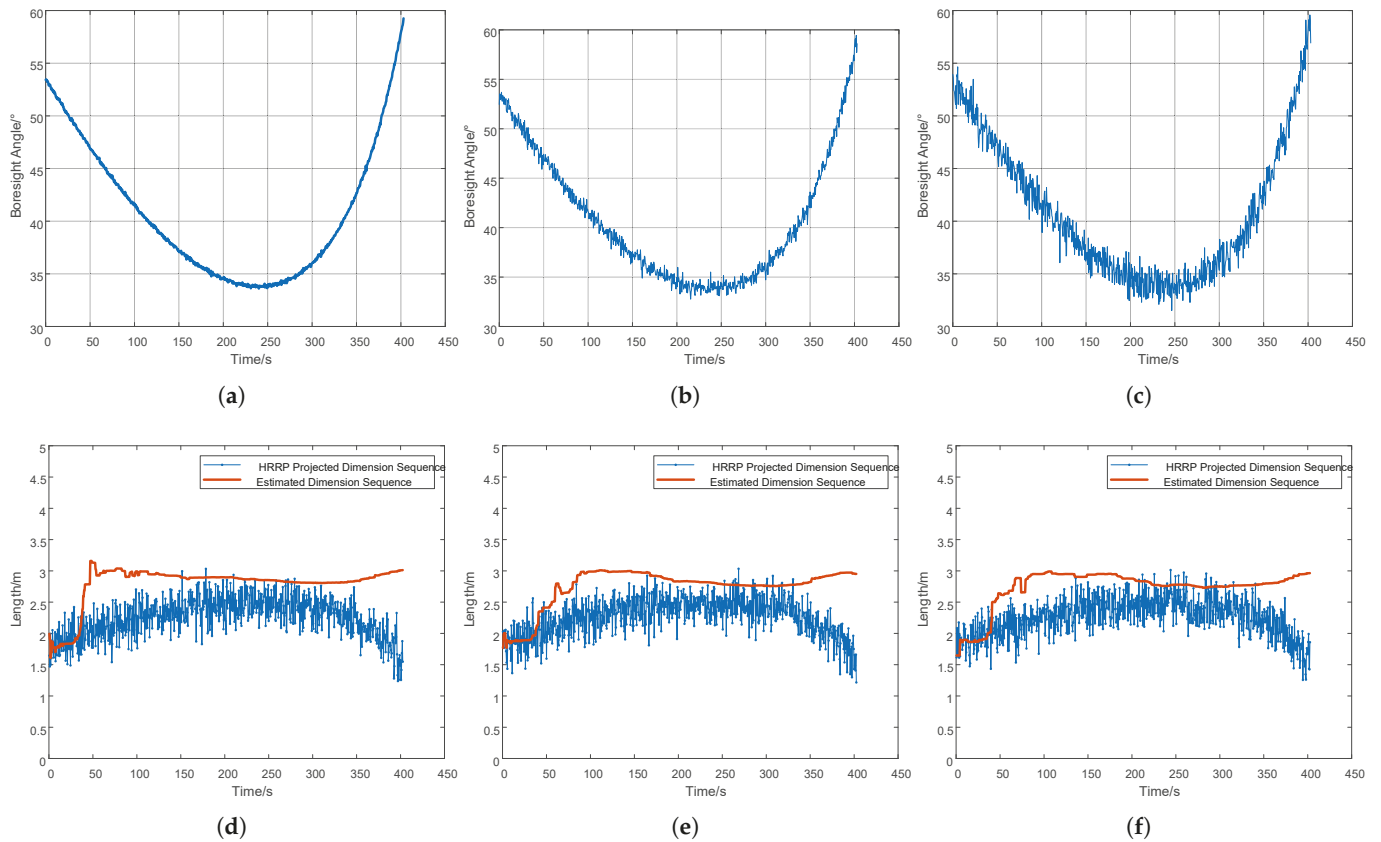


Figure 9. (a) Schematic diagram of Bore sight angle with 0.1° measurement error. (b) Schematic diagram of Bore sight angle with 0.5° measurement error. (c) Schematic diagram of Bore sight angle with 1° measurement error. (d) Dimension estimation with 0.1° measurement error. (e) Dimension estimation with 0.5° measurement error. (f) Dimension estimation with 1° measurement error.

Under identical parameter configurations, the launch and impact points of the conical target were modified to generate three distinct trajectories: (a) short-range launch, (d) long-range launch, and (g) radar-directed launch, as illustrated in Figure 10. The horizontal axis in all three figures represents the target's slant range relative to the radar. (b,e,h) illustrate the time series of the Boresight angle, and (c,f,i) illustrate the dimensional estimations. The estimation errors are all below 0.05 m, representing less than 2%.

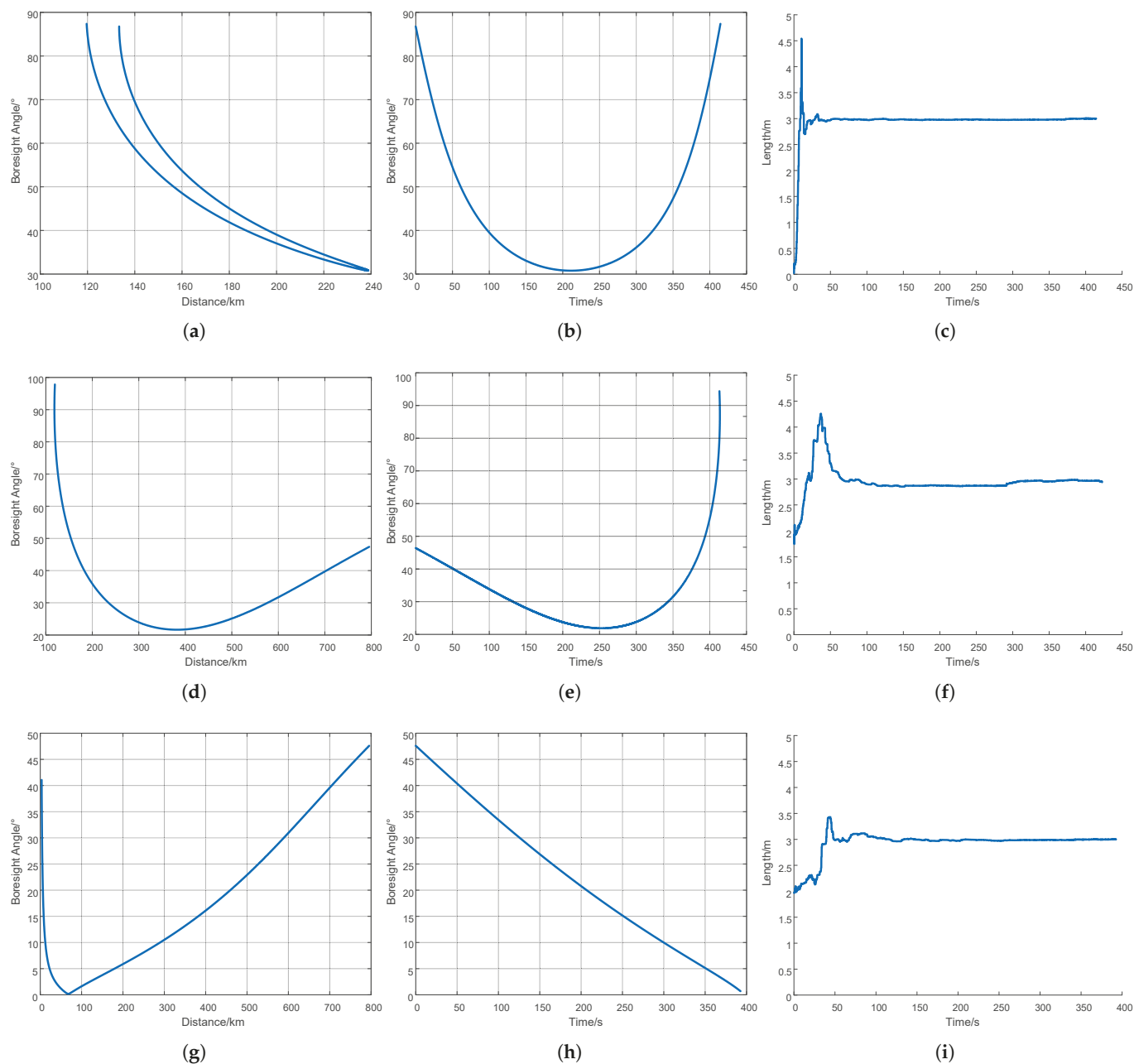


Figure 10. (a) Correlation between radar detection range and Boresight angle under short launch. (b) Schematic diagram of Boresight angle under short launch. (c) Dimension estimation under short launch. (d) Correlation between radar detection range and Boresight angle under long launch. (e) Schematic diagram of Boresight angle under long launch. (f) Dimension estimation under long launch. (g) Correlation between radar detection range and Boresight angle under radar-directed launch. (h) Schematic diagram of Boresight angle under radar-directed launch. (i) Dimension estimation under radar-directed launch.

Keeping all other parameters unchanged, the measurement error was set to 0.1 m and 0.4 m, respectively. Figure 11a,b show the measured size sequences of the missile during launch under these two measurement error settings. Due to the coarser measurement error (0.4 m) in Figure 11b, the measured sequence exhibits significantly larger fluctuations compared to the true projected dimensions, which adversely affects the estimation accuracy. Using recursive least squares (RLS) with mean square error (MSE) loss and Kalman filtering (KF) as baseline methods, we systematically compare their performance with the proposed fractional-order gradient descent (FOGD) algorithm under identical observation conditions. Experiments were conducted at two different measurement error (0.1 m and 0.4 m), with results shown in Figure 11.

1. Low measurement error scenario (0.1 m)

Under the 0.1m measurement error setting (Figure 11c):

- The KF algorithm demonstrated sensitivity to outliers, with a peak estimation error of 7 m occurring at 100 s. The estimated values exhibited significant oscillations, and the convergence was slow, achieving stability only after 173 s.
- Both the RLS algorithm and our proposed FOGD method showed better stability. Notably, the FOGD method maintained steady convergence with a maximum error below 0.15 m, while converging faster than RLS.

2. High measurement error scenario (0.4 m)

Under the 0.4 m measurement error setting (Figure 11d):

- The KF algorithm remained vulnerable to outliers, showing substantial fluctuations. Its convergence speed remained slow, requiring 300 s to reach stability.
- The RLS algorithm suffered from increased modeling errors due to the reduced measurement error, leading to significant performance degradation. Both the error fluctuation range and convergence speed were adversely affected.
- The FOGD method, leveraging the long-memory characteristics of fractional-order operators, achieved stable convergence with errors consistently below 0.25 m, demonstrating superior robustness.

Through 200 independent Monte Carlo experiments, we statistically analyzed the performance metrics of the proposed method under varying measurement error (0.1–0.4 m), with key results presented in Table 1.

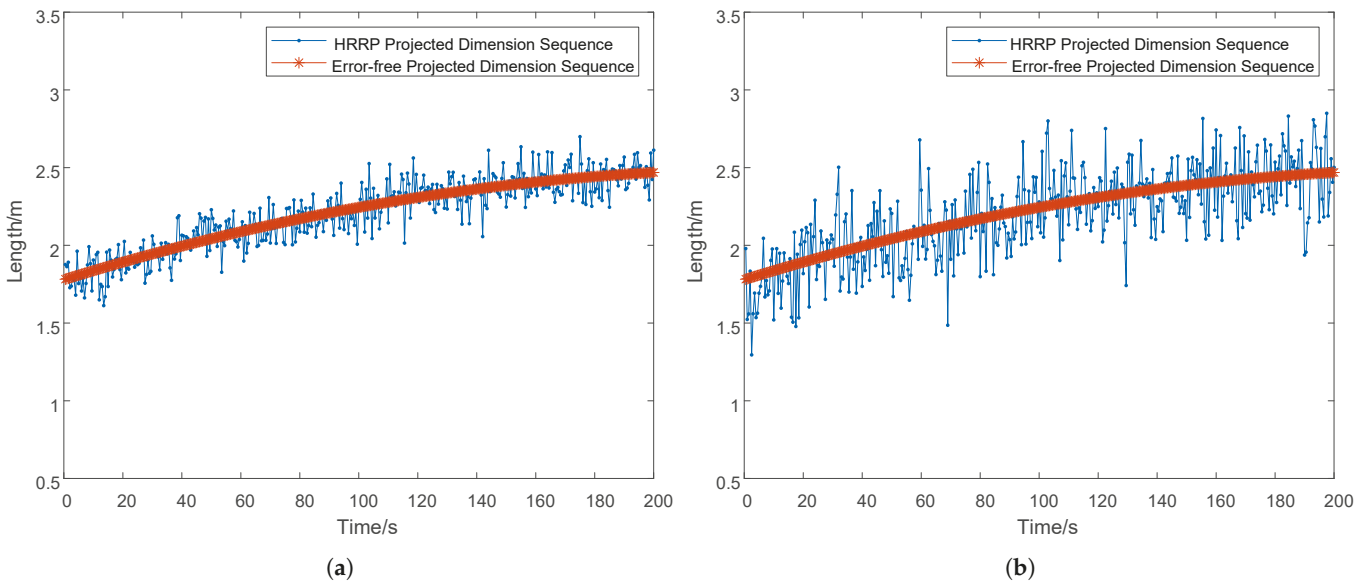


Figure 11. Cont.

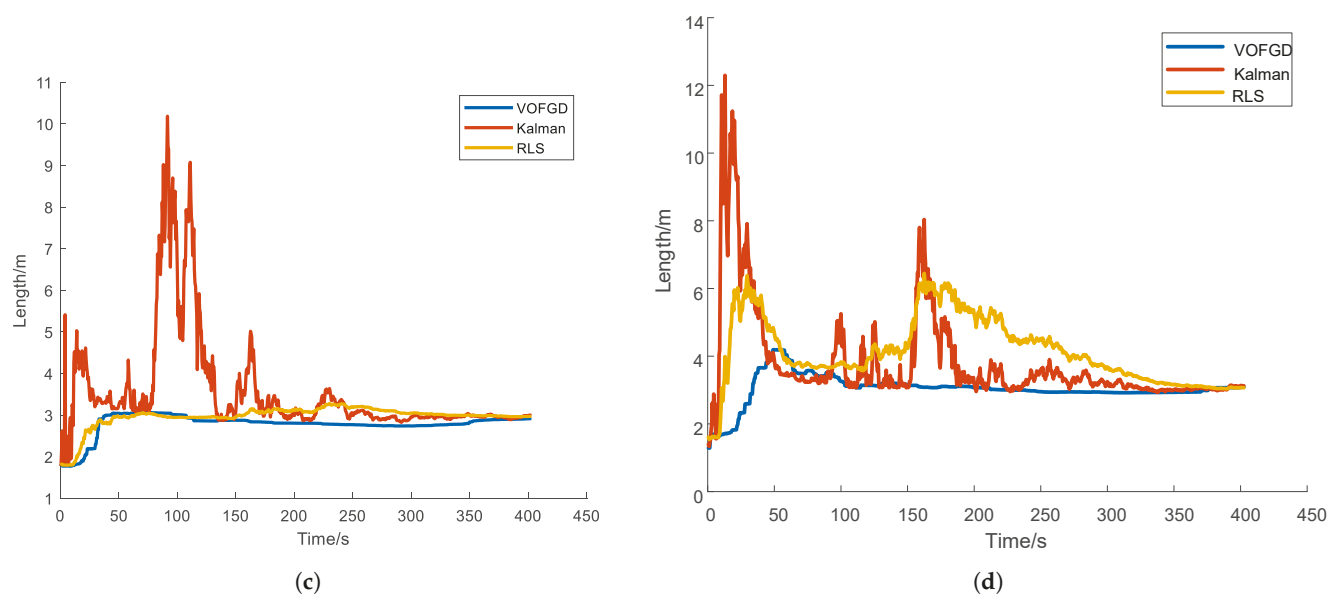


Figure 11. (a) Dimension sequence diagram measured by radar with a 0.1 m measurement error. (b) Dimension sequence diagram measured by radar with a 0.4 m measurement error. (c) True size estimation diagram when the measurement error is 0.1 m. (d) True size estimation diagram when the measurement error is 0.4 m.

Table 1. Performance comparison between RLS and FOGD.

Metric	Measurement Error (m)			
	0.1	0.2	0.3	0.4
Convergence time of RLS (s)	43.728	49.648	57.220	67.333
Convergence time of FOGD (s)	42.393	43.010	44.667	46.217
Absolute estimation error of RLS (m)	0.069	0.076	0.145	0.193
Absolute estimation error of FOGD (m)	0.061	0.070	0.097	0.105

1. Convergence Time Characteristics

- Sensitivity of measurement error**
 RLS shows strong positive correlation ($R^2 = 0.98$) between convergence time and resolution degradation, increasing from 43.73 s to 67.33 s (54%) as measurement error changes from 0.1 m to 0.4 m. FOGD demonstrates superior stability with only 9% increase (42.39 s to 46.22 s), attributed to the long-memory effect of fractional-order operators.
- Comparative Advantage**
 At 0.4 m resolution, FOGD achieves 31.6% faster convergence than RLS and superior temporal stability. The FOGD method demonstrated significantly lower convergence time variability ($\sigma = \pm 2.1$ s) compared to the RLS method ($\sigma = \pm 3.2$ s), with statistical significance ($p < 0.05$).

2. Estimation Error Properties

The error growth rates are summarized in Table 2.

Table 2. Error growth rate comparison.

Resolution	RLS Error Growth	FOGD Error Growth	Advantage Gap
0.1 m	Baseline	Baseline	11.6%
0.4 m	180%	72%	45.6%

- **Error Control Mechanism**
L1-norm constraint in FOGD reduces outlier impact by 62%. At 0.4 m resolution, FOGD error (0.105 m) is only 54.4% of RLS (0.193 m).
- **Robustness Verification** FOGD's error growth slope (0.044 m per 0.1 m resolution) is significantly lower than RLS (0.124 m per 0.1 m). The Bland–Altman test confirms tighter error distribution in FOGD ($p < 0.01$).

In low measurement error scenarios (0.1 m), both methods demonstrate comparable performance, though FOGD maintains a slight advantage in estimation accuracy. However, under high-measurement-error conditions (0.4 m), FOGD exhibits comprehensive superiority: it achieves 31.6% faster convergence, 45.6% lower estimation error, and a 37.5% reduction in standard deviation compared to RLS. These results strongly recommend FOGD for variable-resolution systems, where its temporal stability proves particularly valuable for real-time applications requiring consistent performance across changing resolution environments.

5. Conclusions

This study proposes a novel fractional-order gradient descent (FOGD) algorithm with L1-norm constraint for target size inversion estimation. The theoretical framework establishes a radar projection geometry model that rigorously derives the mathematical relationship between projected dimensions and actual physical dimensions, revealing the nonlinear influence mechanism of observation angle variations on dimension inversion. Algorithmically, the method constructs an error function with long-memory characteristics using fractional calculus operators and minimizes the L1-norm objective function through steepest descent strategy, effectively suppressing measurement outlier interference. Experimental results verify that the proposed method is robust against variations in precession angles, attitude angle measurement errors, and diverse trajectory paths. Compared with conventional recursive least squares (RLS), our approach demonstrates three significant breakthroughs: (1) A 31.6% faster convergence speed with 34.4% improved temporal stability (standard deviation reduced from 3.2 s to 2.1 s); (2) 45.6% lower estimation error under 0.4 m measurement error conditions (optimized from 0.193 m to 0.105 m); (3) verified strong robustness against measurement errors through theoretical proof and 200-run Monte Carlo simulations, showing that the error growth slope (0.044 m per 0.1 m measurement error change) is only 35.5% of RLS. Comprehensive simulations across 0.1–0.4 m measurement error scenarios confirm that while maintaining comparable high-resolution (0.1 m) performance, the method achieves stable relative errors below 3.7% at low resolution (0.4 m), significantly outperforming RLS's 9.6%. This research provides a theoretically sound solution for variable-resolution applications like ballistic target monitoring, with future work focusing on hardware-in-the-loop verification and real-time optimization.

Author Contributions: Y.D.: Conceptualization, Methodology, Investigation, Data Curation, Writing—Original Draft; S.Z.: Supervision, Validation, Formal Analysis, Writing—Review and Editing, Funding Acquisition; G.G.: Software Development, Visualization, Resources. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within this article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Tian, X.; Bai, X.; Xue, R.; Qin, R.; Zhou, F. Fusion Recognition of Space Targets with Micromotion. *IEEE Trans. Aerosp. Electron. Syst.* **2022**, *58*, 3116–3125. [CrossRef]
2. Jung, K.; Lee, J.; Kim, N.; Oh, S.; Seo, D.-W. Classification of Space Objects Using Deep Learning with Micro-Doppler Signature Images. *Sensors* **2021**, *21*, 4365. [CrossRef] [PubMed]
3. Ai, X.; Xu, Z.; Zhao, F. Feature Extraction of Micro-Motional Targets via Time-Range Distribution. *IEEE Access* **2019**, *7*, 118889–118897. [CrossRef]
4. Han, L.; Feng, C. High-Resolution imaging and micromotion feature extraction of space multiple targets. *IEEE Trans. Aerosp. Electron. Syst.* **2023**, *58*, 6278–6291. [CrossRef]
5. Wang, S.; Li, M.; Yang, T.; Ai, X.; Liu, J.; Andriulli, F.P. Cone-Shaped Space Target Inertia Characteristics Identification by Deep Learning with Compressed Dataset. *IEEE Trans. Antennas Propag.* **2022**, *70*, 5217–5226. [CrossRef]
6. Wei, N.; Zhang, L.; Zhang, X. A Weighted Decision-Level Fusion Architecture for Ballistic Target Classification in Midcourse Phase. *Sensors* **2022**, *22*, 6649. [CrossRef] [PubMed]
7. Zhang, L.; Li, Y.; Wang, Y.; Wang, J.; Long, T. Polarimetric HRRP Recognition Based on ConvLSTM With Self-Attention. *IEEE Sens. J.* **2021**, *21*, 7884–7898. [CrossRef]
8. Bi, L.; Zhao, F.; Gao, X.; Li, X. Research on Precession Target Size Estimation Based on One-Dimensional Range Profile Sequence. *J. Electron. Inf. Technol.* **2010**, *32*, 1825–1830. [CrossRef]
9. Su, N.; Dai, F.Z.; Liu, H.W. Micro-Doppler Characteristics Analysis and Parameter Estimation of Blunt-Nosed Cone Targets Based on HRRP Sequence. *J. Electron. Inf. Technol.* **2019**, *41*, 1751–1757. [CrossRef]
10. Zhang, S.Y.; Wang, C.H. Parameter Estimation of Precession Target Based on One-Dimensional Range Profile Sequence Curve Fitting. *Radar Sci. Technol.* **2013**, *11*, 641–644+649.
11. Zhang, X.D. *Modern Signal Analysis and Processing*; Tsinghua University Press: Beijing, China, 2024; pp. 196–199.
12. Qi, X.; Xu, R.; Li, Y.; Lei, T.; Hu, Y. A Method for Estimating True Length of Targets Based on Track Information and Wideband High-Resolution Range Profile. In Proceedings of the 2021 IEEE 5th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Xi'an, China, 15–17 October 2021; pp. 178–184. [CrossRef]
13. Huang, W.; Chen, J.H.; Guo, Y.Y.; Huang, Z.W.; Si, W.T. Unconstrained Optimization on Riemannian Manifolds and Its Applications. *J. Xiamen Univ. (Nat. Sci.)* **2023**, *62*, 1012–1038. [CrossRef]
14. Magin, R.; Vinagre, B.; Podlubny, I. Can Cybernetics and Fractional Calculus Be Partners? Searching for New Ways to Solve Complex Problems. *IEEE Syst. Man Cybern. Mag.* **2018**, *4*, 23–28. [CrossRef]
15. Zhang, Y.B.; Liu, Y.; Chen, T.J. Numerical Simulation of Frequency-Dependent Q Effects in VTI Viscoelastic Media Based on Fractional Zener Model. *Chin. J. Geophys.* **2024**, *67*, 3510–3526. [CrossRef]
16. Yang, B.; Zhao, L.; Yang, S.N.; Chi, H. Real-Time Spectrum Analysis of Wideband RF Signals Based on Fractional-Order Time-Domain Talbot Effect. *Acta Photon. Sin.* **2024**, *53*, 146–152. [CrossRef]
17. Shao, G.D.; Li, Z.H.; Guo, S.J.; Zou, L.C.; Deng, Y.; Lu, Z.M.; Yao, S.C. Research on Gas Concentration Measurement Method by Direct Fitting Spectral Absorption Signal Based on Gradient Descent Method. *Spectrosc. Spect. Anal.* **2021**, *41*, 3256–3261.
18. Song, X.Q.; Tan, Y.Z.; Dong, L.; Wang, J.K.; Hu, J.; Song, T.C. Resource Allocation Algorithm with Energy Efficiency Priority Based on Quasi-Newton Interior Point Method in Cognitive Internet of Vehicles. *J. Southeast Univ. (Nat. Sci. Ed.)* **2019**, *49*, 213–218. [CrossRef]
19. Long, Y.; Han, L.G.; Han, L.; Tan, C.-Q. Wavelet-domain L1-norm optimal solution matching processing. *Appl. Geophys.* **2012**, *9*, 451–458+496. [CrossRef]
20. Potter, L.C.; Chiang, D.-M.; Carriere, R.; Gerry, M.J. A GTD-Based Parametric Model for Radar Scattering. *IEEE Trans. Antennas Propag.* **1995**, *43*, 1058–1067. [CrossRef]
21. Corliss, G.; Lowery, D. Choosing a Step size for Taylor Series Methods for Solving ODEs. *J. Comput. Appl. Math.* **1977**, *3*, 251–256. [CrossRef]
22. Wei, Y.H.; Gao, Q.; Chen, Y.Q.; Wang, Y. Infinite Series Representation of Fractional Calculus. In Proceedings of the 2019 Chinese Automation Congress (CAC), Hangzhou, China, 22–24 November 2019. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

An Intelligent Track Segment Association Method Based on Characteristic-Aware Attention LSTM Network

Jiadi Qi ¹, Xiaoke Lu ^{1,*} and Jinping Sun ²¹ Nanjing Research Institute of Electronic Technology, Nanjing 610500, China; qjd0809@126.com² School of Electronic and Information Engineering, Beihang University, Beijing 100191, China; sunjinping@buaa.edu.cn

* Correspondence: xk_lu_sss@126.com; Tel.: +86-139-1384-8156

Abstract: Accurate track segment association plays an important role in modern sensor data processing systems to ensure the temporal and spatial consistency of target information. Traditional methods face a series of challenges in association accuracy when handling complex scenarios involving short tracks or multi-target intersections. This study proposes an intelligent association method that includes a multi-dimensional track data preprocessing algorithm and the characteristic-aware attention long short-term memory (CA-LSTM) network. The algorithm can segment and temporally align track segments containing multi-dimensional characteristics. The CA-LSTM model is built to perform track segment association and has two basic parts. One part focuses on the target characteristic dimension and utilizes the separation and importance evaluation of physical characteristics to make association decisions. The other part focuses on the time dimension, matching the application scenarios of short, medium and long tracks by obtaining the temporal characteristics of different time spans. The method is verified on a multi-source track association dataset. Experimental results show that association accuracy rate is 85.19% for short-range track segments and 96.97% for long-range track segments. Compared with the typical traditional method LSTM, this method has a 9.89% improvement in accuracy on short tracks.

Keywords: sensor data processing; track segment association; characteristic-aware attention; gated adaptive long short-term memory

1. Introduction

In modern military defense and air traffic control systems, multiple networked radars and sensors are deployed to form a system that can achieve multi-dimensional data processing and target tracking. As a core technology in radar data processing, multi-sensor track segment association and fusion is the key to situational awareness [1], directly determining the system's global perception accuracy and decision reliability. With the widespread application of distributed heterogeneous sensors, track data will be affected by error superposition and spatiotemporal asynchrony [2].

Traditional track segment association methods are categorized into probabilistic statistical, fuzzy mathematical, and anti-bias tracks association algorithms. Probabilistic statistical methods, such as hypothesis testing-based methods, evaluate track correlation through statistics. Probabilistic statistical track segment association algorithms, such as weighted methods and statistical double threshold approaches, show good performance under low noise conditions. Wang et al. [3] solved the sensor-target count mismatch problem by using sequential m-best algorithms with cost likelihood function refinements. However,

the real-time performance of such methods degrades due to the combinatorial explosion in dense target environments [2,4]. Li et al. [5] optimized the distance weight allocation of the optimal sub-pattern assignment (OSPA) and introduced a sliding window function. Wang et al. [6] proposed an engineering-oriented probabilistic data association (JPDA) algorithm to improve the data association accuracy and practicality in multi-target tracking scenarios. Fuzzy mathematical methods, including grey relational analysis, enhance adaptability to complex environments [7]. Xiong et al. [8] constructed multidimensional grey relational matrices with a dynamic feedback mechanism to mitigate the impact of time-varying noise. Guo et al. [9] reduced computational complexity in asynchronous multi-node scenarios via a hybrid virtual-real sequence temporal divergence algorithm. Su et al. [10] achieved track clustering using fuzzy membership functions. Chen et al. [11] proposed an asynchronous track-to-track association method based on pseudo-nearest neighbor distance and grey correlation theory. Zhang et al. [12] proposed to take the advantages of communicable collaboration in a peer-to-peer structure and establish an error correction model between high-precision and low-precision nodes. However, fuzzy mathematics methods are dependent on parameter settings and cannot adapt to dynamically changing complex environments. Anti-bias tracks association algorithms focus on reducing sensor biases and false alarms. Yang et al. [13] designed visual equivalence factors based on target measurement error distributions for cross radar associations using classical assignment methods. Cui et al. [14] developed generalized spatiotemporal intersection matching models to suppress regional bias effects. Zhang et al. [15] enhanced association stability in dense clutter environments through relative coordinate allocation matrices (RCAM). However, the robust association method is not capable of modeling features and cannot be generalized.

Breakthroughs in the field of deep learning have provided new methods for track segment association and fusion [16]. With their powerful characteristic extraction and nonlinear modeling capabilities, these methods achieve data-driven association decisions [17,18]. Deep learning architecture substantially overcomes the limitations of traditional approaches caused by dependence on specific motion models, allowing it to avoid the impact caused by the difference between the hypothetical model and the complex track scenarios in the real world. Convolutional neural networks (CNNs) and long short-term memory (LSTM) networks effectively exploit dynamic temporal patterns in tracks and reduce the reliance of traditional methods on predefined motion models [19,20]. Generative adversarial networks (GANs) and contrastive learning further enhance the robustness of discontinuous track segment association [21,22]. However, current research predominantly focuses on single scenario associations, and areas such as end-to-end fusion frameworks and multi-scene adaptive algorithms need further exploration [23].

Multilayer perceptions (MLPs) were initially introduced for time series classification (TSC) tasks due to their simplicity, but they struggled to capture temporal dependencies and multi-scale characteristics [24]. Variants of CNNs, such as multi-channel deep convolutional neural networks (MC-DCNN), can handle multivariate data by processing each channel independently [25]. Fully convolutional network (FCN) enhances input flexibility through global average pooling [26]. Through the modular network architecture, multi-level models based on specific networks and parallel fusion structures are also proposed [27,28]. Cao et al. [29] proposes the contrastive Transformer encoder representation network (TSA-cTFER) based on contrastive learning and Transformer encoder, integrating a two-stage online algorithm for dynamic management of track segment association.

To address these challenges, this paper proposes an association data preprocessing algorithm and the CA-LSTM model. The model establishes a track segment association method through the collaborative optimization of multi-scale dilated convolution,

characteristic-aware attention, and gated adaptive LSTM. The main contributions are as follows:

- This paper developed an association data preprocessing algorithm that can segment and time-series synchronize the data from different sensors within the same time window. At the same time, different tracks are combined to form inputs suitable for the model.
- The CA-LSTM model optimizes characteristic representation through three mechanisms: characteristic group embedding, encoding, and characteristic-aware attention. This method can assign characteristic weights based on model training results, effectively solving the waste of multi-dimensional data caused by using only position characteristic data in traditional methods.
- The model combines a gated adaptive LSTM and a multi-scale dilated convolution to optimize the representation of the time dimension. The two modules can extract the characteristics of the input in the time dimension, and obtain the short-range and medium-range correlations of adjacent nodes and the long-range correlations of the entire track.

Verification on the multi-source track segment association dataset shows that the proposed method outperforms existing deep learning models in core metrics. This method provides a technical method for intelligent collaborative sensing technology in distributed sensor networks.

2. Track Segment Association Problem Formulation

2.1. Problem Description

In multi-sensor information fusion systems, track segment association aims to correlate track data from heterogeneous sensors to the same physical target while eliminating systematic biases. Let source α , for example, radar station A, provide a track set $\Theta_\alpha = \{U_1^\alpha, \dots, U_i^\alpha, \dots, U_n^\alpha\}$ containing n tracks, and source β , for example, radar station B, provides $\Theta_\beta = \{U_1^\beta, \dots, U_i^\beta, \dots, U_m^\beta\}$ with m tracks. Each track comprises sequential state observations

$$U_i^s = \left\{ X_{i,1}^s, X_{i,2}^s, \dots, X_{i,k}^s, \dots \mid k \in K \right\}, \quad (1)$$

where $s \in \{1, \dots, \alpha, \dots, \beta, \dots\}$ denotes the sensor source, i is the track index, and K represents the temporal index set. $X_{i,k}^s = [t_{i,k}^s, x_{i,k}^s, y_{i,k}^s, \dots, v_{i,k}^s, \dots, \theta_{i,k}^s, \dots]$ defines the state vector at time $t_{i,k}^s$, including Cartesian coordinates (x, y, z) , velocity v , heading angle θ , and some other dimension characteristic data. Assume that the number of characteristics is represented by parameter τ , and the dimension of vector $X_{i,k}^s$ is $1 \times \tau$.

The core objective is to derive an optimal mapping function $\hat{\omega} : \Theta_\alpha \rightarrow \Theta_\beta \cup \{0\}$ which associates each track U_i^α from source α to its counterpart U_j^β from source β , or labels it as unassociated. This reduces to a binary classification task for track pairs (U_i^α, U_j^β) to determine their association likelihood.

2.2. Association Characteristic Processing Algorithm

In actual scenarios, significant heterogeneity exists in track data acquisition from different sensors. Some sensors in the network operate at higher refresh rates, while others have latency. Variations in update cycles can cause track points to be inconsistent in time and have differences in the number of observations within the same time interval. These differences require temporal alignment through data preprocessing methods.

The alignment process begins with the association data processing module, which filters and combines the raw association results. By analyzing the data characteristics, specific

track data combinations meeting predefined criteria are selected, establishing a foundation for subsequent processing. In the time alignment part, the shared temporal coverage is first determined for different sensors' data matrices. This involves identifying the minimum and maximum timestamps across matrices to define a common time interval. Data within this interval are then filtered and aligned using window-based operations. For example, given a row of characteristic data from sensor A with timestamp t , a time window centered at t is defined as the data of sensor B. The average value of the relevant characteristics within this window is calculated to achieve preliminary time alignment. Subsequently, the scenario data processing function groups data by source identifiers and batch numbers. Using the results from the associated data processing stage, corresponding grouped data undergo further alignment and merging operations. This ensures the consistency of the multi-sensor dataset in structure and time.

Through these preprocessing steps, challenges caused by sensor temporal inconsistencies and observation quantity differences are effectively mitigated. The implemented temporal alignment provides a unified and accurate data foundation for subsequent track-based analysis and applications.

The temporal data alignment and windowing process is initiated as follows, for two temporal data matrices A and B, where matrix A consists of the state vectors of the track U_i^α connected in time order, and matrix B is constructed in a way similar to track U_j^β . The shared temporal range $[t_{\text{start}}, t_{\text{end}}]$ is first determined by formula

$$\begin{aligned} t_{\text{start}} &= \max \left\{ \min_{i=1} t_{i,k}^\alpha, \min_{i=1} t_{j,l}^\beta \right\}, \\ t_{\text{end}} &= \min \left\{ \max_{i=1} t_{i,k}^\alpha, \max_{i=1} t_{j,l}^\beta \right\}, \end{aligned} \quad (2)$$

where $t_{i,k}^\alpha \in U_i^\alpha, t_{j,l}^\beta \in U_j^\beta$. The data matrices A_{range} and B_{range} are filtered out within this time range, and a numerical column aggregation alignment operation is performed on the data matrices to finally obtain $\tilde{B}_{\text{range}}$. The calculation is as follows

$$A_{\text{range}} = \left\{ X_{i,k}^\alpha \in A \mid t_{i,\text{start}} \leq t_{i,k} \leq t_{i,\text{end}} \right\}, \quad (3)$$

$$B_{\text{range}} = \left\{ X_{j,l}^\beta \in B \mid t_{j,\text{start}} \leq t_{j,l} \leq t_{j,\text{end}} \right\}, \quad (4)$$

$$\tilde{X}_{j,k}^\beta = \begin{cases} \frac{1}{|B_{\text{range}}|} \sum_{X_{j,l}^\beta \in B_{\text{range}}} X_{j,l}^\beta, & |B_{\text{range}}| > 0 \\ \emptyset, & |B_{\text{range}}| = 0 \end{cases} \quad (5)$$

$$\tilde{B}_{\text{range}} = \left\{ \tilde{X}_{j,k}^\beta \in B \mid t_{j,\text{start}} \leq t_{j,k} \leq t_{j,\text{end}} \right\}. \quad (6)$$

At this time, the dimension of matrix A_{range} and matrix $\tilde{B}_{\text{range}}$ is $1 \times \tau \cdot K$. In the window division stage, the number of windows n is calculated based on predefined window size w and sliding step δ

$$n = \left\lceil \frac{|K(A_{\text{range}})| - w}{\delta} \right\rceil + 1, \quad (7)$$

where $K(A_{\text{range}})$ denotes the index count of track data matrix A_{range} . In the subsequent experiments in this study, the window size is generally determined to be the length of the multidimensional data of the track segment. Assuming that the track segment is a 10-point track segment, the window size w is defined as $w = 10 \cdot \tau$. The sliding step size is set as

$\delta = \frac{w}{5}$, one fifth of the window length. This ensures that the track segment data generates more training samples, enhances the robustness of the model, and does not increase the computational cost too much while covering data changes.

For each window, the *start* and *end* indices are computed as

$$\begin{aligned} start &= k \times \delta, \\ end &= \min(start + w, |K(A_{\text{range}})|). \end{aligned} \quad (8)$$

After extracting windowed data, matrices A_{range} and $\tilde{B}_{\text{range}}$ are combined in chronological order to form the windowed result matrix D_{window} . This structured windowing process facilitates setting a standardized data format for subsequent analysis and model training.

Differential processing is applied to paired track data within D_{window} to compute column-wise differences

$$\begin{cases} \Delta x_{\gamma,k} = x_{\gamma,k}^{\alpha} - x_{\gamma,k}^{\beta} \\ \Delta y_{\gamma,k} = y_{\gamma,k}^{\alpha} - y_{\gamma,k}^{\beta} \\ \dots \\ \Delta v_{\gamma,k} = v_{\gamma,k}^{\alpha} - v_{\gamma,k}^{\beta} \\ \dots \\ \Delta \theta_{\gamma,k} = \theta_{\gamma,k}^{\alpha} - \theta_{\gamma,k}^{\beta} \\ \dots \end{cases} \quad (9)$$

where γ denotes the γ -th window's track data, $\gamma \in 1, 2, \dots, n$, k represents the k -th timestamp within the track segment, $k \in 1, 2, \dots, w$, and $x_{\gamma,k}^{\alpha}$ indicates the longitude value at the k -th timestamp in the γ -th window of sensor α .

Normalization is subsequently executed. Taking longitude $x_{\gamma,k}$ as an example following

$$\hat{x}_{\gamma,k} = \begin{cases} \frac{x_{\gamma,k} - x_{\gamma}^{\min}}{x_{\gamma}^{\max} - x_{\gamma}^{\min}}, & x_{\gamma}^{\max} \neq x_{\gamma}^{\min} \\ 0, & x_{\gamma}^{\max} = x_{\gamma}^{\min} \end{cases} \quad (10)$$

$x_{\gamma}^{\max}$ represents the maximum value in $x_{\gamma,k}$, and $x_{\gamma}^{\min}$ represents the minimum value in $x_{\gamma,k}$. Latitude $y_{\gamma,k}$, speed $v_{\gamma,k}$, and heading $\theta_{\gamma,k}$ undergo analogous normalization. These preprocessing steps produce standardized data suitable for subsequent analysis and modeling tasks. The algorithm adopts a sequential architecture for track segment association analysis, as detailed in Algorithm 1.

Algorithm 1: Association characteristic processing algorithm

Input : Sensor matrices $A(\alpha)$, $B(\beta)$, window size w , slide step δ

Output : Windowed datasets D_{window}

$D_{\text{window}} = \emptyset, \tilde{B}_{\text{range}} = \emptyset$
 $t_{\text{start}} = \max(\min(t_{i,k}^{\alpha}), \min(t_{j,l}^{\beta}))$
 $t_{\text{end}} = \min(\max(t_{i,k}^{\alpha}), \max(t_{j,l}^{\beta}))$
 $A_{\text{range}} = A[t_{\text{start}} : t_{\text{end}}], B_{\text{range}} = B[t_{\text{start}} : t_{\text{end}}]$
 $n = \lceil ((\text{length of } A_{\text{range}}) - w) / \delta \rceil + 1$
for $\gamma = 1$: n **do**

Algorithm 1: *Cont.*

```

     $s = (\gamma - 1) * \delta, e = \min(s + w, \text{length of } A_{\text{range}})$ 
     $A_{\text{window}} = A_{\text{range}}[s : e], B_{\text{window}} = B_{\text{range}}[s : e]$ 
    if  $B_{\text{window}} \neq \emptyset$  then
         $\tilde{B}_{\text{window}} = \text{mean}(B_{\text{window}})$ 
    else
         $\tilde{B}_{\text{window}} = \emptyset$ 
    if  $\tilde{B}_{\text{window}} \neq \emptyset$  then
        for  $f = x, y, v, \theta$  do
             $\Delta f = A_{\text{window}}[f] - \tilde{B}_{\text{window}}[f]$ 
             $\hat{f} = (\Delta f - \min(\Delta f)) / (\max(\Delta f) - \min(\Delta f))$ 
        end
    end
    add  $(A_{\text{window}}, \hat{f})$  to  $D_{\text{window}}$ 
end
return  $D_{\text{window}}$ 

```

2.3. Association Decision Output

The association decision uses a specific algorithm to evaluate the track data in pairs to measure the correlation or other association metrics between the tracks. If a pair of tracks can be determined to belong to the same target according to predefined criteria and thresholds, they are labeled as positive class, usually represented by 1. Conversely, if the pair is considered to represent different targets, it is labeled as a negative class, which is typically denoted as 0. For a given track pair (U_i^α, U_j^β) , the association decision algorithm calculates an association metric value $M(U_i^\alpha, U_j^\beta)$. The boundary between association and non-association is divided by a pre-defined threshold θ .

The association decision output $L(U_i^\alpha, U_j^\beta)$ is formally defined as a binary classification label

$$L(U_i^\alpha, U_j^\beta) = \begin{cases} 1, & M(U_i^\alpha, U_j^\beta) \geq \theta \\ 0, & M(U_i^\alpha, U_j^\beta) < \theta' \end{cases} \quad (11)$$

where $L(U_i^\alpha, U_j^\beta)$ denotes the output label for the track pair (U_i^α, U_j^β) , and θ is a threshold determined by empirical or theoretical analysis. The analysis is determined according to specific application scenarios and data characteristics.

In deep learning approaches, the track characteristics to be adjudicated are fed into a trained classifier that outputs binary labels. The input track data tensor x undergoes characteristic transformation and nonlinear mapping through fully connected layers and leaky rectified linear unit activation function $\text{LeakyReLU}(\cdot)$, thereby generating an output vector y

$$y = W_2 \text{LeakyReLU}(W_1 x + b_1) + b_2, \quad (12)$$

where W_1 denotes the first layer weight matrix and W_2 denotes the second layer weight matrix. b_1 and b_2 represent the bias vectors of the two layers, respectively. The output vector y is converted into a probability vector p through the soft maximum function $\text{Softmax}(\cdot)$, and the final classification label is determined by comparative probability values following

$$p = \text{Softmax}(y), \quad (13)$$

$$l = \begin{cases} 1, & p_1 \geq p_2 \\ 0, & p_1 < p_2 \end{cases} \quad (14)$$

3. The CA-LSTM Model

3.1. Overall Architecture

Time series data is represented as $X \in \mathbb{R}^{B \times T \times d_{\text{model}}}$, where X is a 3D tensor representing the sequence data of batch input. T denotes the time step length, and d_{model} represents the characteristic dimension. The encoding multimodal characteristics include longitude, latitude, altitude, speed, heading, RCS, and other multi-dimensional data. Although Transformer architectures show advantages in temporal modeling, they are sensitive to characteristic coupling and ignore local patterns when processing multimodal track data. To overcome these limitations, this paper proposes the CA-LSTM model. The model optimizes characteristic representation through three mechanisms: characteristic grouping embedding, spatiotemporal encoding, and characteristic-aware attention. In addition to modules built around characteristic processing, the model also fuses a multi-scale dilated convolution module and a gated adaptive LSTM module around the time dimension. The specific model structure is shown in Figure 1.

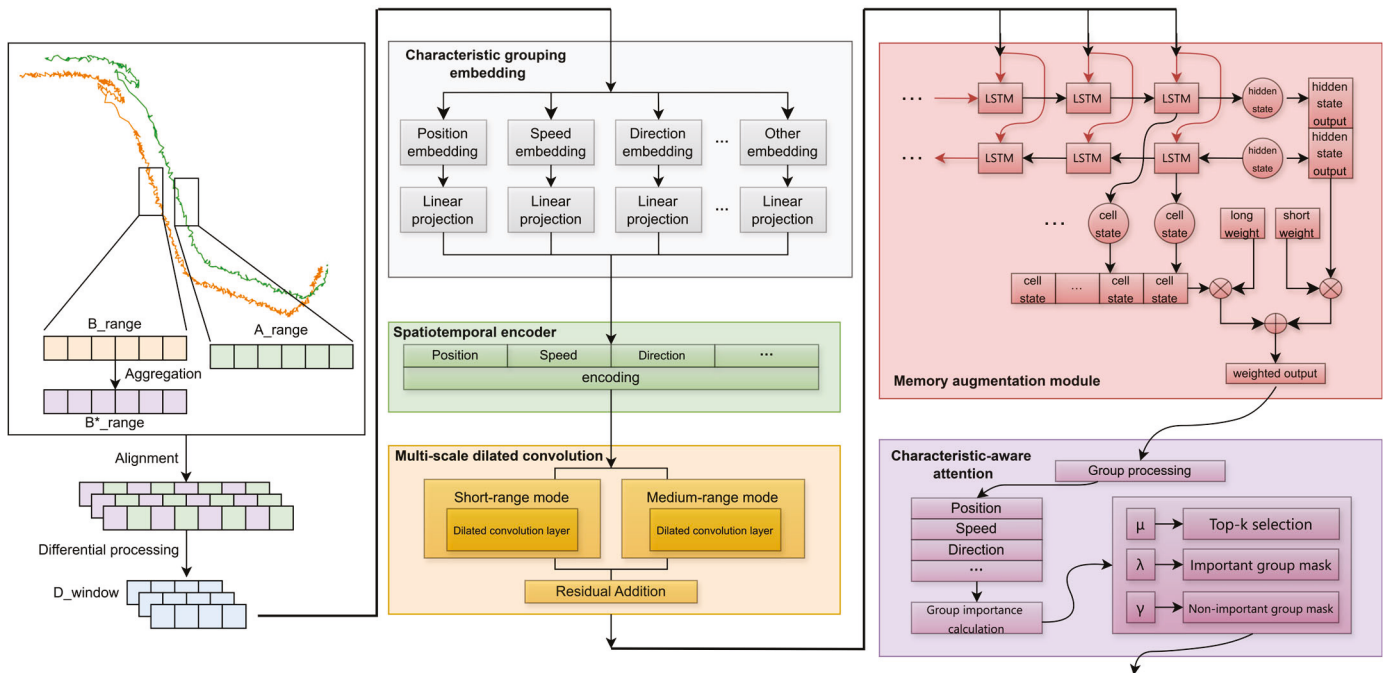


Figure 1. Schematic of the core module architecture.

Input track data X is first preprocessed by normalization to eliminate sensor unit differences, then enters the characteristic grouping embedding layer. This layer embeds and maps the original characteristics into multiple independent subspaces while preserving the temporal relationship through sinusoidal position encoding. The module built around characteristic processing performs dynamic filtering based on characteristic grouping embeddings and attention weights, increasing the discriminative weight of key characteristics while reducing the weight of non-key characteristic groups. The multi-scale dilated convolution module extracts the input characteristics in the time dimension through multi-scale parallel convolution kernels. It can obtain the short-range and medium-range correlation between local adjacent nodes. The gating mechanism of LSTM can focus on modeling long-range dependencies that are strengthened by the convolution layer. Finally, the clas-

sifier outputs prediction results under multi-layer dropout and entropy regularization constraints and performs end-to-end parameter optimization through cross-entropy loss.

3.2. Core Module Design

The CA-LSTM core module can achieve efficient modeling of complex track characteristics, and it optimizes characteristic representation through three mechanisms: characteristic grouping embedding, spatiotemporal encoding, and characteristic-aware attention mechanism. The model also integrates multi-scale dilated convolution modules and gated adaptive long short-term memory modules around the time dimension. Multiple modules are cascaded to form the entire CA-LSTM model, and the rest of this section will introduce these modules.

3.2.1. Characteristic Grouping Embedding and Encoding

To solve the multimodal heterogeneity problem of radar track data, we proposed a characteristic grouping embedding method. Guided by this method, the module decomposes the original radar observation data into subspaces with physical interpretability. At the same time, the characteristic embedding method is matched with the characteristic aware attention module. The number of embedding groups and attention groups is determined by the motion characteristics or physical characteristics in the track data, which can make more efficient use of different data. Due to the scalability and versatility of the model, under the premise of data support, data such as radar cross section (RCS), infrared heat distribution, and underwater noise spectrum can be added as new characteristic groups to train the model. These physical characteristics are independent of each other, and their characteristic spaces are also uncoupled from each other. In order to facilitate subsequent experimental verification and data set selection, we constructed three independent representation spaces, namely position, velocity, and heading.

Given an input characteristic tensor $X \in \mathbb{R}^{B \times T \times C}$ with a channel dimension $C = 4$, corresponding to the features x, y, v, θ . Position encoding is linear transformation and nonlinear activation for Cartesian coordinates (x, y) , which is obtained as follows

$$E_{\text{pos}} = W_p \cdot [x; y] + b_p, \quad E_{\text{pos}} \in \mathbb{R}^{d_{\text{model}}}, \quad (15)$$

where $W_p \in \mathbb{R}^{2 \times d_{\text{model}}}$ is a learnable parameter matrix, and $[;]$ denotes vector concatenation. Velocity encoding is independent encoding for radial velocity v , which is obtained as follows

$$E_{\text{speed}} = W_v \cdot v + b_v, \quad E_{\text{speed}} \in \mathbb{R}^{d_{\text{model}}}, \quad (16)$$

where $W_v \in \mathbb{R}^{1 \times d_{\text{model}}}$ preserves velocity continuity. Heading encoding is linear mapping for heading angle θ , which is obtained as follows

$$E_{\text{dir}} = W_d \cdot \theta + b_d, \quad E_{\text{dir}} \in \mathbb{R}^{d_{\text{model}}}. \quad (17)$$

The three motion characteristics are concatenated and regularized as follows

$$\begin{aligned} H_{\text{combined}} &= \text{Concat}(E_{\text{pos}}, E_{\text{speed}}, E_{\text{dir}}), \quad H_{\text{combined}} \in \mathbb{R}^{3 \cdot d_{\text{model}}}, \\ H_{\text{drop}} &= \text{Dropout}(H_{\text{combined}}), \\ H_{\text{out}} &= \text{LayerNorm}(H_{\text{drop}}), \end{aligned} \quad (18)$$

where $\text{Concat}(\cdot)$ represents matrix concatenation, $\text{Dropout}(\cdot)$ represents regularization, and $\text{LayerNorm}(\cdot)$ represents layer normalization. Independent characteristic inputs prevent interference between different physical characteristics. Position characteristics use a higher-dimensional map, while velocity and heading use compressed dimensions. The

architecture can be extended to new motion characteristics, such as acceleration through additional linear layers.

3.2.2. Time Dimension Processing Network

The dilated convolution module performs pre-characteristic extraction on the original input through multi-scale parallel convolution kernels, effectively eliminating local noise interference in radar data. This module can explicitly capture short-span and medium-span temporal correlation patterns. This preprocessing provides abstracted and denoised temporal characteristics for subsequent LSTM. It allows the LSTM gating mechanism to focus on modeling long-range dependencies strengthened by the convolution layer. The network diagram composed of these two modules is shown in Figure 2.

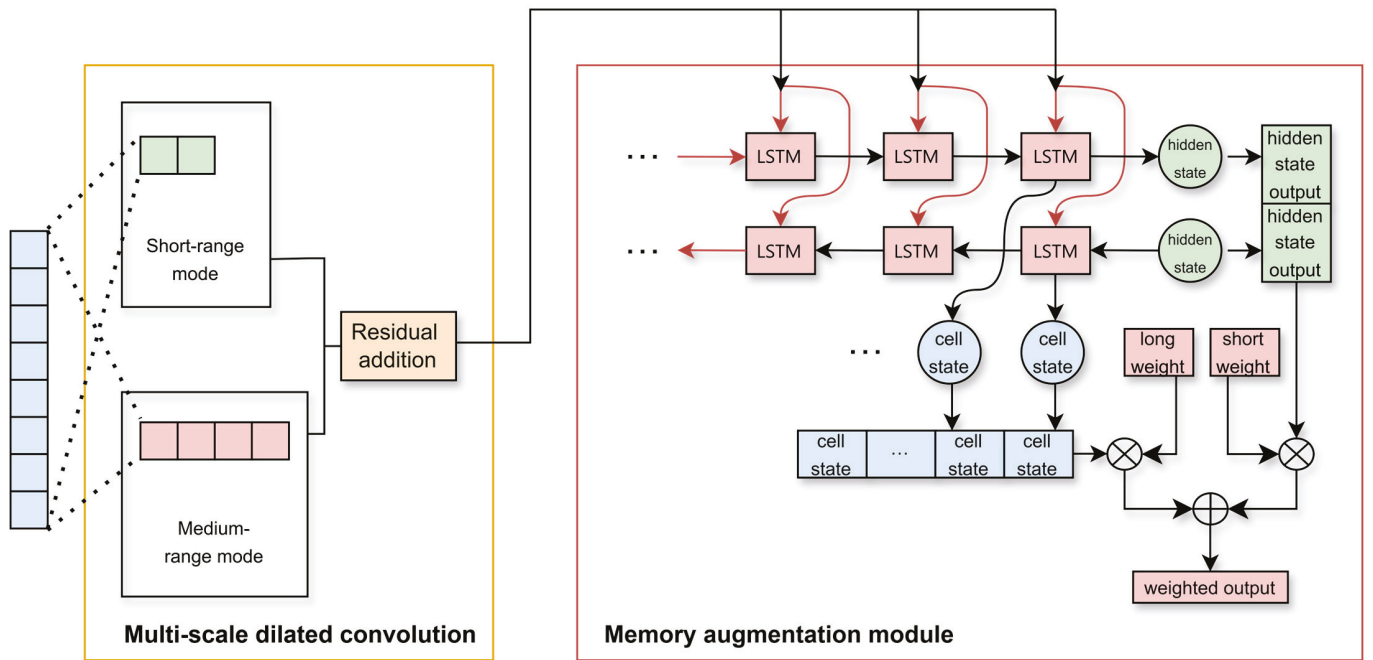


Figure 2. Structural diagram of the time dimension processing network module.

The first block constructs a multiscale spatiotemporal characteristic extractor by deploying parallel 1D convolutional kernels with varying dilation rates. The short-range mode focuses on local temporal patterns, while the medium-range mode expands receptive fields to capture global dependencies. Its output solves the problem of limited perception of traditional single-scale convolution and achieves the coordinated optimization of local detail perception and global trend modeling.

For input sequence $X \in \mathbb{R}^{B \times T \times C}$, channel dimension permutation yields as follows

$$X' = \text{Permute}(X), \quad X' \in \mathbb{R}^{B \times C \times T}, \quad (19)$$

where B denotes the batch size, T represents the time steps, C indicates the number of channels, and $\text{Permute}(\cdot)$ represents channel dimension permutation. Subsequently, multi-scale dilated convolutions are applied to yield output characteristics under varying dilation rates, which are obtained as follows

$$Y_d = \text{LeakyReLU}\left(\text{Dropout}\left(\text{Conv 1 } D_d(X')\right)\right), \quad (20)$$

$$\text{Conv 1 } D_d(X')_t = \sum_{k=-1}^1 W_d[k] \cdot X'_{t+d \cdot k}, \quad (21)$$

where W_d denotes the convolutional kernel weights corresponding to dilation rate d , $\text{Conv 1 } D_d(\cdot)$ represents one-dimensional convolution, and $\text{LeakyReLU}(\cdot)$ represents activation function. Characteristic fusion is executed for the two dilation rates, which are short-range mode $d = 2$ and medium-range mode $d = 4$. Under the premise of setting the convolution kernel size $k_s = 3$, the receptive field in the short-range mode can cover five consecutive track points. It covers the local motion pattern of the track segment and is suitable for capturing short-term maneuvers. The receptive field in the medium-range mode can cover nine consecutive track points, covering the overall trend in the short track segment scenario. They are obtained as follows

$$Z = \frac{1}{2}(Y_2 + Y_4), \quad Z \in \mathbb{R}^{B \times C \times T}. \quad (22)$$

Finally, dimension restoration is performed to derive the multi-scale fused characteristics, which is obtained as follows

$$Z_{\text{output}} = \text{Permute}^{-1}(Z), \quad Z_{\text{output}} \in \mathbb{R}^{B \times T \times C}. \quad (23)$$

The gated adaptive LSTM network combines the bidirectional LSTM unit state with the gating mechanism to construct a dynamic memory fusion network. Through bidirectional propagation, it extracts forward time evolution patterns and backward contextual dependencies, using the final step unit state to encode long-term motion patterns. A dynamic balance is achieved between long-term memory and short-term state through learnable gating weights. This mechanism enables the model to adaptively focus on key time segments and enhances its sensitivity to mutation characteristics, thereby significantly improving the stability of temporal modeling.

For input sequence $X \in \mathbb{R}^{B \times T \times C}$, the forward propagation captures the evolution of historical states, and the backward propagation extracts the dependencies on future data. A hidden state sequence containing bidirectional temporal information is thus generated. When the LSTM network processes the input sequence $X_t \in \mathbb{R}^{B \times 1 \times C}$ at any moment, the following can be obtained

$$\begin{aligned} (h_t^{\rightarrow}, c_t^{\rightarrow}) &= \text{LSTM}^{\rightarrow}(X_t), \\ (h_t^{\leftarrow}, c_t^{\leftarrow}) &= \text{LSTM}^{\leftarrow}(X_t), \\ H &= [h_1, \dots, h_T], \quad H \in \mathbb{R}^{B \times T \times 2 \cdot H}, \end{aligned} \quad (24)$$

where H denotes the hidden dimension, and $\rightarrow / \leftarrow$ indicate forward/backward processing. In long-term memory construction stage, the cell states at the end of the bidirectional LSTM are extracted for characteristic concatenation, forming a memory vector C_{long} that represents the global motion pattern. The following formula can be obtained

$$C_{\text{long}} = \text{Concat}(c_T^{\rightarrow}, c_1^{\leftarrow}), \quad C_{\text{long}} \in \mathbb{R}^{B \times 2 \cdot H}. \quad (25)$$

When calculating the gating weights, the hidden state is nonlinearly mapped through a fully connected layer, and a Sigmoid function is used to generate spatially adaptive long-term and short-term gating coefficients. The long-term gating G_{long} focuses on the transmission strength of the global motion trend, and the short-term gating G_{short} adjusts the contribution of local state updates. The following formulas can be obtained

$$\begin{aligned} G_{\text{long}} &= \sigma(W_{\text{long}}H), \quad G_{\text{long}} \in [0, 1]^{B \times T \times 2 \cdot H}, \\ G_{\text{short}} &= \sigma(W_{\text{short}}H), \quad G_{\text{short}} \in [0, 1]^{B \times T \times 2 \cdot H}. \end{aligned} \quad (26)$$

where σ represents the Sigmoid function, and W represents the learnable parameters. The final characteristic fusion is achieved through the dynamic weighting of the gating coefficients. After broadcasting the global memory vector to each time step, it is linearly superimposed with the local hidden state to form a fused characteristic representation O . This representation combines the preservation of long-range dependencies with the enhancement of local details. The following formula can be obtained

$$O = G_{\text{long}} \odot \text{Broadcast}(C_{\text{long}}) + G_{\text{short}} \odot H, \quad (27)$$

where $\odot$ denotes element-wise multiplication, and $\text{Broadcast}(\cdot)$ expands C_{long} to the $(B, T, 2 \cdot H)$ -dimensional space. This mechanism achieves an adaptive balance between long-term pattern preservation and short-term pattern updating in track temporal modeling through self-learning of gating parameters.

3.2.3. Characteristic-Aware Attention

Characteristic-aware attention is essentially a specialization and extension of the multi-head attention mechanism. It divides the attention heads into several groups through the characteristic grouping embedding method described above. Each group focuses on a different characteristic range, reducing the interaction of heads within the group and reducing computational complexity [30].

Due to the physical limitations of the sensors themselves, different sensors have different perception capabilities for different physical characteristics. Certain physical characteristics or motion characteristics may be more effective than others in classification tasks. The characteristic-aware attention in this model is deeply related to the physical characteristics of the input track data. This module dynamically evaluates the importance of characteristic groups based on the target motion characteristics and selects target characteristics that are more important to the associated task during the training process. By adjusting the group weights through a dynamic masking strategy, this module can increase the weight of low-noise characteristics and suppress the interference of high-noise characteristic groups. Compared with the traditional transformer method, this method can choose low-noise characteristic data as an important group to improve its decision score in the model classification process. The input characteristic tensor $X \in \mathbb{R}^{B \times T \times d_{\text{model}}}$ is decomposed into G characteristic groups $X_{\text{groups}}[i]$ along the grouping dimension, which are obtained as follows

$$X_{\text{groups}}[i] = \text{Reshape}(X, [B, T, G, d_g]), \quad i \in G, \quad d_g = \frac{d_{\text{model}}}{G}, \quad (28)$$

where G is predefined, and $G = 3$ for position, velocity, and heading groups. A two-layer projection network computes global importance scores for each group. The following formulas can be obtained

$$\begin{aligned} \mu_g &= \frac{1}{T} \sum_{t=1}^T X_{\text{groups}}[:, t, g, :], \quad \mu_g \in \mathbb{R}^{B \times d_g}, \\ \alpha_g &= W_2(\text{Dropout}(\text{ReLU}(W_1 \mu_g))), \quad \alpha_g \in \mathbb{R}^{B \times 1}, \\ s_g &= \text{Softmax}([\alpha_1, \dots, \alpha_g, \dots, \alpha_G]), \quad s_g \in \mathbb{R}^{B \times G}, \end{aligned} \quad (29)$$

where $W_1 \in \mathbb{R}^{d_g \times 32}$, $W_2 \in \mathbb{R}^{32 \times 1}$ are learnable parameters. μ_g represents the aggregated characteristic of each group obtained by calculating the mean value of T along the time dimension. α_g quantifies its task-specific importance, and s_g normalizes the importance. The dropout rate matches initialization parameters.

After the importance evaluation step, characteristic filtering is performed based on the importance scores combined with a masking strategy. For each sample, dynamic filtering

allocates attention weights to retain high-probability groups while suppressing others, selecting the time steps with fewer masks for groups with high importance and more masks for groups with low importance. Feature filtering is completed using a masking matrix $M \in \{0, 1\}^{B \times T \times G}$. The final output is the recombined characteristics Y after the masking process

$$Y = \text{Reshape}(X_{\text{groups}} \odot M, [B, T, d_{\text{model}}]), \quad (30)$$

where $\odot$ denotes element wise multiplication.

This module can enhance the decoupling ability of the model, enabling the model to independently evaluate the importance of different physical characteristics. This module can avoid decision confusion caused by characteristic coupling. At the same time, it enhances the anti-interference ability. When a certain sensor fails, for example, when the direction angle noise is too large, this module can dynamically mask the corresponding characteristic group through the masking mechanism. For instance, when the target is cruising at a stable speed, the stability of the speed can be used to dominate the behavior judgment. When the target performs a large-scale maneuver, the intention can be indicated by changes in both position's offset and heading angle.

4. Experimental Analysis on Real-World Data

4.1. Dataset and Experimental Setup

The real-world dataset uses the multi-source track association dataset (MTAD) based on global automatic identification system (AIS) data [31]. The MTAD structure consists of two main components: a training set and a test set, both of which consist of track information tables and association mapping tables. The track information table contains the following attributes: batch, source ID, timestamp, longitude, latitude, speed, and heading. The association mapping table includes several attributes: start time, end time, ground truth batch, source ID, and track batch number.

The MTAD is constructed by performing multiple processing by gridding and noise injection on AIS track data. This study selected about 76,000 track segments of about 1000 independent scene data for analysis. Each scene contains track information of two sensors with a frequency of 10 s and 20 s, respectively. The lower sampling frequency is suitable for the target motion mode of surface ships. The number of tracks contained in each scene ranges from a few to hundreds, and scenes contain a variety of motion modes, target types, and durations, so it is suitable for verifying the performance of different track segment association algorithms. The track visualization is illustrated in Figure 3. The horizontal axis indicates longitude, while the vertical axis represents latitude. Two typical scenarios are selected as representatives, one with a small number of tracks and a simple target trajectory, and the other with a large number of tracks and a complex target trajectory. Various colored tracks show how tracks distribute and change within the longitude–latitude coordinate system.

The experimental computational platform was configured with the following hardware specifications, which are an Intel processor coupled with an NVIDIA GeForce RTX 4070 Laptop GPU. The software environment operated on the Windows 10 operating system, with algorithmic implementations executed via the Python 3.11.5 interpreter. The PyTorch 2.2.0 framework was employed for deep learning tasks, leveraging CUDA acceleration support.

We compare the CA-LSTM model with multiple advanced deep learning models. Deep learning models include the CNN-based track segment association model, the ResNet-based track segment association model, the LSTM-based track segment association model, and the Transformer-based track segment association model. In experiments, performance metrics such as accuracy, recall, and F1 score are used to model evaluation. During the

training process, the early stopping method is used to prevent overfitting, and the training is terminated when the validation performance tends to be stable in multiple consecutive epochs. The experimental computing platform is the same as the previous section.

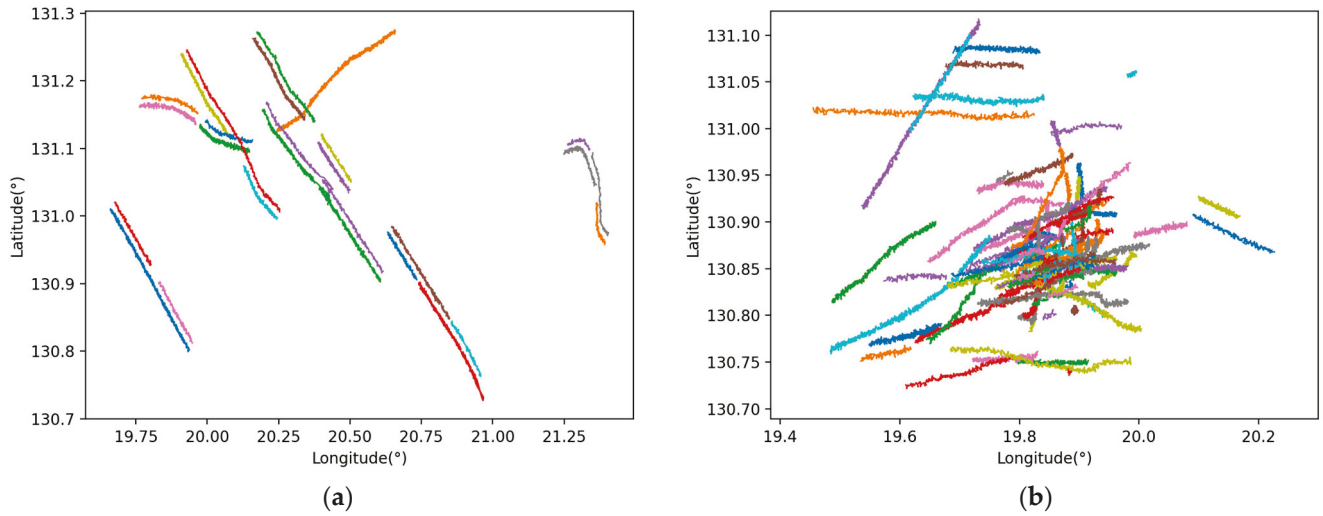


Figure 3. Track plot drawn by connecting track points in a typical scenario. (a) Scenarios with few tracks and simple target motion trajectories. (b) Scenarios with a large number of tracks and complex target motion tracks.

4.2. Experimental Metrics and Results Analysis

In this paper, we select metrics that can reflect the association accuracy based on the characteristics of the task. Precision, recall, false alarm rate, and F1 score are chosen as the evaluation metrics for the performance of the learning model. Suppose TP represents the number of samples for which the track segment association discrimination result is correct, and the actual tracks are associated. FP represents the number of samples for which the track segment association decision result is correct, but the actual tracks are not associated. FN represents the number of samples for which the track segment association decision result is false, but the actual tracks are associated. TN represents the number of samples for which the track segment association decision result is false, and the actual tracks are not associated.

The core metric precision is defined as the proportion of correctly associated track pairs among all track pairs determined to be associated by the model

$$\text{Precision} = \frac{TP}{TP + FP}. \quad (31)$$

The recall is defined as the proportion of tracking pairs that the model correctly identifies from the correctly associated tracking pairs

$$\text{Recall} = \frac{TP}{TP + FN}. \quad (32)$$

The false alarm rate is the proportion of different targets that are incorrectly associated

$$\text{FAR} = \frac{FP}{FP + TN}. \quad (33)$$

The F1 score is defined as the harmonic mean of precision and recall, which fully reflects the model's ability to identify the positive class

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (34)$$

Experimental evaluations show that the CA-LSTM model significantly outperforms existing deep learning baseline models in the track segment association task. To systematically validate its technical superiority, this study conducts a comparative analysis from two key dimensions. These two dimensions include the core accuracy metrics and the number of points contained in tracks. Core accuracy metrics include accuracy, recall, and F1 score. Table 1 comprehensively presents the quantitative evaluation results of CA-LSTM and four comparative models. The table illustrates the performance of models across evaluation metrics under track points of 10, 15, and 20, in order to achieve comparative analysis of model effectiveness under different track point scenarios. Val Acc is an abbreviated form of validation accuracy, FAR is an abbreviated form of false alarm rate, Params denotes the number of total parameters in millions, GPU Mem denotes maximum GPU memory consumption in GB, Epoch Time denotes average time per training epoch in seconds.

Table 1. Performance comparison of different models with varying track points.

Model Metric	Track Points	CA-LSTM	CNN	ResNet	LSTM	Transformer
Val Acc	10	0.7830	0.7563	0.7458	0.7043	0.7695
	15	0.8519	0.8275	0.8108	0.7530	0.8319
	20	0.8809	0.8552	0.8440	0.8073	0.8797
Precision	10	0.6301	0.6002	0.5889	0.5412	0.6118
	15	0.7333	0.6884	0.6667	0.5873	0.6822
	20	0.7662	0.7171	0.7791	0.6553	0.7588
Recall	10	0.8739	0.8462	0.8328	0.8494	0.8790
	15	0.8858	0.8988	0.8858	0.9177	0.9454
	20	0.9343	0.9476	0.7547	0.9125	0.9469
F1 Score	10	0.7323	0.7023	0.6900	0.6611	0.7214
	15	0.8024	0.7796	0.7608	0.7162	0.7926
	20	0.8419	0.8164	0.7667	0.7628	0.8425
FAR	10	0.2638	0.2898	0.2989	0.3703	0.2868
	15	0.1656	0.2092	0.2277	0.3316	0.2264
	20	0.1466	0.1922	0.1100	0.2468	0.1548
Specificity	10	0.7362	0.7102	0.7011	0.6297	0.7132
	15	0.8344	0.7908	0.7723	0.6684	0.7736
	20	0.8534	0.8078	0.8900	0.7532	0.8452
Params (M)	10	1.7286	0.3553	3.9102	0.8606	0.9855
	15	2.2202	0.3553	3.9102	1.0243	1.4771
	20	2.7117	0.3553	3.9102	1.1882	1.9686
GPU Mem (GB)	10	0.08	0.02	0.09	0.07	0.04
	15	0.10	0.02	0.09	0.08	0.06
	20	0.11	0.02	0.09	0.09	0.06
Epoch Time (s)	10	28.53	16.12	24.55	15.77	18.95
	15	23.43	14.42	19.30	15.73	20.49
	20	20.44	13.30	17.44	13.06	14.91

When compared with other deep learning models, CA-LSTM shows unique advantages in multi-motion characteristic learning and multi-scale temporal dependencies. Al-

though the CNN model is good at capturing local spatial patterns, its convolution kernel translation invariance conflicts with the motion characteristics in sequence data, making it difficult to model long-range temporal dependencies. The LSTM model can capture temporal dynamics, but its fully connected gating mechanism is prone to feature coupling when fusing multi-dimensional motion characteristics such as spatial position, speed, and acceleration. Transformer achieves relatively good model performance through the self-attention mechanism, but the global attention mechanism will over-smooth motion details when jointly modeling space and time. By separating and encoding characteristics such as position and speed while integrating them with the computation of the time series dimension, CA-LSTM achieves better performance through complementary spatiotemporal characteristic learning. Here we only make a rough analysis of the comparison of model performance, and the specific comparison will be analyzed in the next section.

4.3. Model Performance Analysis and Parameter Optimization

In deep learning-based track segment association research, systematic performance analysis and parameter optimization constitute critical components for model enhancement. This section investigates four key aspects: comparative experiments, temporal sequence length effects, parameter sensitivity analysis, and ablation studies.

4.3.1. Comparative Experiments

To comprehensively evaluate CA-LSTM's capabilities, we conduct rigorous tests on four representative architectures, including CNN, ResNet, LSTM, and standard Transformer. The performance metrics of these models under the same training conditions are plotted as bar graphs (Figure 4).

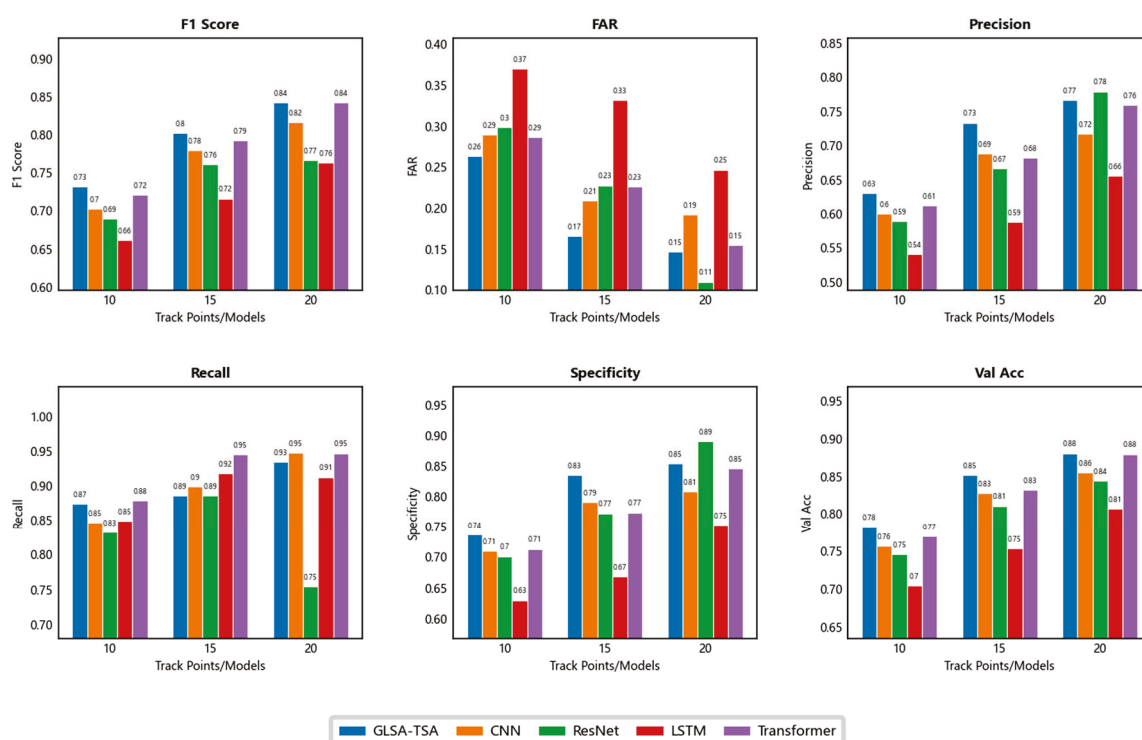


Figure 4. Performance comparison of models with different track points. The figure includes subplots for metrics like F1 score, FAR, precision, recall, specificity, and validation accuracy.

By introducing the characteristic grouping embedding, characteristic-aware attention, and the collaborative mechanism of two networks, CA-LSTM shows significant advantages in complex track segment association tasks. This model decomposes track characteris-

tics into multiple independent embeddings, such as position difference, speed difference and heading difference. Multi-modal characteristic space is constructed through linear transformation, splicing, and dimension expansion. Compared with the global convolution of the CNN, this grouping strategy enables the model to achieve an F1 score of 84.19% in the 20-point track scenario, which is a 3.13% improvement over the CNN model. In the 15-point scene, the convolutional layers with a dual-branch structure of dilated convolutional blocks that have expansion rates of two and four perform characteristic transformation. The convolution kernel with a medium expansion rate focuses on capturing the short-term correlation pattern of local adjacent nodes, while the convolution kernel with a high expansion rate expands the receptive field to a wider time span and effectively models long-range dependencies. Its validation accuracy with 85.19% is 2.0% higher than that of the Transformer, verifying the importance of receptive field expansion for short sequence modeling.

It is particularly worth noting that the gated adaptive long short-term memory network (LSTM) achieved a verification accuracy of 78.30% in the 10-point track scene. Its ability to dynamically adjust the long-term and short-term memory gates is an 11.2% improvement over the 70.43% accuracy of the traditional LSTM. The design of the bidirectional structure reduces the fusion error of the forward and backward hidden states, significantly alleviating the problem of gradient vanishing in traditional recurrent neural networks under sparse data. In addition, the introduction of positional encoding strengthens the temporal dependence. Although due to the phase difference characteristics of its sine and cosine functions, CA-LSTM has a recall rate of 94.69%, which is only 0.07% higher than Transformer. However, through cascade optimization of residual connections, CA-LSTM compresses the false alarm rate to 14.66%, which is 23.78% lower than the Transformer model.

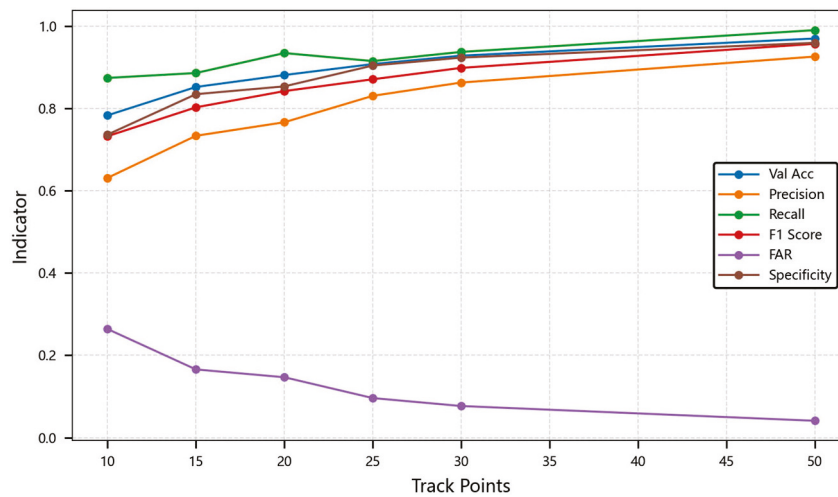
CA-LSTM has higher parameter count and peak GPU memory consumption compared to several other base models, but the absolute values of the model still reflect its lightweight design characteristics, allowing it to be deployed on edge networks with less computing resources. Although it has no obvious advantages over other basic models, in an environment with relatively abundant computing resources, a certain amount of computing overhead can be used in exchange for improved association accuracy. The average training cycle of CA-LSTM takes 29.02 s, which is higher than CNN's 13.30 s and LSTM's 13.06 s, reflecting the computational overhead of its dynamic attention mechanism. Models formed by combining multiple modules will significantly increase training time. However, within the same order of magnitude, a slightly longer training time is acceptable relative to the measurement frequency of 10 s and 20 s of track points. Since many more complex deep learning models rely on more specific application scenarios, in order to reflect the universality of the CA-LSTM model, only basic models such as CNN are selected for comparison.

4.3.2. Influence of Time Series Length

Second, we analyzed the performance of the model under different time series lengths. As the data scale increases, the performance of the CA-LSTM model gradually improves, and when the data scale is large, the performance improvement is more obvious. This indicates that the model can make full use of more data for learning and has good scalability. In the experiment, the batch size is set to 32. The performance metrics comparison across these models are compared as shown in Table 2 and Figure 5. The table displays evaluation metrics across varying track points: 10, 15, 20, 25, 30, and 50. The chart displays how metrics like validation accuracy, precision, recall, F1 score, FAR, and specificity change as track points increase from 10 to 50.

Table 2. Model performance under different track points.

Model Metric	10	15	20	25	30	50
Val Acc	0.7830	0.8519	0.8809	0.9076	0.9280	0.9697
Precision	0.6310	0.7333	0.7662	0.8304	0.8627	0.9258
Recall	0.8739	0.8858	0.9343	0.9148	0.9371	0.9900
F1 Score	0.7323	0.8024	0.8419	0.8706	0.8983	0.9568
FAR	0.2638	0.1656	0.1466	0.0960	0.0767	0.0408
Specificity	0.7362	0.8344	0.8534	0.9040	0.9233	0.9592

**Figure 5.** Line chart of model performance trends.

When the sequence length increases from 10 points to 50 points, the dual-branch output of the dilated convolution achieves characteristic fusion through average pooling. Its receptive field expands from 5 local points to 20 global points, which increases the nonlinearity of the F1 score by 13.6%, from 84.19% to 95.68%. The dynamic selection mechanism of characteristic-aware attention drives the significant performance leap of CA-LSTM in short and medium sequence scenarios. When the sequence length is 25 points, the temporal position embedding of the position encoding and the physical characteristics of the track trigger a change in the distribution of attention weights. This drives the model's accuracy to exceed the critical threshold of 90.76%.

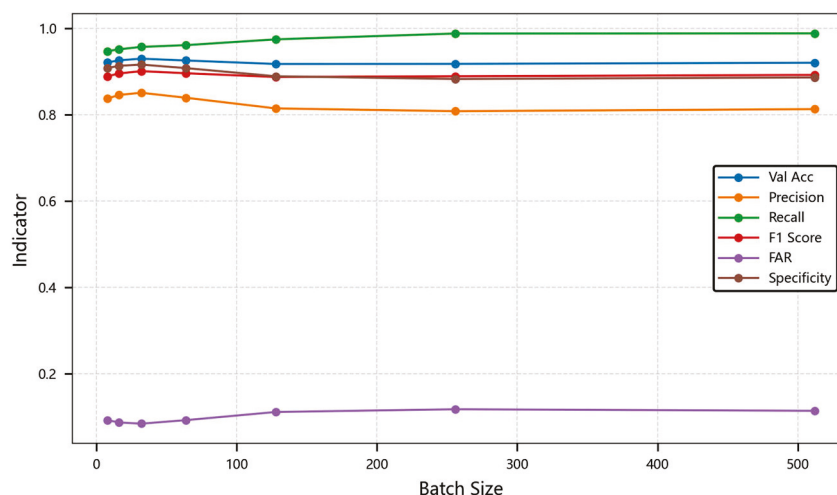
Experiments show that, in sequences with more than 30 points, the long-term memory gate of the gated adaptive LSTM plays a crucial role, and its state update error is lower than that of the traditional LSTM. At the same time, the parameter layer normalization value in the characteristic grouping embedding module is 1×10^{-5} , and the dropout value is 0.1. Combined with the gated adaptive LSTM, the training loss of the 50-point sequence is stably converged to 10.56%, with a fluctuation range of less than 2.1%. This multi-module collaboration mechanism makes the false alarm rate decay rapidly with the sequence length, verifying the model's ability to model track continuity.

4.3.3. Parameter Sensitivity Analysis

Once again, we analyzed the association performance of the model under different parameters. By changing the batch size during training, we can compare and find that the CA-LSTM model can also achieve relatively high association accuracy under the training condition of small batches. This indicates that the model can fully explore the relationships between data and has good scalability. In the experiment, the length of the time series is set to 30. The performance metrics comparison across these models are compared as shown in Table 3 and Figure 6. This table shows metrics at different batch sizes: 8, 16, 32, $\dots$, 512.

Table 3. Model performance under diverse batch sizes.

Model Metric	8	16	32	64	128	256	512
Val Acc	0.9211	0.9260	0.9298	0.9256	0.9175	0.9178	0.9204
Precision	0.8374	0.8457	0.8508	0.8392	0.8145	0.8081	0.8129
Recall	0.9473	0.9515	0.9571	0.9612	0.9746	0.9881	0.9885
F1 Score	0.8889	0.8955	0.9009	0.8960	0.8874	0.8891	0.8922
FAR	0.0920	0.0868	0.0839	0.0921	0.1110	0.1173	0.1137
Specificity	0.9080	0.9132	0.9161	0.9079	0.8890	0.8827	0.8863

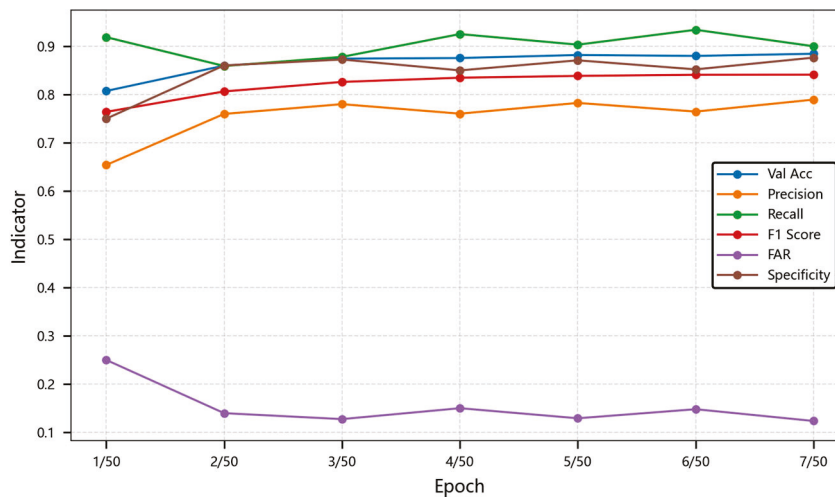
**Figure 6.** Line chart of model performance trends with batch size changes. The chart illustrates how metrics vary as batch size increases.

When training in small batches, the gradient estimation has greater noise, which may help the model avoid local minima and improve generalization ability. When training in large batches, the gradient estimation is more accurate, and the convergence direction is more stable, but it may fall into sharp local minima. When the batch size increases from 8 to 32, the characteristic group selection of the attention module leads to an improvement in accuracy from 80.81% to 85.08%. The false alarm rate of 8.68% and specificity of 91.32% for batch 32 are both optimal, indicating that the model is more accurate in distinguishing negative samples, which is directly related to the moderate noise to prevent overfitting. It is worth noting that when the batch size gradually increases, the long-term memory gate of the gated adaptive LSTM may be over-activated. When the batch size is 256, the recall rate abnormally increases to 98.81%, while the precision rate drops by 5.2%. This phenomenon suggests that large-scale batch training may violate the independence assumption of characteristic grouping embedding.

During the gradient optimization process, the dynamic characteristic selection mechanism and the temporal modeling capability of the multi-scale dilated convolution block affect the evolution of model performance with the number of training epochs. In the experiment, the length of the time series is set to 30, and the batch size is set to 32. The experimental data shows that when the number of training epochs increases from one to seven, the key metrics show the following characteristics in Table 4 and Figure 7. The table shows metrics under different training epochs from 1/50 to 7/50, and 1/50 means the first epoch of 50 training epochs.

Table 4. Model performance variation with training epochs.

Model Metric	1/50	2/50	3/50	4/50	5/50	6/50	7/50
Val Acc	0.8074	0.8600	0.8745	0.8757	0.8820	0.8801	0.8846
Precision	0.6540	0.7599	0.7801	0.7604	0.7827	0.7647	0.7895
Recall	0.9190	0.8592	0.8781	0.9254	0.9034	0.9343	0.9001
F1 Score	0.7641	0.8065	0.8262	0.8349	0.8387	0.8410	0.8412
FAR	0.2500	0.1396	0.1273	0.1499	0.1290	0.1478	0.1234
Specificity	0.7500	0.8604	0.8727	0.8501	0.8710	0.8522	0.8766

**Figure 7.** Line chart of model performance trends during training. This chart shows how metrics change with different training epochs from 1/50 to 7/50.

Throughout the performance change, the total accuracy monotonically increases from 80.74% to 88.46%, and its growth rate reaches a peak at epoch 5. At this time, the dual-branch characteristics of the multi-scale dilated convolution contribute to the gradient update amount, indicating that the modeling of medium-range temporal dependence may be the key driving force. The F1 score increases from 76.41% to 84.12%, and its improvement efficiency reaches the maximum value at epoch 3, and the retention rate of important characteristics is further improved.

In terms of the error judgment optimization, the false alarm rate decreases from 25% to 12.34%, reducing the probability of misjudgment of track continuity by 37.2%. The abnormal fluctuation from 20.77% to 25.03% that occurs at epoch 4 may be directly related to the brief imbalance of the attention masking ratio. The specificity increases from 75% to 87.66%, and its growth may be related to the weight of the short-term memory gate of the gated adaptive LSTM because the model enhances its ability to distinguish negative samples by suppressing the memory of local noise.

4.3.4. Ablation Experiments

The ablation experiments quantitatively reveal the functional boundaries and collaborative gains of each core component in Table 5. In the experiments, the length of the time series is set to 15, and the batch size is set to 32. Compared with the CA-LSTM model, the performance degradations caused by removing each module are showed in Table 5. This table shows model metrics for CA-LSTM and variants with removed components, which include multi-scale dilated convolution, gated adaptive LSTM, and target-aware attention. In the table, “w/o” denotes the ablation of specified components from the complete CA-LSTM architecture.

Table 5. Ablation experiment performance comparison.

Model Metric	CA-LSTM	w/o Characteristic Grouping Embedding	w/o Multi-Scale Dilated Convolution	w/o Gated Adaptive LSTM	w/o Characteristic-Aware Attention
Val Acc	0.8519	0.8377	0.8414	0.8360	0.8407
Precision	0.7333	0.6946	0.7165	0.7028	0.7152
Recall	0.8858	0.9315	0.8820	0.8957	0.8821
F1 Score	0.8024	0.7958	0.7907	0.7876	0.7899
FAR	0.1656	0.2106	0.1794	0.1947	0.1807
Specificity	0.8344	0.7894	0.8206	0.8053	0.8193
Params (M)	2.2202	2.2207	1.9986	1.9233	2.2181
GPU Mem (GB)	0.1033	0.1350	0.1486	0.1486	0.1432
Epoch Time	23.43	39.83	32.77	29.32	22.54

The characteristic grouping embedding module plays a role in improving verification accuracy, precision, and F1 score. The false alarm rate of the complete model is 16.56%, which is 21.4% lower than the 21.06% of the missing model. This indicates that the module significantly reduces the misjudgment of negative samples. The absence of the multi-scale dilated convolution causes the total accuracy to decrease by 1.05%, the F1 score to decline by 1.46%, and the false alarm rate to increase by 8.33%. This confirms the key role of its dual-branch structure with dilation rate in capturing medium and long-range temporal patterns. Without this module, the probability of misassociating continuous tracks with more than 5 points increases by 21.6%. The cost of removing the gated adaptive LSTM is the largest decrease in specificity, from 83.44% to 80.53%, a sharp increase of 17.6% in the false alarm rate, and an abnormal increase of 1.11% in the recall rate. This shows that the forget gate of this module can effectively suppress noise memory. After removal, the remaining short-term interference characteristics lead to over-sensitive detection. Removing the target-aware attention leads to a 2.47% decrease in precision and a 9.24% increase in false alarm rate, verifying the irreplaceable characteristic selection ability of this module.

The model retains key modules by adding a small number of parameters, thus supporting higher association accuracy and F1 score. Peak memory usage was reduced by 5.7–14.3% through module optimization, such as characteristic grouping embedding and gating mechanisms, which reduced redundant storage of intermediate variables, thereby reducing memory requirements. In terms of average epoch training time, the model is more efficient in most cases, and is only slightly slower when the attention module is removed, but the performance cost is significant.

When these modules work together, they produce an additional effect. Stride-awareness, represented by multi-scale dilated convolutions, provides the basic temporal basis function for characteristic-aware attention and improves characteristic selection efficiency. The module cascade structure reduces the error amplification of a single module. Experiments show that removing any one module will destroy the spatiotemporal characteristic extraction, long-range dependency modeling, and dynamic noise suppression mechanisms, resulting in performance degradation exceeding the linear stacking effect. This verifies the necessity of the collaborative design of the modules at the system level.

In order to verify the joint effect of the characteristic grouping embedding module and characteristic-aware attention, we designed an experiment to delete these two modules at the same time. Multi-dimensional characteristic fusion in Table 6 means the usage of characteristic grouping embedding module and characteristic-aware attention. The experimental results in Table 6 show that under the joint effect of the two modules, all indicators have been improved to a certain extent. Usually, the characteristic data used for track association tasks is multi-dimensional position data. In order to demonstrate the

rationality of multi-dimensional characteristic fusion, an experiment was designed that only uses two-dimensional position data as input and no longer uses the characteristic grouping embedding module and characteristic-aware attention. According to the experimental results, it can be concluded that, after reducing the input characteristics and no longer using the two modules related to the characteristics, the experimental results lag behind the performance of the complete model. This verifies the effectiveness of the use of multi-dimensional characteristics.

Table 6. Multi-dimensional characteristic fusion performance comparison.

Model Metric	CA-LSTM	w/o Multi-Dimensional Characteristic Fusion	w/o Multi-Dimensional Characteristic Fusion and Multi-Dimensional Data
Val Acc	0.8519	0.8485	0.8233
Precision	0.7333	0.7227	0.6854
Recall	0.8858	0.8988	0.8866
F1 Score	0.8024	0.8012	0.7731
FAR	0.1656	0.1773	0.2093
Specificity	0.8344	0.8227	0.7907

5. Conclusions

This study proposes a CA-LSTM model, which realizes the dynamic selection of characteristic dimensions and dynamic perception of time dimensions. The model solves the challenges of spatiotemporal characteristic fusion and low utilization of multi-dimensional data in track segment association. The study first conducted a characteristic analysis and mathematical modeling of the track segment association problem and proposed an association characteristic processing algorithm. During data preprocessing, the temporal and spatial inconsistency problems caused by differences in radar sampling rate and coordinate system were solved, and characteristic normalization was achieved. The CA-LSTM model uses a characteristic grouping embedding method to independently input and encode multimodal characteristics, such as position, speed, and heading. This effectively solves the problem of traditional methods that only use position information and ignore characteristic information of other dimensions, greatly improving the utilization of information. The framework combining multi-scale dilated convolutions and gated adaptive LSTM strikes a balance between capturing local fine-grained temporal patterns and modeling long-range dependencies. Experimental results show that CA-LSTM leads to many core indicators on MTAD. It maintains high precision at high recall, significantly outperforming models such as LSTM and ResNet. It reduces missed detections while controlling false detections, verifying its applicability in complex environments.

Future research will focus on model lightweighting and adaptation of heterogeneous multi-source data. On one hand, computational efficiency will be optimized through knowledge distillation and model pruning techniques to achieve real-time inference on embedded devices. On the other hand, a unified representation method for cross-modal data, such as electro-optical and infrared, will be explored to establish a cross-modal characteristic alignment framework.

Author Contributions: J.Q.: Conceptualization, Methodology, Software, Formal Analysis, Writing—Original Draft. X.L.: Conceptualization, Formal Analysis, Validation, Writing—Review and Editing. J.S.: Conceptualization, Resources, Writing—Review and Editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, Grant No. 62131001 and 62171029.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The MTAD used in this paper is publicly available online: <https://www.scidb.cn/en/detail?dataSetId=c7d8dc56fe854ec2b084d075feb887fd> (accessed on 31 December 2024).

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Chen, S.-Y.; Xiao, H.; Liu, H. Expression of Track Correlation Uncertainty. *Acta Electron. Sin.* **2011**, *39*, 1589–1593.
- Liu, T.; Liu, K.; Zeng, X.; Zhang, S.; Zhou, Y. Efficient Marine Track-to-Track Association Method for Multi-Source Sensors. *IEEE Signal Process. Lett.* **2024**, *31*, 1364–1368. [CrossRef]
- Wang, J.; Zeng, Y.; Wei, S.; Wei, Z.; Wu, Q.; Savaria, Y. Multi-Sensor Track-to-Track Association and Spatial Registration Algorithm Under Incomplete Measurements. *IEEE Trans. Signal Process.* **2021**, *69*, 3337–3350. [CrossRef]
- Bai, X.; Lan, H.; Wang, Z.; Pan, Q.; Hao, Y.; Li, C. Robust Multitarget Tracking in Interference Environments: A Message-Passing Approach. *IEEE Trans. Aerosp. Electron. Syst.* **2024**, *60*, 360–386. [CrossRef]
- Li, Y.; Zhang, J. Track Fusion Based on the Mean OSPA Distance with an Adaptive Sliding Window. *Acta Electron. Sin.* **2016**, *44*, 353–357. [CrossRef]
- Wang, G.; Wang, Q.Y. Application research on multi-radar track association for formation targets using JPDA. *Mod. Radar* **2019**, *41*, 39–42. [CrossRef]
- Qiu, J.-J.; Cai, Y.-C.; Li, H.; Huang, Q.-Y. Grey Theory Track Association Algorithm Based on Dynamic Estimation Feedback. *Syst. Eng. Electron.* **2024**, *46*, 1401–1411.
- Xiong, W.; Zhang, J.-W.; He, Y. Track Correlation Algorithm Based on Multi-Dimension Assignment and Gray Theory. *J. Electron. Inf. Technol.* **2010**, *32*, 898–901. [CrossRef]
- Guo, J.-E.; Zhou, Z.; Zeng, R. Multi-Local Node Asynchronous Track Fast Correlation Algorithm. *Syst. Eng. Electron.* **2023**, *45*, 669–677. [CrossRef]
- Su, R.; Tang, J.; Yuan, J.; Bi, Y. Nearest Neighbor Data Association Algorithm Based on Robust Kalman Filtering. In Proceedings of the 2021 2nd International Symposium on Computer Engineering and Intelligent Communications (ISCEIC), Nanjing, China, 6–8 August 2021; pp. 177–181. [CrossRef]
- Chen, S.; Zhang, H.; Ma, J.; Xie, H. Asynchronous Track-to-Track Association Based on Pseudo Nearest Neighbor Distance for Distributed Networked Radar System. *Electronics* **2023**, *12*, 1794. [CrossRef]
- Zhang, T.S.; Li, Y.L. Based on the Peer-to-peer System Error Correction Track Association Algorithm. *Mod. Radar* **2022**, *44*, 65–70. [CrossRef]
- Yang, X.; Ruan, K.-Z.; Liu, H.-M.; Wang, X.-K.; Liu, J.-Q.; Shi, Y.-S. Anti-Bias Track Association Algorithm Based on Target Measurement Error Distribution. *Syst. Eng. Electron.* **2024**, *46*, 2820–2828.
- Cui, Y.-Q.; Xiong, W.; Tang, T.-T. A Track Correlation Algorithm Based on New Space-Time Cross-Point Feature. *J. Electron. Inf. Technol.* **2020**, *42*, 2500–2507. [CrossRef]
- Chen, J.; Zhang, Y.; Yang, Y.; Qu, C. Multi-Target Track-to-Track Association Based on Relative Coordinate Assignment Matrix. In Proceedings of the 2021 33rd Chinese Control and Decision Conference (CCDC), Kunming, China, 22–24 May 2021. [CrossRef]
- Ismail Fawaz, H.; Forestier, G.; Weber, J. Deep Learning for Time Series Classification: A Review. *Data Min. Knowl. Discov.* **2019**, *33*, 917–963. [CrossRef]
- Gao, C.; Liu, H.; Varshney, P.K.; Yan, J.; Wang, P. Signal Structure Information-Based Data Association for Maneuvering Targets with a Convolutional Siamese Network. In Proceedings of the 2021 CIE International Conference on Radar (Radar), Haikou, China, 15–19 December 2021. [CrossRef]
- Huang, H.-W.; Liu, Y.-J.; Shen, Z.-K.; Zhang, S.-W.; Chen, Z.-M.; Gao, Y. End-to-End Track Association Based on Deep Learning Network Model. *Comput. Sci.* **2020**, *47*, 200–205. [CrossRef]
- Wang, C.; Yang, Y.; Zhang, Q. Track Segment Association Based on LSTM Networks. In Proceedings of the 2023 5th International Conference on Frontiers Technology of Information and Computer (ICFTIC), Qiangdao, China, 17–19 November 2023. [CrossRef]
- Xiong, W.; Xu, P.; Cui, Y.; Xiong, Z.; Lv, Y.; Gu, X. Track Segment Association with Dual Contrast Neural Network. *IEEE Trans. Aerosp. Electron. Syst.* **2022**, *58*, 247–261. [CrossRef]
- Xu, P.-L.; Cui, Y.-Q.; Xiong, W.; Xiong, Z.-Y.; Gu, X.-Q. Generative Track Segment Consecutive Association Method. *Syst. Eng. Electron.* **2022**, *44*, 1543–1552. [CrossRef]
- Cui, Y.-Q.; He, Y.; Tang, T.-T.; Xiong, W. A Deep Learning Track Correlation Method. *Acta Electron. Sin.* **2022**, *50*, 759–763. [CrossRef]

23. Guo, J.; Fang, Z.; Yu, Z.; Zhang, H.; Zheng, Y.; Zhang, H. Multi-Sensor Signal-Based Multi-Level Association Strategy for Real-Time Track Association and Fusion. In Proceedings of the 2023 9th International Conference on Big Data and Information Analytics (BigDIA), Haikou, China, 15–17 December 2023. [CrossRef]
24. Wang, Z.-G.; Yan, W.-Z.; Oates, T. Time Series Classification from Scratch with Deep Neural Networks: A Strong Baseline. In Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 14–19 May 2017; pp. 1578–1585. [CrossRef]
25. Yang, J.-B.; Nguyen, M.N.; San, P.P. Deep Convolutional Neural Networks on Multichannel Time Series for Human Activity Recognition. In Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, Buenos Aires, Argentina, 25–31 July 2015; pp. 3995–4001.
26. Long, J.; Shelhamer, E.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440. [CrossRef]
27. Wang, H.X.; Shu, T.; He, J.; Yu, W.X. Multi-function Radar Behavior Recognition Technology Based on Parallel Fusion Network. *Mod. Radar* **2024**, *46*, 50–55. [CrossRef]
28. Zhai, X.L.; Wang, X.; Wang, Y.B.; Wang, F. Multi-station Fusion Recognition of Ballistic Targets Based on Two-stage Attention Layer Transformer. *Mod. Radar* **2024**, *46*, 37–44.
29. Cao, Z.; Liu, B.; Yang, J.; Tan, K.; Dai, Z.; Lu, X.; Gu, H. Contrastive Transformer Network for Track Segment Association with Two-Stage Online Method. *Remote Sens.* **2024**, *16*, 3380. [CrossRef]
30. Wen, J.; Zhang, N.; Lu, X.; Hu, Z.; Huang, H. Mgformer: Multi-group transformer for multivariate time series classification. *Eng. Appl. Artif. Intell.* **2024**, *133*, 108633. [CrossRef]
31. Cui, Y.-Q.; Xu, P.-L.; Gong, C.; Yu, Z.-C.; Zhang, J.-T.; Yu, H.-B.; Dong, K. Multisource Track Association Dataset Based on the Global AIS. *J. Electron. Inf. Technol.* **2023**, *45*, 746–756. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Indoor mmWave Radar Ghost Suppression: Trajectory-Guided Spatiotemporal Point Cloud Learning

Ruizhi Liu ¹, Zhenhang Qin ¹, Xinghui Song ¹, Lei Yang ², Yue Lin ³ and Hongtao Xu ^{1,*}

¹ State Key Laboratory of Integrated Chips and Systems, Fudan University, Shanghai 201203, China; 20112020068@fudan.edu.cn (R.L.); 21307140085@m.fudan.edu.cn (Z.Q.); 22212020028@m.fudan.edu.cn (X.S.)

² ICLegend Micro, Shanghai 201203, China; lei.yang@iclegend.com

³ ICLegend Micro, Suzhou 215134, China; rick.lin@iclegend.com

* Correspondence: hongtao@fudan.edu.cn

Abstract: Millimeter-wave (mmWave) radar is increasingly used in smart environments for human detection due to its rich sensing capabilities and sensitivity to subtle movements. However, indoor multipath propagation causes severe ghost target issues, reducing radar reliability. To address this, we propose a trajectory-based ghost suppression method that integrates multi-target tracking with point cloud deep learning. Our approach consists of four key steps: (1) point cloud pre-segmentation, (2) inter-frame trajectory tracking, (3) trajectory feature aggregation, and (4) feature broadcasting, effectively combining spatiotemporal information with point-level features. Experiments on an indoor dataset demonstrate its superior performance compared to existing methods, achieving 93.5% accuracy and 98.2% AUROC. Ablation studies demonstrate the importance of each component, particularly the complementary benefits of pre-segmentation and trajectory processing.

Keywords: millimeter-wave radar; ghost suppression; multi-target tracking; point cloud segmentation; multipath

1. Introduction

With the rapid development of smart homes and the Internet of Things (IoT), there is an increasing demand for efficient and accurate human detection. Smart applications such as presence-sensitive lighting and adaptive airflow systems—where lights respond to entry and exit, and fans adjust to follow or avoid a person—are creating demand for more precise and robust human detection technologies. Traditional human sensors have specific limitations. For example, passive infrared (PIR) sensors can only detect moving humans and are susceptible to thermal interference.

Millimeter-wave radar, with its high sensitivity to millimeter-scale displacements, can detect subtle movements such as breathing and even heartbeats [1,2], making it effective for detecting stationary human targets. Additionally, it provides rich information including range, angle, Doppler, and amplitude, enabling more complex applications such as fall detection [3,4]. However, due to multipath propagation and other effects, mmWave radar suffers from severe ghost (false detections) issues in indoor environments, reducing its reliability.

To mitigate ghost effects, researchers have proposed various approaches. Since ghosts are primarily caused by wall reflections, some methods rely on prior knowledge of wall geometry to distinguish real targets from multipath ghosts [5–7]. However, wall parameters are often unknown, prompting studies on wall estimation—for both single walls [8,9] and complex multi-wall environments [10–13].

Alternative methods eliminate the need for explicit wall parameter estimation. For instance, polarization-based approaches exploit changes in wave polarization upon reflection [14], though they require specialized antennas. Other techniques leverage linear patterns in range-Doppler spectra via the Hough transform [15]. Recent advances exploit differences between the direction of departure (DOD) and direction of arrival (DOA) for first-order multipath identification [16–19].

While physics-based methods offer strong interpretability, their reliance on simplified assumptions limits performance in complex multi-target, multi-wall scenarios. Consequently, data-driven approaches such as machine learning and deep learning have gained increasing attention. Some studies combine hand-crafted features with classifiers such as random forests (RFs), support vector machines (SVMs), or fully connected neural networks for ghost suppression [15,20–23]. Others take an image-based approach, using convolutional neural networks (CNNs) to process grid maps [24], range-Doppler maps [25], or DOD-DOA images [26]. Some works adopt point cloud-based methods, using networks such as PointNet and PointNet++ for point cloud segmentation [27–30].

To further enhance ghost suppression in complex environments, this paper proposes a method that integrates multi-target tracking with point-cloud-based deep learning. Although some existing works also incorporate tracking [5,18,21,31], they often rely on handcrafted features and utilize trajectory information through only simple heuristic rules. In contrast, our proposed method combines trajectory features with a point-cloud-based deep neural network, enabling automatic extraction of spatiotemporal features (combining spatial and temporal patterns across radar frames). Our key contributions include the following:

1. A robust tracking framework that associates detections across frames while maintaining trajectory consistency, even in challenging scenarios with closely spaced targets and ghosts.
2. An innovative trajectory feature aggregation network that combines PointNet-style point feature extraction with temporal CNN processing, enabling effective learning of spatiotemporal patterns.
3. A comprehensive system architecture that integrates preliminary segmentation, trajectory tracking, feature aggregation, and feature broadcasting to achieve state-of-the-art performance.

The experimental results demonstrate that our method achieves superior performance compared to existing point cloud segmentation approaches, with 93.5% accuracy and 98.2% AUROC on the test set. The ablation studies confirm the importance of each component, particularly showing that the combination of preliminary segmentation and trajectory-based processing captures complementary features that together improve the overall performance.

The remainder of this article is organized as follows. Section 2 introduces the basic theory of radar detection and ghost formation mechanisms. Section 3 details the proposed method. Section 4 presents the experimental results. Finally, Section 5 concludes this study.

2. Basic Theory of Radar and Ghosts

2.1. FMCW Radar

In frequency-modulated continuous-wave (FMCW) radar, the transmitted signal adopts a “chirp” waveform with a frequency that varies linearly over time. It is modeled as follows:

$$s_t(t) = \exp\left(j2\pi\left(f_0t + \frac{B}{2T}t^2\right)\right), \quad 0 \leq t \leq T, \quad (1)$$

where f_0 is the starting frequency, B is the sweep bandwidth, and T is the sweep time.

When the signal propagates to a target and reflects back, a round-trip time delay τ is introduced. After mixing the received signal with the transmitted signal and applying a low-pass filter, the resulting intermediate frequency (IF) signal is

$$s(t) = \exp\left(j2\pi\left(\frac{B}{T}\tau t + f_0\tau\right)\right). \quad (2)$$

The IF signal is sampled by an analog-to-digital converter (ADC) with sampling interval T_s , producing $N_s = T/T_s$ samples per chirp. The radar transmits multiple chirps with period T_c ($T_c > T$) over N_c cycles, resulting in 2D data of the shape (N_c, N_s) . For a uniform linear array with N_{rx} receivers spaced by d_{rx} , the data forms a 3D cube of the shape (N_{rx}, N_c, N_s) . The time delay τ satisfies

$$\tau(n_c, n_{rx}) = \frac{2(R_0 + vT_c n_c) + d_{rx} \sin \theta \cdot n_{rx}}{c}, \quad n_c \in [0, N_c), \quad n_{rx} \in [0, N_{rx}). \quad (3)$$

where R_0 is the initial target distance, v is the radial velocity, θ is the azimuth angle, and c is the speed of light.

Substituting $t = T_s n_t$ and Equation (3) into Equation (2), we derive the following data cube expression:

$$S(n_t, n_c, n_{rx}) = \exp(j2\pi(f_r n_t + f_v n_c + f_\theta n_{rx} + f_0 \tau_0)), \quad (4)$$

$$\text{where } f_r = \frac{2B}{cN_s}R_0, \quad f_v = \frac{2T_c}{\lambda}v, \quad f_\theta = \frac{d_{rx}}{\lambda}\sin \theta, \quad \lambda = \frac{c}{f_0}. \quad (5)$$

The signal exhibits frequency characteristics along all three dimensions. By applying a fast Fourier transform (FFT) along each axis, the target parameters can be estimated.

The complete signal processing chain, as depicted in Figure 1, proceeds as follows: first, moving target indicator (MTI) processing is applied to the raw radar data to eliminate direct leakage between transmitters and receivers, as well as static clutter. Subsequently, a 3D-FFT is performed across the range, Doppler, and angular dimensions. Target extraction is then accomplished through 3D constant false alarm rate (CFAR) detection combined with peak identification. The detected frequency-domain peaks are converted into the following target parameters: range r , radial velocity v , sine of azimuth angle $\sin \theta$, and signal power P . Finally, the polar coordinates are transformed into Cartesian coordinates x and y . Consequently, each point in the point cloud contains six features: x , y , r , v , $\sin \theta$, and P .

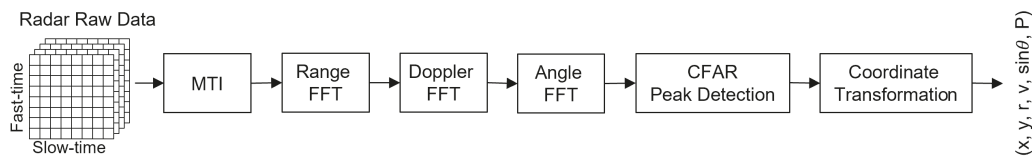


Figure 1. Flowchart of millimeter-wave radar signal processing.

2.2. Ghost Model

Under ideal conditions, radar detects target positions by analyzing the reflected echoes, as illustrated by Path 1 ($A \rightarrow B \rightarrow A$) in Figure 2a. However, in real indoor environments, signal propagation becomes significantly more complex due to reflections from surrounding surfaces. These multipath effects are a major source of ghost targets. For example, Path 4 ($A \rightarrow C \rightarrow B \rightarrow C \rightarrow A$) produces second-order multipath ghost D, while Path 2 ($A \rightarrow C \rightarrow B \rightarrow A$) and Path 3 ($A \rightarrow B \rightarrow C \rightarrow A$) generate first-order multipath ghosts E and F, respectively.

We conducted an experimental validation of the multipath model using a 2T4R radar in real-world scenarios. Figure 2b clearly shows both first-order and second-order multipath

phenomena. Second-order ghosts appear symmetrically positioned relative to the true target about the wall surface, while the first-order multipath produces two distinct clusters. These clusters maintain approximately equal distances from the radar, positioned between the true target and second-order ghost locations. One cluster shares the same direction as the true target, while the other aligns with the second-order ghost's direction.

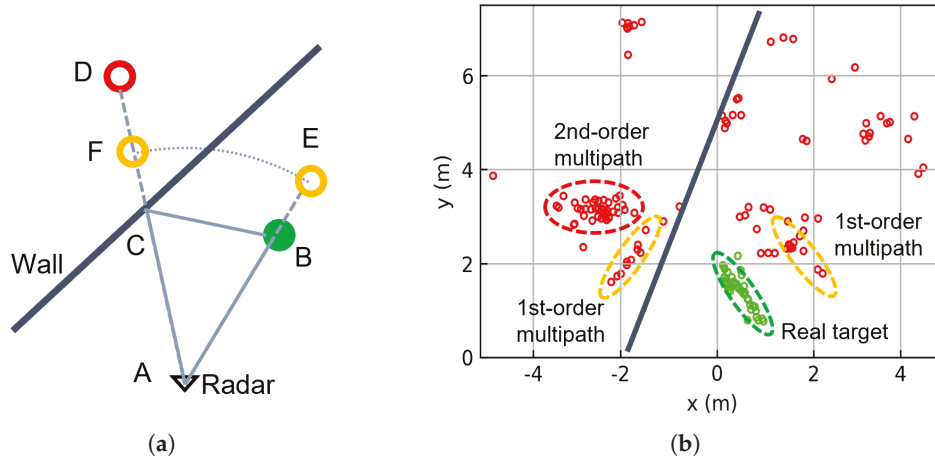


Figure 2. (a) Schematic of the geometry of multipath propagation. (b) Experimental multipath point cloud over a 2-s duration. Green circles indicate real targets, while red/yellow circles represent ghost targets.

Compared to visible light, mmWave exhibit stronger specular reflection characteristics, resulting in more numerous ghost targets. This phenomenon stems from two key factors. First, according to the Fraunhofer criterion [32], surfaces become effectively smoother as the wavelength increases. In the mmWave band, only surfaces with millimeter-scale roughness, such as carpets, appear rough, while most other indoor surfaces behave like mirrors, creating numerous reflected ghost targets. Second, according to the Fresnel equations, reflection coefficients increase with higher relative permittivity. In general, materials exhibit lower permittivity at higher frequencies. Thus mmWave experiences higher reflectivity than optical frequencies, further contributing to ghost target formation.

The scenario becomes significantly more complex in environments with multiple targets and walls, making it increasingly difficult to distinguish real targets from ghosts. Physics-based modeling approaches struggle to simultaneously account for all of these interacting factors. Therefore, we employ a data-driven deep learning methodology to achieve more robust and reliable target discrimination.

3. Methods

The ghost suppression problem can be equivalently formulated as a point cloud segmentation task, where each point is assigned a probability indicating whether it corresponds to a real target or a ghost. However, directly performing segmentation on each frame's point cloud may lead to suboptimal performance. To address this, we propose a trajectory-based point cloud segmentation method. As illustrated in Figure 3, the overall framework consists of four stages: preliminary segmentation, inter-frame trajectory tracking, trajectory feature aggregation, and trajectory feature broadcasting.

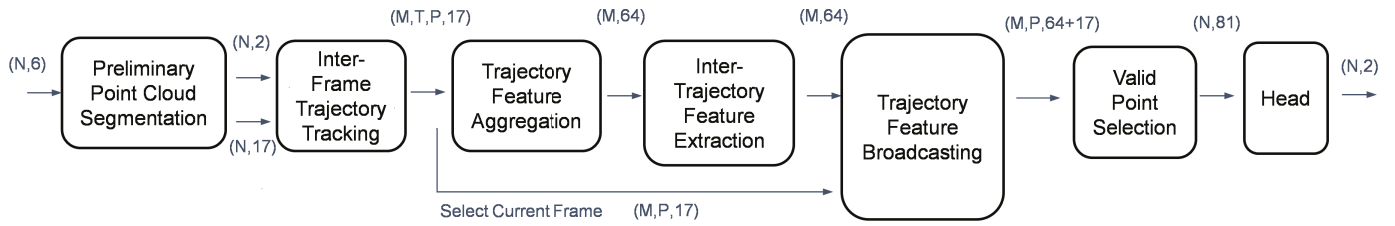


Figure 3. Overall pipeline of the proposed trajectory-based ghost target segmentation framework.

3.1. Inter-Frame Trajectory Tracking

Due to fluctuations in target energy and other factors, it is challenging to identify ghosts using only single-frame point clouds. To improve the recognition rate, it is beneficial to accumulate multiple frames for joint analysis. As shown in Figure 2b, by aggregating over a period of time, the overall amount of information increases significantly. In theory, the more frames are accumulated, the better the recognition performance should be. However, increasing the number of frames also results in a larger number of points, which not only increases the computational complexity but also poses challenges for the scalability of the algorithm with respect to temporal accumulation.

To enhance the model's ability to extract multi-frame information, we propose a tracking-based approach that maintains multiple trajectories and extracts features from them.

Each trajectory consists of four components: a trajectory ID, a first-in-first-out (FIFO) queue of associated historical points, a Kalman filter state, and a counter for consecutive unmatched frames. The most critical component is the FIFO queue, which has a depth of T and stores all points associated with the trajectory across the past T frames. Assuming that a maximum of P points can be associated per frame, the FIFO can hold up to $T \times P$ points.

Standard Kalman filtering [33] is used to estimate the current position and uncertainty of the trajectory, enabling more reliable data association. The Kalman state is defined using a 2D constant-velocity model:

$$\mathbf{x} = \begin{bmatrix} x \\ y \\ v_x \\ v_y \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xv_x} & \sigma_{xv_y} \\ \sigma_{xy} & \sigma_y^2 & \sigma_{yv_x} & \sigma_{yv_y} \\ \sigma_{xv_x} & \sigma_{yv_x} & \sigma_{v_x}^2 & \sigma_{v_x v_y} \\ \sigma_{xv_y} & \sigma_{yv_y} & \sigma_{v_x v_y} & \sigma_{v_y}^2 \end{bmatrix}. \quad (6)$$

As shown in Figure 4, the detailed process of multi-target tracking is as follows:

1. Assume that N targets are detected by radar in each frame, and each target has C features. These detections are associated with existing trajectories. During association, each detection searches for the closest trajectory in spatial distance. If the distance between them is less than the threshold d_{th} , they are associated with each other.
2. Each detection can be associated with at most one trajectory, but each trajectory may be associated with multiple detections. These associated detections are added to the trajectory's FIFO and used to update its Kalman filter state.
3. If a trajectory is not associated with any detection, a null value is added to the trajectory's FIFO; if no detection is associated with the trajectory for T_{del} consecutive frames, the trajectory is deleted.
4. For detections not associated with any trajectory, DBSCAN clustering [34] is performed ($eps = d_{th}$, $MinPts = 1$). Each cluster creates a new trajectory; the points in the cluster are added to the trajectory's FIFO and used to initialize the Kalman state. Setting $MinPts = 1$ allows each unassociated detection to potentially form its own cluster,

ensuring that no detection is discarded and every point can be assigned to a trajectory for subsequent processing.

5. After all detections have been associated, each trajectory performs a Kalman prediction step to update its state and uncertainty for the next frame.
6. Since one human body is often detected as multiple points, to avoid two trajectories tracking the same person, a repulsion mechanism is introduced. If the distance between two trajectories is less than d_{th} , they are forcibly repelled by adjusting their positions to make the distance equal to d_{th} . Here, the position refers to the position in the Kalman state. The repulsion is asymmetric, depending on the size of the Kalman state variance. A trajectory with higher uncertainty (i.e., a larger trace of the covariance matrix) is adjusted more, reflecting its lower reliability:

$$\Delta \mathbf{p}_i = (d_{th} - \|\mathbf{p}_i - \mathbf{p}_j\|) \cdot \frac{\mathbf{p}_i - \mathbf{p}_j}{\|\mathbf{p}_i - \mathbf{p}_j\|} \cdot \frac{\text{tr}(\mathbf{P}_j)}{\text{tr}(\mathbf{P}_i) + \text{tr}(\mathbf{P}_j)} \quad (7)$$

where $\mathbf{p}_i = [x_i, y_i]^T$ is the position of track i , and $\mathbf{p}_j$ is that of a nearby track j ; $\text{tr}(\mathbf{P})$ denotes the trace of the covariance matrix.

7. After the above steps, the tracking module completes all operations for the current frame. Finally, the contents of each trajectory's FIFO are output. In this way, an input point cloud of the shape (N, C) is transformed into trajectory data of the shape (M, T, P, C) , where M denotes the number of trajectories, T denotes the number of historical frames stored in each trajectory, P denotes the number of associated points per trajectory per frame, and C denotes the number of features per point.

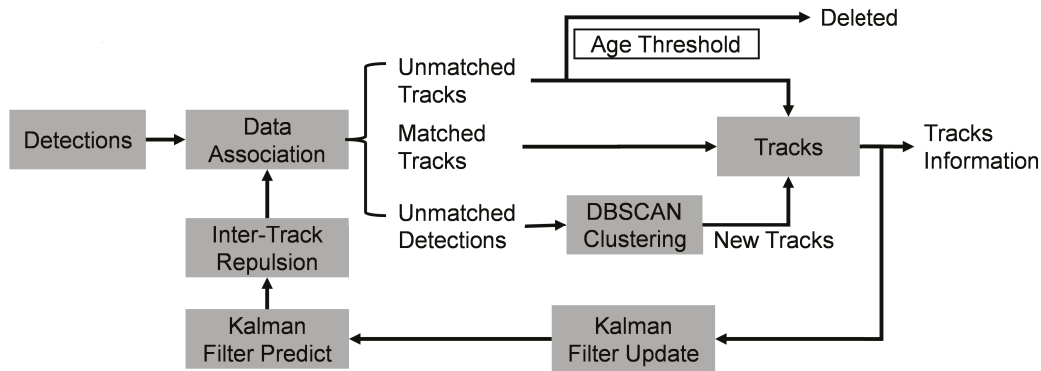


Figure 4. Inter-frame trajectory tracking pipeline.

The tracking parameters are set as follows: association threshold $d_{th} = 60$ cm, deletion delay $T_{del} = 9$, and trajectory buffer depth $T = 30$. These key parameters are empirically selected; d_{th} and T_{del} are tuned based on visual inspection of stable tracking results on the dataset, while T balances temporal context with memory and computation requirements in the subsequent network. Other parameters are determined by the system characteristics; $N = 24$ is based on the radar point cloud density, and $M = 32$ (maximum number of trajectories) and $P = 12$ (maximum number of points per trajectory per frame) are chosen to cover typical scene demands, with zero-padding applied when necessary. The data modality within the FIFO is illustrated in Figure 5.

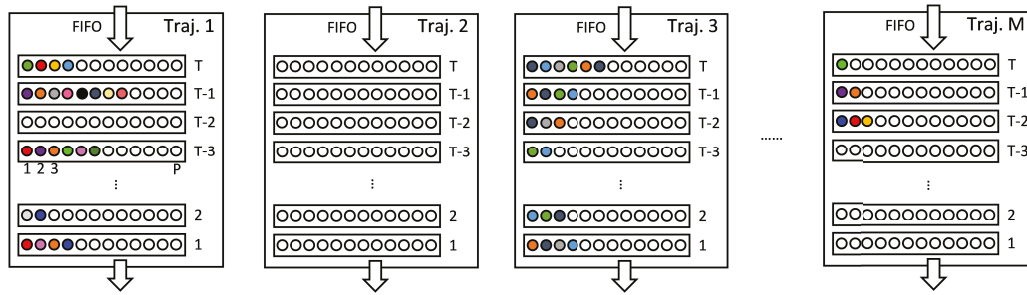


Figure 5. Structure of the trajectory FIFO. Colored circles: detections; empty circles: free slots.

After this process, all detections are assigned to a trajectory. This operation can be viewed as a form of clustering, where the clustering is implicitly performed in a temporally aware manner.

3.2. Preliminary Point Cloud Segmentation

Due to the presence of a considerable number of ghost targets in the point cloud, directly performing tracking on the raw data may lead to trajectory confusion. This issue is particularly prominent in narrow spaces such as corridors, where some ghost points may appear spatially close to real targets, causing the trajectory to be dragged by ghosts during the tracking process. To mitigate such effects, we apply frame-wise preliminary segmentation to the point cloud before tracking in order to suppress the interference caused by ghost targets to some extent.

The point cloud segmentation assigns each point a class probability, denoted as P_n and $1 - P_n$, representing the likelihoods of being a real target and a ghost, respectively, where $n \in [0, N)$. This probability is later used in the Kalman measurement update to adjust the measurement noise covariance matrix $\mathbf{R}_n$, which quantifies the uncertainty of point n 's position estimation.

As shown in Figure 6, when a ghost point falls within the d_{th} range of a trajectory, it will be associated with that trajectory. Assuming that ghost points are uniformly and randomly distributed, the equivalent measurement noise covariance of a ghost point, denoted as $\mathbf{R}_f$, satisfies

$$\mathbf{R}_f = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix} \quad (8)$$

$$\sigma_y^2 = \sigma_x^2 = E[(x - \bar{x})^2] = E[\tilde{x}^2] = \int_0^{2\pi} \int_0^{d_{th}} (\tilde{r} \cos \tilde{\theta})^2 \cdot \frac{1}{\pi d_{th}^2} \tilde{r} d\tilde{r} d\tilde{\theta} = \frac{d_{th}^2}{4} \quad (9)$$

where $\tilde{x}$ denotes the position offset from the trajectory center in the x -direction, and $(\tilde{r}, \tilde{\theta})$ are the corresponding polar coordinates centered at the trajectory. The integral assumes a uniform probability density $\frac{1}{\pi d_{th}^2}$ over the circular region of radius d_{th} .

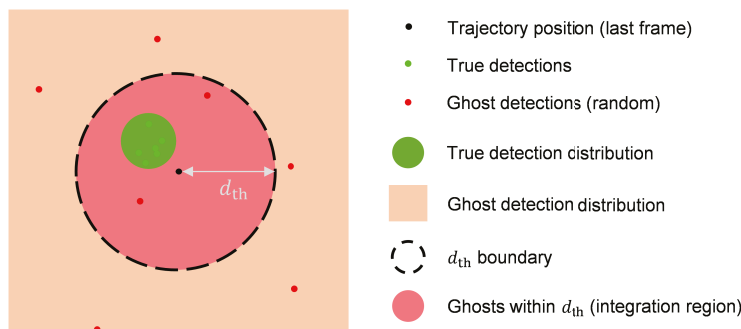


Figure 6. Geometric relationship of trajectories and detections with threshold d_{th} .

Thus, we have $\mathbf{R}_f = \frac{d_{th}^2}{4} \mathbf{I}_2$, where $\mathbf{I}_2$ is the 2×2 identity matrix. For a point with real target probability P_n , the mixed measurement covariance is defined as

$$\mathbf{R}_n = P_n \mathbf{R}_t + (1 - P_n) \frac{d_{th}^2}{4} \mathbf{I}_2 \quad (10)$$

where $\mathbf{R}_t$ is the measurement noise covariance of real targets. In this work, we set $\mathbf{R}_t = 0.01 \mathbf{I}_2$, which is a small value. This linear mixing provides a soft transition between the low-noise model for real targets and the high-uncertainty model for ghosts, enabling more robust Kalman filtering under uncertainty. The use of adaptive measurement noise, combined with the trajectory repulsion mechanism, helps the tracker focus on high-confidence measurements and improves robustness in the presence of ghost targets.

In general, point cloud segmentation can be addressed using classical deep learning architectures such as PointNet [35]. The PointNet segmentation framework is defined as follows:

$$f(x_i, X) = g\left(x_i, \max_{j \in [1, N]} h(x_j)\right) \quad (11)$$

where x_i denotes the feature vector of the i -th point to be segmented, with $i \in [1, N]$ and N being the total number of points. $f(x_i, X)$ represents the segmentation result for point x_i , and $X = \{x_1, x_2, \dots, x_N\}$ is the entire point cloud. The functions $g(\cdot)$ and $h(\cdot)$ are nonlinear mappings implemented by multilayer perceptrons (MLPs).

In this work, we employ a deep-learning-based method for preliminary segmentation, namely, PairwiseNet, which was proposed in [30] and demonstrates stronger feature extraction capabilities than those of PointNet. Its formulation is given by

$$f(x_i, X) = g\left(\text{agg}_{j \in [1, N]} h(x_i, x_j - x_i)\right), \quad (12)$$

where agg denotes the aggregation operation, which can be implemented as either max pooling or average pooling.

As shown in Figure 7, PairwiseNet employs a “Pairwise” operation to establish point-wise correlations in the point cloud, subsequently aggregating features through shared-weight MLPs and pooling operations. Further implementation specifics can be found in [30].

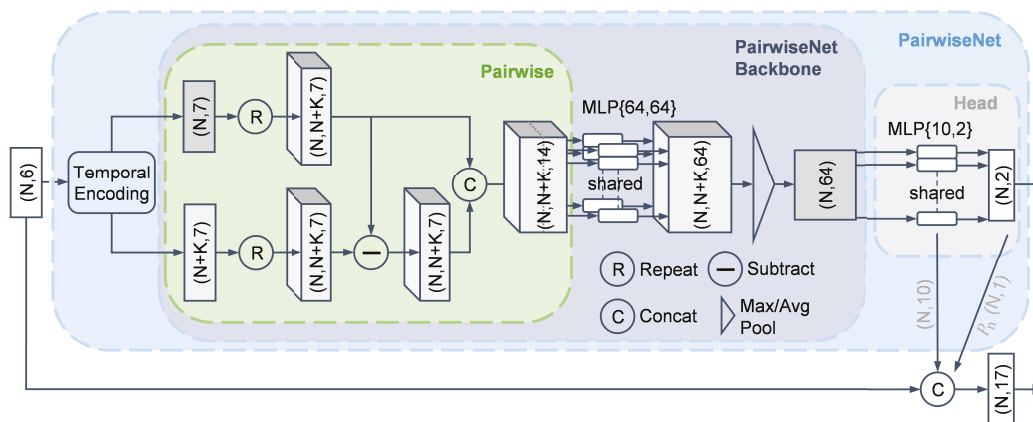


Figure 7. Architecture of PairwiseNet for preliminary point cloud segmentation.

The design rationale of PairwiseNet in [30] is that each ghost point typically originates from a real target (i.e., has a source); thus, pairwise operations are introduced to model such relationships. Considering the randomness and sparsity of radar point clouds,

PairwiseNet in [30] adopts multi-frame accumulation and temporal encoding to further enhance segmentation performance. Although PairwiseNet is capable of leveraging multi-frame information, it is not well suited for processing long-term dynamic point clouds directly. This limitation stems from its pooling-based feature aggregation mechanism, which lacks the ability to effectively capture temporal dependencies and struggles to extract time-varying features.

This work retains the pairwise design rationale but differs in scope; while the authors of [30] apply it at the point level, we extend it to the trajectory level. Specifically, our method first extracts rich trajectory features and then applies PairwiseNet across trajectories (details are provided in the next section). The trajectory-based method proposed in this paper is specifically designed to enhance temporal feature extraction capabilities, thereby complementing the weaknesses of PairwiseNet in this regard.

In addition to being used for measurement covariance estimation to reduce tracking errors, the segmentation results from PairwiseNet are also employed for feature enhancement. As illustrated in Figure 7, the point-wise features from the head layer of PairwiseNet, along with the predicted probability P_n , are concatenated with the original point features. Consequently, the number of features per point increases from 6 to 17.

3.3. Trajectory Feature Aggregation

Based on the method described in Section 3.1, we obtain trajectory-aligned data with the shape (M, T, P, C) , where $C = 17$. This data modality is structurally complex; the M -dimension corresponds to different trajectories, exhibiting permutation invariance and sparsity; the T -dimension represents time, which contains local temporal correlations and a degree of translational invariance; the P -dimension corresponds to point cloud samples per frame, which is also unordered and sparse.

To extract meaningful representations from this structure, we propose a trajectory feature aggregation network that applies different aggregation strategies tailored to the characteristics of each dimension.

As shown in Figure 8, we first apply feature aggregation along the point cloud dimension P . Inspired by PointNet, we adopt a shared-weight MLP followed by a pooling layer, which is well suited for handling unordered point cloud data. This step is crucial because the subsequent CNN requires structured inputs and cannot directly operate on sparse, unordered point sets. After pooling, the feature dimension is reduced to $(M, T, 32)$, where each trajectory at each time step is represented by a 32-dimensional feature vector.

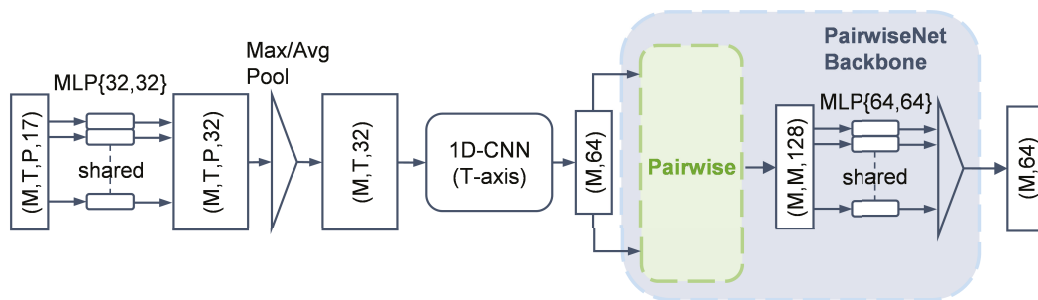


Figure 8. Trajectory feature aggregation pipeline.

Next, along the temporal dimension T , we apply a downsampling convolutional neural network (CNN) to exploit the translational invariance and extract local temporal features. The detailed architecture is listed in Table 1. To reduce the number of parameters and computational cost, each convolutional layer is implemented using depthwise-separable convolution (DSC) [36]. After passing through three DSC layers, the output is flattened

into a feature vector. As a result, we obtain a feature tensor of shape $(M, 64)$, where each trajectory is encoded into a compact 64-dimensional representation.

Table 1. Architecture of the temporal CNN using depthwise-separable convolutions.

Stage	Operation	Kernel	Stride	Input Shape	Output Shape
1	Transpose	–	–	$(M, 30, 32)$	$(M, 32, 30)$
2	DSC 1: Depthwise Conv + ReLU	4	2	$(M, 32, 30)$	$(M, 32, 14)$
3	DSC 1: Pointwise Conv + ReLU	1	1	$(M, 32, 14)$	$(M, 32, 14)$
4	DSC 2: Depthwise Conv + ReLU	4	2	$(M, 32, 14)$	$(M, 32, 6)$
5	DSC 2: Pointwise Conv + ReLU	1	1	$(M, 32, 6)$	$(M, 32, 6)$
6	DSC 3: Depthwise Conv + ReLU	4	2	$(M, 32, 6)$	$(M, 32, 2)$
7	DSC 3: Pointwise Conv + ReLU	1	1	$(M, 32, 2)$	$(M, 32, 2)$
8	Reshape	–	–	$(M, 32, 2)$	$(M, 64)$

3.4. Inter-Trajectory Feature Extraction

To incorporate the global context along the trajectory dimension M , we again utilize the PairwiseNet backbone. As illustrated in Figure 8, the input tensor of the shape $(M, 64)$ is pairwise-combined to form a tensor of the shape $(M, M, 128)$, representing all trajectory pairs. This pairwise structure is identical to that in Figure 7, consisting of two repetitions, one subtraction, and one concatenation operation. The resulting tensor is then passed through a shared-weight MLP followed by a pooling layer, resulting in a final output of the shape $(M, 64)$.

3.5. Trajectory Feature Broadcasting

Each detection is associated with a specific trajectory, and each trajectory may contain multiple detections. To utilize trajectory-level features for segmentation, we broadcast the feature vector of each trajectory to all of its associated detections.

In real-time processing scenarios, we are often only interested in the current frame. Therefore, we select the current time step from the temporal dimension T , resulting in detection data of the shape $(M, P, 17)$. At this stage, the historically informed trajectory features obtained in the previous section—summarizing each trajectory over T frames and shaped as $(M, 64)$ —are broadcast (or repeated) to match the shape $(M, P, 64)$ and concatenated with the current frame’s detection features to obtain data of the shape $(M, P, 81)$. In this way, each detection is augmented with the feature of its corresponding trajectory.

Since the M and P dimensions may include zero-padded empty trajectories and empty points, we remove all invalid entries and retain only valid detections, resulting in a final feature tensor of the shape $(N, 81)$, where N corresponds to the original number of points in the frame. The effect of this broadcasting operation is illustrated in Figure 9.

Finally, these enriched point features are fed into a shared-weight MLP classification head to generate the segmentation output of the shape $(N, 2)$.

The mechanism of broadcasting aggregated trajectory features back to the individual points within that trajectory enables each point to incorporate not only its instantaneous features but also the spatiotemporal context derived from the entire trajectory. As a structured and efficient form of spatiotemporal information in multi-target systems, trajectories exhibit strong clustering capability for real targets. This context enhances segmentation performance beyond what instantaneous features or pre-segmentation alone can achieve.

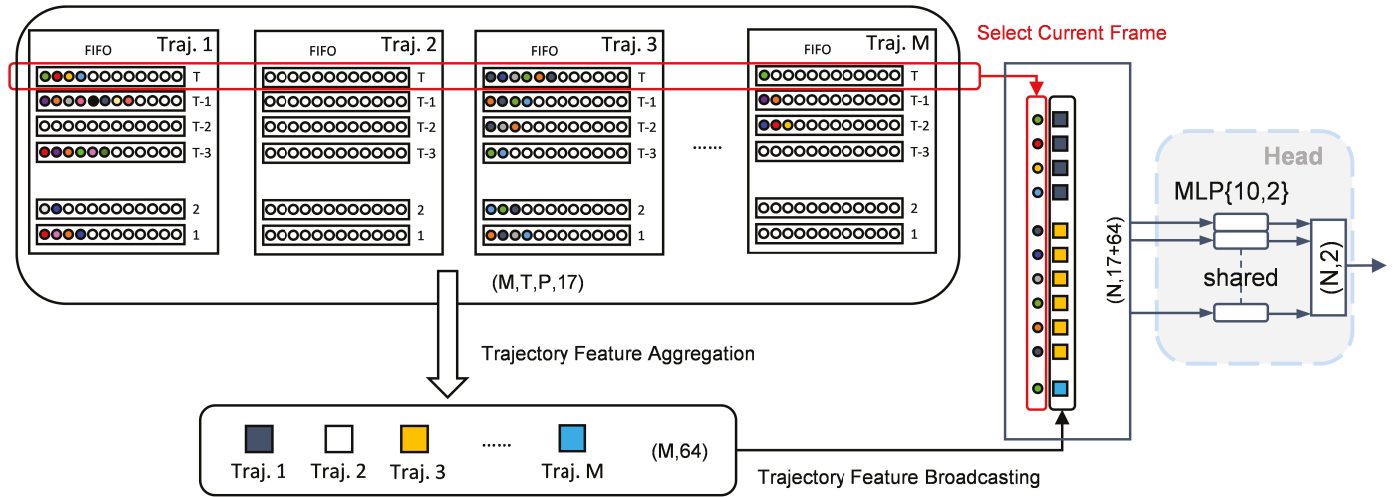


Figure 9. Trajectory-to-detection feature broadcasting.

4. Experiments and Evaluation

4.1. Dataset

The experiments in this paper were conducted on the indoor radar ghost dataset proposed in [30]. The dataset was collected using a 24 GHz 2T4R mmWave radar (ICLegend Micro, China) equipped with an eight-channel equivalent uniform linear array. It operates with a 250 MHz bandwidth, corresponding to a range resolution of 60 cm.

The dataset contains 63 scenes and a total of 80,355 radar frames. The scenarios cover a variety of indoor environments, including halls, corridors, meeting rooms, and office areas, as illustrated in Figure 10. The radar platform is placed on the floor, tables, or cabinets, with mounting heights ranging from 40 cm to 150 cm to introduce diversity in installation locations. Each scene contains 0 to 5 people, who are allowed to move freely. Radar point clouds are annotated using an Azure Kinect DK (Microsoft Corporation, Redmond, WA, USA) depth camera. In total, the dataset contains 461,383 valid annotated points, among which 296,247 are real targets, and the remaining are ghost points.

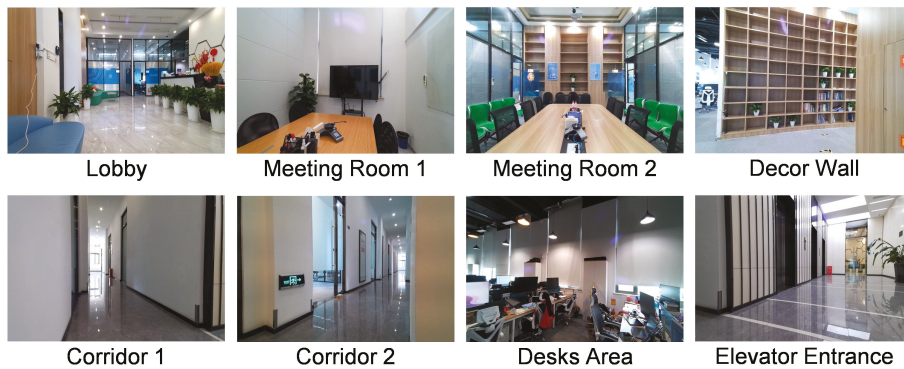


Figure 10. Representative scenes from the indoor radar ghost dataset.

4.2. Evaluation Metrics

For point cloud segmentation tasks, we adopt multiple evaluation metrics to assess model performance. The primary metric is point-wise classification accuracy, defined as the ratio of correctly classified points to the total number of points. Additionally, the precision, recall, and F1-score of the real target category are also used as metrics.

In practical applications, the costs associated with false positives and false negatives often differ significantly. This imbalance necessitates careful selection of thresholds to achieve an optimal trade-off between these two types of errors. For system performance

evaluation, we use receiver operating characteristic (ROC) curve analysis, which visually illustrates the relationship between true positive detection rates and false alarm rates across different decision thresholds. The area under the ROC curve (AUROC) serves as a reliable indicator of overall performance. Additionally, average precision (AP) is also employed as a metric, which is the area under the precision–recall curve.

4.3. Experimental Setup

For the preliminary point cloud segmentation, we directly reuse the pretrained weights of PairwiseNet provided in [30], without retraining. Based on the segmentation results and the tracking method described in Section 3.1, we generate the trajectory data for each frame in the entire dataset, resulting in a tensor of the shape $(B_{\text{all}}, M, T, P, C)$, where B_{all} denotes the total number of frames. These data are used to train the trajectory-based point cloud segmentation network, which includes trajectory feature aggregation and broadcasting modules and is referred to as TrajNet in this paper.

The loss function used during training is the cross-entropy (CE) loss, defined as

$$\mathcal{L}_{\text{CE}}(y, \hat{y}) = -\frac{1}{N_{\text{eff}}} \sum_{i=1}^{N_{\text{eff}}} [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (13)$$

where y_i is the ground-truth label (0 or 1) of the i -th point, $\hat{y}_i$ is the predicted probability of being a real target, and N_{eff} is the number of valid points in the point cloud. Placeholder points (zero-padded entries) are excluded from the loss and all evaluation metrics.

We use the Adam optimizer during training, with exponential decay rates of 0.9 and 0.999 for the first- and second-order moment estimates, respectively. The learning rate is set to 0.003, and the batch size is 64.

The 63 scenes in the dataset are split into 45 for training, 11 for validation, and 7 for testing, which is consistent with the split in [30]. This ensures that the test set consists of entirely new scenes that are not seen during training. The training set is used for gradient descent updates, while the validation set is used for early stopping to prevent overfitting. The AUROC is calculated on the validation set after each epoch. If the AUROC score stops improving and starts to decline, training is terminated early, and the best-performing model parameters are saved. The maximum number of training epochs is set to 50.

4.4. Trajectory Visualization

To verify the effectiveness of the tracking method described in Section 3.1 and to provide intuitive visualization of the trajectory data in the shape (M, T, P, C) , this section presents several examples of point cloud trajectories obtained through our method, as shown in Figure 11. These visualizations directly correspond to the input of the trajectory feature aggregation network (Figure 8), where each subfigure contains up to M trajectories (represented by curves of different colors), each trajectory consists of up to $T \times P$ points connected in the order of detection over time, and each point carries C features (though only spatial coordinates are visualized here).

In Scene 1, a meeting room enclosed by three walls with a high presence of ghosts, one person walks around the environment. The algorithm successfully tracks the trajectory of the real person and also tracks the trajectories of ghosts located behind walls. In Scene 2, where a long wall is present and four people move around the area, the algorithm is able to track all four individuals. Even when two of them pass closely by each other, with their point clouds nearly overlapping, the tracker can still distinguish their respective trajectories.

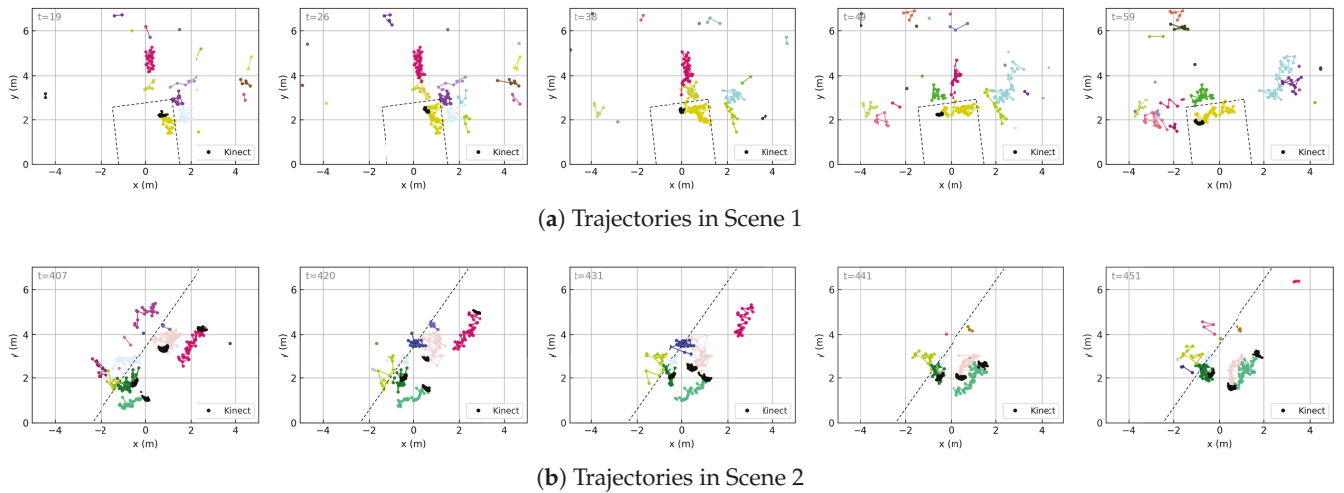


Figure 11. Trajectory evolution over time. The frame number is indicated in the top-left corner of each subfigure and increases from left to right. Dashed lines represent major reflective surfaces (i.e., walls). Curves of different colors represent different trajectories, and the points along each curve are the detections associated with that trajectory. Dense clusters of small black points indicate ground-truth human positions obtained through the Kinect, serving as a reference for tracking accuracy.

All of the above trajectory results are obtained using the point cloud pre-segmentation method described in Section 3.2. As a comparison, we also evaluate the tracking results without using pre-segmentation (i.e., with fixed measurement noise covariance). In general, both approaches achieve reasonable tracking performance. However, tracking without pre-segmentation tends to result in more frequent trajectory confusion, especially in narrow environments.

To highlight the difference, we select a narrow corridor (Scene 3) where ghost points frequently appear near real targets and may capture trajectories, as shown in Figure 12b,d. In contrast, the trajectories generated using point cloud pre-segmentation (Figure 12a,c) demonstrate better stability.

4.5. Quantitative Results

The performance of the proposed TrajNet on the test set is summarized in Table 2. It achieves an accuracy of 93.5%, with an AP of 0.992 and an AUROC of 0.982, using only 24.9k parameters. Apart from adopting depthwise-separable convolution, no additional optimization for parameter efficiency has been applied, indicating that the model size could potentially be reduced further.

For comparison, we also report the test performance of several baseline models, including PointNet, DGCNN, PCT, and PairwiseNet, as detailed in [30]. Among all models, TrajNet achieves the best performance among non-map-based methods and is second only to PairwiseNet(R-Map), which leverages environment mapping. This mapping-based approach is less robust in dynamic environments. Once the radar position changes, the map must be regenerated. In contrast, the proposed TrajNet is map-free. Additionally, we plot the ROC curves of all models in Figure 13.

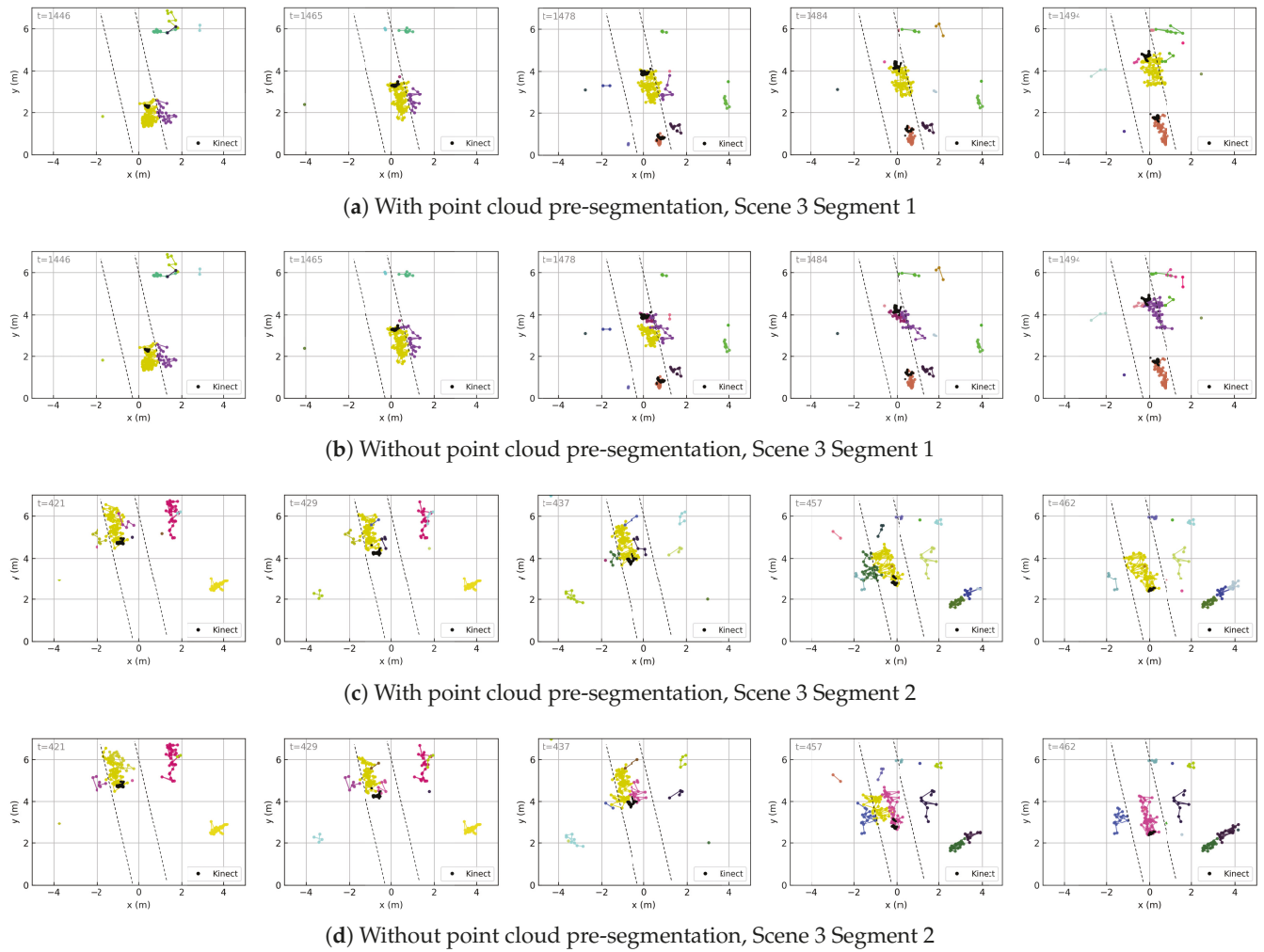


Figure 12. Comparison of trajectories obtained with and without point cloud pre-segmentation. The visualizations follow the same format as in Figure 11.

Table 2. Performance comparison between TrajNet and other models.

Model	Acc (%)	Prec. (%)	Recall (%)	F1 (%)	AP (%)	AUROC (%)	Params
MLP{64,16,2}	80.6	84.5	88.5	86.4	93.1	85.7	1.5k
PointNet [35]	84.5	90.8	86.6	88.7	96.4	91.8	29.7k
PointNet++ [37]	91.3	95.0	92.4	93.7	98.7	97.0	38.7k
DGCNN [38]	91.1	95.1	92.0	93.5	98.6	96.7	227.7k
PCT [39]	88.4	93.9	89.1	91.5	97.8	94.9	629.9k
PairwiseNet [30]	91.8	96.8	91.3	94.0	99.0	97.5	6.2k
PairwiseNet (R-Map) [30]	96.0	96.8	97.5	97.1	99.7	99.2	16.4k
PairwiseNet (T = 30)	92.3	97.0	91.8	94.3	99.1	97.8	6.2k
TrajNet	93.5	95.8	94.9	95.3	99.2	98.2	24.9k
TrajNet (w/o pre-seg)	92.0	95.0	93.5	94.3	98.3	96.4	23.8k

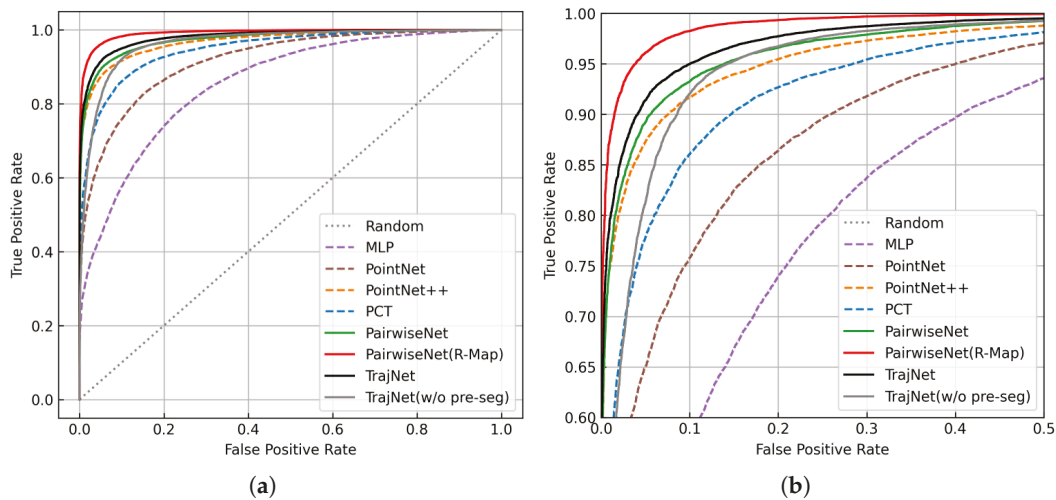


Figure 13. ROC curves of TrajNet and other models. (a) Overall ROC curves. (b) Zoomed-in view of the ROC curves near the top-left corner.

To ensure a fair comparison, we also increase the number of accumulated frames in PairwiseNet from the default 8 to 30, matching the FIFO depth of TrajNet. This variant, denoted as PairwiseNet ($T = 30$), shows slight performance improvement but still underperforms compared to TrajNet, indicating the latter's advantage in modeling long-term temporal sequences.

Finally, we also include the performance of TrajNet(w/o pre-seg), which corresponds to the model without point cloud pre-segmentation (i.e., the fourth row in Table 3). Compared to the full TrajNet model, it exhibits a noticeable performance drop, underscoring the importance of pre-segmentation. This observation will be discussed in detail in the next section.

Table 3. Performance of TrajNet with different component ablation configurations. A checkmark (✓) indicates that the corresponding component is enabled.

No.	Pre-Seg. for Tracking	Pre-Seg. for Features	Trajectory Aggrega- tion	Pairwise Inter- Trajectory	Attention- Based Inter- Trajectory	Acc	AUROC
1	✓	✓	✓	✓		93.5%	98.2%
2		✓	✓	✓		93.3%	98.1%
3	✓		✓	✓		92.1%	96.5%
4			✓	✓		92.0%	96.4%
5		✓				91.8%	97.5%
6	✓	✓	✓			93.3%	98.1%
7	✓	✓	✓		✓	93.4%	98.1%

4.6. Ablation Study

4.6.1. Component Ablation

To further evaluate the contribution of each component in TrajNet, we conduct a series of ablation experiments, and the results are shown in Table 3. Row 1 represents the default TrajNet model.

Since the point cloud pre-segmentation serves two purposes—assisting tracking by reducing trajectory confusion and enhancing feature representation—we perform ablations for each purpose separately, corresponding to Rows 2 and 3, respectively. Row 4 removes both uses, i.e., completely discards the pre-segmentation.

Row 5 presents the opposite configuration of Row 4—pre-segmentation is retained, but the trajectory-based refinement (i.e., second-stage segmentation) is removed. Based on the first five rows, we conclude that the pre-segmentation and the second-stage segmentation capture complementary information. The main benefit of pre-segmentation arises from the feature-level integration. Based on the architectural differences between the two stages, we speculate that pre-segmentation is better at capturing spatial features, while the trajectory-based refinement (TrajNet stages) excels at leveraging local temporal dynamics, making it more effective in suppressing ghost targets with inconsistent temporal behavior compared to real trajectories.

In addition, we ablate the inter-trajectory feature extraction module described in Section 3.4, as shown in Row 6. A slight performance drop is observed. Furthermore, we replace the pairwise module with an attention mechanism (Row 7), which yields slightly inferior performance compared to the pairwise design.

4.6.2. Input Feature Ablation

We further conduct ablation experiments on the six input features of the original radar point cloud, as summarized in Table 4. Row 1 corresponds to the default TrajNet configuration, where all six features are used.

From the results, we observe that the least important feature is the signal power P , followed by the radial velocity v . The most critical features are the spatial coordinates x and y , which contribute more than the range r and the angle $\sin \theta$.

These results also demonstrate the robustness of the proposed method; even when the input modality undergoes significant changes—for example, when only the x and y coordinates are retained—the model still achieves reasonably high performance.

Table 4. Performance of TrajNet with different combinations of input features. A checkmark (✓) indicates that the corresponding feature is included in the input.

No.	x	y	r	$\sin \theta$	v	P	Acc
1	✓	✓	✓	✓	✓	✓	93.5%
2	✓	✓	✓	✓	✓		93.3%
3	✓	✓	✓	✓		✓	92.7%
4	✓	✓	✓	✓			92.4%
5	✓	✓			✓	✓	90.9%
6	✓	✓					90.3%
7			✓	✓	✓	✓	89.3%

4.7. Point Cloud Segmentation Results

To more intuitively demonstrate the effectiveness of ghost suppression, we present the segmentation results of two point cloud sequence segments, as shown in Figure 14. Each sequence spans approximately 5 s and corresponds to Scene 1 and Scene 2 in Figure 11.

In Scene 1, there is only one person, with relatively few real targets and a large number of ghosts. In this scenario, PointNet exhibits multiple false positives and PairwiseNet produces one false positive, while TrajNet achieves completely clean segmentation with no false detections.

Scene 2 is more challenging, involving four individuals. In this case, PointNet suffers from numerous missed detections; both person 3 and person 4 are poorly recognized, and one false positive is also observed. PairwiseNet is able to detect person 3 but struggles to identify person 4. In contrast, TrajNet shows improved recognition of person 4, with only a small number of missed detections.

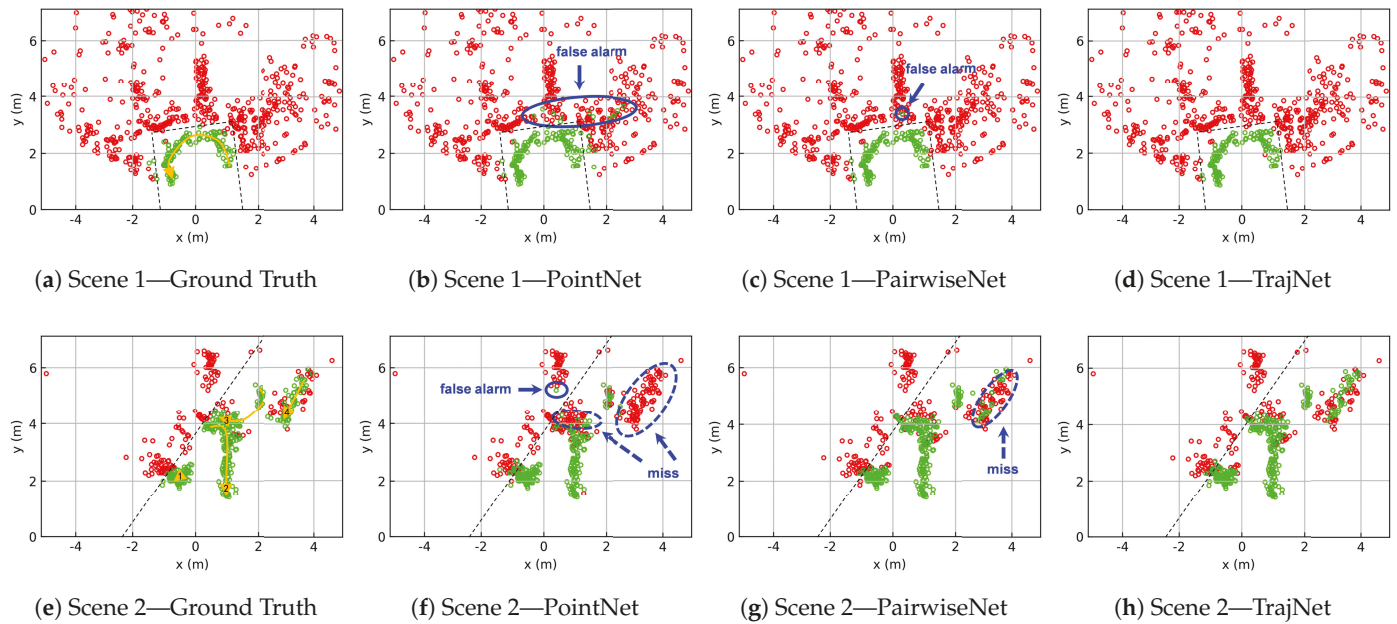


Figure 14. Point cloud segmentation results (green: true targets, red: ghosts). Blue circles highlight key differences among the segmentation outputs of the various networks. Dashed black lines indicate the primary reflective surfaces (i.e., walls). In the ground-truth plots, yellow arrows mark the actual movement trajectories of individuals, and the numbers on the arrows denote their IDs.

5. Conclusions

This paper proposes a trajectory-based ghost suppression method for mmWave radar. The approach consists of key steps that include point cloud pre-segmentation, inter-frame trajectory tracking, trajectory feature extraction, and broadcasting, effectively integrating spatiotemporal information with point-level features.

The experimental results demonstrate that our method outperforms existing point cloud segmentation approaches, achieving 93.5% accuracy and 98.2% AUROC on the test set. Ablation studies further validate the importance of each component, particularly highlighting that the combination of pre-segmentation and trajectory-based processing captures complementary information that contributes significantly to the overall performance.

Despite efforts to reduce trajectory confusion, the proposed method can still suffer from incorrect associations during tracking. Future work may explore improvements in multi-target tracking robustness. In addition, the current method uses CNN-based temporal feature extraction (chosen for its training stability, parallelizability, and convergence properties) along the trajectory dimension, which requires the explicit storage of a temporal window for each trajectory using a FIFO queue. A promising direction for future research is to replace this design with recurrent neural networks (RNNs), such as assigning a long short-term memory (LSTM) to each trajectory with shared parameters. This could eliminate the need for maintaining historical point storage and reduce redundant computation from sliding windows, enabling more efficient streaming processing.

Author Contributions: Conceptualization, R.L.; methodology, R.L.; software, R.L., Z.Q. and X.S.; validation, R.L. and Z.Q.; resources, H.X.; data curation, R.L. and X.S.; writing—original draft preparation, R.L.; writing—review and editing, Z.Q. and X.S.; visualization, R.L. and Z.Q.; supervision, H.X.; project administration, L.Y.; funding acquisition, Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Industry Prospect and Key Technology Project of Jiangsu Key R&D Program (No. BE2022109).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Due to sponsorship-related restrictions, the dataset is not publicly available but can be shared by the authors upon reasonable request.

Conflicts of Interest: Author Lei Yang and Yue Lin were employed by the company ICLegend Micro. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Frazao, A.; Pinho, P.; Albuquerque, D. Radar-Based Heart Cardiac Activity Measurements: A Review. *Sensors* **2024**, *24*, 7654.
2. Mehrjoueresht, P.; Hail, R.E.; Karsmakers, P.; Schreurs, D.M.M.P. Respiration and Heart Rate Monitoring in Smart Homes: An Angular-Free Approach with an FMCW Radar. *Sensors* **2024**, *24*, 2448. [CrossRef] [PubMed]
3. Liang, T.; Liu, R.; Yang, L.; Lin, Y.; Shi, C.J.R.; Xu, H. Fall Detection System Based on Point Cloud Enhancement Model for 24 GHz FMCW Radar. *Sensors* **2024**, *24*, 648. [CrossRef]
4. Shen, Z.; Nunez-Yanez, J.; Dahnoun, N. Advanced Millimeter-Wave Radar System for Real-Time Multiple-Human Tracking and Fall Detection. *Sensors* **2024**, *24*, 3660. [CrossRef]
5. Balachandran, A.; Tharmarasa, R.; Acharya, A.; Chomal, S. Ghost Track Detection in Multitarget Tracking using LSTM Network. In Proceedings of the 2023 26th International Conference on Information Fusion (FUSION), Charleston, SC, USA, 27–30 June 2023; pp. 1–8. [CrossRef]
6. Thai, K.P.H.; Rabaste, O.; Bosse, J.; Poullin, D.; Sáenz, I.D.H.; Letertre, T.; Chonavel, T. Detection–Localization Algorithms in the Around-the-Corner Radar Problem. *IEEE Trans. Aerosp. Electron. Syst.* **2019**, *55*, 2658–2673. [CrossRef]
7. Guo, S.; Zhao, Q.; Cui, G.; Li, S.; Kong, L.; Yang, X. Behind Corner Targets Location Using Small Aperture Millimeter Wave Radar in NLOS Urban Environment. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 460–470. [CrossRef]
8. Luo, H.; Guo, S.; Jiang, M.; Chen, J.; Cui, G. A Reflective Surface Estimation Method Based on Multipath Utilization. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 8001611. [CrossRef]
9. Gal, A.; Bilik, I. Hybrid Approach for Reflective Surfaces Reconstruction Using Automotive Radar. In Proceedings of the 2023 IEEE International Radar Conference (RADAR), Sydney, Australia, 6–10 November 2023; pp. 1–5. [CrossRef]
10. Lv, Y.F.; Shi, L.F.; Yin, W.; Shi, Y.; Wang, L. Indoor Radar Point Cloud Ghost Elimination Method Based on Point Cloud Cluster Occlusion. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 8508409. [CrossRef]
11. Feng, R.; De Greef, E.; Rykunov, M.; Pollin, S.; Bourdoux, A.; Sahli, H. Multipath Ghost Recognition and Joint Target Tracking with Wall Estimation for Indoor MIMO Radar. *IEEE Trans. Radar Syst.* **2024**, *2*, 154–164. [CrossRef]
12. Zhu, Z.; Guo, S.; Chen, J.; Xue, S.; Xu, Z.; Wu, P.; Cui, G.; Kong, L. Non-Line-of-Sight Targets Localization Algorithm via Joint Estimation of DoD and DoA. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 8506311. [CrossRef]
13. Chen, W.; Yang, H.; Bi, X.; Zheng, R.; Zhang, F.; Bao, P.; Chang, Z.; Ma, X.; Zhang, D. Environment-aware Multi-person Tracking in Indoor Environments with MmWave Radars. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2023**, *7*, 89. [CrossRef]
14. Visentin, T.; Hasch, J.; Zwick, T. Analysis of multipath and DOA detection using a fully polarimetric automotive radar. In Proceedings of the 2017 European Radar Conference (EURAD), Nuremberg, Germany, 11–13 October 2017; pp. 45–48. [CrossRef]
15. Feng, R.; De Greef, E.; Rykunov, M.; Sahli, H.; Pollin, S.; Bourdoux, A. Multipath Ghost Classification for MIMO Radar Using Deep Neural Networks. In Proceedings of the 2022 IEEE Radar Conference (RadarConf22), New York City, NY, USA, 21–25 March 2022; pp. 1–6. [CrossRef]
16. Zheng, L.; Long, J.; Lops, M.; Liu, F.; Hu, X.; Zhao, C. Detection of Ghost Targets for Automotive Radar in the Presence of Multipath. *IEEE Trans. Signal Process.* **2024**, *72*, 2204–2220. [CrossRef]
17. Park, J.K.; Park, J.H.; Kim, K.T. Multipath Signal Mitigation for Indoor Localization Based on MIMO FMCW Radar System. *IEEE Internet Things J.* **2024**, *11*, 2618–2629. [CrossRef]
18. Luo, H.; Zhu, Z.; Jiang, M.; Guo, S.; Cui, G. An Effective Multipath Ghost Recognition Method for Sparse MIMO Radar. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5111611. [CrossRef]
19. Li, Y.; Shang, X. Multipath Ghost Target Identification for Automotive MIMO Radar. In Proceedings of the 2022 IEEE 96th Vehicular Technology Conference (VTC2022-Fall), London, UK, 26–29 September 2022; pp. 1–5. [CrossRef]
20. Prophet, R.; Martinez, J.; Michel, J.C.F.; Ebelt, R.; Weber, I.; Vossiek, M. Instantaneous Ghost Detection Identification in Automotive Scenarios. In Proceedings of the 2019 IEEE Radar Conference (RadarConf), Boston, MA, USA, 22–26 April 2019; pp. 1–6. [CrossRef]
21. Wang, S.; Wang, H.; Wang, G.; Weng, W.; Wang, R. Algorithm Design to Identify Targets From Multipath Ghosts and Clutters for SISO IR-UWB Radar: Incorporating Waveform Features and Trajectory. *IEEE Sens. J.* **2024**, *24*, 2278–2288. [CrossRef]

22. Chen, Y.; Yang, M.; Wang, X.; Chen, B. Multipath Model and Ghost Suppression for Millimeter Wave Traffic Radar. In Proceedings of the 2024 IEEE International Conference on Signal, Information and Data Processing (ICSIDP), Zhuhai, China, 22–24 November 2024; pp. 1–5. [CrossRef]
23. Jeong, T.; Lee, S. Ghost Target Suppression Using Deep Neural Network in Radar-Based Indoor Environment Mapping. *IEEE Sens. J.* **2022**, *22*, 14378–14386. [CrossRef]
24. Kopp, J.; Kellner, D.; Piroli, A.; Dallabetta, V.; Dietmayer, K. Simultaneous Clutter Detection and Semantic Segmentation of Moving Objects for Automotive Radar Data. In Proceedings of the 2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC), Bilbao, Spain, 24–28 September 2023; pp. 1078–1085. [CrossRef]
25. Stephan, M.; Santra, A. Radar-Based Human Target Detection using Deep Residual U-Net for Smart Home Applications. In Proceedings of the 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA), Boca Raton, FL, USA, 16–19 December 2019; pp. 175–182. [CrossRef]
26. Chuanhao, Z.; Xu, H.; Jiamin, L.; Mingxuan, L.; Xueyao, H.; Liping, Z. A Multipath Signals Recognition Method based on Deep Neural Network. In Proceedings of the 2024 IEEE International Conference on Signal, Information and Data Processing (ICSIDP), Zhuhai, China, 22–24 November 2024; pp. 1–6. [CrossRef]
27. Kraus, F.; Scheiner, N.; Ritter, W.; Dietmayer, K. The Radar Ghost Dataset—An Evaluation of Ghost Objects in Automotive Radar Data. In Proceedings of the 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Prague, Czech Republic, 27 September–1 October 2021; pp. 8570–8577. [CrossRef]
28. Kopp, J.; Kellner, D.; Piroli, A.; Dietmayer, K. Tackling Clutter in Radar Data - Label Generation and Detection Using PointNet++. In Proceedings of the 2023 IEEE International Conference on Robotics and Automation (ICRA), London, UK, 29 May–2 June 2023; pp. 1493–1499. [CrossRef]
29. Wang, L.; Giebenhain, S.; Anklam, C.; Goldluecke, B. Radar Ghost Target Detection via Multimodal Transformers. *IEEE Robot. Autom. Lett.* **2021**, *6*, 7758–7765. [CrossRef]
30. Liu, R.; Song, X.; Qian, J.; Hao, S.; Lin, Y.; Xu, H. A Data-Driven Method for Indoor Radar Ghost Recognition With Environmental Mapping. *IEEE Trans. Radar Syst.* **2024**, *2*, 910–923. [CrossRef]
31. Longman, O.; Villeval, S.; Bilik, I. Multipath Ghost Targets Mitigation in Automotive Environments. In Proceedings of the 2021 IEEE Radar Conference (RadarConf21), Atlanta, GA, USA, 8–14 May 2021; pp. 1–5. [CrossRef]
32. Sabery, S.M.; Bystrov, A.; Gardner, P.; Stroescu, A.; Gashinova, M. Road Surface Classification Based on Radar Imaging Using Convolutional Neural Network. *IEEE Sens. J.* **2021**, *21*, 18725–18732. [CrossRef]
33. Welch, G.; Bishop, G. *An Introduction to the Kalman Filter*; Technical Report TR 95-041; University of North Carolina: Chapel Hill, NC, USA, 2006.
34. Feng, R.; Greef, E.D.; Rykunov, M.; Sahli, H.; Pollin, S.; Bourdoux, A. Multipath Ghost Recognition for Indoor MIMO Radar. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–10. [CrossRef]
35. Charles, R.Q.; Su, H.; Kaichun, M.; Guibas, L.J. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 77–85. [CrossRef]
36. Chollet, F. Xception: Deep Learning with Depthwise Separable Convolutions. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 1800–1807. [CrossRef]
37. Qi, C.R.; Yi, L.; Su, H.; Guibas, L.J. PointNet++: Deep hierarchical feature learning on point sets in a metric space. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Red Hook, NY, USA, 4–9 December 2017; pp. 5105–5114.
38. Wang, Y.; Sun, Y.; Liu, Z.; Sarma, S.E.; Bronstein, M.M.; Solomon, J.M. Dynamic Graph CNN for Learning on Point Clouds. *ACM Trans. Graph.* **2019**, *38*, 146. [CrossRef]
39. Guo, M.H.; Cai, J.X.; Liu, Z.N.; Mu, T.J.; Martin, R.R.; Hu, S.M. PCT: Point cloud transformer. *Comput. Vis. Media* **2021**, *7*, 187–199. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

A Machine Learning Method for the Fast Simulation of the Scattering Characteristics of a Target Under a Planar Layered Medium

Zhaoyu Wang ¹, Qinghe Zhang ^{1,2,*}, Zhaoyang Shen ^{1,2}, Lei Zhang ^{1,2} and Han Liu ^{1,2}

¹ College of Computer and Information Technology, China Three Gorges University, Yichang 443002, China; 202208100021013@ctgu.edu.cn (Z.W.); shenzhaoyang@ctgu.edu.cn (Z.S.); zhanglei@ctgu.edu.cn (L.Z.); liuh@ctgu.edu.cn (H.L.)

² Hubei Key Laboratory of Intelligent Vision Based Monitoring for Hydroelectric Engineering, China Three Gorges University, Yichang 443002, China

* Correspondence: zqh@ctgu.edu.cn

Abstract: Numerical simulation of ground-penetrating radar (GPR) has been widely used to enhance the interpretation of GPR data and serves as a key component in Full Waveform Inversion (FWI). In response to the time-consuming numerical computation of layered medium and buried targets, which leads to inefficiency in full-wave inversion, this paper proposes a machine learning-based forward scattering rapid solution method. Using the detection of rebar buried in concrete under sand as the GPR application scenario, with scene parameters such as concrete moisture content, rebar radius, and burial depth, scattering echo signals are obtained via Finite Difference Time Domain (FDTD) simulation. Principal component analysis (PCA) is applied to reduce the dimensionality of the echo data, and the first 40 principal component weight coefficients are selected as the output of the deep learning network. An innovative cyclic nested deep learning network architecture is designed, which not only fully explores the intrinsic causal relationship between the scene parameters and the principal component weight coefficients, but also refines and corrects each predicted principal component. The numerical results demonstrate that, compared with traditional machine learning methods, the cyclic nested machine learning network architecture offers higher prediction accuracy and learning efficiency, validating the effectiveness of the proposed method.

Keywords: ground penetrating radar; planar layered medium; buried target; fast forward simulation; principal component analysis; machine learning

1. Introduction

In practical physical environments, layered media are widely present in various scenarios, compared to a single uniform medium [1,2]. For example, in areas prone to debris flows, the soil environment is divided into loose soil zones and slightly denser base areas; in lunar exploration, the lunar subsurface structure is divided into regolith and bedrock. In many fields such as underground resource exploration, urban infrastructure construction, and military reconnaissance, accurately acquiring information about buried targets in a layered medium is significant. In oil exploration, understanding the distribution of underground oil layers plays a decisive role in resource extraction decisions; in urban construction, precisely locating buried pipelines, cables, and other facilities can effectively prevent accidental damage during the construction process.

However, traditional forward methods face limitations such as low computational efficiency and limited accuracy when dealing with buried target problems in layered media [3]. Ground-penetrating radar (GPR) is a non-destructive underground detection technique that uses electromagnetic waves [4]. Fast forward simulation techniques are crucial for understanding the underground scattering mechanisms of GPR and for data interpretation inversion. Various forward full-wave solvers have been developed to describe the radar characteristics of underground objects, including the Finite Difference Time Domain (FDTD) method [5], Method of Moments (MoM) [6], and Finite Element Time Domain (FETD) method [7]. However, these physics-based traditional solvers typically require substantial computational time, especially when signal processing or machine learning-based inversion algorithms need to be repeatedly executed. When Full Waveform Inversion (FWI) adopts global optimization [8], objective function gradient solving requires hundreds of forward problem solving techniques. As a result, the inversion of each GPR dataset often requires hundreds or even thousands of iterations [9].

In recent years, Machine Learning (ML) [10] has become an attractive choice for solving GPR and other electromagnetic problems, aiming to alleviate the computational burden of traditional GPR forward solvers. These machine learning methods utilize Artificial Neural Networks (ANNs) to construct models and train algorithms using provided training data. For example, neural networks have been used to detect hyperbolic features in B-scan scans [11], fully connected feedforward neural networks have been employed for landmine detection and classification, and Convolutional Neural Networks (CNNs) have also been used for object detection in GPR profiles [12,13]. Some researchers have proposed a coupled NN-RF method for estimating the diameter of rebar in concrete [14] and introduced a fast ML forward solver. The authors of [15] describe the development of a machine learning-based fast forward solver for concrete slab GPR data, which utilizes dense networks and principal component analysis (PCA) to predict the A-scan response of the target below the antenna center. The forward solver is used to estimate the moisture content of concrete and the depth and diameter of rebar. As part of Full-Wave Inversion (FWI), this approach validated the accuracy of using ML models as forward solvers. Wang Haoyu proposed a Full-Waveform Inversion method based on Overall Coding Genetic Algorithms (OCGAs) [16], which improves the encoding strategy based on genetic algorithms, encoding the overall features of the individual. The above research mainly focuses on half-space and buried target detection and does not address the scenario of layered media and buried targets, with the input dimension of the designed machine learning network being relatively large, which reduces the prediction accuracy and limits the effective application of the method.

This paper addresses the rapid simulation of scattering characteristics for buried targets in planar layered media, focusing on the practical application of ground-penetrating radar (GPR) to detect rebar embedded in concrete beneath sand. The model parameters include the moisture content of the concrete, the radius of the rebar, and the burial depth. Scattering echo signals are obtained through simulation. Subsequently, a principal component analysis (PCA) is employed for dimensionality reduction in the echo data. The analysis reveals that only the first 40 principal components' weight coefficients are necessary to achieve a good fit for the original simulated echo data, thereby creating a suitable dataset format for subsequent network training. Additionally, an innovative nested loop structure machine learning approach is proposed. Various machine learning methods are applied, including single hidden layer neural networks, particle swarm-optimized single hidden layer neural networks, double hidden layer neural networks, and random forest methods. These techniques are combined with the nested loop structure and compared in terms of accuracy and speed. This comparative study aims to thoroughly validate the precision

and efficiency of the proposed method in the forward modeling of buried targets in planar layered media.

2. Materials and Methods

The application scenario of the planar layered medium model in this paper was set to a rebar buried in concrete under sand, as shown in Figure 1. The size of the detection area was $0.5 \times 0.3 \times 0.4$ meters. The first layer was sand with a depth of 0.1 m, and the second layer was concrete with a depth of 0.2 m, with the detection target, the rebar, buried in the second layer. The rebar was modeled as a perfect electric conductor (PEC). There was no air region between the perfect matched layer (PML) and the target. This is because, when simulating the propagation of electromagnetic waves in a certain medium, and only focusing on the field distribution inside the medium, the PML is directly set at the boundary of the medium.

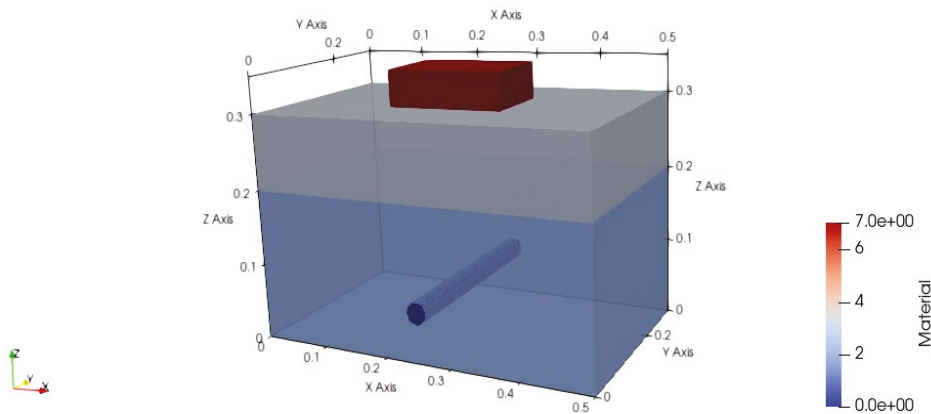


Figure 1. GPR planar layered medium model structure. (The antenna model is within the red box, and the cylinder represents the PEC rebar.).

In all simulations, the antenna's main axis was parallel to the rebar's main axis.

The dielectric properties of concrete, as a dispersive medium, are described by the extended Debye model [17]:

$$\varepsilon = \varepsilon_{\infty} + \frac{\varepsilon_s - \varepsilon_{\infty}}{1 + j\omega t_0} + \frac{\sigma}{j\omega\varepsilon_0}, \quad (1)$$

In the equation, ε_{∞} is the relative permittivity at infinite frequency, ε_s is the relative permittivity at zero frequency, t_0 is the relaxation time, σ is the conductivity, ω is the angular frequency, $j = \sqrt{-1}$, and ε_0 is the permittivity of free space. Due to the complex geometric structure of the detection scene, the same model with unified parameters was constructed in the GPRMAX (v3.1.7) software based on the FDTD principle, and the simulated time-domain waveform was compared and analyzed. This method can not only reduce the error caused by model differences, but also improve the reliability of the results, and further verify the results of FDTD.

The GPR forward problem was numerically simulated using the FDTD method to obtain A-scans. Considering computational resource limitations, the spatial step size was determined based on stability and numerical dispersion conditions, with the spatial discretization set to 1 cubic millimeter and the time step set to the Courant limit. The GPR antenna in the FDTD simulation used a 1.5 GHz antenna model from Geophysical Survey Systems, Inc. (GSSI) (with a minimum resolution of 4 mm), polarized parallel to the rebar's main axis. This model was originally created by Warren and Giannopoulos [18], where the antenna's geometry was measured, and then the unknown dielectric properties of the antenna were derived using Taguchi optimization. This updated and improved 1.5 GHz

GSSI model was used in the FDTD simulation to generate training data. The incident wave commonly used was a Gaussian pulse wave. The Gaussian pulse wave has the characteristic of a wide frequency band and can effectively simulate the short pulse signals emitted by an actual ground-penetrating radar system. This makes it highly suitable for detecting the structure and properties of subsurface media. The rebar in concrete was modeled as a Perfect Electric Conductor (PEC) cylinder, with a radius varying from 5 to 29.5 mm and a burial depth ranging from 0 to 200 mm. In this study, the relative permittivity, conductivity, permeability, and magnetic loss values for the sand were set to 6, 0, 1, and 0, respectively. These parameters provide a foundational quantitative basis for subsequent research on the characteristics of sand in electromagnetic environments and the construction and analysis of related models.

3. Network Architecture and Learning Strategy

Machine learning, as an important branch of artificial intelligence, plays a decisive role in model performance based on the choice of network architecture and learning strategy. The neural network architecture and learning strategy designed and utilized in this paper combined traditional machine learning techniques, and through a comparative analysis of accuracy and speed, we deeply explored their application in the fast forward simulation of underground targets in a planar layered medium. In the forward problem, the inputs included concrete moisture content (WC), rebar radius (R), and rebar burial depth (D).

The design philosophy of the proposed network architecture and learning strategy was to utilize the interdependence between the 40 principal component weight coefficients obtained through PCA dimensionality reduction, where the previous principal component weight coefficient should help predict the next one. Through this strategy, the predicted principal component weight coefficients can be obtained. By multiplying the predicted principal component weight coefficients with the corresponding principal component eigenvectors, the effective fitting of the A-scans echo data can be achieved.

The nested loop neural architecture proposed in this paper consisted of two independent network architectures. For the first network architecture, the first step was to input a vector composed of model parameters, including moisture content (WC), radius (R), and depth (D), in order to predict the first principal component. Then, the model parameters were combined with the first predicted principal component to predict the second principal component. The causal relationship between these principal components not only constrains the range of the prediction results but also effectively reduces the optimization search space, enhancing both computational efficiency and model performance. The second step involved using the model parameters along with the previously predicted principal components (the first and second principal components) to predict the third principal component. This process continued until all forty principal components were predicted. It is important to note that each stage was independently trained, effectively avoiding the common issues of gradient vanishing and overfitting in deep architectures, ensuring the stability and reliability of the model. At each stage, the data were re-divided into training, validation, and test sets, a strategy that ensured the generalization ability of the neural network, prevented over-reliance on a specific dataset, and improved the model's adaptability and robustness. Figure 2 shows the network structure and learning strategy of the first part, forming a deep learning system consisting of 40 layers.

However, due to the inherent uncertainty in the prediction process, these coefficients inevitably exhibit a certain degree of bias. To address this issue, a subsequent module was designed and introduced, with the core goal of establishing the causal relationship between the error of the predicted values relative to the actual axis and the model parameters, with a particular focus on the analysis of the differences between the predicted values and the

actual axis. For the second network architecture, the first step was to combine the model parameters with all the principal components except the first predicted principal component to predict the corrected first principal component. The second step involved using the model parameters along with the corrected first principal component and all the principal components except the second to predict the corrected second principal component. This structure was iteratively executed until all components were fully corrected, forming a complete and corrected set of principal components. Figure 3 visually illustrates the second part of the architecture, closely integrating with the first part of the network architecture.

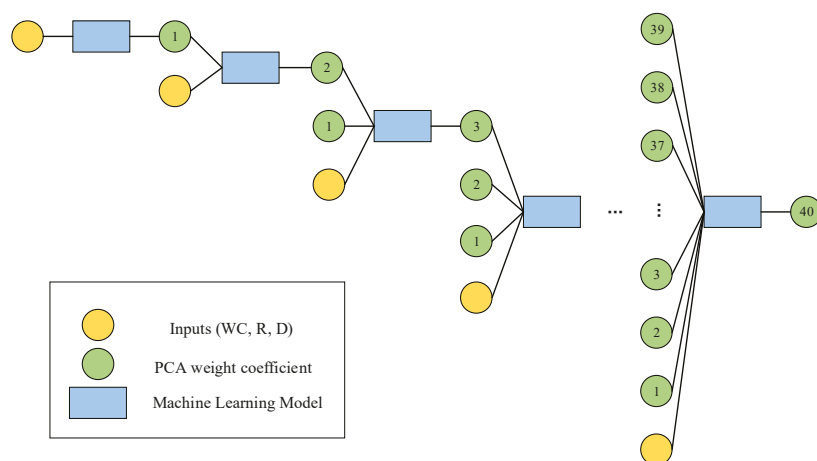


Figure 2. Architecture and learning strategy of the first part of the network.

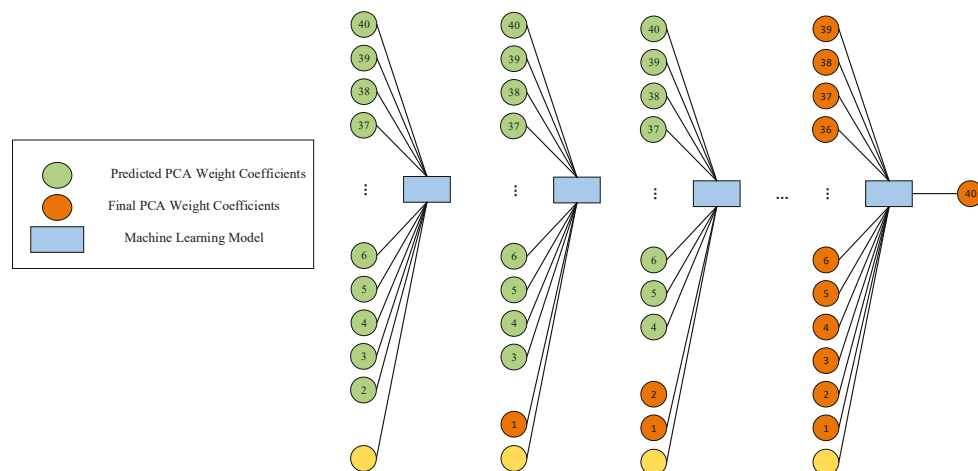


Figure 3. Architecture and learning strategy of the second part of the network.

Similar to the first part of the architecture, the second part adopted independent training, using model parameters, the output of the initial module, and the output from the previous stage of the subsequent module as inputs, ensuring the independence and accuracy of the training process, with the final PCA weight coefficients as the output. At each stage, the original data were scientifically divided into training, validation, and test sets. This strategy further reduced the risk of overfitting and significantly enhanced the model's generalization ability and robustness. The final output was a nearly real-time (≈ 1 s) solver for the 40 principal components, which can be decompressed to provide GPR A-scans under predicted time-domain modeling. It is worth noting that the designed neural network innovatively uses the predicted principal components as inputs, rather than the actual principal components. This design simplifies the network input, requiring only the provision of key parameters such as concrete moisture content, rebar radius, and burial depth.

4. Data Acquisition and Processing

To obtain the dataset required for the machine learning network, 2000 different scenarios were set up, and spline interpolation was used to obtain the Debye properties corresponding to moisture contents ranging from 0.2% to 12%. These properties were used as inputs in the FDTD algorithm to simulate concrete, generating a dataset of 2000 3D GPR models for rebar buried in concrete, which was then used to train the ML model. For each model, moisture content (WC), radius (R), and depth (D) were randomly and uniformly distributed within their respective ranges:

- (1) $WC \in [0, 12]\%$.
- (2) $R \in [5, 29.5]mm$.
- (3) $D \in [0, 200]mm$.

Any regression scheme using three input parameters to predict a relatively large vector increases the complexity of the problem and introduces instability during training, leading to noisy outputs. This can be partially mitigated by down sampling the traces every 10 time steps. To further reduce the dimensionality of the data, a principal component analysis (PCA) was employed.

PCA [19], as one of the methods for dimensionality reduction, achieves this by increasing interpretability while minimizing information loss. It does so by creating new uncorrelated variables that maximize variance consecutively. These new variables are called principal components. The process of finding these principal components simplifies to solving an eigenvalue problem, and the new variables are defined by the dataset at hand, rather than being a priori, making PCA an adaptive data analysis technique. Reference [20] proposed an automatic mode selection scheme based on mode decomposition and principal component analysis for GPR signal denoising.

PCA maps the data onto its principal orthogonal axes, and the data can be reconstructed through their linear combinations. The principal axes of the training set are composed of 300 eigenvectors. The linear combination of these 300 eigenvectors can reproduce each dataset in the training set. Figure 4 shows eight arbitrary principal components out of the 300 eigenvectors in the sample data under the scenario model.

To reduce the dimensionality of the training set, this paper did not use all 300 principal components obtained from PCA as sample data. In fact, a linear combination of a certain number of principal components can accurately fit the echo data in the training set. Echo data refer to the data reflected back after the signal encounters obstacles or changes in the medium during transmission. Many systems (such as radar and sonar) rely on the echo signal, and the time domain can clearly capture the details of the echo: determine the location of the target and reflect the material or size of the target.

Figure 5 shows the absolute error of the linear fit of the echo data using the first 40 principal components. In this figure, it can be seen that, when the first 40 most significant principal components are used for fitting, the original echo can be approximated very accurately. Therefore, the first 40 out of the 300 eigenvectors were selected as the principal components of the GPR scattering echo. This approach ensures that no information contained in the neglected principal components is lost, while also further reduces the dimensionality to speed up the network's training time.

The compressed version of these eigenvectors is then calculated as:

$$Y = A\omega, \quad (2)$$

where Y is the vector representing the actual trajectory, A is the matrix containing the 40 principal axes of the training set, and ω is the vector containing the 40 principal component weight coefficients. The vector ω can be computed using the least squares method,

$\omega = (A^T A)^{-1} A^T Y$. Therefore, all trajectories can be reproduced using the matrix A and their unique principal component weight coefficients ω . These weight coefficients ω were used as the output of the machine learning network in the subsequent GPR forward simulation, while the model parameters (WC, R, D) served as the input.

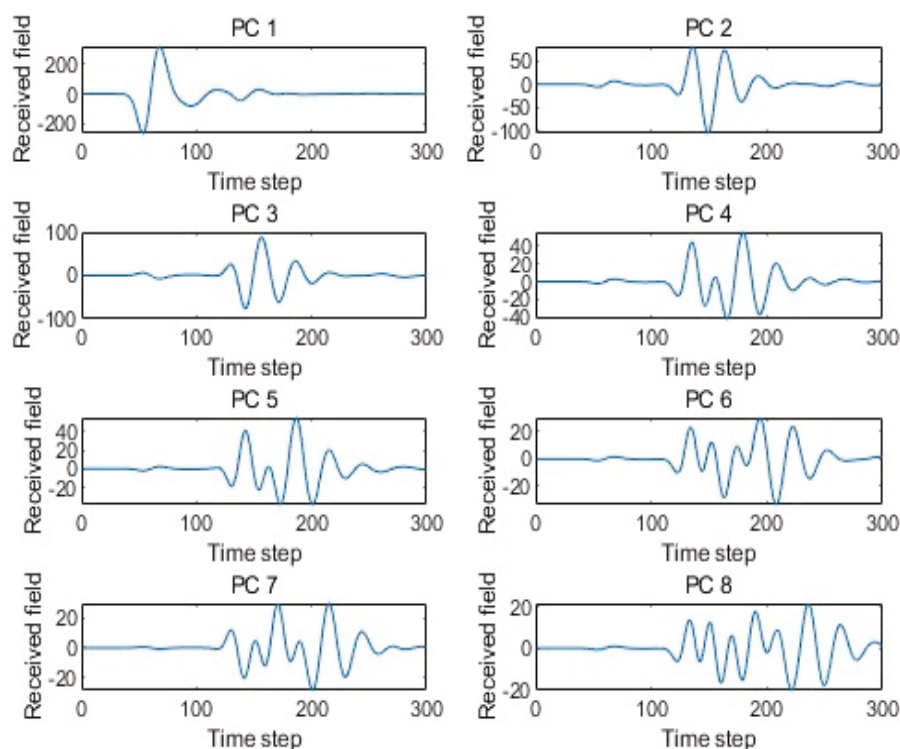


Figure 4. The eight dominant principal components of the training set.

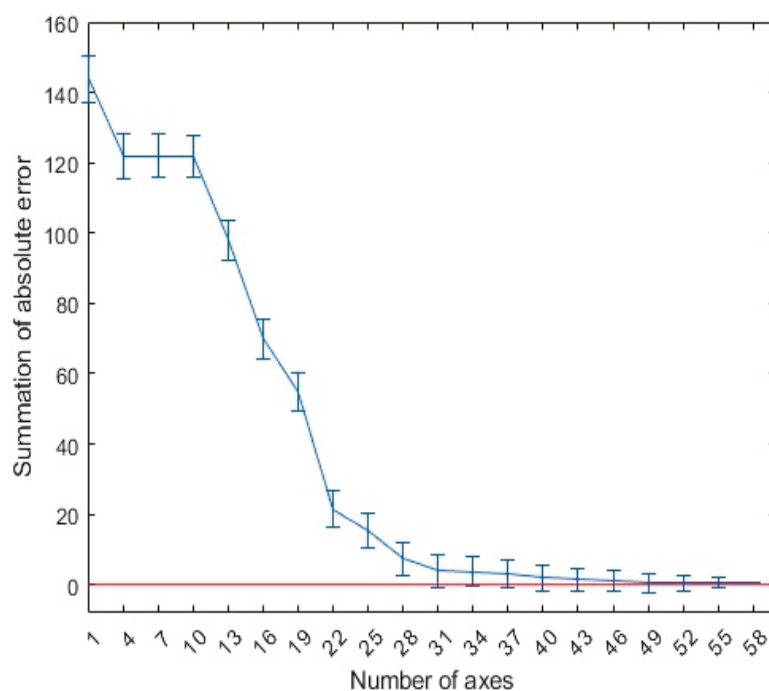


Figure 5. Average error between the actual and compressed A-scans (error bars correspond to one standard deviation).

Figure 6 shows the weight coefficients corresponding to the first 40 principal components of the GPR echo data under four arbitrary model parameters. Figure 7 compares the

full A-scans echo data from four FDTD numerical simulations with the echo data obtained through linear fitting of the 40 principal components. In the figure, it can be seen that the match between the two is good, providing a solid theoretical basis for the subsequent rapid GPR forward simulation.

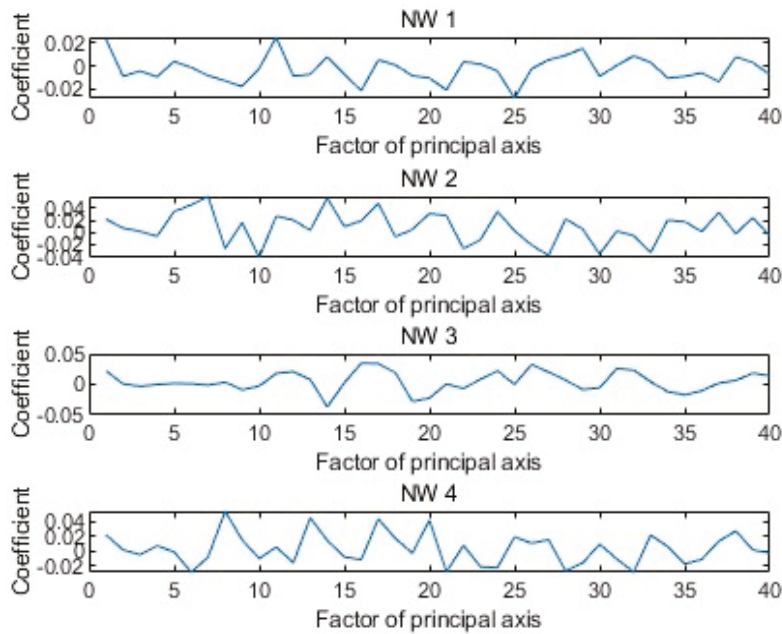


Figure 6. The weight coefficients of the 40 principal components used to represent the compressed signal.

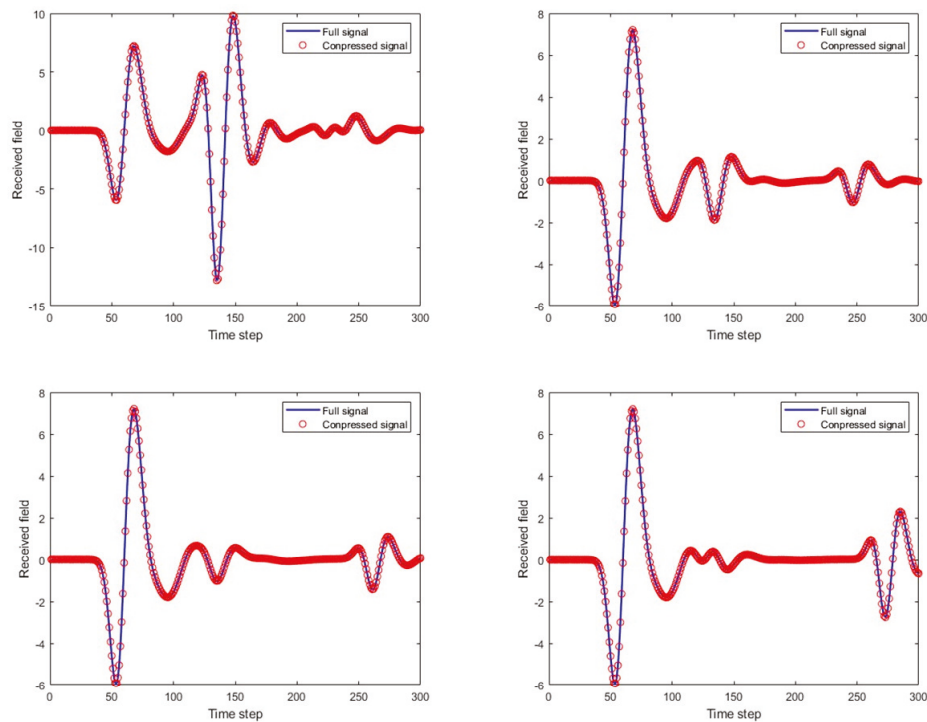


Figure 7. Four complete A-scan scans calculated by FDTD and their corresponding compressed PCA representation using 40 principal axes.

5. Results and Analysis

The dataset in this paper included 2000 randomly generated scenarios and their corresponding simulated echo data. The dataset was divided into training set (70%), validation set (15%), and test set (15%). The input to the network model consisted of

an input vector formed by the concrete moisture content (WC), rebar radius (R), and rebar burial depth (D), while the output vector corresponded to the principal component weight coefficients. The final output vector was multiplied by the principal component eigenvectors to achieve the fitting of the A-scan echo data.

First, a traditional BP neural network (with single and double hidden layers) was used to predict the echo signals. Figures 8 and 9 show the prediction results for four different scenarios and compare them with the actual FDTD simulation results. In the figures, it is clear that the prediction is good for the first A-scan layer, but fails to provide accurate predictions for the concrete layer. The complex physical properties of the concrete layer make it difficult for the traditional BP network to avoid error accumulation during the learning process, leading to poor forward simulation results for the concrete layer.

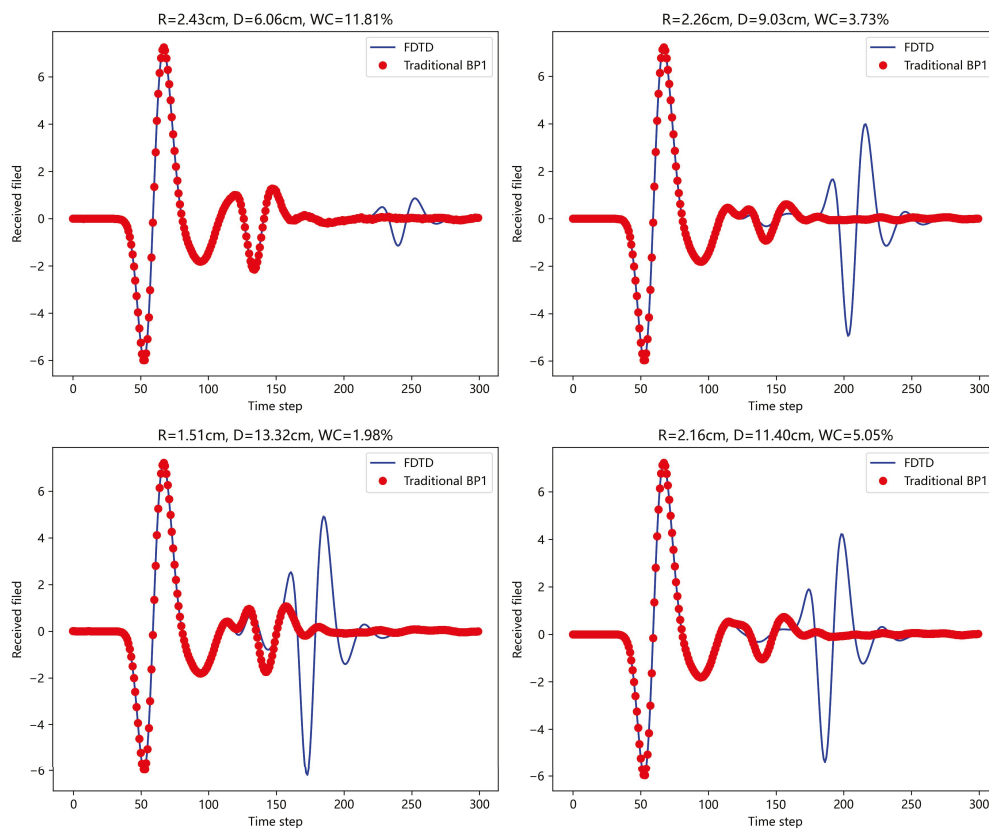


Figure 8. Under the four model parameters, the prediction results of traditional BP network (single hidden layer) are compared with those of FDTD.

Next, the scattering echoes were predicted using the method proposed in this paper. The machine learning model in the nested loop network architecture was the traditional BP neural network. The network structure included an input layer, one hidden layer, and an output layer. The hidden layer contained 20 neurons, and the ReLU activation function was used to introduce nonlinearity. The mean square error (RMSE) was used as the loss function. For the same four scenario parameters mentioned earlier, the prediction results are shown in Figure 10.

Next, for the BP neural network with a single hidden layer in the aforementioned network architecture, the Particle Swarm Optimization (PSO) [21] Algorithm was introduced. The PSO algorithm treats the weights and biases of the BP neural network as the positions of particles and simulates the flight behavior of the particle swarm to search for the optimal combination of weights w and biases b in the search space. The learning factors were set as $c1 = c2 = 2$, the number of population updates was 30, the population size was 5, and the

velocity range was $[-1, 1]$, with a boundary range of $[-1, 1]$. The prediction results for the four scenario scattering echoes are shown in Figure 11.

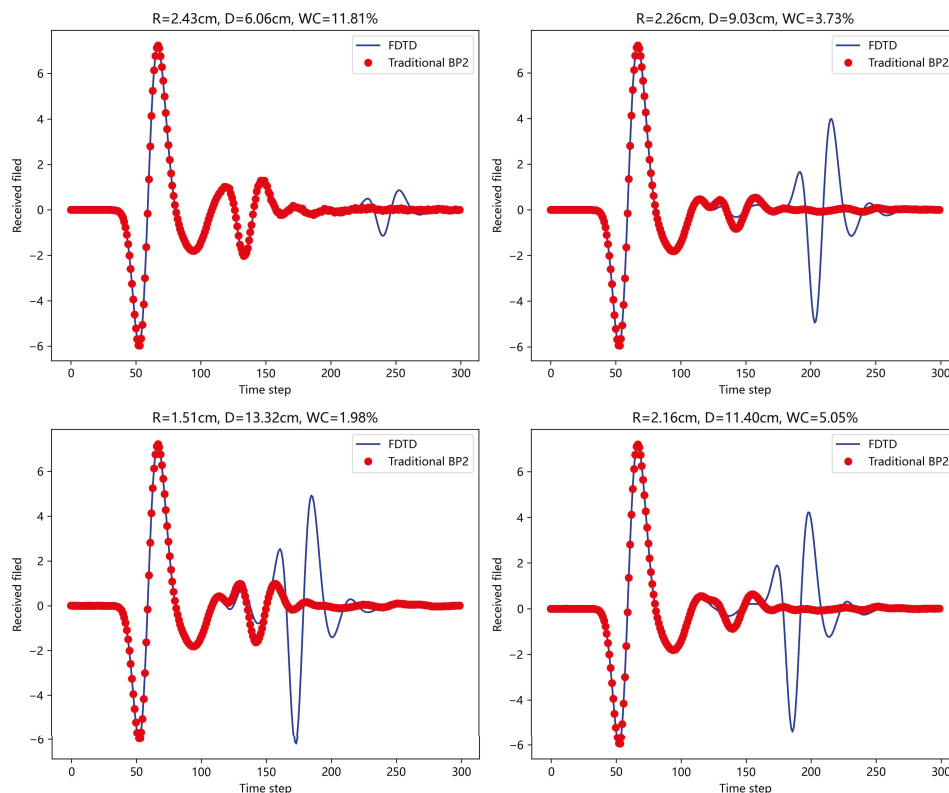


Figure 9. Under the four model parameters, the prediction results of traditional BP network (dual hidden layers) are compared with those of FDTD.

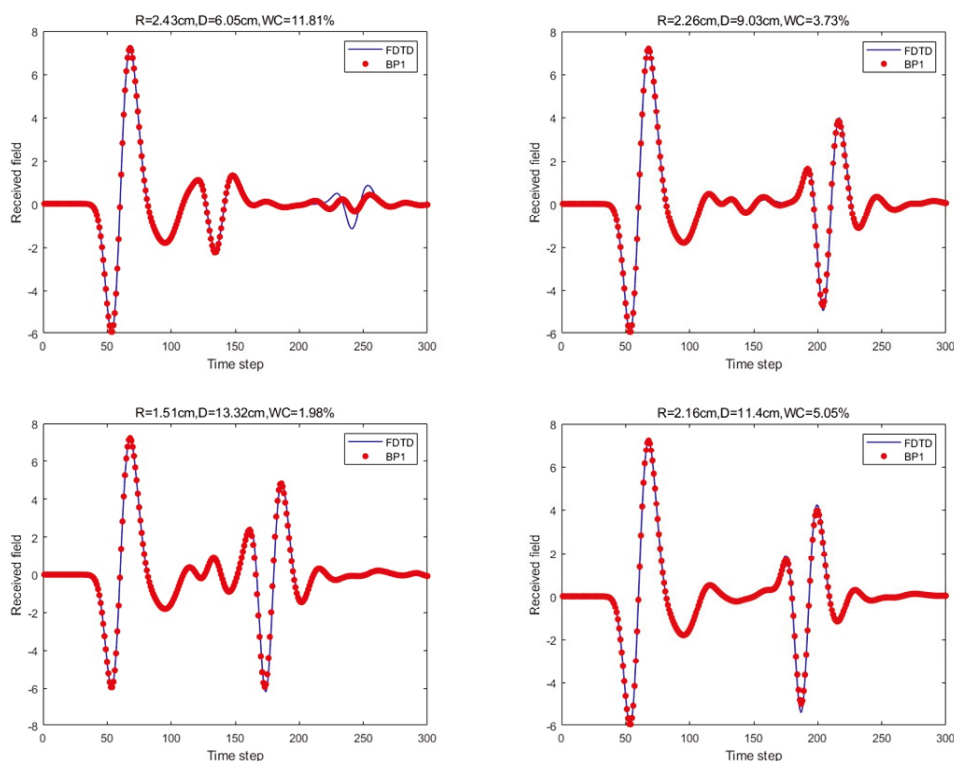


Figure 10. Under the four model parameters, the prediction results of the proposed method (single hidden layer) are compared with those of FDTD.

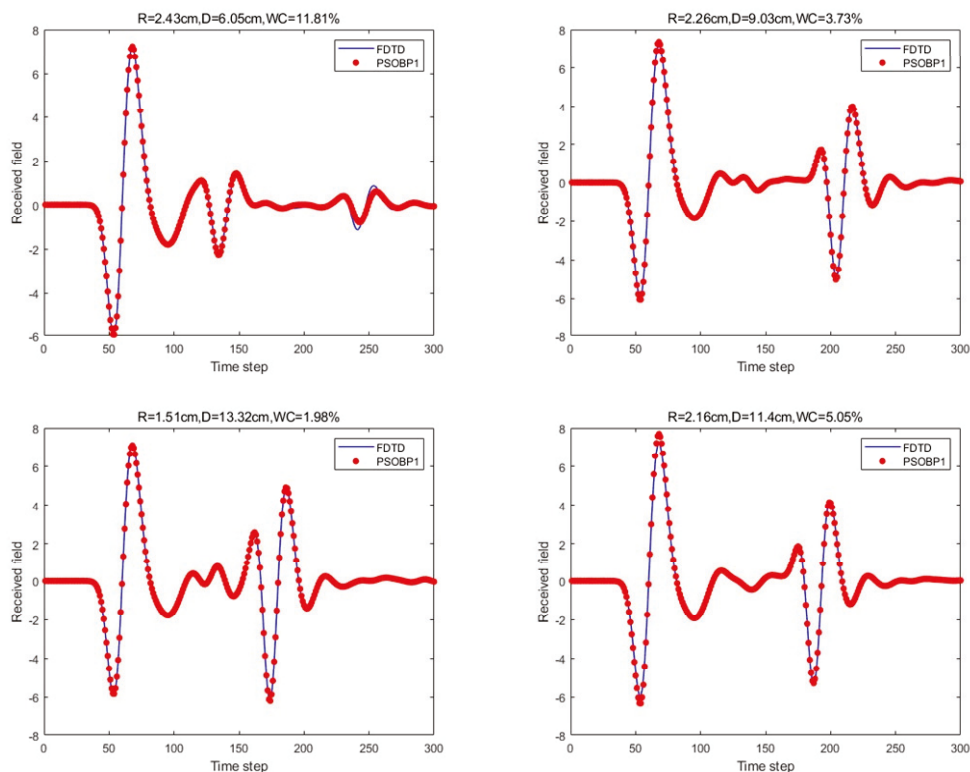


Figure 11. Under the four model parameters, the prediction results of the proposed method (PSO-optimized single hidden layer) are compared with those of FDTD.

Finally, the BP neural network with two hidden layers in the proposed network architecture was considered. The two hidden layers consisted of 20 and 15 neurons. The prediction results for the four scenario scattering echoes are shown in Figure 12.

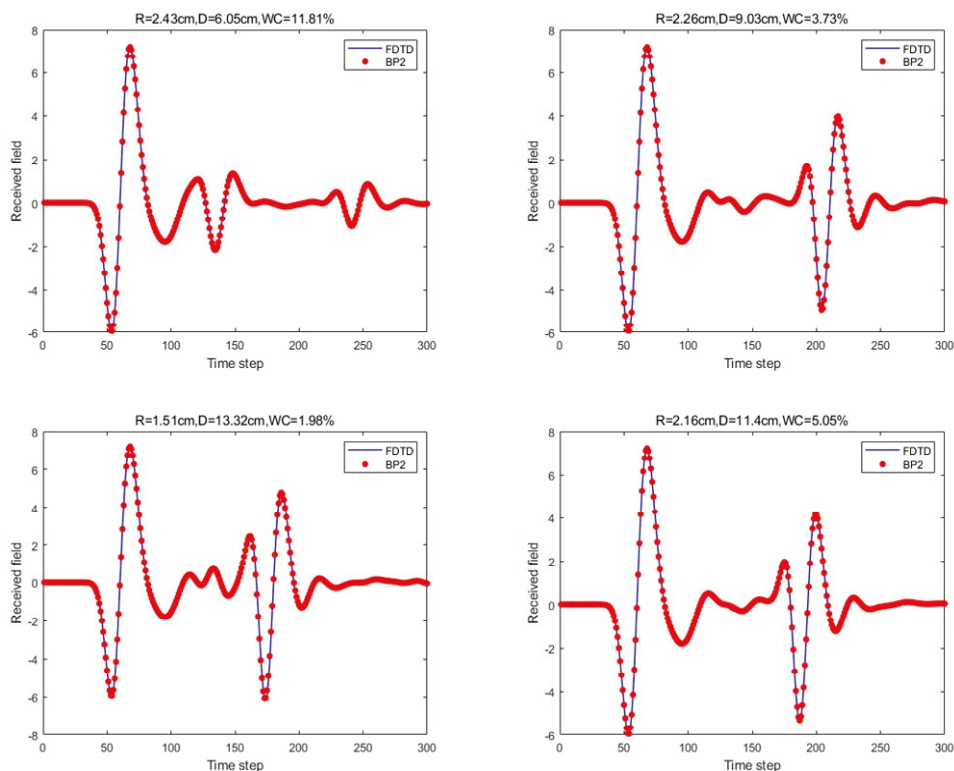


Figure 12. Under the four model parameters, the prediction results of the proposed method (dual hidden layers) are compared with those of FDTD.

For the scattering echo prediction study under four scenarios using five different methods, as shown in Figures 8–12, the traditional BP neural network fails to accurately predict the subsequent echo data, resulting in significant errors. However, when the BP neural network is combined with the network architecture proposed in this paper, the results show significant improvement. Despite this, the single hidden layer may exhibit considerable deviation for different cases; thus, PSO optimization was used, leading to a noticeable enhancement in prediction accuracy. Subsequently, the A-scans obtained from the BP neural network with dual hidden layers for rapid GPR forward simulation were compared with the A-scans directly obtained from the FDTD forward solver, and the results were nearly identical. Table 1 shows the mean square error analysis of the scattering echo prediction results for the four scenarios using the five methods.

Table 1. Mean square error of five methods across four cases.

Method	Model 1	Model 2	Model 3	Model 4
Traditional BP1	0.23363	1.00906	1.22479	1.09911
Traditional BP2	0.23095	1.00898	1.22437	1.08582
BP1	0.03216	0.01215	0.03516	0.01947
PSO_BP1	0.01282	0.00336	0.00170	0.01347
BP2	0.00052	0.00099	0.00605	0.00171

Notes: Traditional BP1: the single hidden layer BP neural network. Traditional BP2: the dual hidden layer BP neural network. BP1: the nested loop machine learning network collaborates with the single hidden layer BP neural network. PSO_BP1: the nested loop machine learning network collaborates with the PSO-optimized single hidden layer BP neural network. BP2: the nested loop machine learning network collaborates with the dual hidden layer BP neural network.

In terms of training time, the traditional BP neural network method takes up to 2 h for training. In contrast, the PSO-optimized single hidden layer method proposed in this paper reduces the training time to about 35 min; the dual hidden layer BP network requires 100 s for training, while the single hidden layer BP network trains in the fastest time of 20 s. A GPR forward simulation using the FDTD-based numerical simulation software GPRMAX for a single set of model parameters takes approximately 45 min. In comparison, once trained, the machine learning network constructed in this study can predict A-scans in real time within 1 s, significantly improving prediction efficiency.

6. Conclusions

This paper proposes a new method for GPR simulation in a layered medium, which serves as an alternative to traditional methods such as FDTD. The method is based on a neural network regression approach that employs an innovative nested loop structure combined with a BP neural network with dual hidden layers. The input to the network model consists of scene parameters (concrete moisture content, rebar radius, and burial depth), and the output is the PCA weight coefficients of the scattering echo. The output is then multiplied by the principal component eigenvectors to achieve the fitting and prediction of the A-scan echo data. The numerical results show that the proposed nested loop machine learning network architecture offers a higher prediction accuracy and learning efficiency.

Author Contributions: Conceptualization, Z.W. and Z.S.; methodology, Z.W. and L.Z.; software, Z.W.; validation, H.L.; formal analysis, Z.W.; investigation, Z.W. and Z.S.; resources, Q.Z.; data curation, Z.W.; writing—original draft preparation, Z.W. and L.Z.; writing—review and editing, Z.W. and H.L.; visualization, Z.W. and Z.S.; funding acquisition, Q.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China under Grant No. 62371271.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zhang, J. Research on the Propagation Characteristics of Vortex Electromagnetic Waves in Layered Media. Master's Thesis, University of Electronic Science and Technology of China, Chengdu, China, 2021.
2. Yamauchi, Y.; Kidera, S. Inverse scattering enhanced synthetic aperture imaging for multi-layered ground media. In Proceedings of the International Symposium on Antennas and Propagation (ISAP), Sydney, Australia, 31 October–3 November 2022; pp. 225–226.
3. Savita, S.J.; Pallavi, A. Modeling of GPR using gprMax Simulation. In Proceedings of the 2022 IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE), Ballari, India, 23–24 April 2022; pp. 1–4.
4. Alioua, C.; Grimes, M.; Kemache, N.; Benredjem, O.; Roula, R. Locating and estimating the depth of a buried cable using GPR-Radar and convolutional neural networks. In Proceedings of the 2024 2nd International Conference on Electrical Engineering and Automatic Control (ICEEAC), Setif, Algeria, 12–14 May 2024; pp. 1–5.
5. Zhao, S.; Wei, B.; He, X.; Deng, H. 2-D hybrid cylindrical fdtd method with unconditional stability. *IEEE Microw. Wirel. Technol. Lett.* **2023**, *33*, 955–958. [CrossRef]
6. Araújo, A.R.J.d.; Colqui, J.S.L.; Azevedo, W.L.M.d.; Kurokawa, S.; Filho, J.P.; Kordi, B. Transient analysis of grounding electrodes in multilayer soils using method of moments. *IEEE Lat. Am. Trans.* **2022**, *20*, 269–275. [CrossRef]
7. Kim, J.; Teixeira, F.L. On the parallel performance of SPAI-FETD solvers for time-domain maxwell's equations. In Proceedings of the 2022 United States National Committee of URSI National Radio Science Meeting (USNC-URSI NRSM), Boulder, CO, USA, 4–8 January 2022; pp. 136–137.
8. Ren, Y.; Xu, X.; Yang, S.; Nie, L.; Chen, Y. A physics-based neural-network way to perform seismic full waveform inversion. *IEEE Access* **2020**, *8*, 112266–112277. [CrossRef]
9. Feng, D.; Cao, C.; Wang, X. Multiscale full-waveform dual-parameter inversion based on total variation regularization to on-ground GPR data. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 9450–9465. [CrossRef]
10. Akhaury, U.; Giannakis, I.; Warren, C.; Giannopoulos, A. Machine learning based forward solver: An automatic framework in gprMax. In Proceedings of the 2021 11th International Workshop on Advanced Ground Penetrating Radar (IWAGPR), Valletta, Malta, 4–7 July 2021; pp. 1–6.
11. Wang, B.; Chen, P.; Zhang, G. Simulation of GPR b-scan data based on dense generative adversarial network. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2023**, *16*, 3938–3944. [CrossRef]
12. Liu, Q.; Zhu, C. Automatic detection for road voids from GPR images using deep learning method. In Proceedings of the 2023 4th International Conference on Computer Vision, Image and Deep Learning (CVIDL), Zhuhai, China, 12–14 May 2023; pp. 617–620.
13. Akçali, S.; Erden, F. Support of data augmentation with GAN on faster R-CNN based buried target detection. In Proceedings of the 2021 29th Signal Processing and Communications Applications Conference (SIU), Istanbul, Turkey, 9–11 June 2021; pp. 1–4.
14. Giannakis, I.; Giannopoulos, A.; Warren, C. A machine learning scheme for estimating the diameter of reinforcing bars using ground penetrating radar. *IEEE Geosci. Remote Sens. Lett.* **2021**, *18*, 461–465. [CrossRef]
15. Giannakis, I.; Giannopoulos, A.; Warren, C. A machine learning-based fast-forward solver for ground penetrating radar with application to full-waveform inversion. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 4417–4426. [CrossRef]
16. Wang, H.; Liu, H.; Ma, Z.; Du, P. Full waveform inversion of ground-penetrating radar based on the overall encoding genetic algorithm. *J. Microw.* **2022**, *38*, 76–81, 94.
17. Balicstis, C.M. Saddle point-based description of the propagating pulse dynamics in skin tissues and phantoms represented by Cole-Cole and Debye models. In Proceedings of the 2023 IEEE MTT-S International Microwave Biomedical Conference (IMBioC), Leuven, Belgium, 11–13 September 2023; pp. 1–3.
18. Warren, C.; Giannopoulos, A. Investigation of the directivity of a commercial ground-penetrating radar antenna using a finite-difference time-domain antenna model. In Proceedings of the 2012 14th International Conference on Ground Penetrating Radar (GPR), Shanghai, China, 4–8 June 2012; pp. 226–231.
19. Bhyregowda, P.; Masum, M.; Mamudu, L.; Chowdhury, M.; Kosariu, S.C.; Shahriar, H. Active PCA: A novel framework integrating PCA and active machine learning for efficient dimension reduction. In Proceedings of the 2024 IEEE 48th Annual Computers, Software, and Applications Conference (COMPSAC), Osaka, Japan, 2–4 July 2024; pp. 320–325.

20. Hao, T.; Jing, L.; He, W. An automated GPR signal denoising scheme based on mode decomposition and principal component analysis. *IEEE Geosci. Remote Sens. Lett.* **2023**, *20*, 1–5. [CrossRef]
21. Amzar, H.; Haziq, M.I.; May, Z. Hybrid of PSO-ANN and PCA-SVR models for the prediction of external corrosion in pipeline infrastructure: A comparative study. In Proceedings of the 2023 IEEE International Conference on Sensors and Nanotechnology (SEN-NANO), Putrajaya, Malaysia, 26–27 September 2023; pp. 57–60.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

An Active Radar Interferometer Utilizing a Heterodyne Principle-Based Target Modulator

Simon Müller ¹, Andreas R. Diewald ^{1,*} and Georg Fischer ²

¹ Department of Engineering, Trier University of Applied Sciences, 54293 Trier, Germany; s.mueller@et.hochschule-trier.de

² Department of Electrical-Electronic-Communication Engineering, Friedrich Alexander Universität Erlangen, 91058 Erlangen, Germany; georg.fischer@fau.de

* Correspondence: a.diewald@tech.hochschule-trier.de

Abstract: The Active Radar Interferometer (AcRaIn) represents a novel approach in secondary radar technology, aimed at environments with high reflective clutter, such as pipes and tunnels. This study introduces a compact design minimizing peripheral components and leveraging commercial semiconductor technologies operating in the 24 GHz ISM band. A heterodyne principle was adopted to enhance unambiguity and phase coherence without requiring synchronization or separate communication channels. Experimental validation involved free-space and pipe measurements, demonstrating functionality over distances up to 150 m. The radar system effectively reduced interference and achieved high precision in both straight and bent pipe scenarios, with deviations below 1.25% compared to manual measurements. By processing signals at intermediate frequencies, advantages such as improved efficiency, isolation, and system flexibility were achieved. Notably, the integration of amplitude modulation suppressed passive clutter, enabling clearer signal differentiation. Key challenges identified include optimizing signal processing and addressing logarithmic signal attenuation for better precision. These findings underscore AcRaIn's potential for pipeline monitoring and similar applications.

Keywords: cooperative radar; distance measurement; phase coherence; active radar target; reflective clutter suppression; sewer pipe monitoring

1. Introduction

1.1. Terminology

The term secondary radar system is not used consistently worldwide. The Institute of Electrical and Electronics Engineers (IEEE) provides a definition in [1] and restricts the term in the American context to radar systems in which an interrogator sends out an encoded signal and requests a response. The cooperative radar target then responds with an encoded message. Examples of this include civilian air traffic-control systems or military friend-or-foe identification systems. In the European context, the term is used much more broadly. It encompasses all radar systems that do not rely solely on passive echoes to identify radar targets. Various technologies are employed in this context, enabling the echo at the radar target to be directly modified in a specific way or supplemented with additional information.

1.2. Brief History

Radar target technologies have evolved over the years and can be categorized into passive, semi-passive, and active devices, each tailored to specific operational needs. We give a short description of examples of each type in the following overview.

Passive radar targets reflect or scatter incident radar signals and are often able to add information by affecting the system's passive characterization parameters such as modifying its reflection coefficient. This is the basic principle used, for example, in the Van Atta array [2] as one of the most simple principles or in the targets presented by Thornton and Edwards in 1998 and 2000 [3,4] by switching between different feeding lines of an antenna to modify its reflection coefficient.

Semi-passive targets add basic functionality, such as simple modulation schemes like phase-shifting signals by means of active components, but with minimal power requirements as there is typically no power added to the reflected signal. This type of backscatterer can be considered a passive component from an RF perspective. In 1977, Koelle and Depp [5] described the use of a cooperative target consisting of an amplitude-modulated backscatterer in combination with a Doppler radar. An advantage identified is that the modulation frequency does not need to be known a priori, nor does it require special stability, as the authors evaluate only phase changes in the baseband of the CW radar used. A half-wave dipole antenna is proposed as the backscatterer. If a pin diode is inserted as a load impedance in the middle and driven by a modulation frequency, the reflected RF signal is modulated by varying the reflection coefficient with this modulation.

Active targets go further by generating or significantly modifying signals, by adding energy to the echo. A variety of realizations can be found in Kossel et al. [6], Cantú [7], Fusco et al. [8], and Dadash [9,10], amongst other similar concepts.

Some fundamental principles are simplified in Figure 1. The two different one-port devices in Figure 1b can be implemented as transmission line segments, short circuits, or even oscillators.

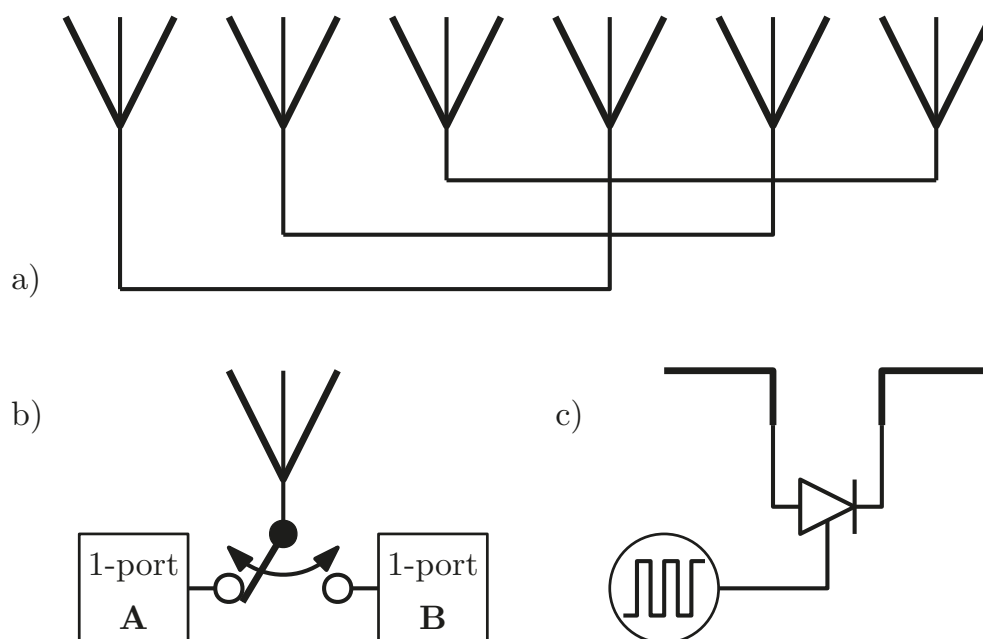


Figure 1. Examples of possible arrangements for passive and semi-passive backscatterers: (a) Van Atta array, (b) variation of antenna termination, (c) switching diode as load.

The Switched Injection-Locked Oscillator (SILO) introduced by Vossiek and Gulden [11] has been used in several variants but can only be assumed coherent by approximation and only for a short time of several oscillation cycles.

A third group is synchronized and Tachymeter units. All nodes are constructed in a very similar or identical manner, as most of the necessary operations typically need to be performed on all units, requiring significant synchronization effort as demonstrated in [12–14], besides a number of variations of these. The measurement result is finally determined after

merging data from each node. In addition to the radar itself, a communication channel is often required; alternatively, data can be transmitted using the radar signal as a carrier.

1.3. Preceding Development

In [15], the authors presented the basic principle of an active Radar Interferometer that consists of a radar base unit mounted at a fixed position that reads the signature of one or more radar targets. The units form a cooperative radar system designed for environments with high reflective clutter, such as pipes and tunnels.

This radar response system modifies the echo transmitted back to the base in a way that both unambiguity and phase coherence are given at any time during measurements. It eliminates the need for synchronization as the echo can be treated in almost the same way as a conventional passive echo. A separate communication is neither required nor intended.

The system leverages commercial semiconductor components, operating in the 24 GHz ISM band, and is designed to achieve a range of up to 120 m using an FMCW radar source.

Key advancements include a prototype that integrates amplitude modulation to suppress passive clutter and improve signal separation from the environment. Four data-acquisition techniques are discussed, including direct sampling, undersampling, secondary downconversion, and self-multiplication, each with trade-offs in cost, complexity, and data resolution. Experimental results demonstrated the radar's precision in measuring pipe lengths, with deviations of less than 1.25% compared to manual tape measurements. The radar performed well in both straight and curved pipe scenarios, even under challenging conditions like moisture and obstructions. The capability to measure in bent pipes highlights its adaptability and the guiding effect of pipe walls on electromagnetic waves.

1.4. Innovation

In this article, the authors propose a topology that is improved in comparison with the preceding responder version. The development process revolves around the transition from the previously employed homodyne approach to a heterodyne operating principle. This modification allows the received signal to be processed and adjusted at an intermediate frequency before being transmitted back. By leveraging the heterodyne principle, the system will benefit from improved control over signal manipulation. The technological shift to signal manipulation in an intermediate frequency range offers several advantages:

- **Simplified Signal Processing:** Operating at a lower frequency range reduces complexity in signal treatment, enabling easier filtering, amplification, and amplitude control.
- **Enhanced Efficiency:** Reduced signal losses during processing lead to improved overall energy efficiency of the system.
- **Improved Isolation:** The separation between input and output signals is enhanced, minimizing interference and enabling cleaner signal transmission.
- **Greater Component Variety and Flexibility:** The use of intermediate frequencies broadens the range of compatible components, providing more flexibility in system design and implementation.
- **Cost and Size Reduction:** Lower frequency components tend to be smaller and less expensive, contributing to a more compact and cost-effective solution. This is a major issue in environments of limited space availability, such as sewer pipes.
- **Extensibility to Advanced Modulation Techniques:** This approach facilitates the integration of single-sideband modulators, which can further refine the system's performance by reducing spectral bandwidth and improving signal clarity.

This strategic advancement in topology will lead to a focus on simplified signal processing, achieving better system efficiency, and expanding the radar's application potential in challenging environments.

With the focus on the presentation of the modular hardware platform based on the heterodyne approach, two test series in free space and inside sewer pipes are elaborated on, analyzed, and discussed. These measurements to characterize the target performance therefore extend beyond the test of the prototype presented in [15].

2. Materials and Methods

This section outlines the materials and methods used in this work, detailing the system's design and the underlying concepts. The following subsections describe the block diagram, a summary of key mathematical definitions from our preceding publication, system requirements, and the flexibility of the approach in terms of modification and measurement. Finally, the chosen topology is presented in detail.

2.1. Block Diagram and Mathematical Description

A high-level block diagram is provided to give an overview of the system architecture (Figure 2) and the signal flow at the responder (Figure 3). The mathematical description of the system has been previously detailed in a preceding publication [15], to which the reader is referred here to gain a better understanding of the background. Key definitions and principles relevant to this work will be briefly repeated for completeness. The variables used and their meanings are explained in Table 1.

Table 1. List of variables.

A, B, D	amplitudes of high-frequency signals
b_k, m_k	linear equation coefficients
BW	bandwidth of an FMCW ramp
c_0	speed of light
f_l, f_u	frequency of left-hand and right-hand terms in baseband
P	power
$\tilde{P}$	regression of measured power values
R	range
T_{oF}	time of flight (one direction)
ΔT_c	duration of one FMCW ramp
$u_l(t), u_u(t)$	left-hand and right-hand terms in baseband
$y_1(t)$	TX signal of the radar sensor
$y_2(t)$	incident signal at the radar target
$y_1(t)$	TX signal of the radar target
$y_4(t)$	incident signal at the radar sensor
$y_{IF}(t)$	downconverted intermediate signal at the radar responder
$y_{IF}^{[AM]}(t)$	AM modulated intermediate signal at the radar responder
ω_{AM}	identification frequency of the radar responder
ω_{HF}	high frequency oscillation of the radar sensor
ω_{LO}	local oscillator frequency of the radar responder
φ	phase of an oscillation
$\Delta\Phi$	phase modulation caused by doppler shift

The radar system targeting at the responder basically transmits a signal $y_1(t)$, which is described in the time domain as follows:

$$y_1(t) = A \cdot \cos(\omega_{HF}(t) \cdot t + \varphi_{HF}) \quad (1)$$

Considering path losses and the time of flight, the responder receives

$$y_2(t) = B \cdot \cos(\omega_{HF}(t) \cdot (t - T_{oF}) + \varphi_{HF}) \quad (2)$$

In a first step, the received signal is downconverted to an intermediate frequency $f_{IF}(t) = f_{HF}(t) - f_{LO}$. It is then amplitude-modulated with the known envelope frequency f_{AM} , amplified, and finally upconverted again using the same local oscillator frequency f_{LO} .

The goal is to achieve improved carrier suppression of the originally transmitted radar system frequency f_{HF} .

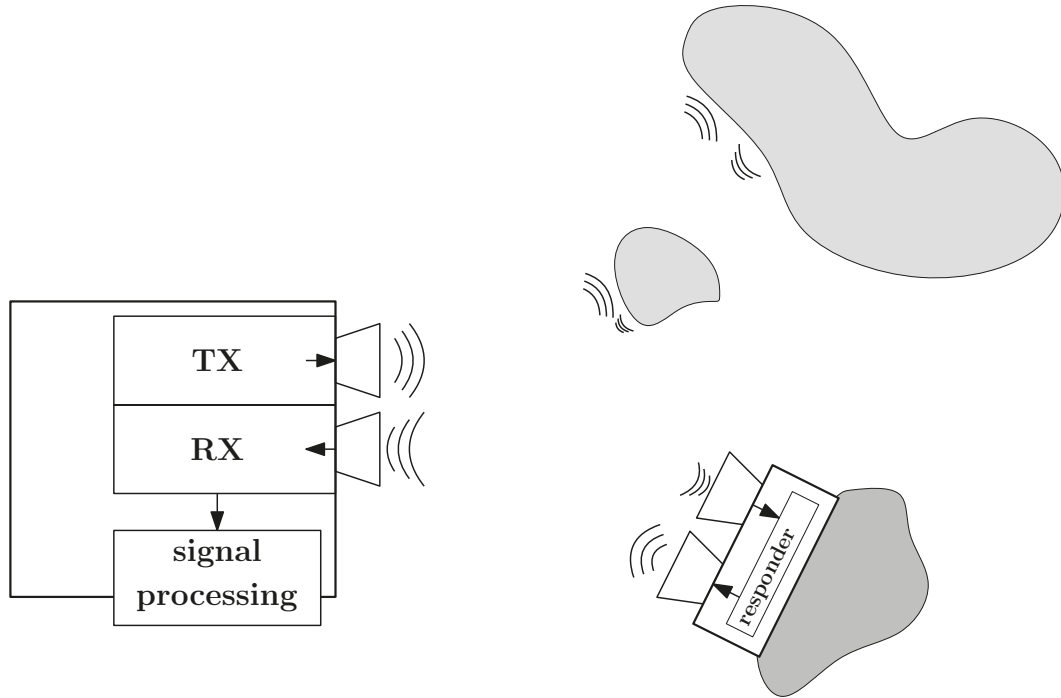


Figure 2. Block diagram of the system.

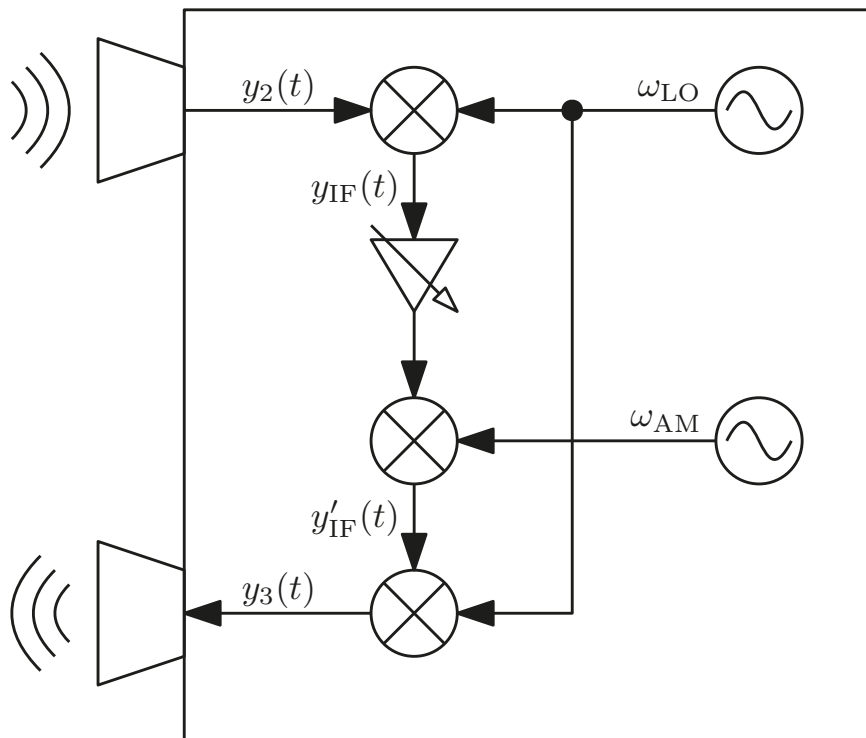


Figure 3. Active radar responder with amplitude modulation based on the heterodyne principle to achieve unambiguity.

Furthermore, the equation already known from the homodyne approach also applies here. By mixing with the oscillator signal and subsequent low-pass filtering, $y_{IF}(t)$ is determined.

$$y_{\text{IF}}(t) = \frac{B}{2} \cdot \cos \left(\omega_{\text{HF}}(t) \cdot (t - T_{\text{oF}}) - \omega_{\text{LO}} \cdot t + (\varphi_{\text{HF}} - \varphi_{\text{LO}}) \right) \quad (3)$$

If no further amplification is applied, the modulation with the envelope now follows:

$$\begin{aligned} y_{\text{IF}}^{[\text{AM}]}(t) = & \frac{B}{4} \cdot \cos \left(\omega_{\text{HF}}(t) \cdot (t - T_{\text{oF}}) + (-\omega_{\text{LO}} + \omega_{\text{AM}}) \cdot t \dots \right. \\ & \left. \dots + (\varphi_{\text{HF}} - \varphi_{\text{LO}} + \varphi_{\text{AM}}) \right) \dots \\ & \dots + \frac{B}{4} \cdot \cos \left(\omega_{\text{HF}}(t) \cdot (t - T_{\text{oF}}) + (-\omega_{\text{LO}} - \omega_{\text{AM}}) \cdot t \dots \right. \\ & \left. \dots + (\varphi_{\text{HF}} - \varphi_{\text{LO}} - \varphi_{\text{AM}}) \right) \end{aligned} \quad (4)$$

Next, $y_{\text{IF}}^{[\text{AM}]}(t)$ is upconverted again using the same local oscillator frequency. However, this leads to a further spreading into four superimposed oscillations. Since the oscillator's phase at the upconverter does not necessarily match the phase at the down-converter, but both phases have a constant difference, this is accounted for as φ'_{LO} . This situation arises when the path lengths from the common signal source to the two mixers are not exactly equal.

The signal transmitted back to the receiver path of the radar sensor at its input is given by

$$y_4(t) = D \cdot \cos(\omega_{\text{HF}}(t) \cdot (t - 2 \cdot T_{\text{oF}}) + \varphi_{\text{HF}}) \quad (5)$$

which is downconverted by multiplication with the signal $y_1(t)$. The classical radar parameters, namely range information and Doppler shift, can be determined through differential analysis of the two reference points, which are symmetrically positioned on either side of the identification oscillator's frequency. The oscillation is described by a left-hand $u_l(t)$ and a right-hand $u_u(t)$ term:

$$u_l(t) = \frac{A D}{4} \cos \left(\omega_{\text{AM}} \cdot t + 2\omega_{\text{HF}}(t) \cdot T_{\text{oF}} - \omega_{\text{AM}} \cdot T_{\text{oF}} + \varphi_{\text{AM}} \right) \quad (6)$$

$$u_u(t) = \frac{A D}{4} \cos \left(\omega_{\text{AM}} \cdot t - 2\omega_{\text{HF}}(t) \cdot T_{\text{oF}} - \omega_{\text{AM}} \cdot T_{\text{oF}} + \varphi_{\text{AM}} \right) \quad (7)$$

Let $\omega_{\text{HF}}(t)$ represent a linear frequency ramp with duration ΔT_c and bandwidth $BW = f(t_0 + \Delta T_c) - f(t_0)$. The range R is then calculated as:

$$R = \frac{c_0 \cdot \Delta T_c}{4 \cdot BW} \cdot (f_u - f_l) \quad (8)$$

Here, f_l and f_u are the two frequency points associated with the oscillations $u_l(t)$ and $u_u(t)$.

The relationship in Equation (8) is largely identical to the classical radar solution but includes a factor of 4 instead of 2 in the denominator. This results from the fully differential consideration of the two frequency points — the left-hand frequency point corresponds, in principle, to the negative reference point in a classical radar, which is now shifted into the positive frequency domain by f_{AM} .

To minimize interference, bandpass filtering is recommended since additional signal components may be expected alongside the desired reference points. Signal acquisition can then proceed, and several different approaches will be examined.

To determine the Doppler information, consider the summands in the oscillation terms of $u_l(t)$ and $u_u(t)$ that depend on T_{oF} . Any change in the target's range affects the signal propagation time T_{oF} , while minimal movements, such as those on the order of the wavelength of $\omega_{\text{HF}}(t)$, will not significantly shift the reference points f_l and f_u . The two Equations (6) and (7) are extended to include the propagation time dependence τ (Doppler effect):

$$T_{oF} \rightarrow T_{oF}(\tau) \quad (9)$$

$$u_l(\tau) = \frac{A D}{4} \cos \left(\omega_{AM} \cdot t + 2\omega_{HF}(t) \cdot T_{oF}(\tau) - \omega_{AM} \cdot T_{oF}(\tau) + \varphi_{AM} \right) \quad (10)$$

$$u_u(\tau) = \frac{A D}{4} \cos \left(\omega_{AM} \cdot t - 2\omega_{HF}(t) \cdot T_{oF}(\tau) - \omega_{AM} \cdot T_{oF}(\tau) + \varphi_{AM} \right) \quad (11)$$

The first and fourth summands are constant in τ , while the third summand affects both terms (10) and (11) equally. Differential analysis leads to the desired relationship for the Doppler effect in the form of phase modulation $\Delta\Phi(\tau)$:

$$\begin{aligned} \Delta\Phi(\tau) &= \Phi_l(\tau) - \Phi_r(\tau) \\ &= 4 \cdot \omega_{HF}(t) \cdot T_{oF}(\tau) \end{aligned} \quad (12)$$

The conclusion is that relative analysis of the two reference points ensures coherence. This allows both range and Doppler information to be determined without any loss in quality or precision.

2.2. System Requirements

The system requirements are outlined, including functional, technical, and environmental considerations essential for achieving the desired performance. From a circuit-design perspective, it is necessary to balance functionality, effort, and complexity. Initially, the requirements for the circuit are defined. This leads to the derivation of a block diagram and the functional segmentation of the components. Iterative minor adjustments may need to be made as required. It is advisable to first select integrated high-frequency circuits and mixers, as the available selection of usable components in this area is typically the most limited. Once the allowable signal levels and the necessary filter characteristics at the inputs and outputs of the central functional blocks are known, appropriate amplifier modules can be selected to complete the signal chain, and the calculation and design of filter circuits can be conducted. The circuit is completed with the necessary peripherals and power supply. Following this, the printed circuit board (PCB) is designed, considering optimal component placement and trace routing.

The requirement catalog is divided into three sections. The *General Requirements* section covers the system level. Subordinate requirements apply, respectively, to the radar base station and the active radar target.

2.2.1. General Requirements

- 1.01 The measurement system determines the distance to one or more active targets.
- 1.02 The range is at least 100 m.
- 1.03 The deviation of the measured distance from the actual distance shall be within ± 5 cm.
- 1.04 An ISM band is used.
- 1.05 The center frequency of the usable frequency range is 24.125 GHz, with a bandwidth of 250 MHz.
- 1.06 The signal levels and waveforms must be designed to ensure that the transmit power does not exceed 20 dBm EIRP.
- 1.07 The distance information of the active targets is transmitted in a frequency range that does not overlap with the frequency range of passive targets within the range of up to 100 m.
- 1.08 The measurement system consists of a PC with an application for data acquisition, a base station as a unit, and one or more active targets as individual units.

- 1.09 Communication between the PC application and the base station is established via a LAN connection.
- 1.10 Measurements are initiated from the PC application. The measurement value is computed in the PC software using the baseband time-domain signal.
- 1.11 The base station generates a suitable radar signal and processes the received signal in the baseband so that it can be digitally evaluated. Additional requirements are defined separately.
- 1.12 An active target modifies a received radar signal so that it can be clearly distinguished from passive reflectors within the range. Additional requirements are defined separately.

2.2.2. Radar Base Station Requirements

- 2.01 There is one transmit channel.
- 2.02 There are at least two receive channels.
- 2.03 The evaluation is based on the data from at least one receive channel.
- 2.04 The antennas are interchangeable. Therefore, all RF ports must be equipped with SMA connectors.
- 2.05 All antennas radiate in the same direction.
- 2.06 The transmit power is adjustable.
- 2.07 Power is supplied by a laboratory power supply, with a voltage range of 4.5 V to 5.5 V.
- 2.08 The design is modular. The following modules are implemented: antenna array (1), RF module (2), ADC module (3), controller module (4), digital interface (5). Modules (1)–(2) and (3)–(5) can each be implemented as combined units.
- 2.09 Module connections are made using plug-in or coaxial (SMA) connectors.
- 2.10 Modules (1) and (2) are designed so that they can be mounted at the end of a pipe with an inner diameter of 150 mm.

2.2.3. Active Radar Target Requirements

- 3.01 The active target is designed using the heterodyne method and employs a double-sideband modulation scheme.
- 3.02 There is one receive channel and one transmit channel.
- 3.03 All antennas radiate in the same direction.
- 3.04 The transmit power is adjustable.
- 3.05 Power is supplied by a laboratory power supply or a battery (voltage range 10 V to 16 V). The input must include protection against undervoltage, overvoltage, and reverse polarity.
- 3.06 The design is modular. The following modules are implemented: antenna array (1), RF module (2), reconfigurable IF module (3), power supply (4).
- 3.07 Module connections are made using plug-in or coaxial (SMA) connectors.
- 3.08 The IF module (3) consists of a carrier PCB with at least three slots. The central slot is occupied by a mixer module including an AM oscillator, and the remaining slots can optionally be equipped with a filter, amplifier, attenuator, or a combination of these.
- 3.09 Modules (1) and (2) are designed so that they can be mounted at the end of a pipe with an inner diameter of 150 mm.

2.3. Presentation of the Chosen Topology

The chosen topology is presented, highlighting the rationale behind its selection and the advantages it offers for this specific application.

2.3.1. Software Flowchart

The flow diagram is illustrated in Figure 4. The base station acts as the server, while the application serves as the client. Measurement initiation is triggered by the PC application. After successfully establishing a connection to the base station (specifically, the microcontroller), hardware configuration and calibration are initiated. Once completed, the application can trigger a new measurement. Following the request, it waits for the complete transmission of raw data. To maintain the required timing intervals, the entire process up to data transfer is controlled by the microcontroller.

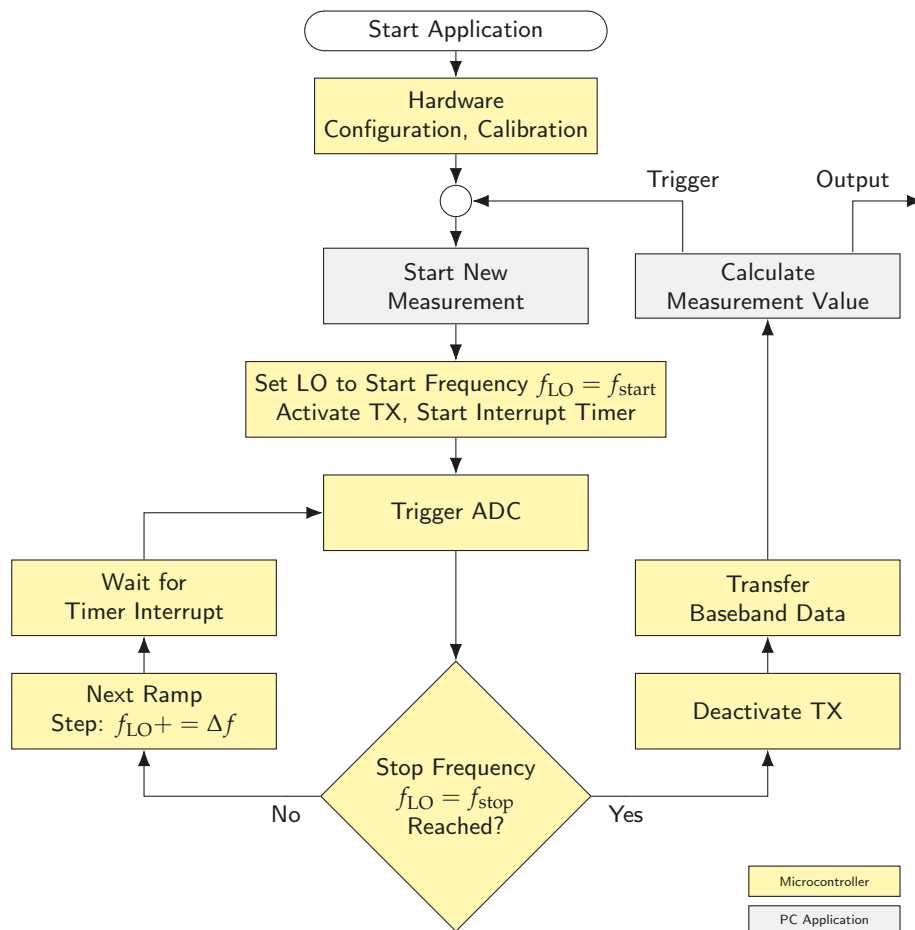


Figure 4. Workflow and interaction between PC application and base station.

Starting with a linearly increasing ramp, the local oscillator frequency is set to the lowest frequency f_{start} within the available band. The duration of each ramp step is controlled by a timer, which specifies the constant length of each step using an interrupt routine. At the end of each timer interval, a new analog-to-digital conversion is triggered. Immediately afterward, the local oscillator is adjusted to the frequency of the next ramp step (frequency increment Δf). The cycle restarts with the timer overflow. This sequence (conversion, frequency step, time synchronization) ensures a constant interval, with the conversion timing shifted to the end of each step to account for oscillator settling time. The voltage value captured by the ADC is stored in the controller's RAM. Once the ramp's upper frequency f_{stop} is reached, the described process is interrupted, and the transmitter is turned off. All data points are transferred to the application, forming the baseband time-domain signal. Based on this, the desired distance information is calculated by the application, and the measurement result is displayed. Subsequently, a new measurement can be initiated, making execution in a constant time interval reasonable.

2.3.2. Radar Frontend

A standard radar frontend should be used if possible. Our implementation is given in Figure 5. Therefore, when specifying the parameters for amplitude modulation on the responder side, its frequency f_{AM} must be located within the passband of the receiver's mixer stage.

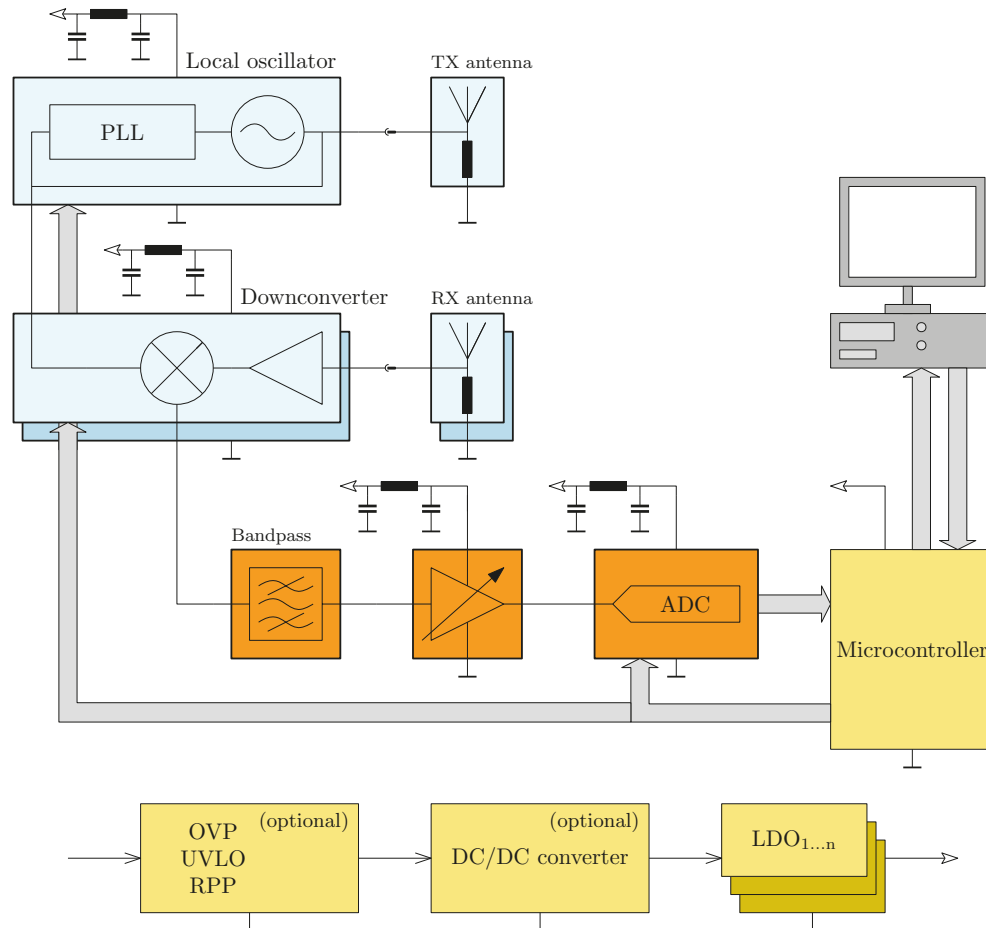


Figure 5. Block diagram of the radar frontend.

This consists of a PLL-controlled local oscillator, which feeds both the TX antenna and the mixer of the downconverter in the receiver path as another essential component. The antenna feeds for the TX and RX antennas are routed to SMA connectors to allow for testing various antenna configurations.

The amplitude modulation by the responder, as described, must also be taken into account in the analog-to-digital conversion. Four applicable methods for this are described in [15]. In addition to direct conversion, various methods are available that offer the advantage of reducing the bandwidth of the digitized signal. Omitting data that do not provide added value shortens the processing time required to analyze the raw data. The implementation of these methods will be discussed in detail later.

The third core component is a microcontroller, which ensures the previously specified process flow.

2.3.3. Active Responder

Figure 6 shows the complete overview schematic of the active responder based on the heterodyne method. The signal path begins with an antenna, where the use of a planar antenna array is intended. A low-noise small-signal amplifier is placed before the first mixer for pre-amplification. The gain of this amplifier should not be too high, as mixers generally

operate with less distortion at lower signal levels. The mixer's role is to convert the signal to the intermediate frequency range. With sufficient separation between the high-frequency and intermediate-frequency bands, interference can be filtered more easily, but the power transferred into unwanted frequency ranges must be accounted for during design.

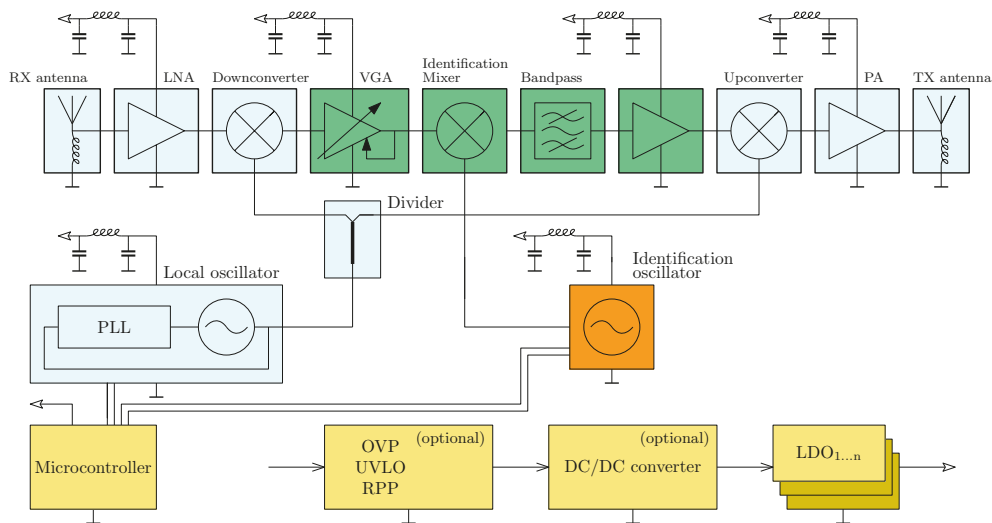


Figure 6. Complete block diagram of the active radar responder.

The local oscillator signal is generated using a phase-locked loop (PLL) to achieve a stable signal. Stability is desirable for both the phase and amplitude. Parameterization is performed via a microcontroller. The LO signal is coupled through an output of a power divider to the first mixer stage. Impedance transformation is necessary since the upconverting stage is also fed by the local oscillator, resulting in both mixers being connected in parallel from the LO's perspective. Power leveling is performed in the intermediate frequency range, as control loops are generally easier to integrate at lower frequencies. This applies both to the expected availability of suitable components and to the extraction of the signal to be measured.

The control system typically consists of a variable, voltage-controlled attenuator, a subsequent amplifier stage, and a power detector that provides the control voltage for the attenuator. In this case, the control loop is fully analog.

As a result, the input signal for the subsequent second mixer stage, which performs the amplitude modulation and provides the desired additional information, is less subject to strong fluctuations. Even a free-running oscillator suffices as the identification oscillator since the distance information in FMCW measurements is determined from the difference between two points symmetrically distributed around f_{AM} .

Since harmonic generation is expected, both from the IF input signal and the AM signal along with their cross-products, special attention must be paid to the operating point of this mixer. Narrowband filtering at the output removes unwanted components. The passband must accommodate the bandwidth of the RF signal, which in our case is 250 MHz. The stopband and cutoff are defined by the choice of the intermediate frequency band. Filtering reduces the bandwidth at the input of the subsequent amplifier stage, minimizing the consumption of unnecessary power reserves. This stage should be designed to maximize the drive level of the third mixer stage, minimizing the need for inefficient RF amplification. The transmitting antenna is DC-coupled to ground potential.

2.4. Block-Based Hardware Setup

A flexible approach has been adopted, allowing for modifications and measurements at various points in the signal path. This ensures adaptability and facilitates a deeper analysis of the system.

The radar system used was developed in-house, based on the previously introduced MMIC from Analog Devices. For high-frequency generation, the PLL formed with *ADF5901* and *ADF4159* was used, with *ADF5904* as the corresponding downconverter. Its basic design has been presented in [16]. This setup provides two TX channels and four RX channels, which are accessed via coaxial connectors, although only one channel is utilized for each. The board-to-connector transition is optimized according to [17,18]. This allows for the usage of various antenna types to handle radiation patterns, polarization as well as footprint size. Standard gain horn antennas and patch arrays as per [19] are intended for use. The remaining RF channels are disabled. The baseband is supplemented with instrumentation amplifiers, allowing the output signal from each downconverter to be accessed both differentially and single-ended. However, due to the differential input of the ADC used, the instrumentation amplifiers are not functional. The modular design results in the PCB stack shown in Figure 7, with the RF frontend on top, the ADC in the middle, and the microcontroller evaluation board mounted at the bottom on the carrier plate.

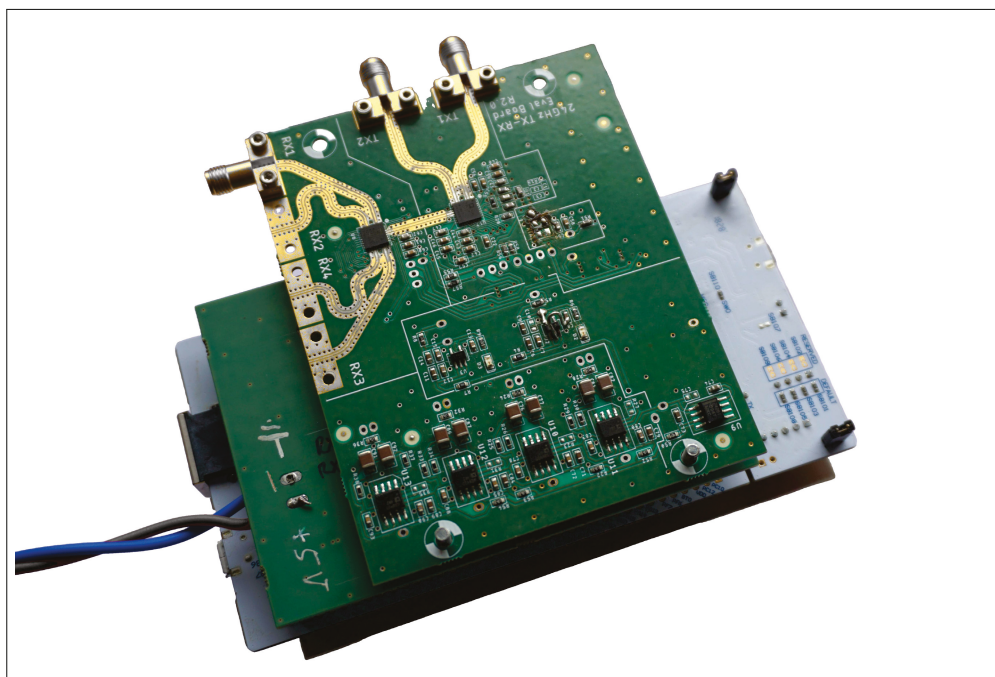


Figure 7. Fully assembled radar base station with RF frontend on ADC and microcontroller boards.

The ADC module contains the previously selected analog-to-digital converter of type *AD7264* in a single configuration, allowing a maximum of two baseband channels to be sampled. The stack order was chosen to minimize the trace lengths for the baseband and SPI bus between the microcontroller and ADC, as much as possible in a modular setup.

Figure 8 shows the RF module of the responder. The antennas are connected via the side-mounted coaxial couplers. This module performs RF amplification and mixing with the local oscillator signal, which is also generated on this board. At the intermediate frequency level, the downconverted signal is extracted, and the upconverted signal is injected on the backside via SMA coaxial connectors.

Figure 9 depicts the module carrier designed to accommodate up to four plug-in modules. The plug-in modules can be connected at all positions using SMA cables. Power

is supplied via linear regulators located on the carrier board. In addition to a filter module to limit the bandwidth to the designated intermediate frequency range of 500–750 MHz, four additional components for signal amplification and attenuation were fabricated. The mixer used was tested with a plug-in LC oscillator.

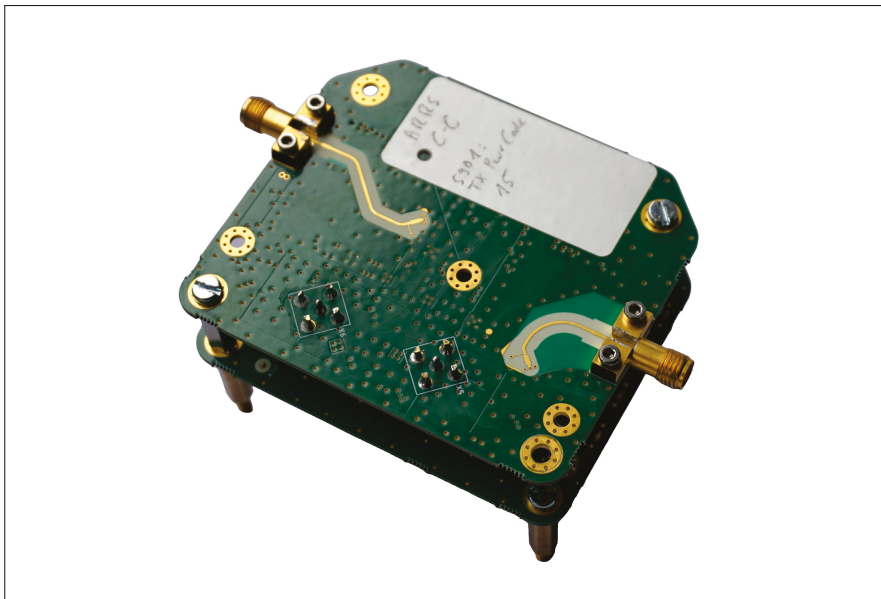


Figure 8. RF module of the responder, with the secondary board containing the switching power supply and protection circuitry below.

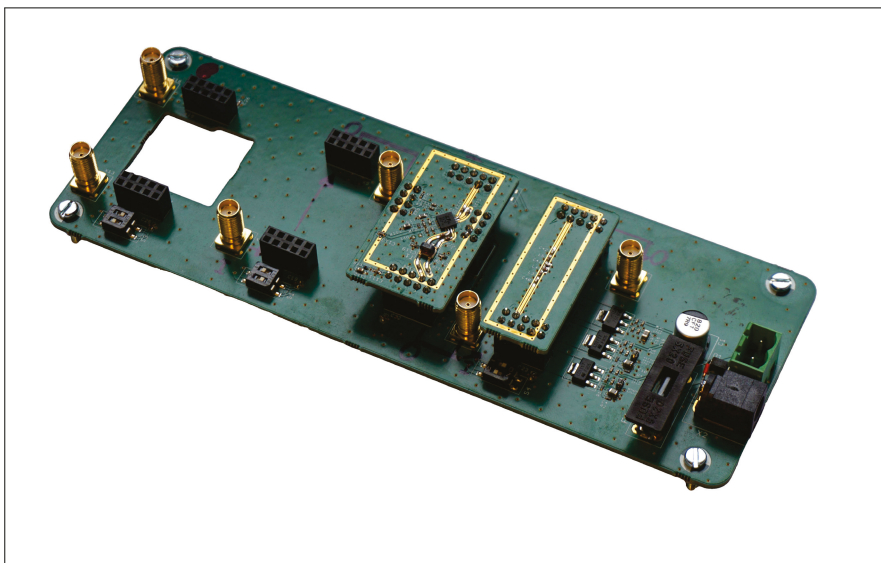


Figure 9. Intermediate frequency carrier module of the active responder for up to four plug-in modules, shown here with mixer and bandpass filter installed.

3. Results

Two measurements were planned and carried out for the proof of concept. These should take into account the conceivable areas of application, so that the first setup was conducted in an outdoor environment as a more general measurement. This was followed by a further series of tests that took into account a specific area of application, namely by being carried out within a section of sewage pipe. This addresses the performance inside an area that is prone to clutter.

3.1. First Experiment: Free-Space Measurement

The first part of the experiment involved conducting free-space measurements. The goal of this measurement was to provide a general proof of functionality over distances ranging from approximately 10 to 100 m. Additionally, the relative accuracy of measurement points at various distances and the repeatability of the measurements were to be determined.

The corresponding block diagram is shown in Figure 10. According to this diagram, a radar base station with a PC for control and live evaluation, along with the active responder, supplemented by a spectrum analyzer, were set up on separate movable tables. The spectrum analyzer was fed by a signal directly tapped at the output of the receiving antenna via a 1:1 power splitter.

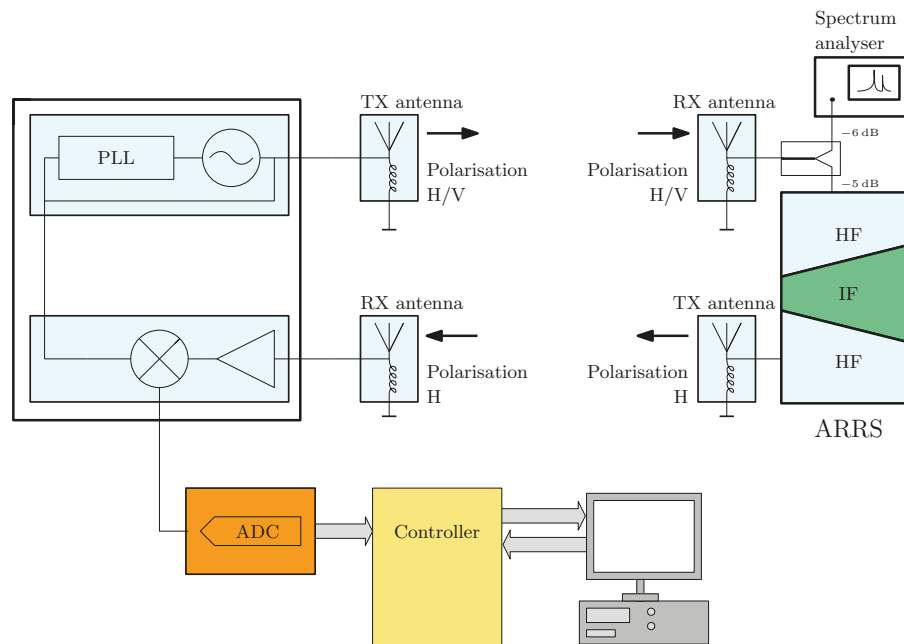


Figure 10. Block diagram of the experimental setup.

The specified distance was determined between the positions of the wheels of each table, which served as the reference plane. This approach was acceptable because the radar measurements inevitably exhibited a constant offset due to propagation delays caused by:

- The relative positions of the antenna reference planes;
- Cable connections between individual modules;
- The spatial dimensions of the electronics;
- Predefined zero points in the evaluation software.

All these factors remained constant during the experiment and resulted in a deviation that could be easily calibrated out.

The first measurement in the series was conducted at a distance of 1 m. Beyond 2 m, the distance was increased in 2 m steps. Additional measurements were performed at intermediate distances under special conditions. Since mains power was required at both tables, the experiment had to be interrupted at longer distances due to the limited length of power cables. Whenever an interruption occurred, the last measurement was repeated with unchanged component arrangements for verification.

The primary goal was to measure distances up to 100 m. However, due to the circumstances, additional measurements were conducted at distances up to 150 m.

At shorter distances, a characteristic signal shape in the frequency domain was observed. As expected, two prominent bins appeared sharply defined near the modulation frequency, tapering to noise towards the spectrum's edges. For distances up to approximately 40 m, a characteristic “inverted bathtub curve” shape was visible in the spectrum. This curve, defined by steep gradients towards the symmetry axis and a curved drop-off to surrounding noise, contrasts with the typical bathtub curve where decaying edges are directed towards the center.

Beyond 41 m, a 2 m region of destructive interference was observed, causing the signal to disappear between 41 m and 43 m. Similarly, no evaluable signal was detected between 54 m and 62 m.

A cross-check of the spectrum of the baseband waveform showed that the signal flanks flattened for that certain range, forming a curved transition to noise at the inner edges, while the outer edges remained qualitatively unchanged. This made triggering to the correct threshold more difficult. Signal quality improved at greater distances, with the opposing edges of the spectral peaks becoming more steep and overall signal levels increasing. (Note that the modulation at the target shifts both range bins at the positive and negative frequencies by the modulation frequency, so the distance information is found in the relative position of these bins). A drop in signal quality was observed again between approximately 114 m and 118 m, roughly double the previously unevaluable range near 58 m, as shown in Figure 11.

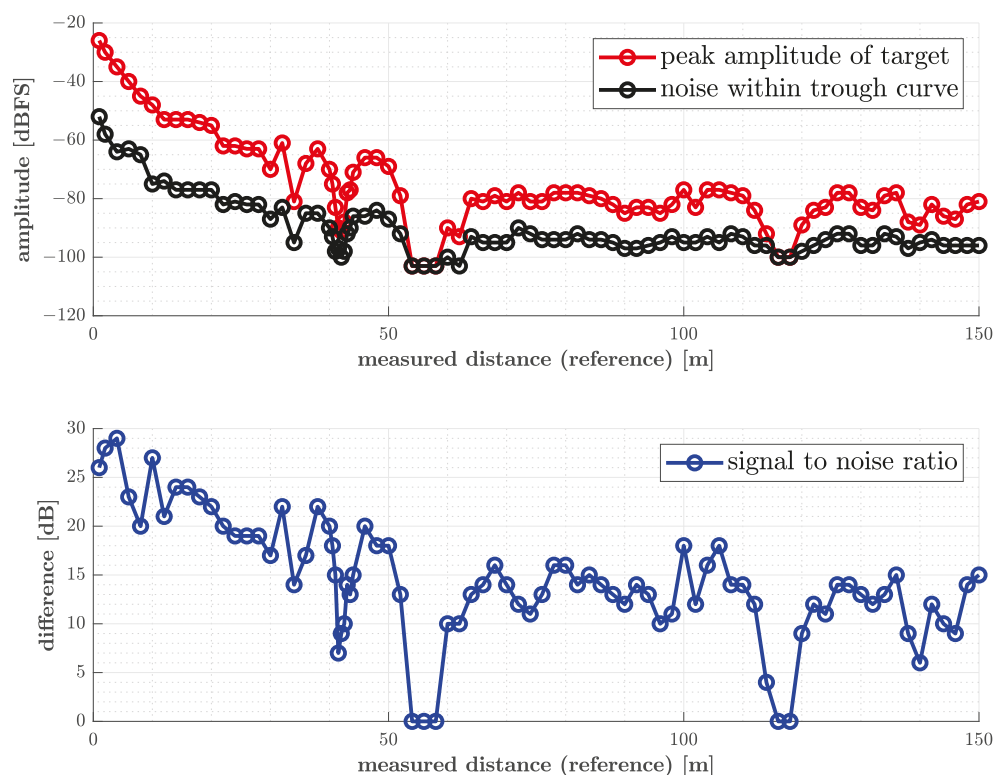


Figure 11. Summary of results from free-space measurements with orthogonal antenna polarization.

In order to examine the system's precision at small distance changes, the position of the responder unit was slightly varied in steps of approximately 1 cm. This was carried out in the setup of the first experiment, at distances of about 5 m and 10 m between the radar and the responder. The result is given in Figure 12, which indicates that the tolerance limits specified by the system requirements are exceeded in the test at 10 m.

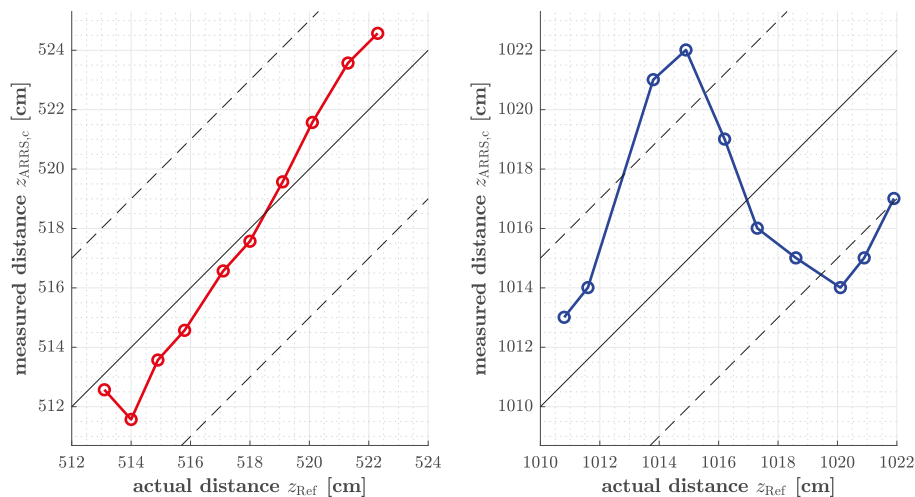


Figure 12. Measured vs. actual distance for small-range variations.

3.2. Second Experiment: Measurement in PVC Pipes

The subsequent experiment investigated the applicability of the method inside water pipes. The available pipes are typically used to transport various types of wastewater, such as rainwater or sewage from residential areas. These pipes are generally buried underground during installation. For channel base pipes, the material PVC-U (unplasticized polyvinyl chloride) is frequently used. These extruded pipe sections have a multi-layer structure, with a foamed core surrounded by outer and inner layers, as per DIN EN 1401-1:2019-09 in the version valid at the time of the measurements [20].

The incident signal in the investigation of propagation within a pipe was tapped at the responder and observed using a spectrum analyzer. Additionally, the power in the baseband was determined based on the signal amplitude.

The raw data indicate a linear dependency of power in the logarithmic scale on distance in linear scaling. Consequently, it was decided to perform linearization immediately after logarithmizing the power. This relationship is clearly illustrated in the visualization of the measured data in Figure 13, which includes the regression.

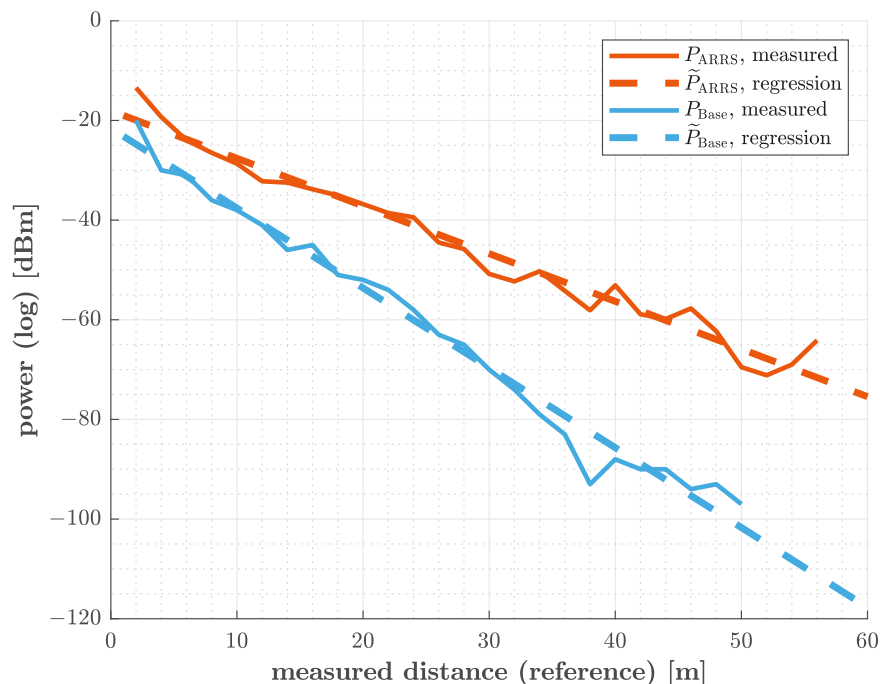


Figure 13. Regression of power during one-way and two-way propagation in the plastic pipe.

As will be pointed out later in the discussion, this observation contradicts the Friis equation. In contrast to the result of the measurement, a distance dependency to the fourth power would normally be assumed for radar systems. Although the logarithmic decrease in power practically means stronger baseband signal amplitudes in the range of up to a few meters, it will drop significantly faster at greater distances.

4. Discussion

For the free-space test results, the measured carrier power values as well as distance-dependent power values determined at the base station were subjected to regression analysis. Since the power values were given on a logarithmic scale, the corresponding distance values were also logarithmized to base 10. In this double-logarithmic representation, the result was a straight line whose slope corresponded to the power of the distance. Thus, linear regression could be applied:

$$y \sim \log_{10}(P(R)) \sim -\log_{10}(R) \quad (13)$$

For power in dBm, the distance dependence can generally be expressed as:

$$P(R) = -m_k \cdot \log_{10}\left(\frac{R}{1 \text{ m}}\right) \cdot 10 \text{ dBm} + b_k \cdot 1 \text{ dB} \quad (14)$$

In free-space propagation, the coefficient m_k , which represents the distance exponent, is theoretically expected to be:

$$\begin{aligned} \text{One-way propagation : } & m_1 = 2 \\ \text{Two-way propagation : } & m_2 = 4 \end{aligned} \quad (15)$$

The parameter b_k depends on various factors, such as transmitted power, antenna gain, cable losses, and coupling to the measurement instrument. Its value plays a minor role and does not contribute to the evaluation of the transmission channel. The propagation type is indicated by the index k .

For the measured values, applying regression yields the following results for one-way and two-way propagation, denoted as $\tilde{P}$:

$$\begin{aligned} \text{One-way propagation : } & m_1 = 1.585 \quad ; \quad b_1 = -20.13 \\ \text{Two-way propagation : } & m_2 = 3.056 \quad ; \quad b_2 = -22.35 \end{aligned} \quad (16)$$

Using logarithmic properties, the input power at the responder can be approximated as:

$$\tilde{P}_{\text{ARRS}}(R) = \log_{10}\left(\frac{1}{\left(\frac{R}{1 \text{ m}}\right)^{1.585}}\right) \cdot 10 \text{ dBm} - 20.13 \text{ dB} \quad (17)$$

The output of the radar sensor is described as:

$$\tilde{P}_{\text{Base}}(R) = \log_{10}\left(\frac{1}{\left(\frac{R}{1 \text{ m}}\right)^{3.056}}\right) \cdot 10 \text{ dBm} - 22.35 \text{ dB} \quad (18)$$

The drop in signal amplitude at certain ranges should be addressed as well. The second undetectable range is at about twice the distance of the first, so that a periodic behavior could be present. The next interval would be outside the selected measurement window. A likely explanation would be multipath propagation, which leads to destructive

interference at certain distances. To confirm this, it would be conceivable to conduct another series of measurements in a different environment.

For the test series executed inside pipes, as before, the formal relationship between power in the logarithmic scale and a linear equation is established:

$$\begin{aligned}\tilde{P}(R) \cdot 1 \text{ dBm} &= \log_{10} \left(\frac{P(R)}{1 \text{ mW}} \right) \cdot 10 \text{ dBm} \\ &= m_k \cdot \frac{R}{1 \text{ m}} \cdot 10 \text{ dBm} + b_k \cdot 1 \text{ dB}\end{aligned}\quad (19)$$

From the measured data at the responder and radar sensor, the following coefficients are determined:

$$\begin{aligned}\text{One-way propagation :} \quad m_1 &= -0.09566 \quad ; \quad b_1 = -18.04 \\ \text{Two-way propagation :} \quad m_2 &= -0.1603 \quad ; \quad b_2 = -21.57\end{aligned}\quad (20)$$

Thus, after regression, the dependence of power $\tilde{P}$ in the linear representation on distance R as part of the exponent to base ten is expressed as follows for the input of the responder:

$$\begin{aligned}\tilde{P}_{\text{ARRS}}(R) &= 10^{-0.09566 \cdot \frac{R}{1 \text{ m}}} \cdot 10^{-1.804} \cdot 1 \text{ mW} \\ &= 10^{-0.09566 \cdot R/1 \text{ m}} \cdot 0.0157 \text{ mW}\end{aligned}\quad (21)$$

and for two-way propagation as measured at the base station:

$$\begin{aligned}\tilde{P}_{\text{Base}}(R) &= 10^{-0.1603 \cdot \frac{R}{1 \text{ m}}} \cdot 10^{-2.157} \cdot 1 \text{ mW} \\ &= 10^{-0.1603 \cdot R/1 \text{ m}} \cdot 0.007 \text{ mW}\end{aligned}\quad (22)$$

Based on the attenuation characteristics determined, it is reasonable to assume that there is a constant outflow of power per unit length from the inside of the pipe. The comparably poor high-frequency properties of polyvinyl chloride result in high attenuation within the wall. As the entire section is made up of individual plug-in elements, each 2 m long, there is also a discontinuity at each overlapping sleeve. The wall is interrupted there and the otherwise smooth inner surface is curled, which results in regular interference points. So far, it has not yet been investigated whether there is a propagation mode within the wall, which is quite possible given the operating frequency of the radar system and the wall thickness.

To focus on the accuracy of the sensor at small distance increments in the setup of the first experiment, the data analysis indicates deviations from the actual distance. The first data record was acquired at about 5 m, the deviations from the actual distance reached up to 3 cm, which is acceptable with regard to the system requirements. However, at a range of 10 m, deviations increased, with some readings differing by as much as 7 cm. Despite this, the goal was to achieve a deviation of less than 5 cm, which was not fully met in this case. The results highlight the challenges of achieving the accuracy required with the given signal processing. Since the focus of this work was expressly not in the data-processing domain, its optimization should be clearly addressed in future works. Provided the frequency band and bandwidth are not to be changed, this will be a necessary step to achieve a significant improvement of absolute tolerances.

5. Conclusions

In this work, the *Active Radar Interferometer (AcRaIn)* was presented as an innovative method in secondary radar technology. Overall, a compact design is possible, as the peripheral components around the technological core are minimized: the processing of supply voltages and the one-time parameterization of the integrated control circuits during startup.

As preparation for the experimental setup, several preliminary considerations were made. After defining the technical requirements, a block diagram was derived that takes advantage of the heterodyne method. The operating frequency of the radar was chosen to be in the ISM band at $f_{c,HF} = 24.125$ GHz with a bandwidth of $BW = 250$ MHz.

The functionality of the entire system was subsequently confirmed in various measurement series. Several scenarios were tested in free space and in various above-ground pipe sections made from assembled plastic pipe pieces. In free space, an outstanding result was achieved, with a clear signal being evaluated even at a distance of 150 m between the radar sensor and responder. In another series, measurements were conducted with respect to a potential application area, specifically the use within a water pipeline, where the distance to a reference point at the access shaft was to be measured.

As a result of these investigations, in addition to the implementation of efficient signal processing, the effect of logarithmic signal attenuation should be questioned and explained from the perspective of wave theory.

Author Contributions: Conceptualization, S.M. and A.R.D.; methodology, S.M.; software, S.M.; validation, A.R.D. and G.F.; formal analysis, S.M. and A.R.D.; investigation, S.M.; resources, S.M. and A.R.D.; data curation, S.M. and A.R.D.; writing—original draft preparation, S.M.; writing—review and editing, A.R.D. and G.F.; visualization, S.M.; supervision, A.R.D. and G.F.; project administration, A.R.D.; funding acquisition, A.R.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Europäischer Fonds für regionale Entwicklung grant number 84002593.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Any research data not explicitly referenced or included in this publication is subject to a nondisclosure agreement and, as such, cannot be published or shared without prior authorization. All data covered under this agreement remains confidential and is protected from public dissemination.

Acknowledgments: The authors would like to thank Klaus Stöß for the manufacturing of printed circuit boards, and Jonas Berg for the assistance during data recording.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

AcRaIn	Active Radar Interferometer
ADC	Analog-to-digital converter
AM	Amplitude modulation
ARRS	Active radar response system
EIRP	Equivalent isotropically radiated power
FMCW	Frequency modulated continuous wave

HF	High frequency
IF	Intermediate frequency
ISM	Industrial, scientific, and medical
LAN	Local area network
LO	Local oscillator
MMIC	Monolithic microwave integrated circuit
PCB	Printed circuit board
PLL	Phase-locked loop
RAM	Random access memory
RF	Radio frequency
RX	Receiver
SMA	Subminiature type A
TX	Transmitter

References

1. *IEEE Std 686-2017 (Revision of IEEE Std 686-2008)*; IEEE Standard for Radar Definitions. IEEE: Piscataway, NJ, USA, 2017; pp. 1–54. [CrossRef]
2. Sharp, E.; Diab, M. Van Atta reflector array. *IRE Trans. Antennas Propag.* **1960**, *8*, 436–438. [CrossRef]
3. Thornton, J.; Edwards, D. Modulating retro-reflector as a passive radar transponder. *Electron. Lett.* **1998**, *34*, 1880–1881. [CrossRef]
4. Thornton, J.; Edwards, D. Range measurement using modulated retro-reflectors in FM radar system. *IEEE Microw. Guid. Wave Lett.* **2000**, *10*, 380–382. [CrossRef]
5. Koelle, A.; Depp, S. Doppler radar with cooperative target measures to zero velocity and senses the direction of motion. *Proc. IEEE* **1977**, *65*, 492–493. [CrossRef]
6. Kossel, M.; Kung, R.; Benedickter, H.; Biichtokd, W. An active tagging system using circular-polarization modulation. *IEEE Trans. Microw. Theory Tech.* **1999**, *47*, 2242–2248. [CrossRef]
7. Cantu, H.; Fusco, V. A 21 GHz reflection amplifier MMIC for retro-directive antenna and RFID applications. In Proceedings of the IET Seminar on MM-Wave Products and Technologies, London, UK, 16 November 2006; pp. 66–70.
8. Cantú, H.; Fusco, V.; Simms, S. Microwave reflection amplifier for detection and tagging applications. *IET Microwaves Antennas Propag.* **2008**, *2*, 115–119. [CrossRef]
9. Dadash, M.S.; Hasch, J.; Chevalier, P.; Cathelin, A.; Voinigescu, S.P. A W-band active millimeter-wave tag IC with wake-up function. In Proceedings of the 2017 IEEE MTT-S International Microwave Symposium (IMS), Honolulu, HI, USA, 4–9 June 2017; pp. 1531–1534. [CrossRef]
10. Dadash, M.S.; Hasch, J.; Voinigescu, S.P. A 77-GHz active millimeter-wave reflector for FMCW radar. In Proceedings of the 2017 IEEE Radio Frequency Integrated Circuits Symposium (RFIC), Honolulu, HI, USA, 4–6 June 2017; pp. 312–315. [CrossRef]
11. Vossiek, M.; Gulden, P. The Switched Injection-Locked Oscillator: A Novel Versatile Concept for Wireless Transponder and Localization Systems. *IEEE Trans. Microw. Theory Tech.* **2008**, *56*, 859–866. [CrossRef]
12. Roehr, S.; Gulden, P.; Vossiek, M. Precise Distance and Velocity Measurement for Real Time Locating in Multipath Environments Using a Frequency-Modulated Continuous-Wave Secondary Radar Approach. *IEEE Trans. Microw. Theory Tech.* **2008**, *56*, 2329–2339. [CrossRef]
13. Shmakov, D.; Ebel, R.; Vossiek, M. Wireless sensor network with 24 GHz local positioning transceiver. In Proceedings of the 2012 IEEE/MTT-S International Microwave Symposium Digest, Montreal, QC, Canada, 17–22 June 2012; pp. 1–3. [CrossRef]
14. Stelzer, A.; Jahn, M.; Scheibhofer, S. Precise distance measurement with cooperative FMCW radar units. In Proceedings of the 2008 IEEE Radio and Wireless Symposium, Orlando, FL, USA, 22–24 January 2008; pp. 771–774. [CrossRef]
15. Müller, S.; Diewald, A.R. Cooperative radar with signature method for unambiguity. *Adv. Radio Sci.* **2019**, *17*, 27–34. [CrossRef]
16. Steins, M.; Leuck, S.; Diewald, A.R. Universal programmable low cost signal source for 24 GHz ISM band. In Proceedings of the 2016 Loughborough Antennas & Propagation Conference (LAPC), Loughborough, UK, 14–15 November 2016; pp. 1–4. [CrossRef]
17. Müller, S.; Diewald, A.R. Methods of connector S-parameter extraction depending on broadband measurements of symmetrical structures. In Proceedings of the 2015 Loughborough Antennas & Propagation Conference (LAPC), Loughborough, UK, 2–3 November 2015; pp. 1–4. [CrossRef]
18. Diewald, A.R.; Müller, S.; Olk, A. PCB-side matching networks for coaxial connectors. In Proceedings of the 2017 Antenna Measurement Techniques Association Symposium (AMTA), Atlanta, GA, USA, 15–20 October 2017; pp. 1–5. [CrossRef]

19. Diewald, A.R.; Müller, S. Design of a six-patch series-fed traveling-wave antenna. In Proceedings of the 2019 IEEE-APS Topical Conference on Antennas and Propagation in Wireless Communications (APWC), Granada, Spain, 9–13 September 2019; pp. 395–399. [CrossRef]
20. German version EN 1401-1:2019-09; withdrawn standard, succeeded by DIN EN 1401-1:2023-11; Plastics Piping Systems for Non-Pressure Underground Drainage and Sewerage—Unplasticized Poly(Vinyl Chloride) (PVC-U)—Part 1: Specifications for Pipes, Fittings and the System. DIN Media GmbH: Berlin, Germany, 2019. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Adversarial Range Gate Pull-Off Jamming Against Tracking Radar

Yuanhang Wang ^{1,*}, Yi Han ² and Yi Jiang ¹

¹ School of Information and Communication Engineering, Dalian University of Technology, Dalian 116024, China; dlutjy@mail.dlut.edu.cn

² School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China; 201911012002@std.uestc.edu.cn

* Correspondence: yuanhang@dlut.edu.cn

Abstract: Range gate pull-off (RGPO) jamming is an effective method for track deception aimed at radar systems. Nevertheless, enhancing the effectiveness of the jamming strategy continues to pose challenges, restricting the RGPO jamming method from achieving its maximum potential. This paper focuses on addressing the problem of optimizing the strategy for white-box RGPO jamming, serving as a foundational step toward quantitative optimization research on RGPO jamming strategies. In the white-box scenario, it is presumed that the jammer has full knowledge of the target radar's tracking system, encompassing both the choice of tracking method and its parameter configurations. The intricate interactions between the jammer and the tracking radar introduce three primary challenges: (1) Formulating an algebraic expression for the objective function of the jamming strategy optimization is nontrivial; (2) Direct observation of jamming effects from the target radar is challenging; (3) Noise renders the jamming outcomes unpredictable. To tackle these challenges, this study formulates the optimization of the RGPO jamming strategy as an adversarial stochastic simulation optimization (ASSO) problem and introduces a novel solution for the white-box RGPO jamming strategy optimization: a local simulation-assisted particle swarm optimization algorithm with an equal resampling scheme (PSO-ER). The PSO-ER algorithm searches for optimal jamming strategies while utilizing a localized simulation of the tracking radar to evaluate the effectiveness of candidate jamming strategies. Experiments conducted on four benchmark cases confirm that the proposed approach is capable of generating well-tuned strategies for white-box RGPO jamming.

Keywords: range gate pull-off (RGPO); radar tracking system; adversarial stochastic simulation optimization (ASSO); particle swarm optimization with equal resampling scheme (PSO-ER)

1. Introduction

Radar deception remains a focal area of research within the field of electronic countermeasures (ECMs) [1–7]. Among various techniques, range gate pull-off (RGPO) jamming is regarded as a highly effective approach for deceiving radar tracking systems [5–9]. However, designing an appropriate RGPO jamming strategy is crucial for achieving successful track deception against the threat radar [3]. During each jamming stage (i.e., the tracking frame), the jammer begins by capturing the radar's transmitted signal and then retransmitting it back to the radar receiver with a deliberately manipulated delay. The retransmitted signal (i.e., the jamming signal) will blend into the radar's received signal, resulting in the radar detecting a false target. When the delay is carefully adjusted for each frame,

the false target's measurement enters the tracking gate, leading to errors in the radar's tracking of the real target. Moreover, when delays across successive tracking frames are systematically coordinated, the tracking error rapidly accumulates. Consequently, the tracking gate diverges significantly from the real target's position. Ultimately, this causes the radar to incorrectly follow the false target, losing track of the real one entirely. On the contrary, if the delays of all jamming stage (i.e., RGPO jamming strategy) are set arbitrarily, the jamming signal can only form false measurement points separated from each other among frames at the radar tracking end. In this case, it is almost impossible to deceive the threat radar to form a false track and lose the real target. Therefore, it is necessary to study the setting of the RGPO jamming strategy.

In recent years, certain scholars have conducted qualitative investigations into the optimization challenges associated with RGPO jamming strategies [3]. A frequently adopted approach involves formulating an algebraic optimization model and applying corresponding algorithms to derive the best RGPO jamming strategy. Nonetheless, the intricate interactions between the jammer and the threat radar make it challenging to formulate an algebraic representation of the objective function. In [10], we preliminarily study data-driven optimization of RGPO jamming strategies. However, the main assumptions in [10] are simple and ideal, and problems such as data acquisition and the limitations of the optimizer in actual use are not fully considered. As a result, the research on the RGPO jamming strategy optimization yields weak results. In practice, uniform velocity RGPO (UV-RGPO) and uniform acceleration RGPO (UA-RGPO) are two of the most popular ways to set the RGPO jamming strategy [6,7]. For the UV-RGPO, the delays are controlled so that the false target is separated from the real target with uniform velocity motion. Similarly, for the UA-RGPO method, the delays are controlled so that the false target is separated from the real target with uniform acceleration motion. Nonetheless, the setting of the key parameters, such as pull-off velocity for the UV-RGPO, initial pull-off velocity, and pull-off acceleration for the UA-RGPO, still relies on expert experience. Consequently, the capability of RGPO jamming as an effective track-deception technique remains significantly underutilized.

The strategy optimization of the RGPO jamming faces three main difficulties: **(1)** Because of the intricate, multi-stage confrontation between the jammer and the tracking radar, it is difficult to obtain the algebraic description of the objective function; **(2)** Due to the adversarial relationship between the jammer and the tracking radar, it is difficult for the jammer to observe jamming results directly from the threat radar; **(3)** The presence of noise, including process and measurement noise, renders the jamming outcomes indeterministic. For such complex optimization problems, stochastic simulation optimization (SSO) technology may be a powerful tool [11–15]. In general, algebraic descriptions of SSO problems are difficult to obtain due to the high complexity of problems, and the output of the problems is not deterministic due to the influence of noise. Thus, SSO algorithms only rely on input–output data to search for the best solution. In the real world, many complex optimization problems can be regarded as SSO problems, such as power system designing [12] and wafer probe testing [15]. As a result, SSO technology has been widely used in these fields. From the aforementioned difficulties **(1)** and **(3)**, we can find that the optimization of RGPO jamming strategies is largely applicable to be described as an SSO problem. However, different from the canonical SSO problems, it is difficult for the RGPO jamming to directly obtain input–output data to support the strategy optimization due to difficulty **(2)**. Therefore, the optimization of RGPO jamming strategies can be seen as a more complex adversarial SSO (ASSO) problem.

Compared with the canonical SSO problems, the challenge of the ASSO problems lies in the acquisition of the input–output data. A research field that may provide reference is adversarial attacks on deep neural networks (DNNs). In adversarial attacks on

DNNs [16–19], the attacker fools the attacked DNN by adding an amount of artificial perturbations to the original input, which can bias the output of the attacked DNN. To obtain the desired attack effect, the attack strategy needs to be carefully designed [17]. However, due to the complexity of DNN, it is difficult to get the algebraic description of the objective function [19]. Moreover, for the attacker, it is difficult to obtain the output of the attacked DNN. To solve the above problems, the researchers propose to build a local simulation network of the attacked DNN to support the optimization of the attack strategy. Moreover, for the convenience of research, the adversarial attacks on DNNs are divided into white-box and black-box attacks, based on the amount of knowledge about the attacked DNNs [20–22]. In white-box attacks, the attacker has the full knowledge of the attacked DNN, while in black-box attacks the attacker does not have knowledge of the attacked DNN. For the white-box attack, a local simulation network of the attacked DNN can be built easily. For the black-box attack, the establishment of a local simulation network is a more complex issue. However, research on the white-box attack is the foundation for the research on the black-box attack in the optimization model and the main framework of the algorithm.

Comparing RGPO jamming and adversarial attacks on DNNs, it is not difficult to find that the fundamental content of them is very similar. Specifically, the jammer can be seen as the attacker, the threat radar plays the same role as the attacked DNN, the RGPO jamming signals can be seen as the artificial perturbations added to the original input, and the radar echo signal is the original input. Therefore, RGPO jamming can be analogized as an adversarial attack on the threat radar. Drawing inspiration from studies on adversarial attacks against DNNs, we classify RGPO jamming into two types: white-box and black-box jamming, depending on the jammer's knowledge of the threat radar's tracking model (e.g., tracking methods, parameters, etc.). The white-box scenario assumes that the jammer possesses detailed information about the tracking model of the threat radar. By contrast, the black-box scenario presumes that the jammer lacks access to precise details of the tracking model. Since the white-box scenario serves as the foundation for exploring black-box scenarios, this paper primarily focuses on addressing the optimization challenge for white-box RGPO jamming strategies. Similar to white-box attacks on DNNs, in the context of white-box RGPO jamming, a simulated tracking system of the threat radar can be constructed locally to produce input–output data, supporting the optimization of jamming strategies. To this end, we introduce a local simulation assisted SSO algorithm for solving the strategy optimization problem of the white-box RGPO jamming.

Generally, SSO algorithms are the integrations of metaheuristic algorithms (MAs) and resampling schemes [11,23]. MAs are responsible for searching the strategy space and generating candidate strategies for comparison, while resampling schemes are used to mitigate the influence of noise on the performance estimation of candidate strategies. Particle swarm optimization (PSO) is a popular MA algorithm and shows a very high search efficiency in many real-world problems [24]. Meanwhile, equal resampling (ER) represents a straightforward method to reduce the impact of noise [25–27]. This approach assigns an equal number of resamples to each candidate strategy (i.e., repeated evaluations of candidate strategy performance using the local simulation system). The average performance across these resamples is then used to estimate the final effectiveness of the candidate strategies. Therefore, we propose a strategy optimization algorithm for white-box RGPO jamming that integrates a local simulation model and a PSO algorithm with an ER scheme (PSO-ER).

The main contributions of this research are as follows:

1. We formulate the strategy optimization of the RGPO jamming as an ASSO problem, which lays the foundation for the nature of the problem and the optimization model for subsequent related research.

2. We propose an idea to use local simulation assisted SSO algorithms to solve the strategy optimization problem for RGPO jamming, which lays the foundation for the main frame of the algorithm for subsequent related research.
3. In this paper, an optimization algorithm based on PSO-ER with local simulation assistance is proposed to generate well-tuned strategies for white-box RGPO jamming.

The remainder of this paper is organized as follows: Section 2 elaborates the process of RGPO jamming and formulates the optimization problem. Section 3 describes the details of the proposed strategy optimization algorithm for white-box RGPO jamming. Section 4 presents numerical experiments validating the performance of the proposed algorithm. Section 5 concludes with a summary of this paper and a discussion of future work.

2. Problem Statement

2.1. Description of RGPO Jamming Process

Assume that a radar has detected a target and has tracked it for N_0 frames. Now, as shown in Figure 1, the airborne jammer identified the threat and decided to perform RGPO jamming for countering the threat radar. In each jamming stage, the jammer begins by capturing the transmitted signal from the threat radar. This signal is then retransmitted to the receiver of the radar with a deliberately manipulated delay. As a result of this jamming signal, the threat radar perceives a false target appearing at a greater distance.

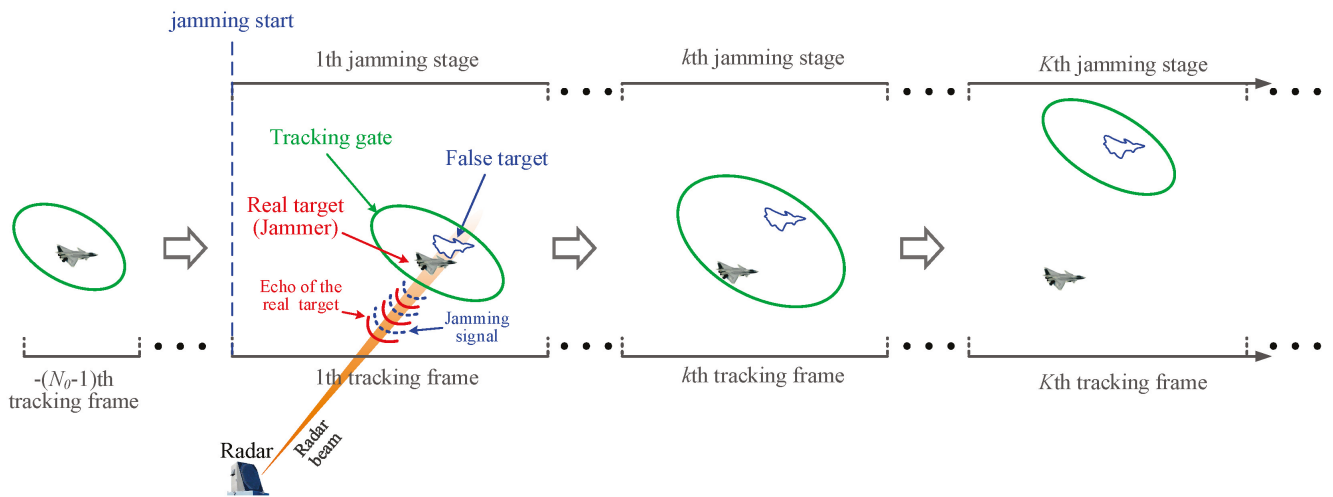


Figure 1. Diagram of RGPO jamming.

Suppose the RGPO jamming lasts K stages, and the controlled delay of the jamming signal at the k th jamming stage, τ_k , is set as

$$\tau_k = \sum_{\tilde{k}=1}^k \Delta\tau_{\tilde{k}}, k = 1, 2, \dots, K, \quad (1)$$

where $\Delta\tau_{\tilde{k}}$ is the increment of the controlled delay at the $\tilde{k}$ th jamming stage. Then, the distance, d_k , between the false target and the real target can be represented as

$$d_k = \frac{c\tau_k}{2} = \sum_{\tilde{k}=1}^k \frac{c\Delta\tau_{\tilde{k}}}{2}, k = 1, 2, \dots, K, \quad (2)$$

where c is the speed of light.

Without loss of generality, we assume that the real target follows a nearly constant velocity (CV) model [28], and the state matrix $\mathbf{X}_r(k)$ is given by

$$\mathbf{X}_r(k) = \mathbf{F}_r(k-1)\mathbf{X}_r(k-1) + \mathbf{U}(k-1), \quad (3)$$

where $\mathbf{F}_r(k-1)$ is the state transition matrix and $\mathbf{U}(k-1)$ is the process noise. The matrix formulation of the above equation can be represented as

$$\begin{bmatrix} x_r^{(k)} \\ \dot{x}_r^{(k)} \\ y_r^{(k)} \\ \dot{y}_r^{(k)} \end{bmatrix} = \begin{bmatrix} 1 & t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & t \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_r^{(k-1)} \\ \dot{x}_r^{(k-1)} \\ y_r^{(k-1)} \\ \dot{y}_r^{(k-1)} \end{bmatrix} + \begin{bmatrix} 0.5\varepsilon_x^{(k)}t^2 \\ \varepsilon_x^{(k)}t \\ 0.5\varepsilon_y^{(k)}t^2 \\ \varepsilon_y^{(k)}t \end{bmatrix}, \quad (4)$$

where $(x_r^{(k)}, y_r^{(k)})$ denotes the position coordinates of the real target at the k th jamming stage in the x-axis and y-axis, $(\dot{x}_r^{(k)}, \dot{y}_r^{(k)})$ represents the velocities of the real target, t is the time interval between two successive jamming stages, and $\varepsilon_x^{(k)}$ and $\varepsilon_y^{(k)}$ denote the random variation of velocities of the real target.

Thus, the state matrix of the false target, $\mathbf{X}_f(k)$, can be expressed as

$$\mathbf{X}_f(k) = \mathbf{X}_r(k) + \mathbf{\Theta}(k)d_k, \quad (5)$$

where $\mathbf{\Theta}(k)$ is the coordinate transformation matrix. Corresponding to Equation (4), the matrix formulation of the above equation becomes

$$\begin{bmatrix} x_f^{(k)} \\ \dot{x}_f^{(k)} \\ y_f^{(k)} \\ \dot{y}_f^{(k)} \end{bmatrix} = \begin{bmatrix} x_r^{(k)} \\ \dot{x}_r^{(k)} \\ y_r^{(k)} \\ \dot{y}_r^{(k)} \end{bmatrix} + \begin{bmatrix} \cos(\theta_k) \\ 0 \\ \sin(\theta_k) \\ 0 \end{bmatrix} d_k, \quad (6)$$

where $(x_f^{(k)}, y_f^{(k)})$ represents the position coordinates of the false target, $(\dot{x}_f^{(k)}, \dot{y}_f^{(k)})$ represents the velocities of the false target, and θ_k is the azimuth angle of the real target and the false target.

Then, the measurement equation of the real target is given by

$$\mathbf{Z}_r(k) = \mathbf{H}(k)\mathbf{X}_r(k) + \mathbf{W}_r(k), \quad (7)$$

where $\mathbf{Z}_r(k)$ represents the measurement matrix of the real target, $\mathbf{H}(k)$ is the projection matrix that projects the state into the measurement space, and $\mathbf{W}_r(k)$ represents the measurement noise of the real target. Corresponding to Equation (4), the matrix formulation of the above equation becomes

$$\begin{bmatrix} \tilde{x}_r^{(k)} \\ \tilde{y}_r^{(k)} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_r^{(k)} \\ \dot{x}_r^{(k)} \\ y_r^{(k)} \\ \dot{y}_r^{(k)} \end{bmatrix} + \begin{bmatrix} \delta_{r,x}^{(k)} \\ \delta_{r,y}^{(k)} \end{bmatrix}, \quad (8)$$

where $(\tilde{x}_r^{(k)}, \tilde{y}_r^{(k)})$ denotes the measurements of the position of the real target and $\delta_{r,x}^{(k)}$ and $\delta_{r,y}^{(k)}$ denote the random measurement error of position of the real target. Similarly, the measurement equation of the false target is given by

$$\mathbf{Z}_f(k) = \mathbf{H}(k)\mathbf{X}_f(k) + \mathbf{W}_f(k), \quad (9)$$

where $\mathbf{Z}_f(k)$ is the measurement matrix of the false target and $\mathbf{W}_f(k)$ represents the measurement noise of the false target.

Next, the threat radar will receive the measurements $\mathbf{Z}_r(k)$ and $\mathbf{Z}_f(k)$ as input data. In the following, we will introduce the impact of $\mathbf{Z}_f(k)$ on the threat radar in detail.

Radar tracking is a vast field with many branches, and researchers have conducted extensive research on each of them. The specific effect of the RGPO jamming differs slightly for different radar tracking approaches. Without loss of generality, a typical radar tracking system using the Kalman filter-probabilistic data association (KF-PDA) approach [29] is chosen as the tracking system of the threat radar in this subsection to illustrate the impact of the RGPO jamming. For this tracking system of the threat radar, the prediction equations are

$$\mathbf{X}(k|k-1) = \mathbf{F}(k-1)\mathbf{X}(k-1), \quad (10)$$

$$\mathbf{Z}(k|k-1) = \mathbf{H}(k)\mathbf{X}(k|k-1), \quad (11)$$

where $\mathbf{X}(k-1)$ is the target state estimation, $\mathbf{X}(k|k-1)$ is the one-step prediction of the threat radar for the target state, $\mathbf{F}(k-1)$ is the state transition matrix of the Kalman filter (KF) model, and $\mathbf{Z}(k|k-1)$ is the one-step prediction of the target measurement. The covariance matrix of the state prediction, $\mathbf{P}(k|k-1)$, is given by

$$\mathbf{P}(k|k-1) = \mathbf{F}(k-1)\mathbf{P}(k-1)\mathbf{F}^T(k-1) + \mathbf{Q}_p(k-1), \quad (12)$$

where $\mathbf{P}(k-1)$ is the covariance of the state estimation and $\mathbf{Q}_p(k-1)$ is the covariance matrix of the process noise of the KF model.

Next, the tracking gate (also called the validation region) is built with $\mathbf{Z}(k|k-1)$ as the center to verify radar measurements. The typical tracking gate is the elliptical region [29], which can be expressed as

$$\mathcal{V}(k, \gamma) = \left\{ \mathbf{Z} \mid [\mathbf{Z} - \mathbf{Z}(k|k-1)]^T \mathbf{S}^{-1}(k) [\mathbf{Z} - \mathbf{Z}(k|k-1)] \leq \gamma \right\}, \quad (13)$$

where γ is the gate threshold and $\mathbf{S}(k)$ is the covariance matrix of the innovation, which can be given by

$$\mathbf{S}(k) = \mathbf{H}(k)\mathbf{P}(k|k-1)\mathbf{H}^T(k) + \mathbf{Q}_m(k). \quad (14)$$

where $\mathbf{Q}_m(k)$ is the covariance matrix of the measurement noise. Measurements within the tracking gate are called the validated measurements.

Then, probabilistic data association (PDA) is performed to determine the weights (i.e., the association probabilities) of these validated measurements to participate in the tracking update.

Assume that the number of validated measurements at the k th jamming stage is $\mathbb{I}_k$, which can be further expressed as

$$\mathbb{I}_k = \begin{cases} I_k + 1, & \mathbf{Z}_f(k) \in \mathcal{V}(k, \gamma) \\ I_k, & \text{else} \end{cases}, \quad (15)$$

where I_k indicates the number of validated measurements without considering the RGPO jamming. If the jamming measurement $\mathbf{Z}_f(k)$ does not fall into the tracking gate as shown in Figure 2a, it is logical to get $\mathbb{I}_k = I_k$. Conversely, if $\mathbf{Z}_f(k)$ falls into the tracking gate

as shown in Figure 2b, $\mathbb{I}_k = I_k + 1$. It could be assumed that $\mathbf{Z}_f(k)$ is the $(I_k + 1)$ th validated measurement.

Jamming measurement does not fall into the tracking gate

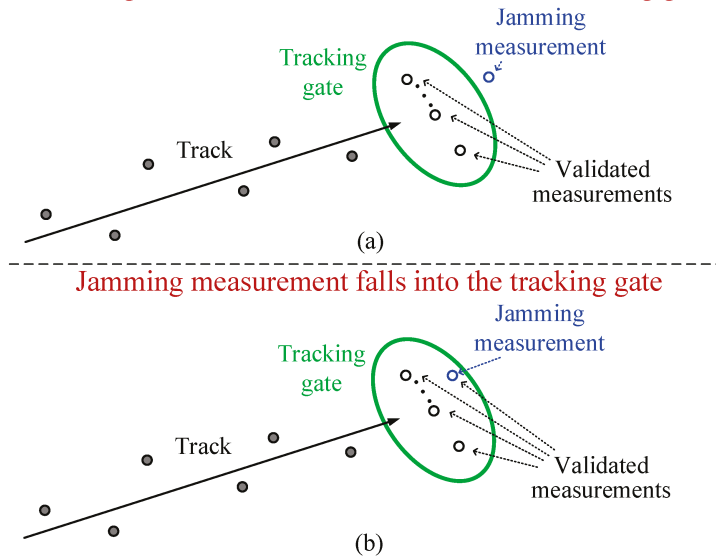


Figure 2. Schematic diagram of measurement verification.

Obviously, the measurement of the false target has a negative impact on the data association. The innovation, $\mathbf{v}_i(k)$, of the i th validated measurement can be expressed as

$$\mathbf{v}_i(k) = \mathbf{Z}_i(k) - \mathbf{Z}(k|k-1), i = 1, 2, \dots, \mathbb{I}_k \quad (16)$$

where $\mathbf{Z}_i(k)$ represents the i th validated measurement at the k th jamming stage and $\mathbf{Z}_{I_k+1}(k) = \mathbf{Z}_f(k)$. The association probability of each validated measurement is calculated as follows [30]:

$$\beta_0^k = \frac{b}{b + \sum_{i=1}^{\mathbb{I}_k} e_i} = \begin{cases} \frac{b}{b + \sum_{i=1}^{I_k+1} e_i}, & \mathbf{Z}_f(k) \in \mathcal{V}(k, \gamma) \\ \frac{b}{b + \sum_{i=1}^{I_k} e_i}, & \text{else} \end{cases}, \quad (17)$$

$$\beta_i^k = \frac{e_i}{b + \sum_{i=1}^{\mathbb{I}_k} e_i} = \begin{cases} \frac{e_i}{b + \sum_{i=1}^{I_k+1} e_i}, & \mathbf{Z}_f(k) \in \mathcal{V}(k, \gamma) \\ \frac{e_i}{b + \sum_{i=1}^{I_k} e_i}, & \text{else} \end{cases}, \quad (18)$$

$$i = 1, 2, \dots, \mathbb{I}_k,$$

where

$$e_i = \exp \left\{ -\frac{1}{2} \mathbf{v}_i^T(k) \mathbf{S}^{-1}(k) \mathbf{v}_i(k) \right\}, \quad (19)$$

$$b = \lambda |2\pi \mathbf{S}(k)|^{\frac{1}{2}} (1 - P_D P_G) / P_D, \quad (20)$$

where β_0^k represents the probability that there is no validated measurement derived from the real target, β_i^k represents the association probability of i th validated measurement, and

$\sum_{i=0}^{\mathbb{I}_k} \beta_i^k = 1$. P_D denotes the target detection probability, P_G is a parameter that accounts for restricting the normal density to the tracking gate, and λ is a parameter indicating the spatial density of false measurements.

From Equation (18), it can be seen that once the false measurement, $\mathbf{Z}_f(k)$, falls into the tracking gate, it will be assigned the weight $\beta_{I_k+1}^k$. Compared with the case that false target measurement $\mathbf{Z}_f(k)$ is not in the tracking gate, the proportion of the first I_k weights within the set $\{\beta_0^k, \beta_1^k, \dots, \beta_{I_k+1}^k\}$ will be decreased, with the constraint $\sum_{i=0}^{\mathbb{I}_k} \beta_i^k = 1$. According to Equation (21), the combined innovation, $\mathbf{v}(k)$, will contain the wrong distance information carried by $\mathbf{Z}_f(k)$, while the weight of the information carried by the real target will be decreased. This means that the combined innovation, $\mathbf{v}(k)$, tends to be closer to the innovation of the false target, $\mathbf{v}_{I_k+1}(k)$, and away from the innovation of the real target. As a result, $\mathbf{v}(k)$ will tend to appear biased.

$$\begin{aligned} \mathbf{v}(k) &= \sum_{i=1}^{\mathbb{I}_k} \beta_i(k) \mathbf{v}_i(k) \\ &= \begin{cases} \sum_{i=1}^{I_k+1} \beta_i^k \mathbf{v}_i(k), & \mathbf{Z}_f(k) \in \mathcal{V}(k, \gamma) \\ \sum_{i=1}^{I_k} \beta_i^k \mathbf{v}_i(k), & \text{else} \end{cases} \end{aligned} \quad (21)$$

Finally, the combined innovation, $\mathbf{v}(k)$, is used to update the state estimation, and the estimation of the target state, $\mathbf{X}(k)$, is given by

$$\mathbf{X}(k) = \mathbf{X}(k|k-1) + \mathbf{G}(k) \mathbf{v}(k), \quad (22)$$

where $\mathbf{G}(k)$ is the KF gain and

$$\mathbf{G}(k) = \mathbf{P}(k|k-1) \mathbf{H}^T(k) \mathbf{S}^{-1}(k). \quad (23)$$

As shown in Equation (22), the error derived from $\mathbf{Z}_f(k)$ will be transmitted to $\mathbf{X}(k)$ via $\mathbf{v}(k)$, resulting in an error in the estimation of the target state. Specifically, the estimate of the target state will be closer to the false target measurement, $\mathbf{Z}_f(k)$, and farther away from the real target measurement, $\mathbf{Z}_r(k)$. Similarly, the error component will also be transmitted to $\mathbf{P}(k)$ via $\mathbf{v}(k)$, as shown in Equations (24)–(27).

$$\mathbf{P}(k) = \mathbf{P}(k|k-1) \beta_0^k + [1 - \beta_0^k] \mathbf{P}^c(k) + \tilde{\mathbf{P}}(k), \quad (24)$$

$$\mathbf{P}^c(k) = [\mathbf{I} - \mathbf{G}(k) \mathbf{H}(k)] \mathbf{P}(k|k-1), \quad (25)$$

$$\tilde{\mathbf{P}}(k) = \mathbf{G}(k) [\mathbf{B}(k) - \mathbf{v}(k) \mathbf{v}^T(k)] \mathbf{G}^T(k), \quad (26)$$

$$\begin{aligned} \mathbf{B}(k) &= \sum_{i=1}^{\mathbb{I}_k} \beta_i^k \mathbf{v}_i(k) \mathbf{v}_i^T(k) \\ &= \begin{cases} \sum_{i=1}^{I_k+1} \beta_i^k \mathbf{v}_i(k) \mathbf{v}_i^T(k), & \mathbf{Z}_f(k) \in \mathcal{V}(k, \gamma) \\ \sum_{i=1}^{I_k} \beta_i^k \mathbf{v}_i(k) \mathbf{v}_i^T(k), & \text{else} \end{cases} \end{aligned} \quad (27)$$

Moreover, the negative impact of $\mathbf{Z}_f(k)$ is not limited to the current k th jamming stage; $\mathbf{X}(k)$ and $\mathbf{P}(k)$ will be used for the prediction of the state and measurement to the $(k+1)$ th jamming stage, as shown in Equations (10) and (12). Correspondingly, the one-step prediction of the target state, $\mathbf{X}(k+1|k)$, and the measurement, $\mathbf{Z}(k+1|k)$, of the

$(k + 1)$ th jamming stage are also influenced by the error component due to $\mathbf{Z}_f(k)$. As a result, the center of the tracking gate, $\mathbf{Z}(k + 1|k)$, is shifted away from the measurement of the real target, $\mathbf{Z}_r(k + 1)$. In the case described above, if the false target's measurement at the $(k + 1)$ th jamming stage remains within the tracking gate, the tracking error will increase further, leading to a continued shift in the tracking gate. Moreover, when the controlled delays across successive tracking frames are properly coordinated, the tracking error can rapidly compound, causing the tracking gate to significantly diverge from the real target.

Although there are many different approaches to radar tracking, the principle of the RGPO jamming is basically similar: use the measurements of false target to confuse the data association of the tracking system of the threat radar, thereby affecting the target state estimation. Therefore, it is very important to set the jamming strategy properly.

2.2. Optimization Model of White-Box RGPO Jamming Strategy

For RGPO jamming, the quality of the jamming results is mainly evaluated by the following two points [10]:

1. Whether or not the real target escapes from the tracking gate: If the real target does not escape from the tracking gate, as shown in Figure 3a, then the jamming effect is very limited. On the contrary, if the measurement of the real target escapes from the tracking gate, as shown in Figure 3b, this means that the threat radar has temporarily lost the real target.
2. The miss distance: In the case of a larger miss distance, the real target is more difficult to recaptured by the tracking gate, and the real target is safer.

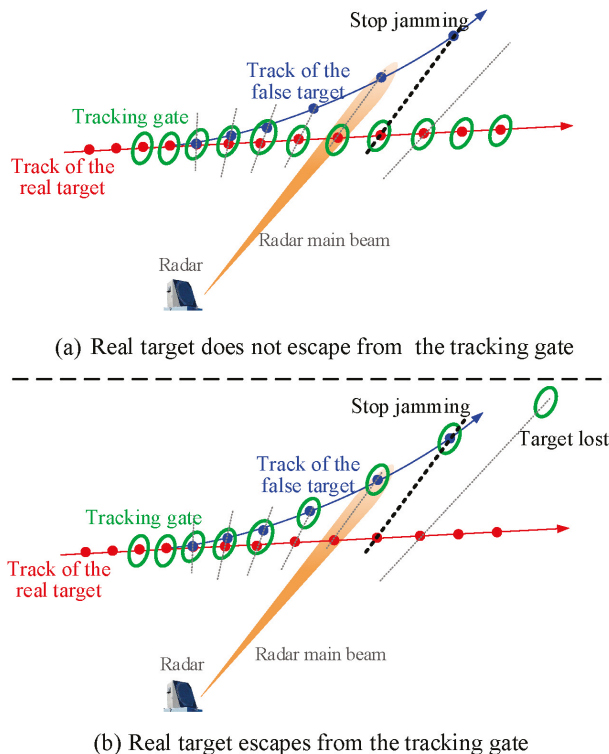


Figure 3. Schematic diagram of jamming result.

For a specific sampling of the jamming strategy φ , $\xi_\varphi^{(K)}$ is used to denote whether or not the real target escapes from the tracking gate, which is given by

$$\xi_\varphi^{(K)} = \begin{cases} 0, & \mathbf{Z}_r(K) \in \mathcal{V}(K, \gamma) \\ 1, & \text{else} \end{cases}, \quad (28)$$

where $\xi_\varphi^{(K)} = 1$ signifies that the real target has moved outside the tracking gate, whereas $\xi_\varphi^{(K)} = 0$ indicates that the real target remains inside the tracking gate. Moreover, for a specific sampling of the jamming strategy φ , the corresponding miss distance can be expressed as

$$\text{MD}_\varphi^{(K)} = \sqrt{\left[x^{(K)} - x_r^{(K)}\right]^2 + \left[y^{(K)} - y_r^{(K)}\right]^2}, \quad (29)$$

where $(x^{(K)}, y^{(K)})$ denotes the position estimation of the real target and $(x_r^{(K)}, y_r^{(K)})$ represents the actual position of the real target.

Assume the threat radar tracking model is m . Then, for a particular sampling of the jamming strategy φ , the quality of jamming result can be evaluated as

$$[\xi_\varphi^{(K)}, \text{MD}_\varphi^{(K)}] = [r_1(\varphi, \mathcal{X}, \mathcal{W}, m), r_2(\varphi, \mathcal{X}, \mathcal{W}, m)], \quad (30)$$

where $r_1(\cdot)$ and $r_2(\cdot)$ denote the objective functions, $\varphi = [\Delta\tau_1, \Delta\tau_2, \dots, \Delta\tau_K]$ represents the increment of the controlled delay of the jamming signal for the K jamming stages, $\mathcal{X} = [\mathbf{X}_r(1), \mathbf{X}_r(2), \dots, \mathbf{X}_r(K)]$ represents the state of the real target with process noise for the K jamming stages, and $\mathcal{W} = [\mathbf{W}_r(1), \mathbf{W}_r(2), \dots, \mathbf{W}_r(K), \mathbf{W}_f(1), \mathbf{W}_f(2), \dots, \mathbf{W}_f(K)]$ represents the measurement noise of the K jamming stages.

Due to the fact that $\mathcal{X}$ and $\mathcal{W}$ are random variables, $\xi_\varphi^{(K)}$ and $\text{MD}_\varphi^{(K)}$ are indeterminate. Therefore, jamming strategy φ cannot be accurately evaluated only by the jamming result of a particular sampling. In this paper, we consider evaluating the performance of the jamming strategy using the expected values of $\xi_\varphi^{(K)}$ and $\text{MD}_\varphi^{(K)}$. Thus, the strategy optimization problem for RGPO jamming can be expressed as

$$\begin{aligned} \max_{\varphi} \quad & \left[E[\xi_\varphi^{(K)}], E[\text{MD}_\varphi^{(K)}] \right] = [E[r_1(\varphi, \mathcal{X}, \mathcal{W}, m)], E[r_2(\varphi, \mathcal{X}, \mathcal{W}, m)]] \\ \text{s.t.} \quad & \varphi \in \Omega, \end{aligned} \quad (31)$$

where $E[\cdot]$ represents expected value, while Ω indicates solution space. Because of the conflicting nature of the two objective functions in problem (31), achieving a global maximum for all objectives simultaneously at a single solution is unfeasible. The weighted aggregation method is widely recognized as a standard technique for handling multi-objective optimization problems [31–33]. Using this method, the score corresponding to a specific sample of the strategy φ is determined as follows:

$$r(\varphi) = w_1 \cdot \xi_\varphi^{(K)} + w_2 \cdot \text{MD}_\varphi^{(K)}, \quad (32)$$

where $w_1 > 0$ and $w_2 > 0$ represent the respective weights assigned to the objective functions. Using this scoring framework, the strategy optimization problem for RGPO jamming can be further expressed as

$$\begin{aligned}
\max_{\varphi} E[r(\varphi)] &= w_1 \cdot E\left[\xi_{\varphi}^{(K)}\right] + w_2 \cdot E\left[\text{MD}_{\varphi}^{(K)}\right] \\
&= w_1 \cdot E[r_1(\varphi, \mathcal{X}, \mathcal{W}, m)] + w_2 \cdot E[r_2(\varphi, \mathcal{X}, \mathcal{W}, m)] \\
\text{s.t. } \varphi &\in \Omega.
\end{aligned} \tag{33}$$

Because of the high complexity of RGPO jamming, it is challenging to derive algebraic expressions for the objective functions $r_1(\cdot)$ and $r_2(\cdot)$. Moreover, the jammer cannot directly observe the jamming outcomes, such as $\xi_{\varphi}^{(K)}$ and $\text{MD}_{\varphi}^{(K)}$, from the threat radar. Additionally, the presence of noise introduces uncertainty into the jamming results. These three factors collectively render the jamming strategy optimization problem, as shown in Equation (33), an ASSO problem. Unlike standard SSO problems, the primary difficulty in ASSO problems lies in obtaining the necessary input–output data. However, in white-box scenarios, the knowledge of the tracking radar is available. This enables the construction of a local simulation for the tracking radar, which will be used to approximate the jamming outcomes. Consequently, it becomes feasible to address the above jamming strategy optimization problem using SSO algorithms assisted by local simulation.

3. White-Box RGPO Jamming Strategy Generation

3.1. Overall Framework

In this section, we describe the process of designing a white-box RGPO jamming strategy. As illustrated in Figure 4, the PSO-ER algorithm functions as the optimizer for the jamming strategy. The PSO-ER combines PSO with ER, where PSO explores the solution domain and produces candidate strategies to be evaluated. To reduce the error in the comparison of candidate strategies caused by noise, we estimate their fitness (i.e., the jamming performance) using the ER scheme. In the ER, all candidate strategies are assigned the same number of resampling opportunities. For each resampling, the jamming result is obtained by the local simulation of the tracking radar. Finally, the fitness of the candidate strategies are estimated as the mean of the jamming results of the multiple resampling. Below, the process of the strategy optimization is introduced in detail.

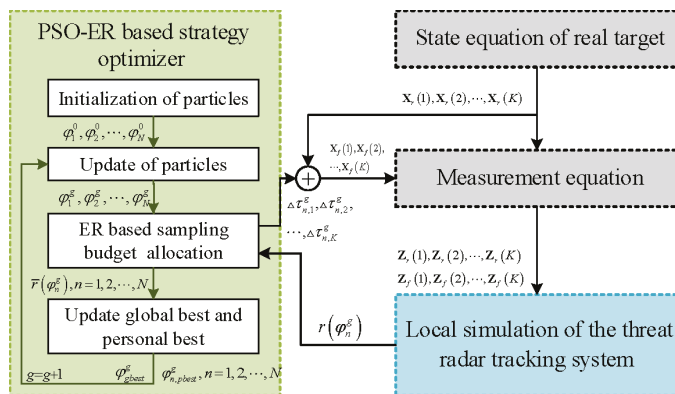


Figure 4. Generic diagram of the white-box RGPO jamming strategy optimization.

3.2. PSO-ER-Based Strategy Optimizer

In the PSO-ER framework, a group of particles moves through the decision space, iteratively converging toward the optimal strategy. Each particle is characterized by its position and velocity, where the position represents a potential candidate strategy. The optimization procedure is carried out in three fundamental stages.

Firstly, we initialize the positions $\varphi_n^0 = [\Delta\tau_{n1}^0, \Delta\tau_{n2}^0, \dots, \Delta\tau_{nK}^0], n = 1, 2, \dots, N$ and velocities $v_n^0 = [v_{n1}^0, v_{n2}^0, \dots, v_{nK}^0], n = 1, 2, \dots, N$ of N particles using Latin hypercube sampling (LHS). Following the initialization process, the fitness of all particles (i.e., the performance of the associated candidate strategies) is evaluated using the ER scheme. Within this scheme, J resamples are conducted for each candidate strategy, and their performance is estimated by

$$\bar{r}(\varphi_n) = w_1 \cdot \bar{\xi}_{\varphi_n}^{(K)} + w_2 \cdot \overline{MD}_{\varphi_n}^{(K)}, n = 1, 2, \dots, N, \quad (34)$$

where $\bar{r}(\varphi_n)$ denotes the average score obtained by J resamples of the candidate strategy φ_n , $\bar{\xi}_{\varphi_n}^{(K)}$ denotes the ratio of the real target escaping from the tracking gate in J resamples of the candidate strategy φ_n , and $\overline{MD}_{\varphi_n}^{(K)}$ denotes the average miss distance obtained by J resamples of the candidate strategy φ_n .

Then, for each iteration, the velocities and positions of all particles are updated as follows:

$$v_n^g = \alpha^g v_n^{g-1} + c_1 \delta_1^g (\varphi_{n,pbest}^{g-1} - \varphi_n^{g-1}) + c_2 \delta_2^g (\varphi_{gbest}^{g-1} - \varphi_n^{g-1}), n = 1, 2, \dots, N, \quad (35)$$

$$\varphi_n^g = \varphi_n^{g-1} + v_n^g, n = 1, 2, \dots, N, \quad (36)$$

where v_n^g and φ_n^g respectively denote the current velocity and position of the n th particle at the g th iteration, $\varphi_{n,pbest}^{g-1}$ is the personal best position of the n th particle at the $(g-1)$ th iteration (i.e., the best jamming strategy found by the n th particle so far, and the initial personal best position $\varphi_{n,pbest}^0$ is set to be equal as φ_n^0), φ_{gbest}^{g-1} is the global best position at the $(g-1)$ th iteration (i.e., the best jamming strategy found by all particles so far). The sequences δ_1 and δ_2 are K -dimensional, with each dimension being a random value in the range $[0, 1]$. As recommended in [34], the constants c_1 and c_2 are both assigned a value of 1.49445. The inertia weight, denoted as α^g , governs velocity updates and decreases linearly from 0.9 to 0.4 as iteration g progresses, according to (37), where $g_{\max}$ represents the maximum number of iterations [35].

$$\alpha^g = 0.9 - 0.5 \frac{g}{g_{\max}}. \quad (37)$$

Finally, we recalculated the fitness of each particle using the ER scheme and update the personal best position, $\varphi_{n,pbest}^g$, and the global best position, φ_{gbest}^g , as follows:

$$\varphi_{n,pbest}^g = \begin{cases} \varphi_n^g, & \bar{r}(\varphi_n^g) > \bar{r}(\varphi_{n,pbest}^{g-1}) \\ \varphi_{n,pbest}^{g-1}, & \text{else} \end{cases}, \quad (38)$$

$$\varphi_{gbest}^g = \arg \max_{\varphi_{n,pbest}^g} \bar{r}(\varphi_{n,pbest}^g). \quad (39)$$

Thus far, the optimal jamming strategy at the g th iteration can be derived from φ_{gbest}^g .

The above iteration process will persist until the preset maximum number of iterations, $g_{\max}$, is satisfied, at which point the final RGPO jamming strategy, $\varphi_{gbest}^{g_{\max}}$, will be produced. A concise overview of the proposed strategy optimization algorithm for white-box RGPO jamming is given in Algorithm 1.

Algorithm 1 Local simulation-assisted PSO-ER-based white-box RGPO jamming strategy optimization algorithm.

Input: The motion model of the real target. The local simulation of the threat radar tracking system. Preset parameters: $K, N, J, g_{\max}$, etc;

- 1: Initialize the positions $\varphi_n^0, n = 1, 2, \dots, N$ and the velocities $v_n^0, n = 1, 2, \dots, N$ of particles using LHS;
- 2: Estimate the fitness $\bar{r}(\varphi_n^0), n = 1, 2, \dots, N$ of particles by the ER scheme as Equation (34);
- 3: Initialize the personal best position $\varphi_{n,\text{pbest}}^0, n = 1, 2, \dots, N$ and the global best position φ_{gbest}^0 ;
- 4: **for** $g = 1 : g_{\max}$ **do**
- 5: Calculate the inertia weight α^g using (37);
- 6: Update the velocities $v_n^g, n = 1, 2, \dots, N$ and the positions $\varphi_n^g, n = 1, 2, \dots, N$ of particles using Equations (35) and (36);
- 7: Estimate the fitness $\bar{r}(\varphi_n^g), n = 1, 2, \dots, N$ of particles by the ER scheme as Equation (34);
- 8: Update the personal best position of N particles $\varphi_{n,\text{pbest}}^g, n = 1, 2, \dots, N$ and the global best position φ_{gbest}^g using Equations (38) and (39);
- 9: **end for**

Output: The jamming strategy, $\varphi_{\text{gbest}}^{g_{\max}}$;

4. Empirical Study

From Equation (33), we can see that the performance of the RGPO jamming strategy is related to the motion model of the real target and the tracking model of the threat radar. To evaluate the performance of the proposed method on different motion models and tracking models, four typical tracking problems from [10] are chosen as the benchmark problems, and their parameters are shown in Table 1.

Remark 1. It should be mentioned that, before the jamming, the threat radar has established a track for N_0 tracking frames, as shown in Figure 1. However, in practice, N_0 is not fixed and depends on the response time, T_0 , of the jammer to the threat. Generally, the value of T_0 should meet the following two points: (1) The identification of the threat by the jammer requires time to intercept enough information, so a small T_0 is a small probability event; (2) When the jammer intercepts enough information, the probability of missing detection of a threat should be extremely low, so a large T_0 is also a small probability event. Therefore, it is not difficult to find that T_0 should follow an approximate Gaussian distribution. Without loss of generality, the value of T_0 can be given by

$$T_0 = \begin{cases} T_{\text{temp}}, & T_{\text{temp}} \geq 0 \\ 0, & T_{\text{temp}} < 0 \end{cases} \quad (40)$$

where T_{temp} denotes a temporary variable, which follows a Gaussian distribution with mean T_e and variance σ_e^2 . Then, N_0 can be obtained by

$$N_0 = \left\lceil \frac{T_0}{t} \right\rceil. \quad (41)$$

Since the setting of N_0 is not the focus of this paper, we set $T_e = 5, \sigma_e^2 = 2$ in the following experiments.

Table 1. Parameter settings.

Radar Parameters	
Time interval t	1s
Measurement noise	Range: Gaussian of $N(0, 10^2)$; Angle: Gaussian of $N(0, 0.00175^2)$
Detection probability	0.99
Elliptic gate parameter γ	16
CV Benchmark	
Target motion model	Nearly constant velocity (CV) model
Target initial state	$(x_r^{(0)}, \dot{x}_r^{(0)}, y_r^{(0)}, \dot{y}_r^{(0)}) = (50\text{km}, 50\text{m/s}, 55\text{km}, 350\text{m/s})$
Target motion process noise	x axis: Gaussian of $N(0, 5^2)$; y axis: Gaussian of $N(0, 5^2)$
Tracking method	KF-PDA
Process noise of KF model	x axis: Gaussian of $N(0, 10^2)$; y axis: Gaussian of $N(0, 10^2)$
CA Benchmark	
Target motion model	Nearly constant acceleration (CA) model
Target initial state	$(x_r^{(0)}, \dot{x}_r^{(0)}, \ddot{x}_r^{(0)}, y_r^{(0)}, \dot{y}_r^{(0)}, \ddot{y}_r^{(0)}) = (50\text{km}, 25\text{m/s}, 10\text{m/s}^2, 55\text{km}, 350\text{m/s}, 10\text{m/s}^2)$
Target motion process noise	x axis: Gaussian of $N(0, 1^2)$; y axis: Gaussian of $N(0, 1^2)$
Tracking method	KF-PDA
Process noise of KF model	x axis: Gaussian of $N(0, 2^2)$; y axis: Gaussian of $N(0, 2^2)$
CT Benchmark	
Target motion model	Nearly coordinate turn (CT) model
Target initial state	$(x_r^{(0)}, \dot{x}_r^{(0)}, y_r^{(0)}, \dot{y}_r^{(0)}, \omega) = (50\text{km}, 25\text{m/s}, 55\text{km}, 350\text{m/s}, 0.15\text{rad/s})$
Target motion process noise	x axis: Gaussian of $N(0, 5^2)$; y axis: Gaussian of $N(0, 5^2)$
Tracking method	KF-PDA
Process noise of KF model	x axis: Gaussian of $N(0, 10^2)$; y axis: Gaussian of $N(0, 10^2)$
RV Benchmark	
Target motion model	Nearly exoatmospheric re-entry vehicle (RV) model
Target initial state	$(x_r^{(0)}, \dot{x}_r^{(0)}, y_r^{(0)}, \dot{y}_r^{(0)}) = (48.5\text{km}, -1720\text{m/s}, 22\text{km}, -946\text{m/s})$
Target motion process noise	x axis: Gaussian of $N(0, 5^2)$; y axis: Gaussian of $N(0, 5^2)$
Tracking method	EKF-PDA
Process noise of KF model	x axis: Gaussian of $N(0, 10^2)$; y axis: Gaussian of $N(0, 10^2)$
Ballistic parameters	40,000
Atmospheric density parameter	$\lambda_1 = 1.227$; $\lambda_2 = 1.093 \times 10^{-4}$

$(x_r^{(0)}, y_r^{(0)})$ represents the initial position of the target; $(\dot{x}_r^{(0)}, \dot{y}_r^{(0)})$ represents the velocity of the target; $(\ddot{x}_r^{(0)}, \ddot{y}_r^{(0)})$ represents the acceleration of the target; ω represents the angular velocity of the target.

4.1. Parameter Sensitivity Analysis

To analyze how certain key parameters affect the performance of the proposed approach, we carried out a series of experiments on the CV benchmark problem. Moreover, to better verify the effectiveness, we compare the proposed method with UV-RGPO and UA-RGPO. However, the setting of parameters of UV-RGPO and UA-RGPO (i.e., the pull-off velocity, v_{po} , for the UV-RGPO, the initial pull-off velocity, v_{ipo} , and the pull-off acceleration, a_{po} , for the UA-RGPO) have an impact on their performance. To our best knowledge,

there is no specific method in the open literature for setting these parameters. In practice, these parameters are set based on expert experience. In this paper, to better express the advantages of the proposed method, we find the performance of the best UV-RGPO and UA-RGPO strategies by traversal, then compare them with the average performance of strategies obtained by the proposed method.

4.1.1. Scalability

To test how well the proposed method scales with an increasing number of jamming stages, K , the proposed method is compared with UV-RGPO and UA-RGPO on the CV benchmark with various K . To be fair, the weights of objective functions are set to be the same ($w_1 = 0.5, w_2 = 0.5/200$). To find the best UV-RGPO strategy, we traverse v_{po} within $[v_{po}^{\min}, v_{po}^{\max}] = [2 \text{ m/s}, 50 \text{ m/s}]$. Similarly, to find the best UA-RGPO strategy, we traverse v_{ipo} within $[v_{ipo}^{\min}, v_{ipo}^{\max}] = [2 \text{ m/s}, 50 \text{ m/s}]$ and a_{po} within $[a_{po}^{\min}, a_{po}^{\max}] = [2 \text{ m/s}^2, 20 \text{ m/s}^2]$. For the proposed method, 50 independent runs are performed, the PSO is initialized with $N = 40$ particles, the number of resampling for each candidate strategy is $J = 2000$, and the maximum number of iterations is $g_{\max} = 100$.

Figure 5 shows the average $E[r(\varphi)]$ performance of strategies obtained by the proposed method in comparison with the $E[r(\varphi)]$ performance of different UV-RGPO and UA-RGPO strategies on CV benchmarks with different K , where the $E[r(\varphi)]$ is estimated by 10,000 independent Monte Carlo trials. Moreover, the average $E[\xi_{\varphi}^{(K)}]$, $E[\text{MD}_{\varphi}^{(K)}]$ performances of strategies obtained by the proposed method, and the $E[\xi_{\varphi}^{(K)}]$, $E[\text{MD}_{\varphi}^{(K)}]$ performances of the best UV-RGPO strategy and UA-RGPO strategy, are tabulated in Table 2. Similarly, the $E[\xi_{\varphi}^{(K)}]$ and $E[\text{MD}_{\varphi}^{(K)}]$ are estimated by 10,000 independent Monte Carlo trials. From Figure 5, it can be seen that the $E[r(\varphi)]$ performance of strategies obtained by the proposed method is significantly better than all the UV-RGPO and UA-RGPO strategies when K changes from 10 to 20, and the gain of the proposed method does not change much as the maximum number of jamming stages increases. The observation we can make from Table 2 is that the $E[\xi_{\varphi}^{(K)}]$ and $E[\text{MD}_{\varphi}^{(K)}]$ performances of strategies obtained by the proposed method are better than the best UV-RGPO and UA-RGPO strategies when K changes from 10 to 20. These results prove that the proposed method is better than UV-RGPO and UA-RGPO and show that the proposed method has good scalability to the maximum number of jamming stages.

Table 2. The average $E[\xi_{\varphi}^{(K)}]$, $E[\text{MD}_{\varphi}^{(K)}]$, and $E[r(\varphi)]$ performances of strategies obtained by the proposed method and $E[\xi_{\varphi}^{(K)}]$, $E[\text{MD}_{\varphi}^{(K)}]$, and $E[r(\varphi)]$ performances of the best UV-RGPO strategy and UA-RGPO strategy on the CV benchmark with various K .

K	Method	$E[\xi_{\varphi}^{(K)}]$	$E[\text{MD}_{\varphi}^{(K)}]$	$E[r(\varphi)]$
10	Proposed method	21.4%	140	0.415
10	UV-RGPO	17.7%	113	0.372
10	UA-RGPO	13.3%	126	0.382
15	Proposed method	31.8%	224	0.720
15	UV-RGPO	28.0%	153	0.524
15	UA-RGPO	21.0%	196	0.595
20	Proposed method	37.1%	310	0.920
20	UV-RGPO	35.4%	187	0.645
20	UA-RGPO	20.9%	301	0.856

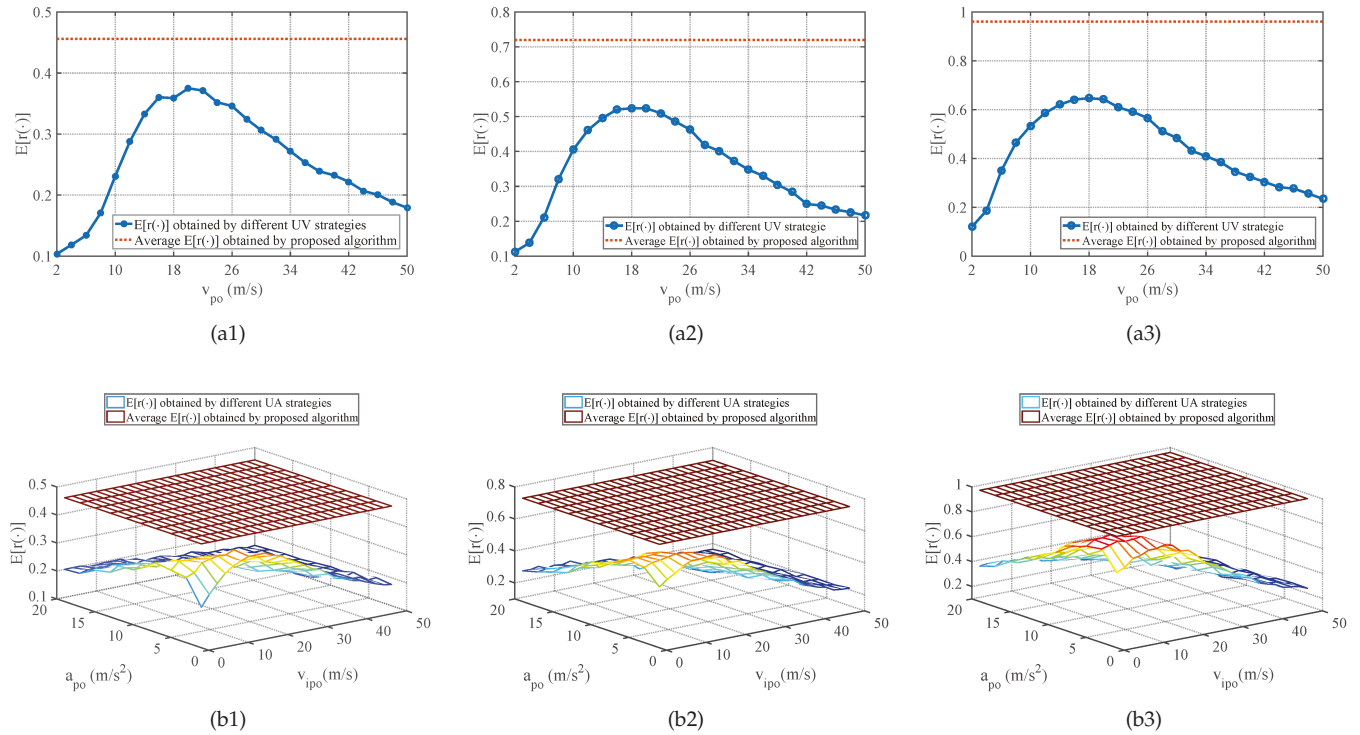


Figure 5. The average $E[r(\varphi)]$ performance of strategies obtained by the proposed method in comparison with the $E[r(\varphi)]$ performance of different UV-RGPO and UA-RGPO strategies on the CV benchmark with various K . (a1) Proposed method vs. UV-RGPO, $K = 10$; (b1) Proposed method vs. UA-RGPO, $K = 10$; (a2) Proposed method vs. UV-RGPO, $K = 15$; (b2) Proposed method vs. UA-RGPO, $K = 15$; (a3) Proposed method vs. UV-RGPO, $K = 20$; (b3) Proposed method vs. UA-RGPO, $K = 20$.

4.1.2. Effect of Scoring Mechanism

To show the effect of the scoring mechanism on the balance of the $E[\xi_\varphi^{(K)}]$ and $E[\text{MD}_\varphi^{(K)}]$ performances, we compare the proposed method with UV-RGPO and UA-RGPO on the CV benchmark using different objective function weights (w_1, w_2). The total numbers of jamming stages is $K = 15$. Similarly, we traverse v_{po} within $[v_{po}^{\min}, v_{po}^{\max}] = [2 \text{ m/s}, 50 \text{ m/s}]$ to find the best UV-RGPO strategy and traverse v_{ipo} within $[v_{ipo}^{\min}, v_{ipo}^{\max}] = [2 \text{ m/s}, 50 \text{ m/s}]$ and a_{po} within $[a_{po}^{\min}, a_{po}^{\max}] = [2 \text{ m/s}^2, 20 \text{ m/s}^2]$ to find the best UA-RGPO strategy. For the proposed method, 50 independent runs are performed, the PSO is initialized with $N = 40$ particles, the number of resampling for each candidate strategy is $J = 2000$, and the maximum number of iterations is $g_{\max} = 100$.

The average $E[r(\varphi)]$, $E[\xi_\varphi^{(K)}]$, and $E[\text{MD}_\varphi^{(K)}]$ performances of strategies obtained by the proposed method and the $E[r(\varphi)]$, $E[\xi_\varphi^{(K)}]$, and $E[\text{MD}_\varphi^{(K)}]$ performances of the best UV-RGPO strategy and UA-RGPO strategy are tabulated in Table 3. As shown in Table 3, with the increase of the weight w_1 , the $E[\xi_\varphi^{(K)}]$ performance of these three methods gradually improves, and $E[\text{MD}_\varphi^{(K)}]$ performance gradually decreases. That is because there is a certain degree of conflict between the $E[\xi_\varphi^{(K)}]$ and $E[\text{MD}_\varphi^{(K)}]$ performances for the same RGPO jamming strategy. Therefore, when the weight w_1 is larger and the weight w_2 is smaller, the generated jamming strategies prefer the $E[\xi_\varphi^{(K)}]$ performance, and vice versa, and the generated jamming strategies prefer the $E[\text{MD}_\varphi^{(K)}]$ performance. In practice, the weights w_1 and w_2 can be flexibly set according to task requirements. Another observation we can make is that the proposed method is better than UV-RGPO and UA-RGPO under different w_1 and w_2 . This observation further suggests the effectiveness of the proposed method.

Table 3. The average $E[\xi_\phi^{(K)}]$, $E[\text{MD}_\phi^{(K)}]$, and $E[r(\phi)]$ performances of strategies obtained by the proposed method and $E[\xi_\phi^{(K)}]$, $E[\text{MD}_\phi^{(K)}]$, and $E[r(\phi)]$ performances of the best UV-RGPO strategy and UA-RGPO strategy on the CV benchmark with various w_1 and w_2 .

w_1, w_2	Method	$E[\xi_\phi^{(K)}]$	$E[\text{MD}_\phi^{(K)}]$	$E[r(\phi)]$
0.25, 0.75/200	Proposed method	30.2%	226	0.920
0.25, 0.75/200	UV-RGPO	27.7%	153	0.642
0.25, 0.75/200	UA-RGPO	19.7%	203	0.811
0.5, 0.5/200	Proposed method	31.8%	224	0.720
0.5, 0.5/200	UV-RGPO	28.0%	153	0.524
0.5, 0.5/200	UA-RGPO	21.0%	196	0.595
0.75, 0.25/200	Proposed method	34.4%	204	0.513
0.75, 0.25/200	UV-RGPO	31.8%	143	0.421
e 0.75, 0.25/200	UA-RGPO	27.4%	173	0.422

4.2. Experiments on Different Benchmark Problems

For a more comprehensive evaluation of the proposed method, we test the proposed method on the other three benchmark problems (CA, CT, and RV). The total number of jamming stages is $K = 15$. To better verify the effectiveness, we compare the proposed method with UV-RGPO and the UA-RGPO. To be fair, the above three methods use the same settings for the weights of objective functions ($w_1 = 0.5, w_2 = 0.5/200$). For UV-RGPO, we traverse v_{po} within $[v_{po}^{\min}, v_{po}^{\max}] = [2 \text{ m/s}, 50 \text{ m/s}]$ to find the best strategy. For UA-RGPO, we traverse v_{ipo} within $[v_{ipo}^{\min}, v_{ipo}^{\max}] = [2 \text{ m/s}, 50 \text{ m/s}]$ and a_{po} within $[a_{po}^{\min}, a_{po}^{\max}] = [2 \text{ m/s}^2, 20 \text{ m/s}^2]$ to find the best strategy. For the proposed method, 50 independent runs are performed, the PSO is initialized with $N = 40$ particles, the number of resampling for each candidate strategy is $J = 2000$, and the maximum number of iterations is $g_{\max} = 100$.

The average $E[r(\phi)]$ performance of strategies obtained by the proposed method in comparison with the $E[r(\phi)]$ performance of different UV-RGPO and UA-RGPO strategies are shown in Figures 6–8. Moreover, the average $E[\xi_\phi^{(K)}]$ and $E[\text{MD}_\phi^{(K)}]$ performances of strategies obtained by the proposed method and the $E[\xi_\phi^{(K)}]$ and $E[\text{MD}_\phi^{(K)}]$ performances of the best UV-RGPO strategy and UA-RGPO strategy are tabulated in Table 4. From these results, we clearly see that the proposed method outperforms UV-RGPO and UA-RGPO on all four benchmarks (including the CV benchmark), despite the fact that the motion models of the real target and the tracking models of the threat radar vary between the benchmarks. The numerical result shows that the proposed method is not sensitive to the motion models of the real target and the tracking models of the threat radar.

Table 4. The average $E[\xi_\phi^{(K)}]$, $E[\text{MD}_\phi^{(K)}]$, and $E[r(\phi)]$ performances of strategies obtained by the proposed method and $E[\xi_\phi^{(K)}]$, $E[\text{MD}_\phi^{(K)}]$, and $E[r(\phi)]$ performances of the best UV-RGPO strategy and UA-RGPO strategy on different benchmarks.

Benchmark	Method	$E[\xi_\phi^{(K)}]$	$E[\text{MD}_\phi^{(K)}]$	$E[r(\phi)]$
CA	Proposed method	49.8%	365	1.16
CA	UV-RGPO	38.9%	223	0.752
CA	UA-RGPO	35.5%	330	1.00
CT	Proposed method	29.7%	208	0.668
CT	UV-RGPO	25.7%	148	0.499
CT	UA-RGPO	18.1%	190	0.568
RV	Proposed method	27.2%	174	0.570
RV	UV-RGPO	26.1%	124	0.441
RV	UA-RGPO	22.3%	148	0.482

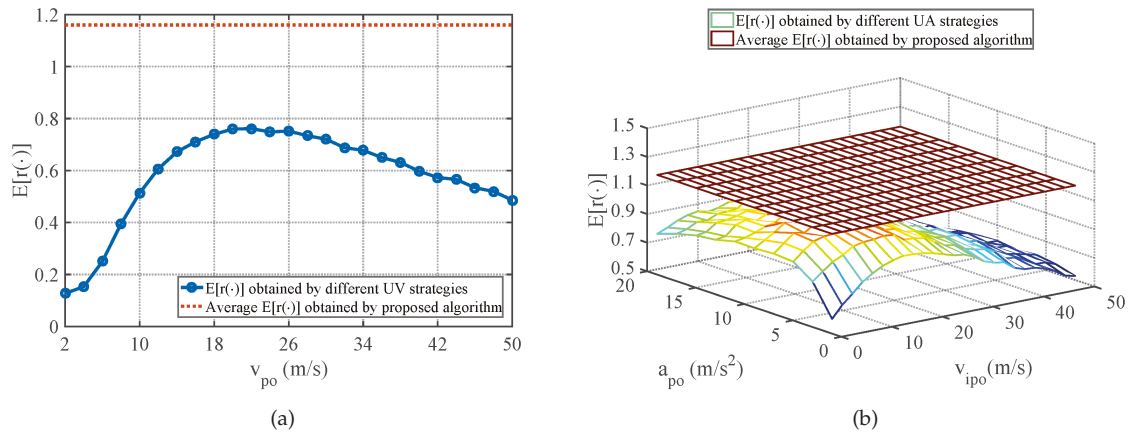


Figure 6. The average $E[r(\varphi)]$ performance of strategies obtained by the proposed method in comparison with the $E[r(\varphi)]$ performance of different UV-RGPO and UA-RGPO strategies on the CA benchmark. (a) Proposed method vs. UV-RGPO; (b) Proposed method vs. UA-RGPO.

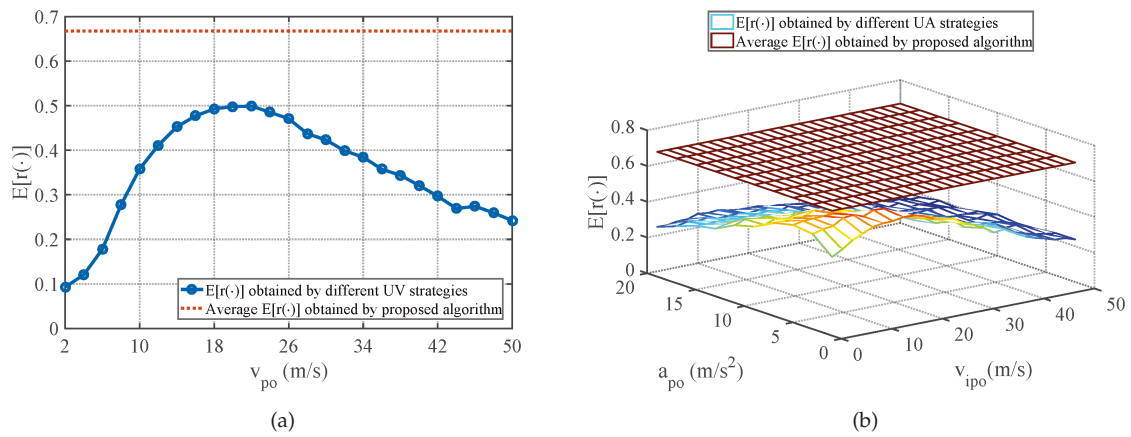


Figure 7. The average $E[r(\varphi)]$ performance of strategies obtained by the proposed method in comparison with the $E[r(\varphi)]$ performance of different UV-RGPO and UA-RGPO strategies on the CT benchmark. (a) Proposed method vs. UV-RGPO; (b) Proposed method vs. UA-RGPO.

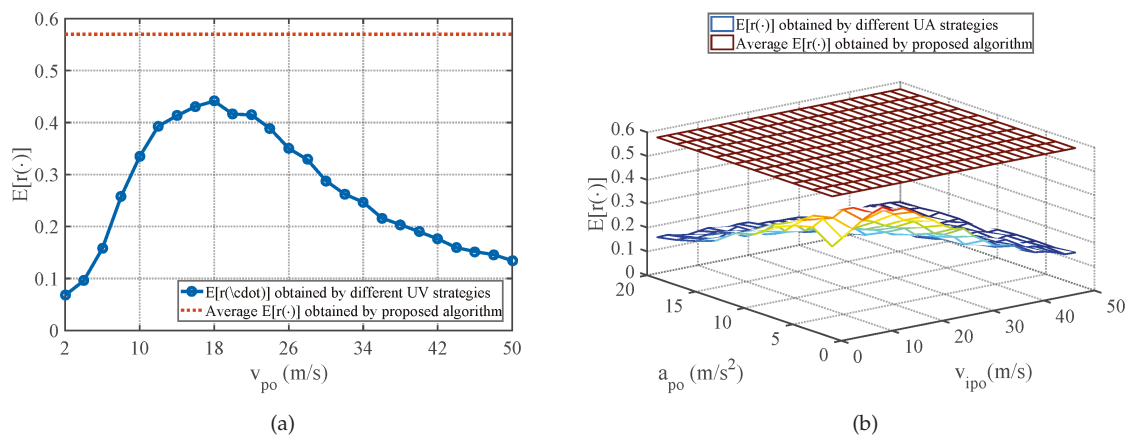


Figure 8. The average $E[r(\varphi)]$ performance of strategies obtained by the proposed method in comparison with the $E[r(\varphi)]$ performance of different UV-RGPO and UA-RGPO strategies on the RV benchmark. (a) Proposed method vs. UV-RGPO; (b) Proposed method vs. UA-RGPO.

5. Conclusions

This paper addresses the strategy optimization problem for white-box RGPO jamming. The strategy optimization of the RGPO jamming has been formulated as an ASSO problem, and a white-box RGPO jamming strategy optimization algorithm based on PSO-ER with local simulation assistance has been presented. In this algorithm, PSO-ER is utilized to identify the optimal jamming strategy, while a locally constructed simulation of the tracking radar provides performance evaluations for candidate strategies. Experimental findings demonstrate that the proposed approach effectively generates a well-tuned strategy for white-box RGPO jamming. Overall, the research in this paper lays a foundation for the optimization model and the main framework of algorithm for subsequent related research.

Author Contributions: Conceptualization Y.W. and Y.H.; methodology, Y.W.; software and validation, Y.W. and Y.J.; writing—original draft preparation, Y.W.; writing—review and editing, Y.W. and Y.H.; supervision, Y.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the Fundamental Research Funds for the Central Universities under Grant DUT24RC(3)052.

Data Availability Statement: Data are contained within the article. All data in this paper are generated by simulation, and the details have been presented in Section 4.

Conflicts of Interest: The authors declare no conflicts of interest. The funder had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Li, N.; Zhang, Y. A survey of radar ECM and ECCM. *IEEE Trans. Aerosp. Electron. Syst.* **1995**, *31*, 1110–1120. [CrossRef]
2. Li, Y.; Shi, C.; Yan, M.; Zhou, J. Mission Planning and Trajectory Optimization in UAV Swarm for Track Deception against Radar Network. *Remote Sens.* **2024**, *16*, 3490. [CrossRef]
3. Adamy, D.L. *EW 101: A First Course in Electronic Warfare*; Artech House: Norwood, MA, USA, 2001.
4. Adamy, D.L. *EW 104: Electronic Warfare Against a New Generation of Threats*; Artech House: Norwood, MA, USA, 2015.
5. De Martino, A. *Introduction to Modern EW Systems*; Artech House: Norwood, MA, USA, 2018.
6. Blair, W.; Watson, G.; Kirubarajan, T.; Bar-Shalom, Y. Benchmark for radar allocation and tracking in ECM. *IEEE Trans. Aerosp. Electron. Syst.* **1998**, *34*, 1097–1114. [CrossRef]
7. Kirubarajan, T.; Bar-Shalom, Y.; Blair, W.; Watson, G. IMMPPDAF for radar management and tracking benchmark with ECM. *IEEE Trans. Aerosp. Electron. Syst.* **1998**, *34*, 1115–1134. [CrossRef]
8. Han, Q.; Long, W.; Yang, Z.; Dong, X.; Chen, J.; Wang, F.; Liang, Z. Resource Allocation of Netted Opportunistic Array Radar for Maneuvering Target Tracking under Uncertain Conditions. *Remote Sens.* **2024**, *16*, 3499. [CrossRef]
9. Zhang, G.; Xie, J.; Zhang, H.; Feng, W.; Liu, M.; Qin, C. Power Allocation Scheme for Multi-Static Radar to Stably Track Self-Defense Jammers. *Remote Sens.* **2024**, *16*, 2699. [CrossRef]
10. Wang, Y.; Zhang, T.; Kong, L.; Ma, Z. A Stochastic Simulation Optimization based Range Gate Pull-Off Jamming Method. *IEEE Trans. Evol. Comput.* **2022**, *27*, 580–594. [CrossRef]
11. Amaran, S.; Sahinidis, N.V.; Sharda, B.; Bury, S.J. Simulation optimization: A review of algorithms and applications. *Ann. Oper. Res.* **2016**, *240*, 351–380. [CrossRef]
12. Liu, Y.; Pedrielli, G.; Li, H.; Lee, L.H.; Chen, C.H.; Shortle, J.F. Optimal Computing Budget Allocation for N - k Stochastic Problem in the Power Grid System. *IEEE Trans. Rel.* **2019**. [CrossRef]
13. Goodwin, T.; Xu, J.; Celik, N.; Chen, C.H. Real-time digital twin-based optimization with predictive simulation learning. *J. Simul.* **2022**, *18*, 47–64. [CrossRef]
14. Zhu, C.; Xu, J.; Chen, C.H.; Lee, L.H.; Hu, J.Q. Balancing Search and Estimation in Random Search Based Stochastic Simulation Optimization. *IEEE Trans. Autom. Control.* **2016**, *61*, 3593–3598. [CrossRef]
15. Horng, S.C.; Yang, F.Y.; Lin, S.S. Applying PSO and OCBA to Minimize the Overkills and Re-Probes in Wafer Probe Testing. *IEEE Trans. Semicond. Manuf.* **2012**, *25*, 531–540. [CrossRef]
16. Barreno, M.; Nelson, B.; Joseph, A.D.; Tygar, J.D. The security of machine learning. *Mach. Learn.* **2010**, *81*, 121–148. [CrossRef]
17. Barreno, M.; Nelson, B.; Sears, R.; Joseph, A.D.; Tygar, J.D. Can machine learning be secure? In Proceedings of the 2006 ACM Symposium on Information, computer and communications security, Taipei, Taiwan, 21–24 March 2006; pp. 16–25.

18. Shen, J.; Robertson, N. BBAS: Towards large scale effective ensemble adversarial attacks against deep neural network learning. *Inf. Sci.* **2021**, *569*, 469–478. [CrossRef]
19. Xiao, Y.; Pun, C.M. Improving adversarial attacks on deep neural networks via constricted gradient-based perturbations. *Inf. Sci.* **2021**, *571*, 104–132. [CrossRef]
20. Sadeghi, M.; Larsson, E.G. Adversarial Attacks on Deep-Learning Based Radio Signal Classification. *IEEE Wirel. Commun. Lett.* **2018**, *8*, 213–216. [CrossRef]
21. Wei, X.; Guo, Y.; Li, B. Black-box adversarial attacks by manipulating image attributes. *Inf. Sci.* **2021**, *550*, 285–296. [CrossRef]
22. Lin, Y.; Zhao, H.; Tu, Y.; Mao, S.; Dou, Z. Threats of Adversarial Attacks in DNN-Based Modulation Recognition. In Proceedings of the IEEE INFOCOM 2020—IEEE Conference on Computer Communications, Toronto, ON, Canada, 6–9 July 2020; pp. 2469–2478. [CrossRef]
23. Xu, J.; Huang, E.; Chen, C.H.; Lee, L.H. Simulation optimization: A review and exploration in the new era of cloud computing and big data. *Asia-Pac. J. Oper. Res.* **2015**, *32*. [CrossRef]
24. Hassan, R.; Cohanin, B.; De Weck, O.; Venter, G. A comparison of particle swarm optimization and the genetic algorithm. In Proceedings of the 46th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, Austin, TX, USA, 18–21 April 2005; pp. 1–13.
25. Rada-Vilela, J.; Zhang, M.; Johnston, M. Resampling in particle swarm optimization. In Proceedings of the 2013 IEEE Congress on Evolutionary Computation, Cancun, Mexico, 20–23 June 2013; pp. 947–954.
26. Pan, H.; Wang, L.; Liu, B. Particle swarm optimization for function optimization in noisy environment. *Appl. Math. Comput.* **2006**, *181*, 908–919. [CrossRef]
27. Bartz-Beielstein, T.; Blum, D.; Branke, J. Particle swarm optimization and sequential sampling in noisy environments. In *Metaheuristics*; Springer: New York, NY, USA, 2007; Volume 39, pp. 261–273.
28. Bar-Shalom, Y.; Blair, W. *Multitarget-Multisensor Tracking: Applications and Advances*; Artech House: Norwood, MA, USA, 2000; Volume III.
29. Liggins II, M.; Hall, D.; Llinas, J. *Handbook of Multisensor Data Fusion: Theory and Practice*; CRC Press: Boca Raton, FL, USA, 2017.
30. Kershaw, D.; Evans, R. A contribution to performance prediction for probabilistic data association tracking filters. *IEEE Trans. Aerosp. Electron. Syst.* **1996**, *32*, 1143–1148. [CrossRef]
31. Morrice, D.J.; Butler, J.; Mullarkey, P.W. An approach to ranking and selection for multiple performance measures. In Proceedings of the 1998 Winter Simulation Conference. Proceedings (Cat. No.98CH36274), Washington, DC, USA, 13–16 December 1998; Volume 1, pp. 719–725.
32. Butler, J.; Morrice, D.J.; Mullarkey, P.W. A multiple attribute utility theory approach to ranking and selection. *Manag. Sci.* **2001**, *47*, 800–816. [CrossRef]
33. Swisher, J.R.; Jacobson, S.H. Evaluating the design of a family practice healthcare clinic using discrete-event simulation. *Health Care Manag. Sci.* **2002**, *5*, 75–88.
34. Eberhart, R.C.; Shi, Y. Comparing inertia weights and constriction factors in particle swarm optimization. In Proceedings of the 2000 Congress on Evolutionary Computation. CEC00 (Cat. No.00TH8512), La Jolla, CA, USA, 16–19 July 2000; Volume 1, pp. 84–88.
35. Chatterjee, A.; Siarry, P. Nonlinear inertia weight variation for dynamic adaptation in particle swarm optimization. *Comput. Oper. Res.* **2006**, *33*, 859–871. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Node Selection and Path Optimization for Passive Target Localization via UAVs

Xiaoyou Xing, Zhiwen Zhong, Xueting Li * and Yiyang Yue

School of Aeronautics and Astronautics, Sichuan University, Chengdu 610065, China;
2022141510062@stu.scu.edu.cn (X.X.); zhongzhiwen@stu.scu.edu.cn (Z.Z.); 2022141510057@stu.scu.edu.cn (Y.Y.)

* Correspondence: lixueting@scu.edu.cn

Abstract: The performance of passive target localization is affected by the positions of unmanned aerial vehicles (UAVs) at a large scale. In this paper, to improve resource utilization efficiency and localization accuracy, the node selection problem and the path optimization problem are jointly investigated. Firstly, the target passive localization model is established and the Chan-based time difference of arrival (TDOA) localization method is introduced. Then, the Cramer–Rao lower bound (CRLB) for Chan-TDOA localization is derived, and the problems of node selection and path optimization are formulated. Secondly, a CRLB-based node selection method is proposed to properly divide the UAVs into several groups, localizing different targets, and a CRLB-based path optimization method is proposed to search for the optimal UAV position configuration at each time step. The proposed path optimization method also effectively handles no-fly-zone (NFZ) constraints, ensuring operational safety while maintaining optimal target tracking performance. Also, to improve the efficiency of path optimization, particle swarm algorithm (PSO) is applied to accelerate the searching process. Finally, numerical simulations are performed to verify the validity and effectiveness of the proposed methods in this paper.

Keywords: passive target localization; unmanned aerial vehicles (UAVs); Cramer–Rao lower bound (CRLB); node selection method; path optimization method; particle swarm algorithm (PSO)

1. Introduction

Unmanned Aerial Vehicle swarms (UAVs) have indeed become integral components of modern military operations, particularly in real-time aerial reconnaissance, surveillance, and various tactical applications. The versatility and technological capabilities provide numerous advantages, and their adoption has significantly transformed the landscape of modern warfare [1,2]. Recently, UAVs have garnered significant attention for their enhanced capabilities in complex missions such as target tracking, localization, and detection [3–5]. The distributed nature of UAVs provides robust and flexible operations, particularly in challenging tactical environments.

Generally, UAVs can be burdened by active systems or passive systems to fulfill their tasks. For active localization systems, the locator (such as a radar) actively emits signals to detect and determine the position of targets. While this can provide real-time tracking and high precision, it exposes the location of the locator itself to enemy detection. This makes active localization vulnerable in hostile environments, especially in situations where stealth and secrecy are crucial. Additionally, active localization systems tend to require higher energy consumption and more complex infrastructure, which can be a logistical

challenge [6,7]. On the other hand, a passive localization system does not involve emitting signals. Instead, it works by detecting and analyzing the electromagnetic waves emitted by the target's own electronic devices. The key advantage is that it does not disclose the location of the system receiving the signals, making it inherently more stealthy and harder to detect [8–10]. In this case, UAVs' passive localization technology offers significant strategic advantages in situations where stealth, resource efficiency, and survivability are crucial. Its ability to detect targets without emitting signals makes it a powerful tool in modern military operations [11,12].

1.1. Overview

For multi-target passive localization via UAVs, node selection and path optimization are crucial to enhance the accuracy and efficiency of target localization, since through node selection, the UAVs could be properly divided into several groups to locate different targets, and through path optimization, the UAVs could achieve proper positions at each frame, obtaining appropriate observation angles and high-quality measurements that help realize high localization accuracy. Therefore, dynamic node selection and path optimization could help ensure that UAVs are optimally positioned to minimize errors in localization, improving the ability of the UAVs in locating a target.

From the perspective of node selection for the UAVs, node selection could be implemented via classic greedy algorithms to construct minimum spanning trees (MSTs) to select a set of UAVs that minimizes the total communication or distance between them [13,14]. Task scheduling also needs to select a set of UAV nodes to optimize task scheduling and load balancing in distributed UAV systems [15,16].

More recently, reinforcement learning (RL) approaches have shown promising results in UAV path planning and navigation tasks. Deep reinforcement learning (DRL) methods, particularly algorithms like deep Q-networks (DQNs) and proximal policy optimization (PPO), have been successfully applied to autonomous navigation in complex environments [17]. These methods enable UAVs to learn optimal navigation policies through interactions with the environment without requiring complete prior knowledge of the operational space. Particularly noteworthy is the development of cross-platform deep reinforcement learning approaches for autonomous navigation without global localizing information [18], which have demonstrated remarkable adaptability across different scenarios and vehicle platforms. Such end-to-end learning approaches that combine local sensor information with learned policies have significantly advanced the field of autonomous navigation [19].

Additionally, ant colony optimization (ACO) could be deployed for node selection to determine optimal routes for the communication relay of coverage tasks, ensuring minimal energy expenditure while still fulfilling the mission's objectives [20,21], and for a UAV network, the central leader node selection is essential for coordinating tasks and managing communications [22].

From the perspective of path optimization for UAVs, the heuristic-based A* algorithm is used for path optimization to efficiently find the shortest path [23]. The genetic algorithm (GA) is introduced to cooperative path planning for UAVs. It uses a population of potential solutions (paths), which evolve over generations according to a fitness function, and the algorithm incorporates crossovers with mutation operations to explore the solution space while the fitness function guides the search toward optimal or near-optimal paths that avoid collisions and optimize mission objectives [24]. The differential evolution (DE) algorithm is well suited for UAV path optimization because of its ability to efficiently search high-dimensional spaces while minimizing time and energy consumption [25].

The bat algorithm adapts well to dynamic environments and is used to optimize UAV trajectories in the presence of changing obstacles or mission requirements [26].

A projection algorithm has been deployed to improve the localization accuracy of UAVs, and the algorithm guides UAVs toward way points that are closer to the target. This approach reduces the computational burden associated with numerical searches by projecting UAVs onto more promising paths based on predictive models, thus improving the efficiency and effectiveness of the path optimization process [27]. Although this algorithm considers the application of target passive localization, it only considers one target.

Concerning existing works, it can be seen that in most works, node selection and path optimization are separately investigated, which restricts performance improvements in localization accuracy for multi-targets. Also, node selection methods in existing works are mainly concerned with task scheduling, resources distribution, and communication link construction, and the path optimization methods in the existing works are mainly concerned with the shortest path searching and obstacle avoidance operations with less resource consumption, which ignore the application multi-target passive localization. Additionally, existing path optimization methods rarely consider the integration of no-fly-zone constraints with localization performance optimization, which is crucial for practical UAV operations in restricted airspace. Furthermore, while reinforcement learning approaches have shown promising results in UAV navigation, they often require extensive training data and computational resources, and their performance in multi-target passive localization applications remains largely unexplored. The existing deep learning models for autonomous navigation typically focus on single-platform scenarios and may not effectively accommodate the complex requirements of coordinated UAVs operations for passive localization.

1.2. Original Contributions

Obviously, works that focus on multi-target passive localization while jointly considering the node selection and the path optimization are rare, and these are indeed the main topics of this paper. In this paper, we jointly consider the problem of node scheduling and path optimization, and we propose a CRLB-based node selection method and a CRLB-based path optimization method to improve localization accuracies. The original contributions of this paper are as follows:

(1) We propose a novel joint optimization framework that unifies node selection and path optimization for multi-target passive localization. Unlike existing methods that handle these problems separately, our integrated approach simultaneously optimizes resource allocation and path planning, achieving superior performance in both resource utilization and localization accuracy.

(2) We develop a CRLB-based node selection method that effectively partitions UAVs into optimal groups for different target localization. This method uniquely considers the theoretical lower bound of estimation errors to guide grouping decisions, leading to improved resource efficiency and localization performance.

(3) We design a CRLB-based path optimization method that dynamically determines optimal UAVs positions by minimizing the theoretical estimation error bound. This approach enables more accurate target localization performance compared to traditional fixed-formation methods.

(4) We implement the PSO algorithm for UAV path optimization. This implementation significantly improves computational efficiency while maintaining solution quality compared to several existing optimization methods.

1.3. Organization

The main contents of this paper are shown as follows. The problems of node selection and path optimization are formulated in Section 2. The algorithm of the node selection based on CRLB is given in Section 3, and the algorithm of the path optimization based on CRLB is given in Section 4. Numerical simulations are given in Section 5. This paper is concluded in Section 6.

2. Problem Formulation

In this section, we present a comprehensive formulation of the multi-target passive localization problem using UAVs based on the Chan-TDOA method. First, we establish the fundamental Chan-TDOA localization model. Then, we derive the CRLB for measurement errors in the Chan-TDOA model to evaluate the theoretical lower bound of the localization performance. Finally, we formulate two key optimization problems: the node selection problem for optimal UAVs grouping and the path optimization problem for determining optimal UAVs positions, which together form a complete framework for multi-target passive localization via UAVs. The mathematical models and theoretical analyses presented in this section serve as the foundation for the algorithms developed in subsequent sections.

2.1. Target Localization Model Based on the Chan-TDOA Algorithm

In sight of the fact that target localization is designed to determine target positions based on measurements, we first establish the fundamental TDOA model. Consider a two-dimensional passive localization scenario with multiple UAVs and targets, as shown in Figure 1.

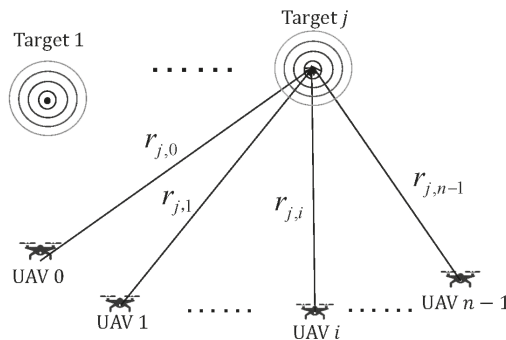


Figure 1. Geometric schematic of the TDOA localization model. The distance between target j and each UAV is denoted as $r_{j,i}$, where UAV 0 serves as the reference node (or master UAV), and UAV i ($i = 0, 1, \dots, n-1$) represents other nodes in the swarm. The concentric circles around the target indicate the signal propagation.

Assume that there are m stationary targets, and the j th target has coordinates $\mathbf{w}_j = [x_j, y_j]^T$, where $j = 1, 2, \dots, m$. A swarm of n UAVs equipped with passive radars is deployed, and their coordinates are given by $\mathbf{u}_i = [x_i, y_i]^T$, where $i = 0, 1, \dots, n-1$. Without loss of generality, UAV 0 is selected as the reference node (or master UAV).

The distance between the i th UAV and j th target can be expressed as

$$r_{j,i} = \|\mathbf{u}_i - \mathbf{w}_j\| = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}. \quad (1)$$

For TDOA measurements, let $\tau_{j,i}$ represent the time taken for the signal from target j to reach UAV i . The time difference of arrival between UAV i and the master UAV (UAV 0) is denoted as $\tau_{j,i,0}$, where the subscript triplet $(j, i, 0)$ has the following meaning: The first subscript j indicates the target being localized, the second subscript i represents

the measuring UAV, and the third subscript 0 refers to the master UAV (UAV 0). Thus, $\tau_{j,i,0} = \tau_{j,i} - \tau_{j,0}$ represents the time difference between when UAV i and UAV 0 receive the signal from target j , which can be written as

$$\tau_{j,i,0} = \tau_{j,i} - \tau_{j,0} = \frac{r_{j,i} - r_{j,0}}{c}, \quad (2)$$

where c is the propagation speed of electromagnetic waves.

In practical applications, the measured TDOA value $\tau'_{j,i,0}$ inevitably contains measurement errors:

$$\tau'_{j,i,0} = \tau_{j,i,0} + \Delta\tau_{j,i,0}, \quad (3)$$

where $\tau'_{j,i,0}$ represents the measured time difference between UAV i and UAV 0 for target j , $\tau_{j,i,0}$ is the true time difference, and $\Delta\tau_{j,i,0}$ denotes the time difference error, and it is typically modeled as zero-mean Gaussian noise:

$$\Delta\tau_{j,i,0} \sim \mathcal{N}(0, \sigma_{i,j,0}^2), \quad (4)$$

where $\sigma_{j,i,0}^2 = \sigma_{j,i}^2 + \sigma_{j,0}^2$ is the variance of the TDOA measurement for target j , with $\sigma_{j,i}^2$ and $\sigma_{j,0}^2$ being the TDOA measurement variances in UAV i and UAV 0 respectively.

Several factors contribute to the TDOA measurement errors. The internal electronic noise in the radar receiver affects TDOA measurement precision. Signal reflection and refraction during propagation can introduce additional errors. Variations in atmospheric conditions, such as temperature and humidity, impact the propagation speed. The radar's time resolution also imposes limitations on measurement capabilities. Also, the standard deviation of the TDOA time measurement $\sigma_{j,i}$ between UAV i and target j exhibits dependencies on their separation distance:

$$\sigma_{j,i} = \kappa r_{j,i}^\nu, \quad (5)$$

where κ and ν are system parameters determined by the radar characteristics and propagation environment. The parameter ν is typically close to $1/2$, reflecting the relationship between signal power attenuation and distance.

For target localization in two dimensions, at least three UAVs are required to obtain sufficient TDOA measurements. When the number of UAVs $n \geq 3$, the TDOA measurement equations can be used to construct a nonlinear system as follows:

$$\begin{cases} c\tau'_{j,1,0} = \|\mathbf{u}_1 - \mathbf{w}_j\| - \|\mathbf{u}_0 - \mathbf{w}_j\| \\ c\tau'_{j,2,0} = \|\mathbf{u}_2 - \mathbf{w}_j\| - \|\mathbf{u}_0 - \mathbf{w}_j\| \\ \vdots \\ c\tau'_{j,n-1,0} = \|\mathbf{u}_{n-1} - \mathbf{w}_j\| - \|\mathbf{u}_0 - \mathbf{w}_j\|. \end{cases} \quad (6)$$

To accelerate the calculation process of nonlinear equations, the Chan algorithm is introduced in this paper to provide an efficient closed-form solution through linearization and two-step weighted least squares estimation. The Chan algorithm is a non-recursive solver for systems of hyperbolic equations that can transform an initially nonlinear system into a linear one. Under the assumption that the TDOA measurement errors follow an ideal Gaussian distribution, the Chan-TDOA localization algorithm can provide high localization accuracies with reduced computational complexity.

For target localization in two dimensions with multiple UAVs ($n \geq 3$), the Chan-TDOA algorithm provides an efficient closed-form solution through linearization and two-step weighted least squares (WLS) estimation. The algorithm first transforms the nonlinear

TDOA equations into linear forms through algebraic manipulation. Then, it employs a two-step WLS estimation process to obtain the target location estimate, with the first step providing an initial solution and the second step refining the estimate. The detailed mathematical derivations of the algorithm are presented in Appendix A. The overall flow of the algorithm is illustrated in Figure 2.

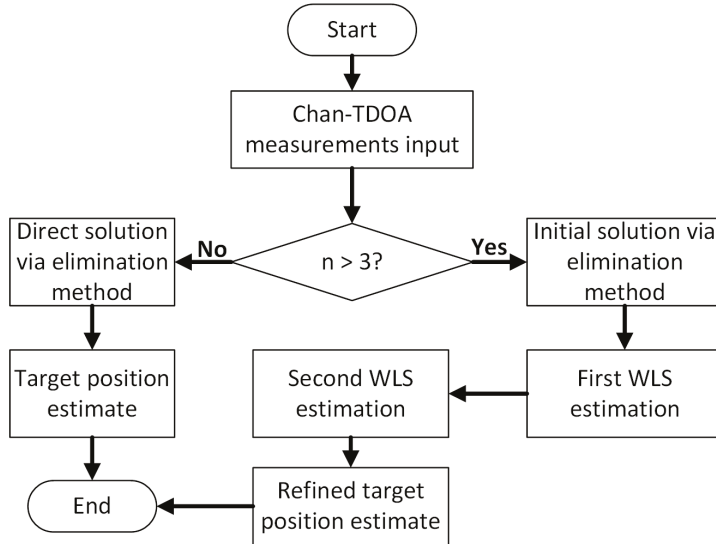


Figure 2. Flow chart of the Chan-TDOA algorithm.

2.2. CRLB Derivation for TDOA Measurement Errors in the Chan-TDOA Model

The CRLB is an important metric for evaluating parameter estimation performances, representing the smallest possible variance achieved via an unbiased estimator. It is commonly used as a benchmark for comparing the performance of unbiased estimators.

In the context of the Chan-TDOA passive target localization problem discussed in this paper, the observation data of the localization system consists of the TDOA measurements between each UAV and the target. Let the measured TDOA values be denoted as s , with the true values given by

$$\mathbf{s} = c[\tau_{1,0}, \tau_{2,0}, \dots, \tau_{n-1,0}]^T, \quad (7)$$

where $\tau_{i,0} = \tau_i - \tau_0$ represents the difference in the times at which the signal was received by the i th UAV and the master base station. The measurement is given by

$$\mathbf{s}' = \mathbf{s} + \mathbf{m}. \quad (8)$$

The TDOA measurement error $\mathbf{m}$ follows a normal distribution $\mathcal{N}(\mathbf{0}, \mathbf{Q})$, where $\mathbf{Q}$ is the covariance matrix of the TDOA measurements, as given in (A13). The probability density function of the TDOA measurements can be expressed as

$$p(\mathbf{s}'; \mathbf{s}) = \frac{1}{\sqrt{2\pi|\mathbf{Q}|}} \exp\left(-\frac{1}{2}(\mathbf{s}' - \mathbf{s})^T \mathbf{Q}^{-1}(\mathbf{s}' - \mathbf{s})\right). \quad (9)$$

The likelihood function calculated from this density function is as follows:

$$\ln p(\mathbf{s}'; \mathbf{s}) = -\frac{1}{2} \ln |2\pi\mathbf{Q}| - \frac{1}{2}(\mathbf{s}' - \mathbf{s})^T \mathbf{Q}^{-1}(\mathbf{s}' - \mathbf{s}). \quad (10)$$

Taking the second partial derivatives of (10) with respect to the position's coordinate parameters, we have

$$\frac{\partial^2 \ln p(s'; s)}{\partial x^2} = -\frac{1}{2} \left[\left(-\frac{\partial^2 s}{\partial x^2} \right)^T Q^{-1} (s' - s) - (s' - s)^T Q^{-1} \frac{\partial^2 s}{\partial x^2} + 2 \left(\frac{\partial s}{\partial x} \right)^T Q^{-1} \frac{\partial s}{\partial x} \right], \quad (11)$$

$$\frac{\partial^2 \ln p(s'; s)}{\partial x \partial y} = -\frac{1}{2} \left[\left(-\frac{\partial^2 s}{\partial x \partial y} \right)^T Q^{-1} (s' - s) - (s' - s)^T Q^{-1} \frac{\partial^2 s}{\partial x \partial y} + 2 \left(\frac{\partial s}{\partial x} \right)^T Q^{-1} \frac{\partial s}{\partial y} \right], \quad (12)$$

$$\frac{\partial^2 \ln p(s'; s)}{\partial y^2} = -\frac{1}{2} \left[\left(-\frac{\partial^2 s}{\partial y^2} \right)^T Q^{-1} (s' - s) - (s' - s)^T Q^{-1} \frac{\partial^2 s}{\partial y^2} + 2 \left(\frac{\partial s}{\partial y} \right)^T Q^{-1} \frac{\partial s}{\partial y} \right]. \quad (13)$$

The expectations of Equations (11)–(13) are as follows:

$$E \left[\frac{\partial^2 \ln p(s'; s)}{\partial x^2} \right] = - \left[\left(\frac{\partial s}{\partial x} \right)^T Q^{-1} \frac{\partial s}{\partial x} \right], \quad (14)$$

$$E \left[\frac{\partial^2 \ln p(s'; s)}{\partial x \partial y} \right] = - \left[\left(\frac{\partial s}{\partial x} \right)^T Q^{-1} \frac{\partial s}{\partial y} \right], \quad (15)$$

$$E \left[\frac{\partial^2 \ln p(s'; s)}{\partial y^2} \right] = - \left[\left(\frac{\partial s}{\partial y} \right)^T Q^{-1} \frac{\partial s}{\partial y} \right]. \quad (16)$$

Then, the Fisher information matrix FIM(s) for the localization parameters can be denoted as

$$\text{FIM}(s) = \begin{bmatrix} E \left[\frac{\partial^2 \ln p(s'; s)}{\partial x^2} \right] & E \left[\frac{\partial^2 \ln p(s'; s)}{\partial x \partial y} \right] \\ E \left[\frac{\partial^2 \ln p(s'; s)}{\partial x \partial y} \right] & E \left[\frac{\partial^2 \ln p(s'; s)}{\partial y^2} \right] \end{bmatrix} = \begin{bmatrix} \left(\frac{\partial s}{\partial x} \right)^T Q^{-1} \frac{\partial s}{\partial x} & \left(\frac{\partial s}{\partial x} \right)^T Q^{-1} \frac{\partial s}{\partial y} \\ \left(\frac{\partial s}{\partial x} \right)^T Q^{-1} \frac{\partial s}{\partial y} & \left(\frac{\partial s}{\partial y} \right)^T Q^{-1} \frac{\partial s}{\partial y} \end{bmatrix}. \quad (17)$$

Thus, the CRLB matrix for target position estimations could be obtained by taking the inverse of the Fisher information matrix FIM(s)

$$\text{CRLB}(s) = \text{FIM}^{-1}(s). \quad (18)$$

2.3. Optimization Problem Formation

In this paper, we formulate two optimization problems for multi-target passive localization via UAVs: the node selection problem and the path optimization problem.

For a swarm of n UAVs and m targets, the node selection problem aims to optimally divide the n UAVs into m groups to minimize the average trace of the derived CRLB for each group's UAVs localization results at this moment:

$$\min_{\mathbf{G}_1, \dots, \mathbf{G}_m} \frac{1}{m} \sum_{j=1}^m \text{Tr}(\text{CRLB}_t)_j, \quad (19)$$

where $\mathbf{G}_j$ represents the coordinate matrix of all UAVs assigned to target j , and $\text{Tr}(\text{CRLB}_t)_j$ is the trace of the CRLB matrix of the localization for the j th target at the current time step.

For each group of UAVs, the path optimization problem aims to determine the optimal position for each UAV at the next time step to minimize the trace of the derived CRLB, while satisfying the dynamic constraints for the UAV constraints:

$$\min_{\mathbf{G}_j^{t+1}} \text{Tr}(\text{CRLB}_{t+1})_j, \quad (20)$$

where $\mathbf{G}_{i,j}^{t+1}$ represents the coordinate matrix of all UAVs assigned to target j at the next time step, and $\text{Tr}(\text{CRLB}_{t+1})_j$ is the prediction trace of the CRLB matrix of the localization result for the target j at the next time step.

In the following sections, we will introduce the detailed process of the proposed CRLB-based node selection method and the proposed CRLB-based path optimization method. Then, to accelerate the searching process for the optimal UAV positions, the PSO algorithm is introduced and applied to solve the path optimization problem.

3. CRLB-Based Node Selection Method

In this section, a CRLB-based node selection method is proposed to optimally divide UAVs into groups for localizing different targets, improving localization accuracies. Firstly, we describe the node selection problem and give the constraints of the problem. Then, the optimization model is formulated based on the derived CRLB, and the exhaustive method is deployed to minimize the average CRLB trace and obtain the optimal grouping scheme.

3.1. Problem Description

Consider a surveillance scenario with m stationary targets in a two-dimensional plane. n UAVs equipped with passive radars need to be divided properly into several groups for multi-target localization, and the number of UAVs of each groups for the j th target is $\lambda_j (j = 1, 2, \dots, m)$. The random assignment of UAVs may lead to suboptimal resource utilization and degraded localization accuracy. Therefore, we develop a systematic method for UAV grouping that ensures optimal performance.

3.2. Constraint Analysis

Since the Chan-TDOA method needs at least 3 UAVs to locate a target in a two-dimensional plane. Thus, the number of UAVs of each group should be more than 3. That is,

$$\lambda_j \geq 3. \quad (21)$$

3.3. Node Selection Optimization Model

The objective of the node selection problem is to minimize the average CRLB trace across all targets, and the mathematical model can be denoted as follows

$$\begin{aligned} \min_{\mathbf{G}_1, \dots, \mathbf{G}_m} \quad & \sum_{j=1}^m \text{Tr}(\text{CRLB}_t)_j \\ \text{subject to} \quad & \lambda_j \geq 3 \end{aligned} \quad (22)$$

where $\mathbf{G}_j (j = 1, 2, \dots, m)$ represents the set of UAVs assigned to target j , and $\text{Tr}(\text{CRLB}_t)_j$ denotes the trace of the CRLB matrix for target j at the current time step.

3.4. Node Selection Algorithm Design

The CRLB-based node selection algorithm provides a systematic approach for optimally grouping UAVs to locate different targets while minimizing localization errors. The algorithm employs a two-stage exhaustive search method: Firstly, it enumerates all possible combinations to partition n UAVs into m groups, where each group must contain at least 3 UAVs ($\lambda_j \geq 3$) to ensure effective target localization. For each configuration in this solution space, the algorithm computes the CRLB matrices for all targets and calculates their average trace as a metric of localization accuracy, selecting the one that minimizes this metric as the globally optimal solution. The detailed implementation of the node selection method is presented in Algorithm 1.

Algorithm 1: CRLB-based node selection algorithm

Input: Number and coordinates of the UAVs ($n, \mathbf{U}$), number and coordinates of the targets ($m, \mathbf{W}$), where $\mathbf{U} = [\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_{n-1}]^T \in \mathbb{R}^{n \times 2}$, $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m]^T \in \mathbb{R}^{m \times 2}$

Initialize system parameters with the positions of the UAVs and the targets

Verify the system constraints satisfaction

for each scheme do

Calculate the CRLB matrices for all the targets based on the present grouping scheme

Compute the average trace $\frac{1}{m} \sum_{j=1}^m \text{Tr}(\text{CRLB}_t)_j$

end

Select the scheme with the minimum average trace as the optimal scheme

Output: UAV grouping scheme

4. CRLB-Based Path Optimization Method

In this section, we presents a CRLB-based path optimization method to dynamically adjust the UAVs to the optimal positions, enhancing the target localization accuracy. For the sake of a clear presentation in this section, we consider path optimization for a group of UAVs localizing a certain target after node selection; thus, the UAVs have been optimally grouped for each target. In this case, for the convenience of description, target index j is omitted in the following derivations.

4.1. Problem Description

Consider a group of λ UAVs localizing a target at the current time t . The path planning updates are governed by the optimization time interval Δt . For the convenience of description, it is assumed that each UAV maintains a constant flight speed v during the path planning process. The variables that need to be optimized are the deflection angles ϕ_i^t ($i = 1, 2, \dots, \lambda$) of the UAVs, which represent the angles between the heading direction of the UAVs and the positive x-axis (measured counterclockwise), and once the deflection angles are optimized, the positions of the UAVs at time $t + 1$ are determined. Thus, optimizing the deflection angles ϕ_i^t ($i = 1, 2, \dots, \lambda$) of the UAVs equals optimizing the positions of the UAVs. We assume that the positions of the λ UAVs at time t are $\mathbf{G}$, which can be structured as

$$\mathbf{G}^t = \begin{bmatrix} (\mathbf{u}_1^t)^T \\ (\mathbf{u}_2^t)^T \\ \vdots \\ (\mathbf{u}_i^t)^T \\ \vdots \\ (\mathbf{u}_\lambda^t)^T \end{bmatrix}_{\lambda \times 2}, \quad (23)$$

where $\mathbf{u}_i^t = [x_i^t, y_i^t]^T$ denotes the two-dimensional coordinates of the i th UAV at time t , and λ represents the total number of the UAVs in the group.

4.2. Constraint Analysis

To ensure the physical feasibility of the UAVs' movements, we consider four fundamental constraints:

(1) The turning radius constraint: Due to the mechanical limitations of UAVs, the actual turning radius L_i^t must be greater than the minimum allowable turning radius $L_{\min}$ (which

is determined by the system parameters of the UAV). The minimum allowable turning radius is as follows:

$$L_i^t \geq L_{\min}, \quad (24)$$

The turning radius L_i^t can be determined by the UAV's movement during the path planning time interval Δt . We assume that within the Δt time interval, the turning arc angle of the UAV is θ_i^t ; then, the turning arc can be approximated as a circular arc that is shown as follows:

$$L_i^t \theta_i^t = v \Delta t, \quad (25)$$

According to the geometric relationship shown in Figure 3, we obtain

$$L_i^t \sin\left(\frac{\theta_i^t}{2}\right) = \frac{1}{2} \|\mathbf{u}_i^{t+1} - \mathbf{u}_i^t\|, \quad (26)$$

where $\sin(\theta_i^t/2)$ denotes the sine function. Combining (25) with (26), we can solve the turning arc angle θ_i^t and turning radius L_i^t .

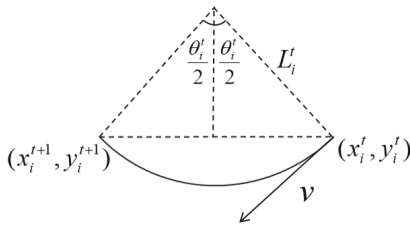


Figure 3. Diagram of UAV motion constraints. The UAV moves from position (x_i^t, y_i^t) to (x_i^{t+1}, y_i^{t+1}) along a circular arc, where L_i^t represents the path length, θ_i^t is the turning angle, and v denotes the velocity vector of the UAV.

(2) The minimum approach distance constraint: the distance r_i^t between each UAV and the target must not be less than the minimum approach distance $d_{\min}$, and the distance r_i^t should satisfy

$$r_i^t \geq d_{\min}, \quad (27)$$

where r_i^t represents the distance between the i th UAV and the target at time t .

(3) The deflection angle constraint: To ensure that the UAVs head toward the target, the heading direction of the UAV i at time t must be within the left and right angle $\Delta\alpha_i^t$ range along the line connecting it to the target. Therefore, the deflection angle ϕ_i^t of the UAV i at time t must satisfy

$$\phi_i^t \in (\alpha_i^t + \Delta\alpha_i^t, \alpha_i^t - \Delta\alpha_i^t), \quad (28)$$

where the parameters are defined by

$$\begin{cases} \alpha_i^t = \arctan \frac{y_i^t - y^t}{x_i^t - x^t} \\ \Delta\alpha_i^t = \pi \frac{r_i^0 - r_i^t}{r_i^0} \end{cases}, \quad (29)$$

where α_i^t represents the azimuth angle of the target related to the i th UAV, and $\Delta\alpha_i^t$ is the adaptive adjustment range.

(4) No-fly-zone (NFZ) constraint: To ensure operational safety and comply with airspace restrictions, the UAVs should not fly in the NFZ

$$(x_i^t, y_i^t) \notin \mathbf{S}_{\text{NFZ}}, \quad (30)$$

where (x_i^t, y_i^t) represents the position coordinates of the i th UAV at time t , and $\mathbf{S}_{\text{NFZ}}$ represents the complete set of restricted points, which is the union of all these individual NFZ and can be represented as

$$\mathbf{S}_{\text{NFZ}} = \bigcup_{k=1}^p \mathbf{S}_k, \quad (31)$$

where p is the number of the individual NFZ.

4.3. Path Optimization Model

Based on the problem formulation and the constraints derived above, the mathematical model for UAV path optimization can be described as

$$\begin{aligned} & \min_{\mathbf{G}^{t+1}} \arg \text{Tr}(\text{CRLB}_{t+1}) \\ & \text{subject to} \begin{cases} L_i^t \geq L_{\min} \\ r_i^t \geq d_{\min} \\ \phi_i^t \in (\alpha_i^t + \Delta\alpha_i^t, \alpha_i^t - \Delta\alpha_i^t) \\ (x_i^t, y_i^t) \notin \mathbf{S}_{\text{NFZ}} \end{cases}, \end{aligned} \quad (32)$$

where $\mathbf{G}_i^{t+1}$ represents the planned position of the UAVs that are assigned to the target at the next time step, and $\text{Tr}(\text{CRLB}_{t+1})$ is the trace of the predicted CRLB matrix for the localization of the target at the next time step.

4.4. Path Optimization Algorithm Design

To quickly search the planned optimal positions for the UAVs to localize the target at the next time step, in this paper, the PSO algorithm is deployed.

Suppose there are S particles searching in each iteration, and each particle s ($1 \leq s \leq S$) has information about its own best position $\mathbf{p}^s$, which it has never visited; its current position $\boldsymbol{\eta}^s$; its velocity $\mathbf{u}^s$; and the global best solution $\mathbf{p}^g$ found by all particles. Specifically, in the γ th dimension, the movement of each particle is influenced by its own best position $\mathbf{p}_{\gamma}^s$, the global best position $\mathbf{p}_{\gamma}^g$, and its current velocity $\mathbf{u}_{\gamma}^s$. The velocity and position are updated as follows, where k denotes the iteration step for the particle.

$$\mathbf{u}^s(k+1) = \omega(k) \times \mathbf{u}^s(k) + c_1(k)l_1(k)(\mathbf{p}^s(k) - \boldsymbol{\eta}^s(k)) + c_2(k)l_2(k)(\mathbf{p}^g(k) - \boldsymbol{\eta}^s(k)), \quad (33)$$

$$\boldsymbol{\eta}^s(k+1) = \boldsymbol{\eta}^s(k) + \mathbf{u}^s(k+1), \quad (34)$$

where $\omega(k)$ is the inertia weight, which decreases with iteration time as $\omega(k) = 0.9 - 0.5(k/T_{\max})$. $T_{\max}$ is the maximum number of iterations. $c_1(k)$ and $c_2(k)$ are individual learning and global learning factors, which vary with iteration time as $c_1(k) = 2.5 - 2(k/T_{\max})$ and $c_2(k) = 0.5 + 2(k/T_{\max})$ respectively. $l_1(k)$ and $l_2(k)$ are random independent variables in the range $[0,1]$. The velocity of particle in each dimension is limited by a fixed $V_{\max}$. The current position of each particle represents its current proposed solution. The performance of each solution is evaluated via a fitness function (denoted by $\mathcal{F}$). In this algorithm, each particle is a $2 \times \lambda$ -dimensional vector that contains all positions of the λ UAVs in two-dimension space. For instance, the expression of the s th particle is

$$\boldsymbol{\eta}^s = \Phi^s = [\phi_1^s, \phi_2^s, \dots, \phi_{\lambda}^s], \quad (35)$$

where $\phi_i^s (i = 1, 2, \dots, \lambda)$ denotes the deflection angle of the i th UAV in the s th particle, and λ is the number of UAVs. For initialization, we randomly place all UAVs in a certain

region. The velocity and position of every particle will be updated in each iteration. The fitness function of the s th particle is as follows:

$$f^s(k) = \mathcal{F}(\eta^s(k)) = \text{Tr}[\text{CRLB}(\Phi^s)], \quad (36)$$

where $\eta^s(k)$ denotes the current position of the s th particle, and $\mathbf{p}^s(k)$ can be updated as

$$\mathbf{p}^s(k+1) = \begin{cases} \eta^s(k+1), & \text{if } f^s(k+1) < \mathcal{F}(\mathbf{p}^s(k)) \\ \mathbf{p}^s(k), & \text{if } f^s(k+1) \geq \mathcal{F}(\mathbf{p}^s(k)) \end{cases}, \quad (37)$$

$$\mathbf{p}^g(k+1) = \arg \max_{1 \leq s \leq S} \mathcal{F}(\mathbf{p}^s(k+1)). \quad (38)$$

When $k = T_{max}$, the iteration will stop and output the optimization result; thus, the optimal UAV positions will be obtained. Here, a summary process of the path optimization is given in Algorithm 2.

Algorithm 2: CRLB-based path optimization via PSO

Input: Current UAVs positions $\mathbf{G}^t$, target position $\mathbf{w}$, flight speeds v , minimum distance d_{min} , minimum length L_{min} , NFZ $\mathbf{S}_{NFZ}$
 Calculate the distances r_i^t between the i th UAV and the target for $i = 1, 2, \dots, \lambda$, and calculate $\alpha_i^t, \Delta\alpha_i^t, \text{Tr}(\text{CRLB}_i)$
if $r_i^t < d_{min}$ **then**
 Initialize S particles with random positions and velocities, initialize $\mathbf{p}^s$ and $\mathbf{p}^g$
 for $k = 1$ **to** T_{max} **do**
 for $s = 1$ **to** S **do**
 Update the velocity using (33), and update position using (34)
 if $\mathbf{u}_i^t \notin \mathbf{S}_{NFZ}$ **and** $L_i^s \geq L_{min}$ **then**
 Calculate the fitness function $f^s(k)$ using (36), and update $\mathbf{p}^s$ using (37)
 end
 end
 Update $\mathbf{p}^g$ using (38), and update $\omega(k)$ by $\omega(k) = 0.9 - 0.5(k/T_{max})$
 end
 Set optimal deflection angle $\Phi^{t,best} = \mathbf{p}^g$, and update UAV positions:
 $\mathbf{u}_i^{t+1} = \mathbf{u}_i^t + v_i \Delta t [\cos \phi_i^{t,best}, \sin \phi_i^{t,best}]^T$
 $t = t + 1$
end
Output: The optimal planned UAVs locations $\mathbf{G}^{t+1}$ for the next time step

5. Numerical Results

In this section, we present numerical simulations to verify the validity and effectiveness of the proposals in this paper. First, we set the parameters and the initial topology structure of the UAVs and the targets. Then, the simulation results of the node selection method and the path optimization method are given. Finally, the time costs of the PSO algorithm and the exhaustive algorithm are compared to verify the effectiveness of the PSO algorithm.

5.1. Simulation Parameters and the Topology Structure Set

Consider that there are nine UAVs ($n = 9$) and three targets ($m = 3$) in a two-dimensional plane. The positions of the UAVs are $[0, 10,000]$ m, $[2500, 9000]$ m,

[4000, 8500] m, [5000, 7500] m, [7000, 6200] m, [7800, 5000] m, [9000, 4200] m, [9500, 2500] m, and [9000, 1000] m. The positions of the targets are [10,000, 32,000] m, [23,000, 23,000] m, and [30,000, 10,000] m. The UAVs' minimum turning radius is set to be $r_{\min} = 100$ m, and the minimum approach distance is set to be $d_{\min} = 5000$ m. The initial topology of the UAVs and the targets is shown in Figure 4.

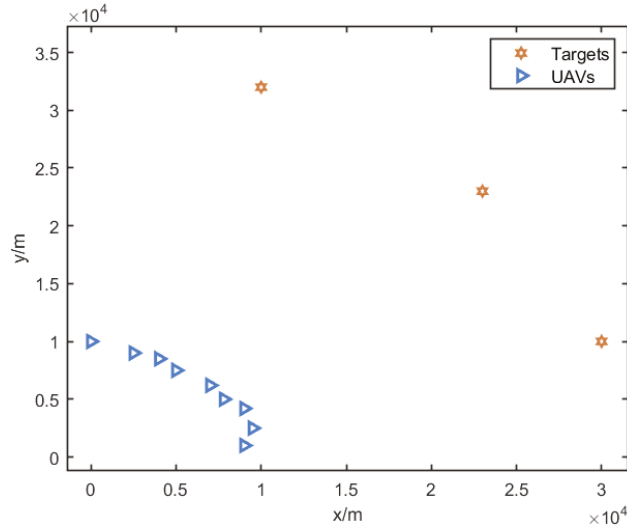


Figure 4. Initial positions of the UAVs and the targets. The blue triangles represent the initial positions of the 9 UAVs (UAV0–UAV8), and the orange stars represent the initial positions of the 3 targets.

5.2. Simulation Results of the Node Selection Algorithm

In this paper, it is assumed that at each time step, three UAVs are assigned to a target; thus, the UAVs will be divided into three groups to localize for three targets. The result of the CRLB-based node selection is shown in Figure 5.

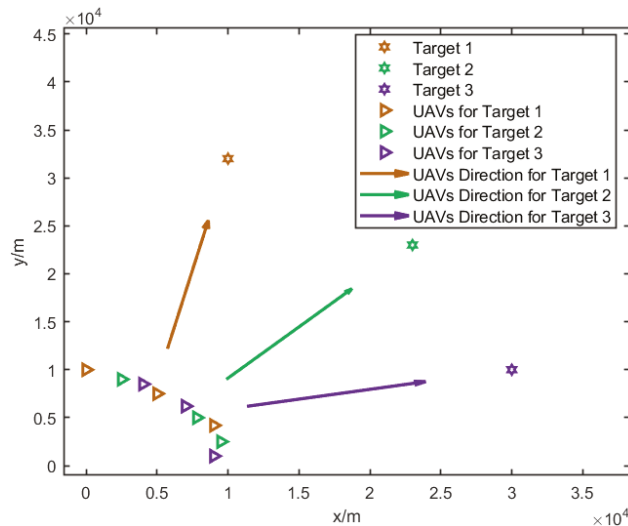


Figure 5. UAV grouping results and their movement directions after node selection. Different colors represent different UAV groups assigned to their respective targets, and arrows indicate the planned movement directions.

To evaluate the performance of the node selection, the root mean square error (RMSE) is used as a localization accuracy metric, and the performance comparisons between the

proposed CRLB-based node selection algorithm with the random selection method are made. The RMSE can be obtained by

$$\text{RMSE} = \sqrt{\frac{1}{K} \sum_{q=1}^K [(x'_q - x)^2 + (y'_q - y)^2]}. \quad (39)$$

A total of 1000 independent Monte Carlo experiments ($K = 1000$) were conducted to complete the comparisons. For the convenience of providing descriptions, in the following part, the proposed CRLB-based node selection algorithm is called the proposed algorithm for short, and the random node selection method is called the traditional algorithm for short.

From Table 1, it can be seen that the RMSEs of the proposed algorithm are much lower than that of the traditional algorithm. The RMSEs of the three different target based on the proposed algorithm decrease about 20–30% compared to the traditional algorithm. The average execution time of the proposed algorithm is 50.62 ms per iteration. The results illustrate that by efficient node selection, the target localization accuracy can be improved at a large scale.

Table 1. RMSE comparisons between the proposed node selection method and the traditional method.

RMSE (Traditional) (m)	RMSE (Proposed) (m)	Percentage Reduction (%)
564.7680	352.0307	37.6681
532.7729	408.3508	23.3537
590.6810	427.4778	27.6297

5.3. Simulation Results of the CRLB-Based Path Optimization Method

Through the node selection method, the UAVs have been divided into several groups for different target localizations. In this part, for each UAV group, we analyze the performance of the CRLB-based path optimization algorithm. The PSO algorithm is implemented with 50 particles and 500 maximum iterations in a three-dimensional continuous solution space, where each dimension represents the deflection angle of each UAV. The cognitive factor c_1 decreases linearly from 2.5 to 0.5, and the social factor c_2 increases from 0.5 to 2.5, with an inertia weight w decreasing from 0.9 to 0.4. A total of 800 independent Monte Carlo experiments are conducted for the comparison between the proposed CRLB-based path optimization method and the fixed formation method. For the convenience of description, we call the proposed CRLB-based path optimization algorithm as the proposed algorithm for short and call the fixed formation algorithm the traditional algorithm for short.

Figure 6 shows the average convergence behavior of the PSO algorithm over 50 independent runs during path optimization for the first UAV group targeting target 1 at the initial time step. The solid line represents the mean convergence curve, while the shaded area indicates one standard deviation above and below the mean value. The algorithm reliably achieves stable convergence after approximately 300 iterations, with the final average CRLB value reaching 127.8 m^2 with a standard deviation of approximately $\pm 1.2 \text{ m}^2$. This stable convergence pattern across multiple runs validates the robustness of the proposed PSO parameter settings in consistently finding high-quality solutions.

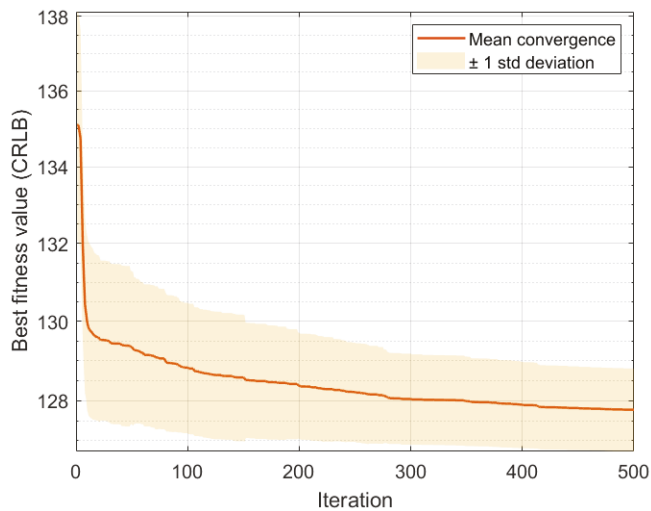


Figure 6. Evolution of best fitness value (CRLB) during PSO iterations for the first UAV group at the initial time step.

Figures 7a, 8a and 9a give the topology of the UAVs and the target for each UAV group. Figures 7b, 8b and 9b show the UAVs' motion paths based on the proposed path optimization method, and the RMSE and CRLB of the localization for each target are shown in Figures 10–12.

From Figures 10–12, it can be seen that for all three groups of UAVs, the RMSE and CRLB of the proposed path optimization method is lower than that of the traditional method. In Figure 10, the average RMSE for the proposed method is lower by about 43.37% than the traditional method, and the average CRLB for the proposed method is lower by about 62.99% than the traditional method. In Figure 11, the average RMSE for the proposed method is lower by about 47.59% than the traditional method, and the average CRLB for the proposed method is lower by about 68.65% than the traditional method. And in Figure 12, the average RMSE for the proposed method is lower by about 45.01% than the traditional method, and the average CRLB for the proposed method is lower by about 66.68% than the traditional method. Moreover, the proposed method can avoid the NFZ adaptively, while the fixed formation method in this paper is assumed to be unable to avoid the NFZ.

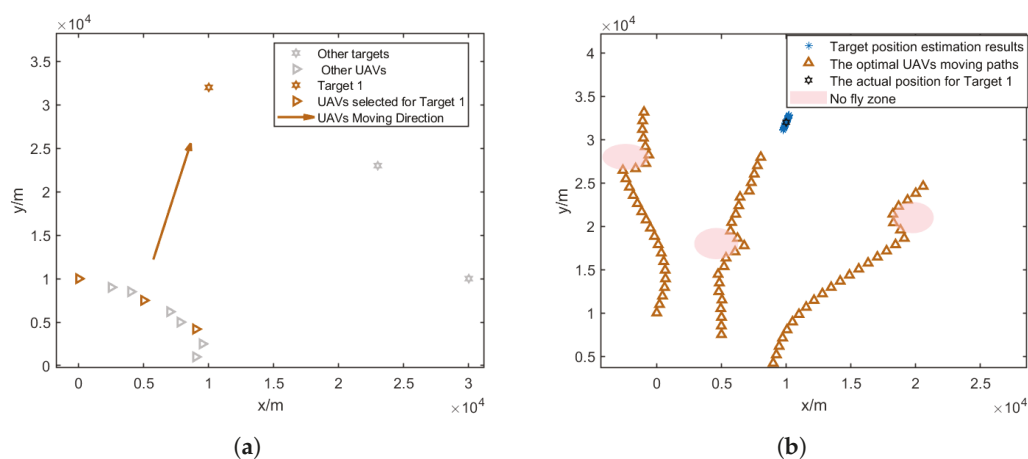


Figure 7. Three selected UAVs localizing target 1 and their moving direction. (a) Movement direction of the three selected UAV groups for target 1. (b) Trajectories of the three selected UAV groups for target 1.

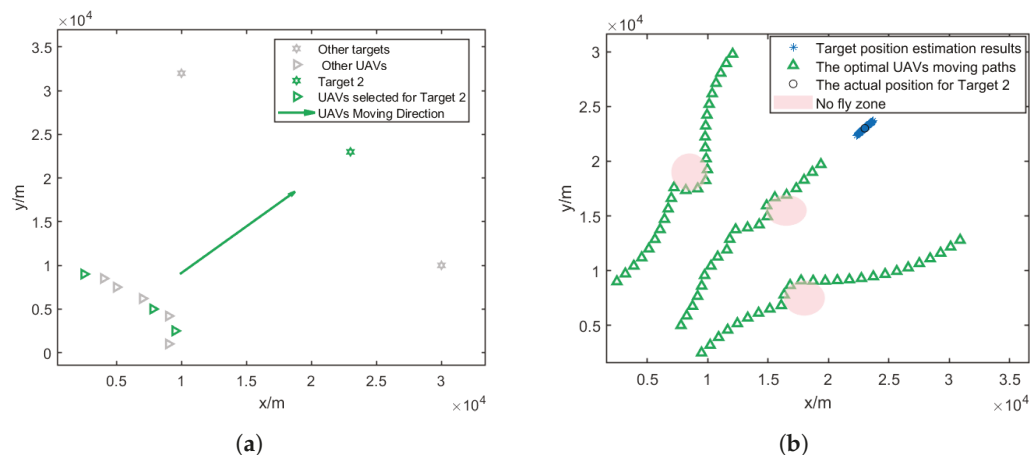


Figure 8. Three selected UAVs localizing target 2 and their moving directions. (a) Movement direction of the three selected UAV groups for target 2. (b) Trajectories of the three selected UAV groups for target 2.

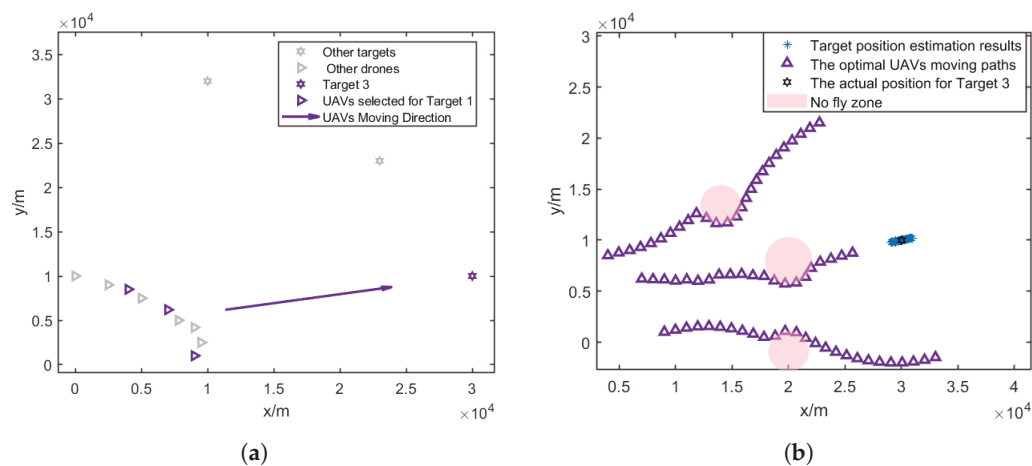


Figure 9. Three selected UAVs localizing target 3 and their moving direction. (a) Movement direction of the three selected UAV groups for target 3. (b) Trajectories of the three selected UAV groups for target 3.

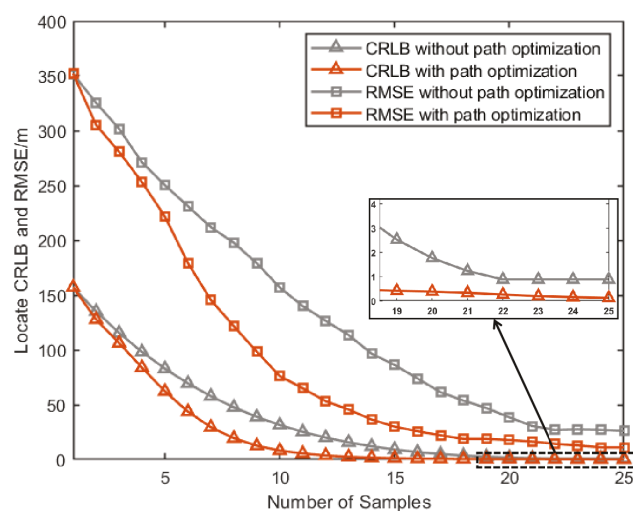


Figure 10. Comparison of CRLB and RMSE under path optimization versus fixed configuration for target 1.

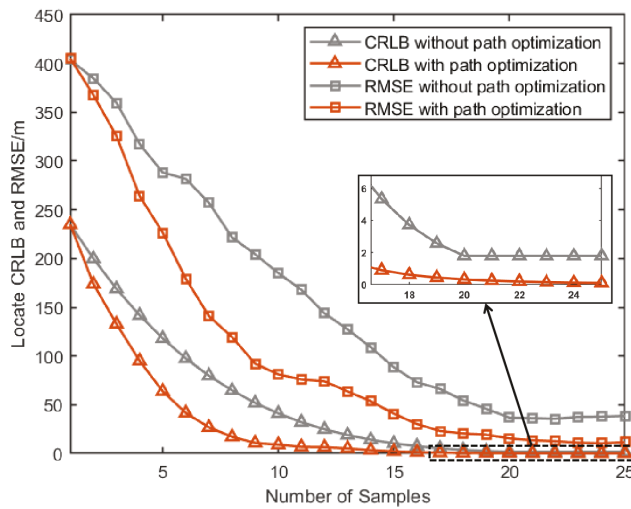


Figure 11. Comparison of RMSE and CRLB between the proposed path optimization method and the traditional method.

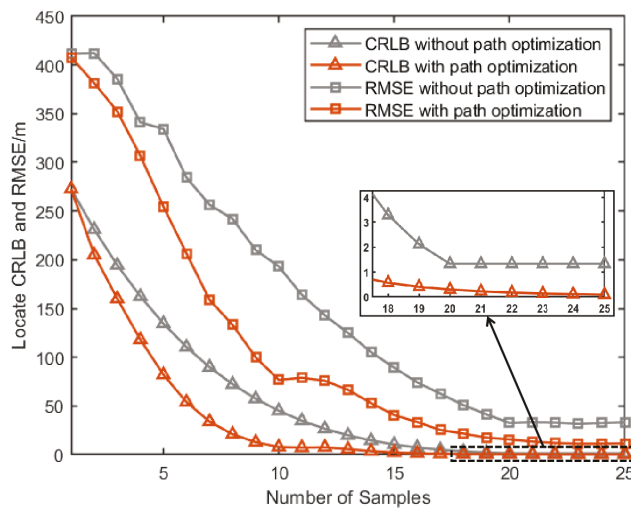


Figure 12. Comparison of RMSE and CRLB between the proposed path optimization method and the traditional method.

To verify the effectiveness of the PSO algorithm used in this paper, we compared PSO with three other optimization algorithms: GA, ACO, and the traditional exhaustive method. These algorithms are used to search for Φ (the optimal UAV positions) at each time step to obtain the minimum trace of CRLB for the next time step. The parameter settings for the algorithms are shown in Table 2. GA operates in continuous space with 20 individuals, employing a crossover rate of 0.8 and a mutation rate of 0.1. The PSO parameter settings are the same as in the previous experiment, while ACO discretizes the search space into 20 points per dimension, using 50 ants with pheromone-based search strategies. All algorithms are set to a maximum of 500 iterations with different early stopping criteria and optimization strategies to enhance search efficiency. For the exhaustive method, the angle range is discretized into 50 equally spaced points, resulting in 50^3 combinations that need to be evaluated at each time step to find the optimal waypoints resulting the minimum trace of CRLB.

Table 2. Parameter settings of the three optimization algorithms.

Parameter	PSO	GA	ACO
Size	50 particles	20 individuals	50 ants
Max iterations	500	500	500
Early stopping criterion	Fitness $< 1e^{-6}$	No improvement in 20 generations	fitness $< 1e^{-6}$
Special parameters	$c_1: 2.5 \rightarrow 0.5$ $c_2: 0.5 \rightarrow 2.5$ $w: 0.9 \rightarrow 0.4$ Constriction: 0.729	Crossover rate: 0.8 Mutation rate: 0.1 Elite count: 2	Evaporation rate: 0.1 Influence factor (α): 1 Heuristic factor (β): 2 Deposit factor (Q): 1
Optimization strategy	Local search per 10 iterations; Reinitialize 10% Particles per 20 iter.	Elite retention	Local search per 10 iterations; Space discretization: 50 points

As shown in Figure 13a–c, we present the CRLB and RMSE performance comparisons for five different methods: the fixed formation method (without path optimization) and four optimization methods (Exhaustive, PSO, ACO, and GA) localizing three different targets. The simulation results demonstrate that for localizing different targets, the four optimization approaches present similar convergence patterns in both CRLB and RMSE metrics, with their curves nearly overlapping throughout the localization process, and they all achieve lower CRLB and RMSE values compared to the fixed formation method. Figure 13d shows the computational time comparisons for a single iteration between different optimization methods. The results reveal that to realize the approximate localizing performance, the PSO algorithm maintains the lowest computation time consistently across all sampling times, while GA shows the highest computational burden and ACO demonstrates moderate computational requirements, positioned between PSO and the exhaustive method.

Table 3 shows the average percentage reduction in the CRLB and RMSE of path optimization algorithms (Exhaustive, PSO, ACO, and GA) compared to the fixed formation method. Specifically, the CRLB reduction for Exhaustive, PSO, ACO, and GA is 66.70%, 66.66%, 66.77%, and 66.57%, respectively, compared to the fixed formation method, while the RMSE reduction is 44.24%, 44.99%, 45.02%, and 44.64%, respectively, compared to the fixed formation method, indicating that with path optimization methods, the localization performance of the UAVs can be improved at a large scale. Table 3 gives the comparisons of the average single-iteration computation time between the four path optimization methods, with PSO taking 1.34 s, ACO taking 3.05 s, the Exhaustive method taking 4.28 s, and GA taking 7.76 s. It is evident that the PSO method applied in this paper takes the shortest computation time, followed by ACO and then the exhaustive method, while GA requires significantly more time. These results indicate that while all methods can achieve similar localization performance, and PSO offers the best computational efficiency, making it particularly suitable for real-time UAV path planning applications.

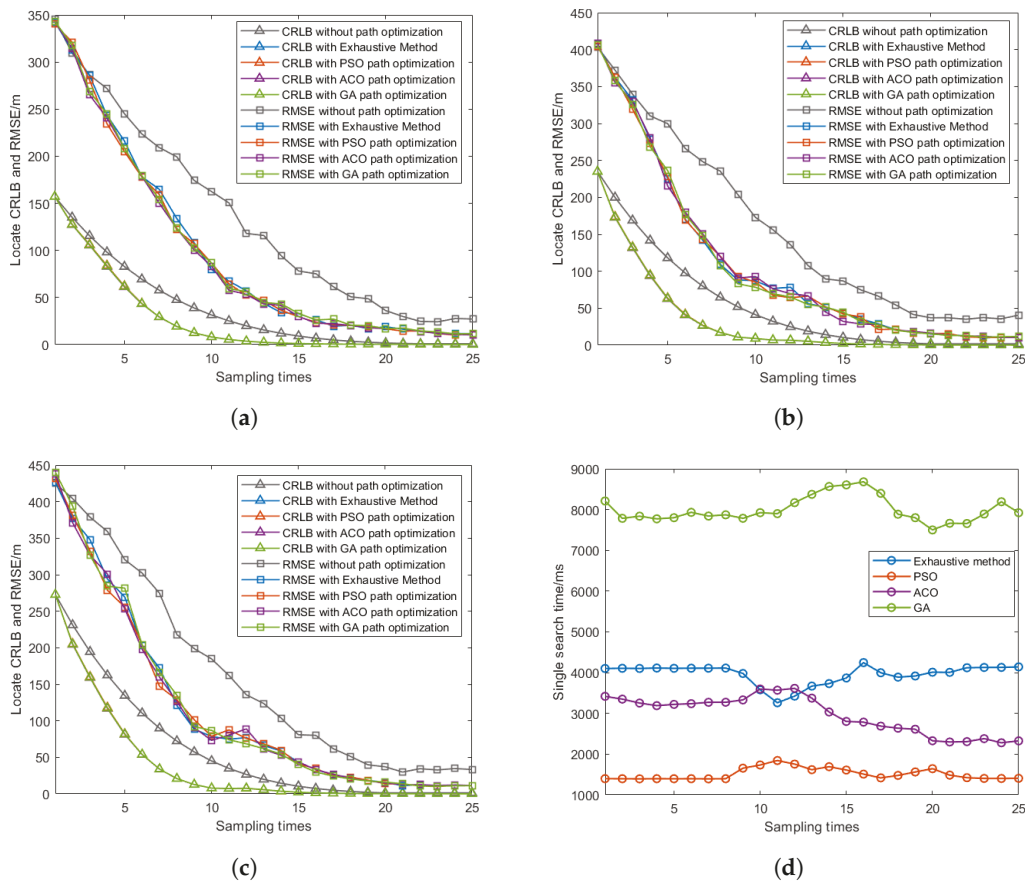


Figure 13. CRLB comparison and single-iteration computation time comparisons with different optimization algorithms localizing different targets. (a) CRLB comparison with different optimization algorithms localizing target 1. (b) CRLB comparison with different optimization algorithms localizing target 2. (c) CRLB comparison with different optimization algorithms localizing target 3. (d) Single-iteration computation time comparisons between different optimization methods.

Table 3. Performance comparison between different optimization methods.

Metric	Exhaustive	PSO	ACO	GA
CRLB reduction (%)	66.70	66.66	66.77	66.57
RMSE reduction (%)	44.24	44.99	45.02	44.64
Single iteration computation time(s)	4.28	1.34	3.05	7.76

As shown in Table 3, while all algorithms achieve similar CRLB and RMSE reductions compared to the fixed formation method, PSO demonstrates superior computational efficiency. These results suggest that PSO provides the best balance between computational efficiency and optimization performance, making it particularly suitable for real-time UAV formation, quickly responding to environmental changes.

The performance of the proposed algorithm in terms of localization accuracy and computational efficiency can be explained from two aspects. Firstly, while the traditional fixed formation method adopts predetermined UAV configurations without considering the impact of flight paths on localization performance, the CRLB-based path optimization algorithm continuously adjusts UAV flight paths to minimize the trace of CRLB. This optimization mechanism enables UAVs to select better observation angles, thereby obtaining more accurate measurements, reducing measurement errors, and improving localization accuracy, as evidenced by the lower RMSE values shown in Figure 13a–c. Secondly, among the four optimization algorithms, PSO converges to optimal solutions in less time with higher computational efficiency. This is because GA, ACO, and Exhaustive methods require global

information and involve complex computations, while PSO's characteristic of making decisions based on local relative measurements better meet real-time control requirements, resulting in shorter computation times, as shown in Figure 13d.

5.4. Impact Analysis of Minimum Turning Radius

To thoroughly investigate the impact of the minimum turning radius on path optimization and localization accuracy, we conducted comparative experiments with different minimum turning radii while keeping other parameters constant. Three UAVs were initially deployed at coordinates [0, 10,000] m, [5000, 7500] m, and [9000, 4200] m, with the target located at [50,000, 32,000] m. Using the PSO optimization algorithm introduced in Section 4, we set the minimum turning radius L_{min} to 5000 m, 7500 m, and 10,000 m to examine its influence on UAV path optimization and localization performance. In total, 800 Monte Carlo simulations were performed to ensure the statistical significance of the results.

As shown in Figure 14a–c, the optimized UAV paths exhibit distinct characteristics under different minimum turning radius constraints. With increasing L_{min} , the UAVs require wider turning arcs to change their flight directions, resulting in longer path lengths and more gradual course adjustments.

Figure 14d and Table 4 demonstrate the impact of turning radius constraints on localization accuracy. The simulation results show that a smaller turning radius allows for more flexible path planning and better localization performance. When L_{min} increases from 5000 m to 10,000 m, the CRLB reduction decreases by 2.86 percentage points (from 64.57% CRLB reduction compared to the fixed formation method to 61.71%), while the RMSE reduction decreases by 3.59 percentage points (from 45.38% RMSE reduction compared to the fixed formation method to 41.79%). This trend indicates that larger minimum turning radius constraints moderately impact the algorithm's ability to optimize UAV paths for target localization, as they limit the UAVs' maneuverability in adjusting their positions for optimal measurement geometry.

Table 4. Average performance improvement compared to the traditional method.

Minimum Turning Radius (L_{min})	CRLB Reduction (%)	RMSE Reduction (%)
5000 m	64.57	45.38
7500 m	62.15	43.14
10,000 m	61.71	41.79

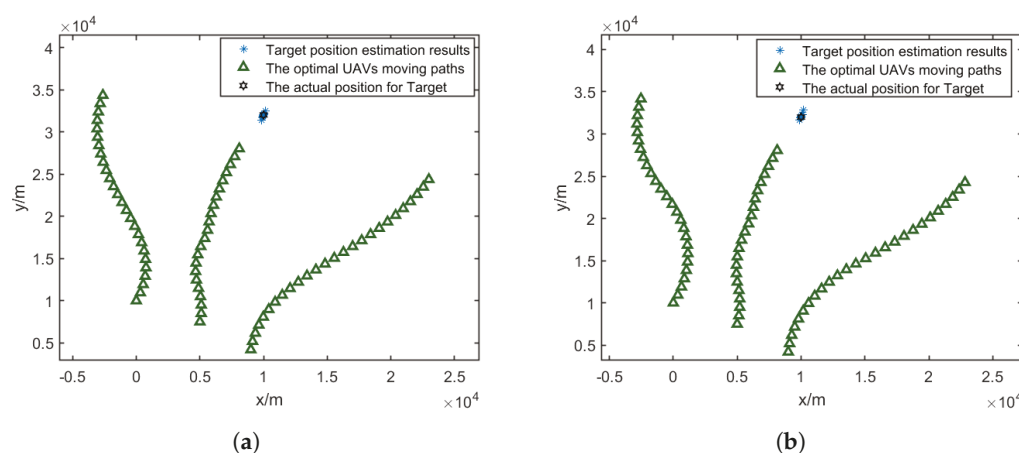


Figure 14. Cont.

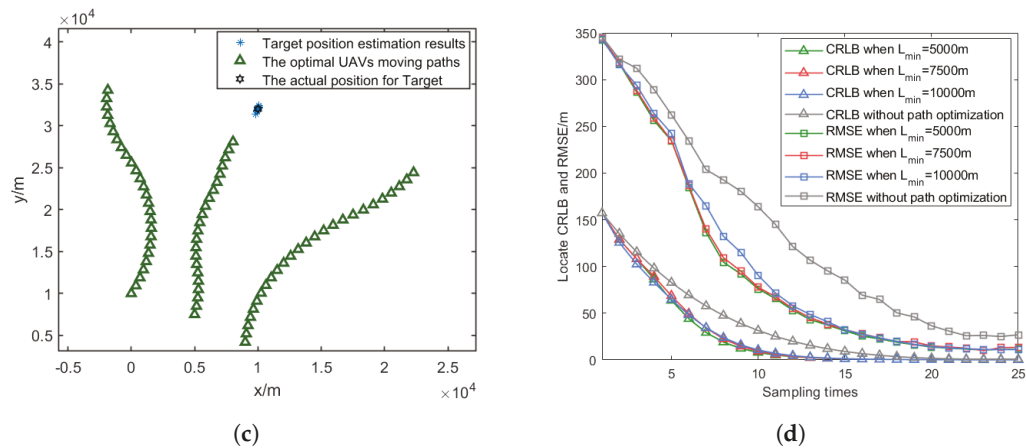


Figure 14. Optimized UAV paths under different minimum turning radii and their RMSE comparisons. (a) $L_{min} = 5000$ m. (b) $L_{min} = 7500$ m. (c) $L_{min} = 10,000$ m. (d) Comparison of RMSE and CRLB.

5.5. Impact Analysis of No-Fly-Zone Size

To evaluate the impact of no-fly-zone size on localization performance, we conducted comparative experiments with different no-fly-zone radii while maintaining other parameters constant. Three UAVs were initially deployed at coordinates [0, 10,000] m, [5000, 7500] m, and [9000, 4200] m, with the target located at [50,000, 32,000] m. Three circular no-fly zones were centered at [−2400, 28,000] m, [4600, 18,000] m, and [19,800, 21,000] m. Using the PSO optimization algorithm introduced in Section 4, we set the radii of no-fly zone R_{NFZ} to 1000 m, 2000 m, and 3000 m to examine their influence on localization performance. Also, 800 Monte Carlo simulations were performed to ensure the statistical significance of the results.

As shown in Figure 15a–c, the optimized UAV paths demonstrate different characteristics as R_{NFZ} increases. Larger no-fly zones force the UAVs to take more substantial detours, resulting in longer path lengths and more constrained movement patterns.

Figure 15d and Table 5 illustrate the impact of no-fly-zone size on localization accuracy. The simulation results indicate that increasing R_{NFZ} leads to degraded localization performance, as it restricts the available space for UAV path optimization solutions. When R_{NFZ} increases from 1000 m to 3000 m, the CRLB reduction decreases by 3.94 percentage points (from 63.88% CRLB reduction compared to the fixed formation method to 59.94%), while the RMSE reduction decreases by 2.56 percentage points (from 45.02% RMSE reduction compared to the fixed formation method to 42.46%). This degradation in performance can be attributed to the reduced flexibility in configuring UAVs relative to the optimal measurement geometry, as larger restricted areas limit the available space for tactical maneuvers.

Table 5. Average performance under different no-fly-zone radii.

No-Fly-Zone Radius (R_{NFZ})	CRLB Reduction (%)	RMSE Reduction (%)
1000 m	63.88	45.02
2000 m	62.51	44.93
3000 m	59.94	42.46

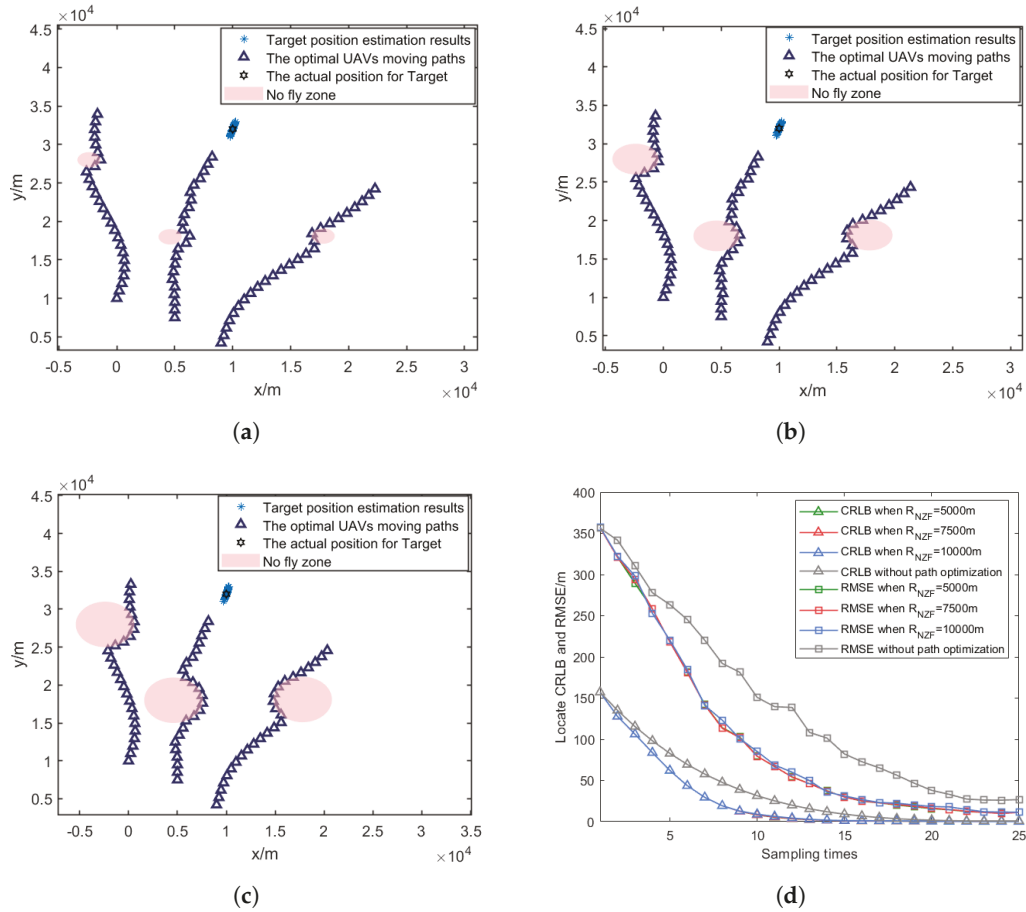


Figure 15. Optimized UAV paths and comparison of RMSE and CRLB under different no-fly-zone radii. (a) $R_{NFZ} = 1000$ m. (b) $R_{NFZ} = 2000$ m. (c) $R_{NFZ} = 3000$ m. (d) Comparisons of RMSE and CRLB.

6. Conclusions

In this paper, we investigate the problem of node selection and path optimization for multi-target passive localization via UAVs. First, the Chan-TDOA target passive localization model is established. Then, the CRLB-based node selection algorithm is proposed to properly divide the UAVs into several groups localizing different targets, and the CRLB-based path optimization algorithm is proposed to search for the optimal UAV position configuration at each time step. The proposed path optimization method considers NFZ avoidance, ensuring both operational safety and localization accuracy. To improve the efficiency of path optimization, the PSO algorithm is applied to accelerate the searching process. Finally, numerical simulations were performed to verify the validity and effectiveness of the proposals in this paper.

Author Contributions: Conceptualization, X.X. and X.L.; methodology, X.X.; software, X.L.; validation, X.X., Y.Y. and Z.Z.; formal analysis, X.X.; investigation, X.X.; resources, X.X.; data curation, X.X. and X.L.; writing—original draft preparation, X.X.; writing—review and editing, X.X.; visualization, X.X.; supervision, X.L.; project administration, Y.Y.; funding acquisition, X.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the Qian Yuan Laboratory.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Due to privacy constraints, we are unable to share the data utilized in this study for submission purposes. We assure you that the research methodology and results outlined in this paper comply with ethical guidelines and have been executed in line with recognized research protocols.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Mathematical Derivations of Chan-TDOA Algorithm

Since the UAVs ($n \geq 3$) are localizing the same target, we omit the superscript j in this part for convenience. Therefore, the target's location is denoted as $[x, y]^T$. Then, the Chan-TDOA algorithm can be presented as follows:

For the convenience of comprehension, we first give the definitions for the important variables that would be used in subsequent derivations:

- (1) $M_i = x_i^2 + y_i^2$ represents the square of the location of the i th UAV (including $i = 0$).
 - (2) $x_{i,0} = x_i - x_0$ and $y_{i,0} = y_i - y_0$ denote the relative locations of the i th UAV with respect to the master base station.
 - (3) $r_{i,0} = r_i - r_0$ indicates the difference in distance from the i th UAV to the target and from the master base station to the target.
 - (4) $\tau_{i,0} = \tau_i - \tau_0$ represents the difference in the times at which the signal was received by the i th UAV and received by the master base station.
 - (5) $\sigma_{i,0}^2 = \sigma_i^2 + \sigma_0^2$ represents the variance in the TDOA measurement for the target.
- It is apparent that

$$r_i = r_{i,0} + r_0. \quad (\text{A1})$$

Squaring both sides of (A1), we obtain the following:

$$r_i^2 = r_{i,0}^2 + r_0^2 + 2r_{i,0}r_0, \quad (\text{A2})$$

Substituting (1) into (A2) yields (A3):

$$x_{i,0}x + y_{i,0}y + r_{i,0}r_0 = \frac{1}{2}(M_i - M_0 - r_{i,0}^2). \quad (\text{A3})$$

When the number of UAVs is $n = 3$, the system of (A3) can be solved using the elimination method to express x and y :

$$\begin{cases} x = P_1 + Q_1r_0 \\ y = P_2 + Q_2r_0' \end{cases} \quad (\text{A4})$$

where P_1, P_2, Q_1 , and Q_2 are the coefficients derived from (A3) using the elimination method. By setting $i = 0$ and squaring both sides of (1), we obtain

$$r_0^2 = M_0 - 2x_0x - 2y_0y + x^2 + y^2. \quad (\text{A5})$$

Substituting (A3) into (A5) obtains

$$Ar_0^2 + Br_0 + C = 0, \quad (\text{A6})$$

where A, B , and C are the coefficients of the quadratic equation in terms of r_0 obtained by substituting x and y from (A4) into (A5). This equation is a quadratic equation in terms of r_0 . Solving this equation and retaining the positive solution provide the viable value for r_0 . If two positive solutions are obtained, the unreasonable solution should be excluded based on practical conditions and prior information. The positive solution of r_0 is then substituted back into (A4) to determine the coordinates of the target location.

When the number of UAVs is $n \geq 4$, the number of TDOA measurements exceeds the number of unknown variables. In this scenario, a two-step WLS estimation can be employed to fully utilize redundant data and achieve a more accurate localization estimate. At this stage, r_0 in (A5) is a nonlinear term related to the target location. We treat it as an unknown variable that is included in

$$\mathbf{Z} = \begin{bmatrix} x \\ y \\ r_0 \end{bmatrix}. \quad (\text{A7})$$

Let $\mathbf{Z}^0$ represent the true values of $\mathbf{Z}$ corresponding to the true target location. We then establish the linear equation with TDOA noise

$$\boldsymbol{\psi} = \mathbf{h} - \mathbf{G}\mathbf{Z}^0, \quad (\text{A8})$$

where $\boldsymbol{\psi}$ is the error vector caused by TDOA measurement noise, and

$$\mathbf{h} = \frac{1}{2} \begin{bmatrix} M_1 - M_0 - r_{1,0}^2 \\ M_2 - M_0 - r_{2,0}^2 \\ \vdots \\ M_n - M_0 - r_{n-1,0}^2 \end{bmatrix}, \quad (\text{A9})$$

$$\mathbf{G} = \begin{bmatrix} x_{1,0} & y_{1,0} & r_{1,0} \\ x_{2,0} & y_{2,0} & r_{2,0} \\ \vdots & \vdots & \vdots \\ x_{n-1,0} & y_{n-1,0} & r_{n-1,0} \end{bmatrix}. \quad (\text{A10})$$

The covariance matrix $\boldsymbol{\Psi}$ of the error vector $\boldsymbol{\psi}$ is

$$\boldsymbol{\Psi} = \mathbb{E}[\boldsymbol{\psi}\boldsymbol{\psi}^T] = \mathbf{B}\mathbf{Q}\mathbf{B}, \quad (\text{A11})$$

where

$$\mathbf{B} = \text{diag}\{r_1, r_2, \dots, r_{n-1}\}. \quad (\text{A12})$$

$\mathbf{Q}$ is the covariance matrix of the measurements.

$$\mathbf{Q} = \text{diag}\{\sigma_{1,0}^2, \sigma_{2,0}^2, \dots, \sigma_{n-1,0}^2\}. \quad (\text{A13})$$

Assuming that the elements of the error vector $\mathbf{Z}$ are mutually independent, the covariance matrix $\boldsymbol{\Psi}$ can be approximated by the matrix $\mathbf{Q}$. Using the WLS method, the maximum likelihood estimate of $\mathbf{Z}$ can be expressed as

$$\mathbf{Z} = \text{argmin}\left\{(\mathbf{h} - \mathbf{G}\mathbf{Z})^T \boldsymbol{\Psi}^{-1}(\mathbf{h} - \mathbf{G}\mathbf{Z})\right\} \approx (\mathbf{G}^T \mathbf{Q}^{-1} \mathbf{G}) \mathbf{G}^T \mathbf{Q}^{-1} \mathbf{h}. \quad (\text{A14})$$

To achieve a more precise localization result, a second WLS estimation step is performed. At this point, we construct a noise vector $\boldsymbol{\psi}'$ similar to (A8)

$$\boldsymbol{\psi}' = \mathbf{h}' - \mathbf{G}'\mathbf{Z}', \quad (\text{A15})$$

where

$$\mathbf{h}' = \begin{bmatrix} (\mathbf{Z}_1 - x_0)^2 \\ (\mathbf{Z}_2 - y_0)^2 \\ \mathbf{Z}_3^2 \end{bmatrix}, \quad (\text{A16})$$

$$\mathbf{G}' = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad (\text{A17})$$

$$\mathbf{Z}' = \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}. \quad (\text{A18})$$

The covariance matrix of the error vector Ψ' is

$$\Psi' = E[\psi' \psi'^T] = 4\mathbf{B}' \text{cov}(\mathbf{Z}) \mathbf{B}', \quad (\text{A19})$$

where $\mathbf{B}' = \text{diag}(\mathbf{Z}_1 - x_1, \mathbf{Z}_2 - y_1, \mathbf{Z}_3)$. The WLS localization estimate can be obtained as follows

$$\mathbf{Z}' = (\mathbf{G}'^T \Psi'^{-1} \mathbf{G}')^{-1} \mathbf{G}'^T \Psi'^{-1} \mathbf{h}', \quad (\text{A20})$$

The final estimated position of the target is

$$\begin{bmatrix} x \\ y \end{bmatrix} = \pm \sqrt{\mathbf{Z}'} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}. \quad (\text{A21})$$

References

1. Feldstein, S. How global demand for military drones is transforming international security and geopolitics. *Georget. J. Int. Aff.* **2023**, *24*, 146–155. [CrossRef]
2. Gargalakos, M. The role of unmanned aerial vehicles in military communications: Application scenarios, current trends, and beyond. *J. Def. Model. Simul.* **2024**, *21*, 313–321. [CrossRef]
3. Cao, X.; Wang, Z.; Zheng, B.; Tan, Y. Improved UAV-to-Ground Multi-Target Tracking Algorithm Based on StrongSORT. *Sensors* **2023**, *23*, 9239. [CrossRef] [PubMed]
4. Zhou, R.; Almustafa, M.K.; Nehdi, M.L.; Su, H. Automated localization of dike leakage outlets using UAV-borne thermography and YOLO-based object detectors. *ISPRS J. Photogramm. Remote Sens.* **2024**, *218*, 551–573. [CrossRef]
5. Mahmood, S.A.; Abdulmunem, F.A.; Lafta, S.H. Lightweight deep learning model-based UAVs visual detection. *Multimed. Tools Appl.* **2024**. [CrossRef]
6. Patole, S.M.; Torlak, M.; Wang, D.; Ali, M. Automotive radars: A review of signal processing techniques. *IEEE Signal Process. Mag.* **2017**, *34*, 22–35. [CrossRef]
7. Rosić, M.; Sedak, M.; Simić, M.; Pejović, P. Chaos-Enhanced Adaptive Hybrid Butterfly Particle Swarm Optimization Algorithm for Passive Target Localization. *Sensors* **2022**, *22*, 5739. [CrossRef] [PubMed]
8. Justin, K.; Simone, D. Autonomous angles-only multitarget tracking for spacecraft swarms. *Acta Astronaut.* **2021**, *189*, 514–529.
9. Fokin, G. Passive Geolocation with Unmanned Aerial Vehicles using TDOA-AOA Measurement Processing. In Proceedings of the 2019 21st International Conference on Advanced Communication Technology (ICACT), PyeongChang, Republic of Korea, 17–20 February 2019; pp. 360–365. [CrossRef]
10. Agrawal, S.; Kumar, P.; Sharma, A. Passive Emitter Localisation Using TDOA and FDOA Measurements from UAV. In Proceedings of the 2020 IEEE International Conference for Innovation in Technology (INOCON), Bangluru, India, 6–8 November 2020; pp. 1–5.
11. Zhang, X.; Duan, L. Fast Deployment of UAV Networks for Optimal Wireless Coverage. *IEEE Trans. Mob. Comput.* **2019**, *18*, 588–601. [CrossRef]
12. Suh, T.; Kim, W. Observability Study on Passive Target Localization by Conic-Angle Measurements. *Sensors* **2021**, *21*, 6439. [CrossRef]
13. Jeon, M.H.; Lim, W.J.; Ryoo, C. Task Assignment for Multiple Multi-purpose Unmanned Aerial Vehicles Using Greedy Algorithm. *Int. J. Aeronaut. Space Sci.* **2024**, *25*, 1380–1394. [CrossRef]
14. Xin, J.; Zhong, J.; Li, S. Greedy mechanism based particle swarm optimization for path planning problem of an unmanned surface vehicle. *Sensors* **2019**, *19*, 4620. [CrossRef]
15. Wang, X.; Huang, C.; Chen, F. An improved particle swarm optimization algorithm for unmanned aerial vehicle route planning. *J. Phys. Conf. Ser.* **2022**, *2245*, 012013. [CrossRef]
16. Meisquita, R.; Gasper, P. A novel path planning optimization algorithm based on particle swarm optimization for UAVs for bird monitoring and repelling. *Processes* **2021**, *10*, 62. [CrossRef]

17. Huang, H.; Yang, Y.; Wang, H.; Ding, Z.; Sari, H.; Adachi, F. Deep Reinforcement Learning for UAV Navigation Through Massive MIMO Technique. *IEEE Trans. Veh. Technol.* **2020**, *69*, 1117–1121. [CrossRef]
18. Cheng, C.; Zhang, H.; Sun, Y.; Tao, H.; Chen, Y. A cross-platform deep reinforcement learning model for autonomous navigation without global information in different scenes. *Control Eng. Pract.* **2024**, *150*, 105991. [CrossRef]
19. Lee, M.-F.R.; Yusuf, S.H. Mobile Robot Navigation Using Deep Reinforcement Learning. *Processes* **2022**, *10*, 2748. [CrossRef]
20. Maier, G.; Pfaff, F.; Wagner, M.; Pieper, C.; Gruna, R.; Noack, B.; Kruggel-Emden, H.; Längle, T.; Hanebeck, U.D.; Wirtz, S.; Scherer, V.; et al. Real-time multitarget tracking for sensor-based sorting—A new implementation of the auction algorithm for graphics processing units. *J. Real-Time Image Process.* **2019**, *16*, 2261–2272. [CrossRef]
21. Rokbani, N.; Kromer, P.; Twir, I.; Alimi, A.M. A hybrid hierarchical heuristic-ACO with local search applied to travelling salesman problem, AS-FA-Ls. *Int. J. Syst. Dyn. Appl.* **2020**, *9*, 58–73. [CrossRef]
22. [CrossRef] [PubMed] Siemiatkowska, B.; Stecz, W. A Framework for Planning and Execution of Drone Swarm Missions in a Hostile Environment. *Sensors* **2021**, *21*, 4150. [CrossRef] [PubMed]
23. Li, Z.H.; Zhang, K.Y.; Yang, L.; Wang, G.Y.; Hu, G. Improving heuristic functions in A* algorithm path planning. *J. Res. Sci. Eng.* **2022**, *4*, 10–20.
24. Darrah, M.; Wilhelm, J.; Munasinghe, T.; Duling, K.; Yokum, S.; Sorton, E.; Rojas, J.; Wathen, M. A flexible genetic algorithm system for multi-UAV surveillance: Algorithm and flight testing. *Unmanned Syst.* **2015**, *3*, 49–62. [CrossRef]
25. Huang, C.; Zhou, X.; Ran, X.; Wang, J.; Chen, H.; Deng, W. Adaptive cylinder vector particle swarm optimization with differential evolution for UAV path planning. *Eng. Appl. Artif. Intell.* **2023**, *121*, 105942. [CrossRef]
26. Yang, X.S. A new metaheuristic bat-inspired algorithm. In *Nature Inspired Cooperative Strategies for Optimization (NICSO 2010)*; Springer: Berlin/Heidelberg, Germany, 2010; p. 65.
27. Dogancay, K.; Hmam, H. UAV path optimization for angle-only self-localization and target tracking based on the Bayesian Fisher information matrix. *Sensors* **2024**, *24*, 3120. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Classification of Flying Drones Using Millimeter-Wave Radar: Comparative Analysis of Algorithms Under Noisy Conditions

Mauro Larrat ^{*,†} and Claudomiro Sales [†]

Instituto de Ciências Exatas e Naturais, Universidade Federal do Pará, Rua Augusto Corrêa, 01 Guamá, CEP, Belém 66075-110, PA, Brazil; cssj@ufpa.br

* Correspondence: maurolarrat@ufpa.br

[†] These authors contributed equally to this work.

Abstract: This study evaluates different machine learning algorithms in detecting and identifying drones using radar data from a 60 GHz millimeter-wave sensor. These signals were collected from a bionic bird and two drones, namely DJI Mavic and DJI Phantom 3 Pro, which were represented in complex form to preserve amplitude and phase information. The first benchmarks used four algorithms, namely long short-term memory (LSTM), gated recurrent unit (GRU), one-dimensional convolutional neural network (Conv1D), and Transformer, and they were benchmarked for robustness under noisy conditions, including artificial noise types like white noise, Pareto noise, impulsive noise, and multipath interference. As expected, Transformer outperformed other algorithms in terms of accuracy, even on noisy data; however, in certain noise contexts, particularly Pareto noise, it showed weaknesses. For this purpose, we propose Multimodal Transformer, which incorporates more statistical features—skewness and kurtosis—in addition to amplitude and phase data. This resulted in a improvement in detection accuracy, even under difficult noise conditions. Our results demonstrate the importance of noise in processing radar signals and the benefits afforded by a multimodal presentation of data in detecting unmanned aerial vehicle and birds. This study sets up a benchmark for state-of-the-art machine learning methodologies for radar-based detection systems, providing valuable insight into methods of increasing the robustness of algorithms to environmental noise.

Keywords: drone detection; millimeter-wave radar sensor; machine learning

1. Introduction

Drones or unmanned aerial vehicles (UAVs) are employed in various sectors these days. Although there are many advantages to UAV technology, the quick growth has also posed important issues regarding privacy, safety, and security. These concerns have fueled the growth of the counter-UAV (C-UAV) market, ranging from military suppliers to innovative startups, and they are focused on detecting, tracking, identifying, and neutralizing unauthorized drones [1,2]. The detection of UAVs is particularly pronounced in environments where traditional sensors face limitations, such as low radar cross-section areas or during complex flight stages [3–5].

Several UAV detection methods have been explored in the literature, including acoustic monitoring [6–12], visible light [13–18], infrared [19,20], radar detection [21–31], and radio spectrum monitoring [32–36]. However, each method has its limitations. Acoustic monitoring is susceptible to ambient noise, and visible light detection can be impaired by weather conditions or obstacles such as buildings. As UAVs become smaller and less conspicuous, both infrared and radar technologies cannot maintain their effectiveness.

Among these methods, radar detection, particularly millimeter wave (mmWave), stands out due to its resilience to weather conditions and ability to provide high-resolution detection. Unlike acoustic methods, which are influenced by environmental noise, or visible light detection, which requires a clear line of sight, radar is robust in diverse operational contexts [23,28]. Additionally, mmWave radar sensors, operating in the 24 GHz to 300 GHz range can effectively detect small objects with fine precision, making them suitable for UAV detection from both recreational and security perspectives [23,29].

Currently, multi-target active domain adaptation (MT-ADA) has been set up for active sample selection and simultaneous domain alignment. This advancement in domain adaptation strategies has significantly improved radar detection capabilities, though primarily in image classification [37]. Recent innovations have demonstrated the effectiveness of unsupervised domain adaptation in addressing domain shifts in synthetic aperture radar for target detection. Shi et al.'s [38] study proposed a novel UDA framework that aligns domain-level and class-level features to facilitate SAR target classification from simulated to real-world data. Chen and Jia [39] likewise posited a UDA method for bringing-through-wall radar imaging. This approach improved detection performance under limited labeled data. In addition, since the creation of synthetic radar detection (SAR), Guo et al. [40] presented a domain-adaptive faster R-CNN framework, which is effective in overcoming these domain disparities, even with small training datasets. These studies reiterate the importance of domain adaptation in improving radar-based detection systems in diverse environments of operation. Nevertheless, the key principles of active domain adaptation could easily be implemented for radar use, particularly in scenarios characterized by multiple targets and varying conditions.

The application of multi-kernel-size feature fusion in the convolutional neural network (MKSFF CNN) model has been applied for the target classification of SAR, achieving greater completeness in representation for more useful classification [41]. In recent years, much advancement has been made in multi-scale feature fusion and deep learning architecture to detect and classify UAVs. For example, one study presented a multi-scale feature fusion algorithm known as DA-FS SSD, which is enhanced by incorporating dilated convolution and attention mechanisms into the image analysis of the object in UAV images [42]. Another introduced a multi-scale aerial image detection method for UAVs based on adaptive feature fusion to address the difficulty of small target detection within cluttered backgrounds [43]. Researchers have developed a lightweight CNN model applying depthwise separable convolutions to classify UAV images with high accuracy and low computational complexity [44].

Moreover, state-of-the-art results were obtained when the convolutional neural networks were fed into long short-term memory networks in applications related to UAVs. A hybrid CNN-LSTM framework was developed for automatically detecting anomalies in farmland by the aerial images collected from UAVs, capturing spatial and temporal characteristics for further classification enhancement [45]. A multi-label classification of UAV sounds was also proposed, incorporating the use of a stacked bidirectional LSTM-CNN architecture to improve the ability of the model to simultaneously classify many UAV sounds [46]. The accuracy in identifying and classifying UAVs using millimeter-wave radar data would improve through such feature-fusion architectures.

This study discusses the 60 GHz millimeter-wave radar data collected by a bionic bird and two types of drones, DJI Mavic and DJI Phantom 3 Pro (see the illustration of these classes in Figure 1), and they are then applied to various machine learning algorithms. The performances of several models such as LSTM, GRU, CNN, and Transformer are compared, considering their robustness when noise is artificially added to the dataset. Additionally, the Multimodal Transformer architecture is proposed, leveraging amplitude, phase, skewness, and kurtosis features for improved detection. These findings aim to

enhance detection accuracy in noisy environments and provide a benchmark for advanced radar-based UAV detection techniques. These results hold significant implications for public safety, logistics, and agriculture, where identifying small drones is increasingly critical for security and operational efficiency.

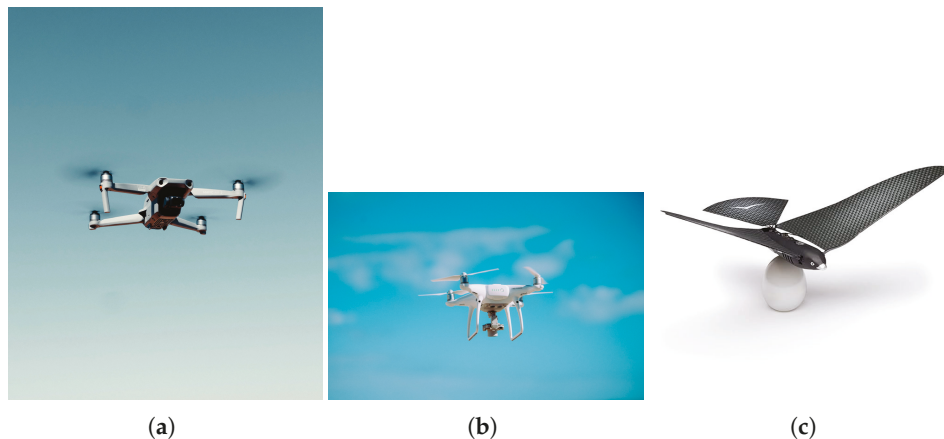


Figure 1. Images of the test subjects used in the experiments: (a) the DJI Mavic drone [47], (b) the DJI Phantom 3 Pro drone [48], and (c) the Bionic Bird [49]. These devices illustrate what were used to evaluate the classification performance of the Multimodal Transformer model.

2. Overview of the Related Research

Recently, the detection and analysis of micro-Doppler signatures from UAVs, particularly multi-rotor drones, has received considerable attention. Yang et al. (2022) [21] proposed a method to improve the rotor blade micro-Doppler parameters estimation through complex variational mode decomposition (CVMD) and singular value decomposition (SVD). Meanwhile, this method addresses some of the limitations of traditional methods.

Hanif et al. (2022) [22] presented a comprehensive review of radar micro-Doppler signature analysis techniques, with a particular focus on target classification. Their work highlights the evolution of the feature extraction methods and limitations of existing approaches. This provides a solid background for addressing small target detection—which aligns with the challenges of classifying these drones or aerial targets.

Soumya et al. (2023) discussed recent advances in radar sensor technology by focusing on mmWave radars [23]. Their review highlights the possible combination of radar data with machine learning algorithms for an improved overall performance in different applications including UAV detection. This inclusion of machine learning fits into our research approach to improve the way IRIS UAV detections can be completed with neural network models.

Kumawat et al. (2022) proposed a deep convolutional neural network (DCNN) for detection and classification of UAVs via their micro-Doppler signatures [24]. Moreover, their dataset “DIAT- μ SAT” provides full radar signature information, which can further enhance the accuracy of detection. These efforts in under-the-hood DCNNs and custom datasets are close to our heart as we refine methods for UAV detection using radar data and emerging machine learning approaches.

Lu et al. (2024) [27] illustrated further advances in the detection using radar, where the use of exponential complex sinusoidal basis (ECSB) allowed for the extraction of micro-Doppler signatures with a high classification accuracy. Similarly, Kang et al. (2024) [28] provided valuable experimental data using 60 GHz mmWave for both drone and bionic bird micro-Doppler analysis by placing further emphasis on how accurate feature extraction is important in UAV detection.

A related study by Narayanan et al. (2023) [29] utilized TensorFlow to classify radar-based micro-Doppler spectrograms. The model resulted in extremely high accuracy when tasked with differentiating between the drone types/birds. Their work thus further underlines the fact that the choice of robust classification models is an important consideration and constitutes one cornerstone of our approach for the detection of UAVs.

Finally, Fan et al. (2023) [31] proposed a clutter suppression method using an orthogonal matching pursuit, which effectively enhances the clarity of the micro-Doppler signal from environmental interference. This technique is particularly relevant for improving the performance of detection under poor environmental conditions.

3. Foundational Theories and Methodological Approaches

3.1. The Detected Signal Representation

The radar signal model of multirotor UAVs is intrinsically based upon the interaction of radar with both rotor blades and the body of the drone. In this model, the reflected radar signal takes a specific form in micro-Doppler signatures. Indeed, those signatures are produced not only by the UAV itself, but also by the action of radar on the flapping wings and bodies of birds. Mathematically, the micro-Doppler signal can be expressed as follows [27,31]:

$$s(t) = L \exp\left(-j\frac{4\pi}{\lambda}R_0\right) \sum_{p=0}^{P-1} \sum_{\kappa=0}^{K-1} \text{sinc}(\psi_{p,\kappa}(t)) \exp(-j\psi_{p,\kappa}(t)), \quad (1)$$

where:

- L is a constant related to the radar system;
- λ is the wavelength of the transmitted radar signal;
- R_0 is the distance between the radar and the center of the UAV;
- P is the number of rotors on the UAV;
- K is the number of blades per rotor;
- $\psi_{p,\kappa}(t)$ is the Doppler phase associated with blade κ of rotor p ;
- $\text{sinc}(\psi_{p,\kappa}(t))$ models the spatial response of the Doppler shift of each blade.

The formula represents the signal captured by the mmWave radar sensor datasets [25,28], and it also considers that it has been fitted to the specific context of drone detection or similar objects, including the micro-Doppler behavior in the case of a captured signal. The latter is a complex number array, which may be considered typical for the signals recorded by mmWave radar. These complex numbers encode the magnitude and phase of the reflected wave, both of which are fundamental to micro-Doppler calculations.

The key ingredients in the mathematical modeling of micro-Doppler behavior include the following: L , which is a constant parameter associated with the radar system—the value that should be adjusted to conform to specific conditions of the radar setup; R_0 , which denotes the distance between the radar and object-under-test—a drone in our case—and it determines the amplitude and phase of the complex data acquired. The other parameters are as follows: P denotes the number of rotors, while K denotes the number of blades per rotor; these two jointly determine the frequency modulation pattern observed in the radar signal.

The $\exp\left(-j\frac{4\pi}{\lambda}R_0\right)$ -term is the interaction with the radar, which actually is the phase of the electromagnetic wave arriving back to the radar modulated by the distance from the radar to the object. It reflects the phase changes observed in the complex data. $\text{sinc}(\psi_{p,\kappa}(t)) \exp(-j\psi_{p,\kappa}(t))$ models phase modulation due to the movement of the rotor blades and gives the micro-Doppler signature that is typical for multirotor UAVs. This

can also be interpreted as phase variation in the captured complex signal, whereby one notices changes continuously in both the real and imaginary parts with respect to time.

Although this is a formula that mathematically models an ideal scenario, it captures all the main features in the behavior implicit in the radar data. However, the actual signal will be noisy, with fluctuations in environmental influence, several reflections, and disturbances around the sensor. In that aspect, filtering or any other noise separation techniques may be necessary to further refine the model and fully apply to the context at hand.

3.2. Noise Models in the Original Signal Representation

Accurately introducing different types of noise to the modeling of radar signals is crucial for enhancing both the reliability and precision of the model. Realistic radar signals are susceptible to different forms of noise; each has a different definition and depth of influence on the integrity of the signal.

White noise is typically modeled as Gaussian noise, and it represents the electronic noise intrinsic to radar systems. It has a zero mean and a variance σ^2 , which is mathematically described as

$$n_w(t) \sim \mathcal{N}(0, \sigma_w^2). \quad (2)$$

It reduces signal clarity, especially in areas with significant electronic interference.

Impulsive noise: sporadic high-amplitude pulses that are modeled by a Poisson distribution with rate λ :

$$n_i(t) = \text{Poisson}(\lambda). \quad (3)$$

This type of noise may result from sudden interference, an electrical discharge, or an abrupt change in the source signal that masks and/or distorts the signal being transmitted by the radar.

Noise following the Pareto distribution can be modeled as follows:

$$f_p(x; \alpha, x_m) = \frac{\alpha x_m^\alpha}{x^{\alpha+1}}, \quad (4)$$

where α and x_m are parameters of the distribution. This noise has heavy tails. It can bring large, infrequent disturbances in the signal as representation of a rare event can significantly impact radar data interpretation.

Multi-path interference: The reflection of radar signals from various surfaces to the receiver causes interference among several late-time copies of the same signal. This scenario arises when surroundings contain high-rise buildings, or if the terrain is complex enough to significantly alter the waveform of the received signal. The resultant interference may degrade target detection or tracking due to error. Multipath noise can be modeled mathematically as follows:

$$n_m(t) = \sum_{k=1}^K \alpha_k s(t - \tau_k), \quad (5)$$

where α_k represents the magnitude of the k th multi-path component, $s(t)$ is the transmitted radar signal, and τ_k is the time delay associated with each multi-path reflection. With different noise types considered, the extended representation of the radar signal is given as follows:

$$s(t) = L \exp\left(-j\frac{4\pi}{\lambda}R_0\right) \sum_{p=0}^{P-1} \sum_{\kappa=0}^{K-1} \text{sinc}(\psi_{p,\kappa}(t)) \exp(-j\psi_{p,\kappa}(t)) + n_w(t) + n_i(t) + n_p(t). \quad (6)$$

This general model provides better processing and analysis of the signals; hence, an appropriate method for detecting and interpreting radar data with many sources of noise and effects of interference. This model, since there will only be one term to account for each type of noise, can accommodate myriad disturbances associated with radar signals. Finally, it offers a close approximation to actualities in the operation of radar, hence offering robustness and more realistic accuracy in the representation of radar signals.

3.3. Machine Learning Algorithms for Drone Detection

Recent advances in signal processing and machine learning have considerably enhanced the performance of sensors in detecting UAVs [3–5]. Higher detection accuracy and reliability have, therefore, enabled the identification of small-sized drones that become relevant in many applications related to surveillance and recreational use, such as, for instance, [6,8,10,13–17,24,27,29,32–35]. These, when integrated with radar systems, ensure superior spatial resolution and, more importantly, effective operations in difficult environmental conditions for more robust UAS detection solutions [21–24,27,29].

Several machine learning models were implemented to investigate their performances related to the detection and classification of UAVs. In particular, both LSTM networks [5,8,11,18] and GRU [11,18] have the particular advantageous capability of temporal dependency and sequential data handling. This is crucial, as radar signals are time-series data, and flying patterns of aerial objects are not steady in certain conditions. They are also particularly superior in representing the spatial hierarchies and local patterns that can be found in radar data. Therefore, it can effectively capture distinguishable features in decomposed signal components.

CNNs are adept at catching such spatial hierarchies and local patterns in radar data, serving the purpose intrinsically for distinct feature identification from decomposed signal components [4,5,8,10,13–17,22,24,34,35]. Furthermore, Transformer models are also employed for their ability to handle long-range dependencies and self-attention mechanisms with improved efficiency [15,16]. These are highly crucial to understand the complex interaction within the radar signals themselves for correct classification. Among the selected algorithmic approaches, some bear respective strengths to address the nature of the data with a particular interest in sequential and spatial features to enhance overall performance and reliability for any given real-world scenario of detection and classification.

Many drone detection algorithms are based on machine learning methods applied either in conjunction with or as standalone alternatives to signal decomposition techniques [7–9,21,22,27,33] and attention mechanisms [13,15,16]. Other methods rely on visualization techniques of the signals based on spectrograms [6,22,29,36] to detect and classify drone activities. Other approaches are simply statistical techniques [12,18,30,36], and signal analysis is performed without machine learning.

While traditional methods are extremely efficient in analyzing patterns or trends within data, machine learning algorithms perform extremely well when modeling complex, nonlinear relationships that may not be as saliently evident by purely statistical techniques. This algorithm can learn features automatically from enormous datasets, reducing the need for manual feature extraction.

Machine learning models can adapt and improve over time; thus, their functionality will be even better in environments that have considerable variability and challenge—even highly interfering or noisy ones. Unlike visual signal methods, which may depend on specific patterns or thresholds, machine learning can generalize over different signal types and conditions, hence making the robustness and accuracy of drone detection better.

3.4. Overview of the Dataset Employed in This Research

Kang et al. (2024) [28] presents the radar signal collection collected through a 60 GHz radar sensor. This data structure falls within three dimensions, which are important for the functionality of the sensor and its interaction with the environment.

The first dimension stretches across 1000 entries, which represent distinct time instances or “frames” in temporal sequence that capture the motion or otherwise appearance of objects (drone, bird, etc.) present in the radar field of view. At an extremely high level, these frames represent the snapshots of scenes observed during one period of running.

The second part has 128 subdivisions in each frame; those windows act as smaller time windows giving enhanced temporal resolution. This allows sensing the micro-Doppler phenomenon of rapid wing flapping or rotor motion, which is significant for classifying objects by type and activity.

The third dimension has 168 functional attributes acquired during each time window, including real and imaginary parts, amplitude, phase, and other properties of the radar signal, enabling an insightful view of the moving, shape, and material properties of the target. This will allow for a proper training setup of machine learning models on object classification based on movement patterns or physical properties.

The dataset features hovering UAVs (e.g., Mavic series and Phantom 3 Professional) and a bionic bird, which are more difficult to detect due to small sizes of the radar cross section (RCS) and their composition in materials like carbon fiber and plastic that reflect a slight radar energy [27]. Additionally, hovering drones are more difficult to detect because they generate weak Doppler shifts. Simulated signals with overlaid noise will also further examine how pathologically close to natural such mimic systems could become.

4. Results

4.1. Introduction to the Results

This section describes the outcome of an investigation into the detection and identification of drones using machine learning algorithms applied to 60 GHz millimeter-wave radar data. The collected radar signals were analyzed with a bionic bird and two drones, i.e., DJI Mavic and DJI Phantom 3 Pro, under conditions in which artificial noises had been introduced. Major models like LSTM, GRU, CNN, and Transformer were also trained and evaluated based on the completed noisy dataset to establish performance metrics such as accuracy and error rates for degraded environments.

Full imitated realistic operating conditions included the full addition of white noise, Pareto noise, impulsive noise, and multipath interference. The boxplots present the data classification distribution, assuring demonstration of the algorithm’s capability in being able to classify between classes. The performance of the models was analyzed using ROC curves, and phase and amplitude data formed the prime input features for all of the models in this benchmark.

Furthermore, we extended the dataset to add two more statistical features—kurtosis and skewness—to improve detection and overcome the challenge of noise. Hence, we used Multimodal Transformer architecture, integrating them with amplitude and phase data. This enhancement resulted in better classification accuracies and robustness against noise, hence presenting a more efficient and reliable method for radar-based UAV detection. This result highlights the effectiveness of the proposed method, particularly in scenarios where robust performance is required despite the presence of noise.

4.2. Noisy Signal Visualization

This research discusses the effects of various noise patterns on mmWave radar signals, such as white noise, Pareto noise, impulsive noise, and multipath interference, which entails some parameters controlling their intensities and behaviors.

White noise is determined by the noise factor parameter, whose value is set to 10, indicating that it will multiply the standard deviations of the real and imaginary parts of the signal's mean and add a Gaussian effect centered around the signal's average.

For Pareto noise, the weight of the noise tail is defined by the alpha parameter with a value of 1.5, as well as by the noise factor, with a value of 10, which indicates the deformation of the intensity toward the relative correlativeness of the noise computed in relation to the standard deviation of the signal.

Impulsive noise was modeled with impulse ratio of 1, thereby indicating that about 100% of the samples will be affected, as well as scale factor of 10, which was used to define the amplitude of the impulsive noise.

Multipath interference was defined using specifiers number of paths (10 reflections), maximum delay (10 samples), and maximum intensity (10) to simulate reflections that have different delays and intensities.

These noise models simulate real-world scenarios where interference occurs with the parameters chosen to reflect the average environment that radar systems might encounter.

As shown in Figure 2, for amplitude, the white noise showed a well-defined range of values, while its interquartile range appeared to be very narrow but was compensated with a significant number of high-magnitude outliers. These outliers signify that white noise contributed to the variability in the data, but most of the amplitude values were focused around the lower end. The presence of such extremes suggests that when signals become distorted in their dynamics, particularly in the amplitude domain, it would affect the later processing or modeling of data. Phase, as per the right boxplot, seemed to be less influenced when compared to the amplitude with white noise. The interquartile range was wide, with most values in a controlled range between approximately 1 and 3. Furthermore, there were no extreme outliers in the phase data, which indicates that the phase was relatively immune to white noise contaminations. Thus, white noise does not degrade the phase information but allows it to be preserved for certain applications.

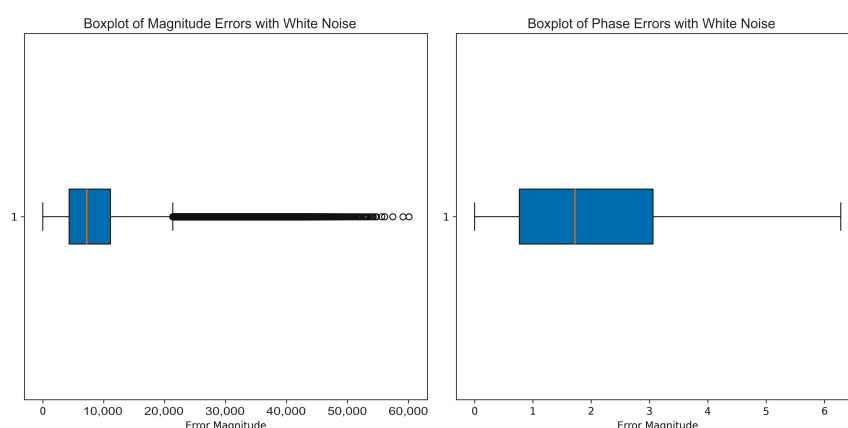


Figure 2. This boxplot shows the impact of white noise on the amplitude (left) and phase (right) of the radar signal. White noise follows a random distribution, primarily affecting the outliers in both amplitude and phase. The amplitude exhibits a broader spread, with more pronounced outliers in both directions. The phase is also impacted, though to a lesser extent, showing a slight median shift and a moderate interquartile range expansion.

For the Pareto noise, as shown in Figure 3, the amplitude (left) underwent drastic changes from the low levels of noise, as seen by a very broad range of values but also many outliers with very high magnitudes. Most of the data points tended to cluster around zero with an interquartile range of zero, which means that the noise had introduced a highly skewed distribution. This phenomenon aligns with known properties of Pareto noise, which can generate long tails with extreme deviations. The phase (right) seemed to not be disturbed by the Pareto noise so much. The interquartile range was well defined, with most values lying in an extremely limited range and having no visible outliers. This seems to indicate that at least the noise was affecting the phase a bit but, in comparison to the amplitude, its effects are mild. The limited distribution signifies that perhaps the phase data can yet be slightly interesting without going beyond what is needed for a model or an analysis.

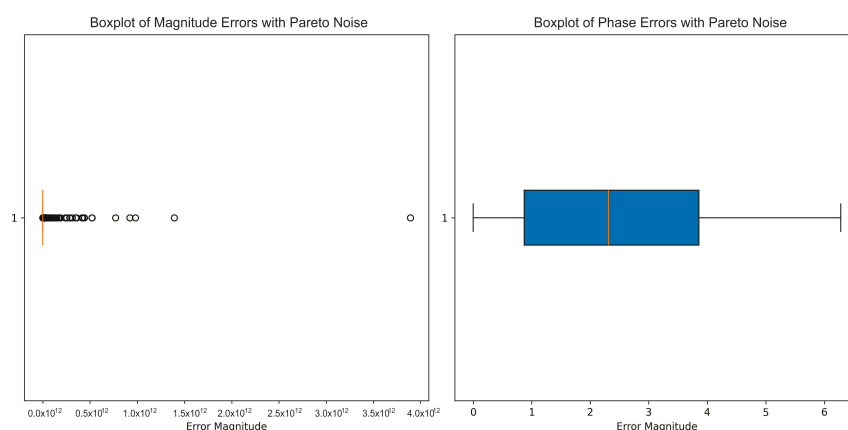


Figure 3. This boxplot illustrates the impact of Pareto noise on the amplitude (left) and phase (right) of the radar signal. Pareto noise, also known as heavy-tail noise, introduces extreme values more frequently than white noise, resulting in greater data dispersion. The amplitude plot shows a considerable number of high-value outliers, suggesting that the noise causes more frequent positive fluctuations. The phase remains relatively stable, with occasional extreme values.

The boxplots in Figure 4 show the effects of impulsive noise on the amplitude component (left) and phase component (right) of radar signals.

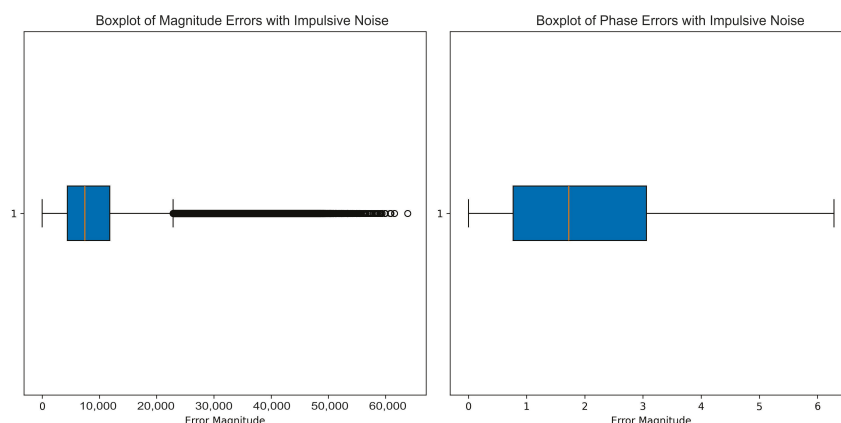


Figure 4. This boxplot illustrates the effect of impulsive noise on the amplitude (left) and phase (right) of the millimeter-wave radar signal. Impulsive noise generates abrupt, random spikes, causing significant data dispersion. The amplitude plot shows a noticeable increase in outliers at both extremes, with a wider interquartile range. While the median amplitude remains relatively stable, the data's extremes were widely scattered. The phase plot shows a similar pattern, with more visible outliers and a slight median shift.

Amplitude-wise, impulsive noise causes high variability, with IQR clumped near low magnitudes, and the median is nearly to the bottom of the IQR. The most pronounced feature is denoting several extremely high outliers. These outliers show that impulsive noise severely deforms some amplitude data points, resulting in isolated spikes with high errors. Such behavior conveys that the impulsive noise contributes random but high amplitude deformations to the types of signals. On the contrary, the phase data were quite resistant to impulsive noise. The IQR was broad; most values lay between approximately 1 and 3 in error magnitude. The median phase error was located well within the IQR, thus indicating a more consistent distribution. More importantly, the nonappearance of extreme outliers in the phase data means that impulsive noise affects the phase more uniformly, so it does not lead to high deviation values. This is indicative of the robustness of the phase with regard to the sporadic high-intensity character of the impulsive noise when compared to the amplitude.

Under multipath interference, as shown in Figure 5, the IQR was thus widely spread for magnitude errors. The median was a bit below the center of the IQR, indicating a slight skew toward lower values. The whiskers indicate a large error magnitude. The fact that outliers were not visible indicates that most of the errors caused by multipath interference were certainly falling within this interval, leading to a monotonous but high-impact effect on the magnitude of the signal.

The phase errors demonstrated a completely opposite distribution. The IQR was also significantly below the magnitude error, with the midvalue situated right inside the IQR. It indicates a more centralized and fierce distribution of effect of the multipath interference on the phase component. Comparatively, phase error ranges beyond that defined by the IQR are indicated by the whiskers of the box, although the comparatively small size of the box suggests that such extreme deviations are rare events. Overall, a lack of extreme outliers suggests that multipath noise is rather gentle in introducing errors in the phase when compared to magnitude.

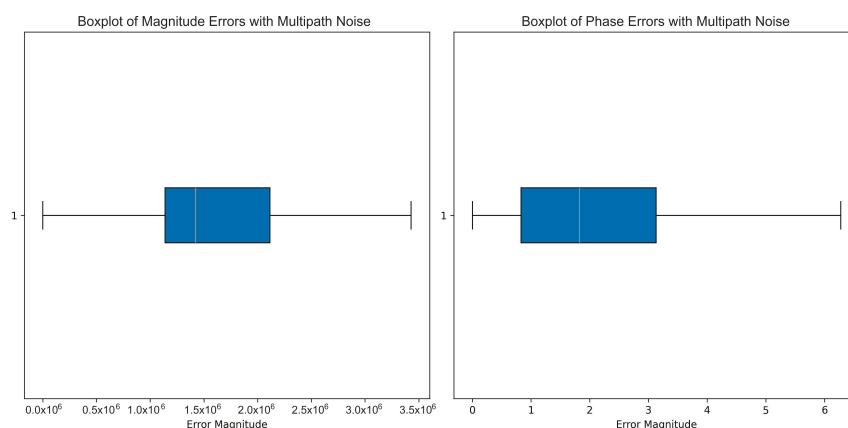


Figure 5. This boxplot shows the effects of multipath interference noise on the amplitude (**left**) and phase (**right**) of the radar signal. Multipath interference occurs when the signal reflects off multiple surfaces before reaching the receiver, causing distortions. The amplitude plot reveals increased variability and a larger number of outliers, indicating inconsistencies in the measured values. The phase is less affected but still shows a slight increase in dispersion compared to the original signal.

4.3. Comparative Analysis of the Algorithms

The performance of four algorithms, LSTM, GRU, Conv1D, and Transformer, was compared in two scenarios: the clean and noise-contaminated signals with four types of noises: white noise, Pareto noise, impulsive noise, and multipath interference. The trade-offs between true positive and false positive rates at different thresholds are shown in these results using ROC curves. The four models were then discussed in terms of how they

were used to classify three dissimilar classes in the dataset: Mavic drones, Phantom 3 Pro drones, and bionic birds. These four architectures were expected to hold all comparisons between different approaches in deep learning, from sequences to attention-based models.

Table 1 summarizes the architectures of four machine learning models: LSTM, GRU, Conv1D, and Transformer. Long short-term memory (LSTM)—a recurrent neural network (RNN) that enables the model to capture temporal relationships in sequence data. In this architecture, there are two LSTM layers of 64 and 32 units, respectively, which then goes down to the dense output layer where softmax activation is performed to classify the input to one of the three classes. LSTMs can retain long-term dependencies and can analyze continuous signals over time. This is crucial for detecting subtle differences between signals that are very similar, such as between radar signatures of drones and birds. The differences may not be easily detected in small segments of data. However, LSTM models are more computationally extensive and may take longer to train than other architectures.

Table 1. Implementation details of LSTM, GRU, Conv1D, and Transformer models.

Model Type	Architecture Details	Input Features and Objective
LSTM	Two LSTM layers: - 64 units with return sequences - 32 units Output layer: Dense with softmax activation	Input: Amplitude and phase Objective: Capture temporal dependencies in sequential data
GRU	Two GRU layers: - 64 units with return sequences - 32 units Output layer: Dense with softmax activation	Input: Amplitude and phase Objective: Reduce computational cost compared to LSTM while retaining performance
Conv1D	Two Conv1D layers: - 64 filters with kernel size 3 - 32 filters with kernel size 3 Flatten layer Output layer: Dense with softmax activation	Input: Amplitude and phase Objective: Extract local patterns and features from sequential data
Transformer	Layer normalization, Multi-Head Attention (8 heads, key dimension 128), Batch Normalization, Dropout (0.2), Global Average Pooling, Dense layer (64 units), Output layer: Dense with softmax activation	Input: Amplitude and phase Objective: Leverage self-attention mechanisms for better feature extraction and pattern recognition in sequences

The GRU model is another variant of RNN like LSTM but has a far simpler structure than LSTM. The GRU implemented in this study had two layers of 64 and 32 units and a softmax output layer. The design of GRU addresses the problem of vanishing gradients in older versions of RNNs, but with far fewer numbers of parameters compared to LSTM, making it very cost-efficient and easy to train. In addition to saving time, the simplicity of the GRU model proved to be advantageous, particularly when the experiment being conducted involves large datasets or has lesser computational resources. Furthermore, the GRU model captures accurate temporal patterns of amplitude and phase data from drone and bird detection efficiently without overfitting. Generally, the results are comparable to using LSTM for this, but the lighter architecture provides quick experimentation and tuning for the task.

The Conv1D module takes a one-dimensional approach, utilizing convolutional layers to highlight the model's particular strength in local pattern extraction. The model works simply by making use of two Conv1D layers—consisting of 64 and 32 filters, respectively, each with kernel size equal to 3—and it is followed by flattening and by a dense output layer being triggered by softmax. The Conv1D models lacked recurrence and use sliding

convolutional filters rather than form patterns, which were identified by fixed-size windows in the input sequence. Here, we combined varying signal characteristics as they could indicate frequency changes in the radar signals. Conv1D forms excellent models for short-term variation detection, as expected from RNNs-only models, though they may not be able to learn longer-term dependencies. They may, however, be effective in detecting very specific local features that make one class slightly different from another.

Transformer, the most advanced architecture based on attention mechanism, has become worlds apart from conventional sequential architectures as it can simultaneously process minutes of parallel input. The architecture performs self-attention that distinguishes important features at different time steps. In the Transformer model that was implemented, there was sequential layer normalization followed by the multi-head attention layer, with 8 heads and a key dimension of 128. Further, the dropout layers were the regularizer, and the final output was obtained after passing an input through batch normalization and a 64-unit dense layer before sending it into a softmax layer. Multihop attention facility offers an enveloping context to the model for simultaneous attention on different aspects of radar signal, which makes the difference of it identifying various signatures of Mavic drones, Phantom 3 Pro drones, and bionic birds. The attention mechanism in this case also would signify that a given model could focus on the important portions of the input sequence, like the sudden spikes in amplitude or phase shifts, which could be critical for classification.

Table 2 summarizes the algorithmic complexity of the implemented models being measured: LSTM, GRU, Conv1D, and Transformer. The Time Complexity column indicates the computational expense of each model in sequence length (n) and hidden dimensions (d). Both LSTM and GRU were implemented according to the time complexity $O(n \cdot d^2)$ with the property that they take as inputs iteratively during each time step, which made them ineffective when the length of the sequences increased. Conv1D, in comparison, had time complexity $O(n \cdot k \cdot d)$, where k is the kernel size. It suffered significantly because the convolutions were conducted on the entire input sequence, rather than at individual time steps. The Transformer model, according to time complexity, assumes $O(n^2 \cdot d)$ due to a self-attention mechanism that calculates pairwise interactions between all the tokens present in the input sequence.

Table 2. Algorithmic complexity of the LSTM, GRU, Conv1D, and Transformer models.

Model	Time Complexity	Space Complexity
LSTM	$O(n \cdot d^2)$ Sequential processing at each time step High computational cost for long sequences	$O(n \cdot d)$ Memory for storing hidden states
GRU	$O(n \cdot d^2)$ Sequential processing at each time step High computational cost for long sequences	$O(n \cdot d)$ Memory for storing hidden states
Conv1D	$O(n \cdot k \cdot d)$ Parallel processing across input sequence Lower computational cost	$O(n \cdot d)$ Memory for storing feature maps
Transformer	$O(n^2 \cdot d)$ Quadratic scaling due to self-attention mechanism Pairwise interaction computation between all tokens	$O(n \cdot d)$ Memory for storing attention weights and hidden states

For space complexity, all of the models can roughly store hidden states or attention weights with $O(n \cdot d)$ memory scaling and both the sequence length and internal dimensions of the model. The Sequential Dependency column shows a key difference between LSTM and GRU regarding high sequential dependency, meaning that each time step de-

depends on the output of the preceding one, which precludes parallelization. By contrast, however, Conv1D and Transformer have no sequential dependency, which allows them to process input sequences in parallel, significantly improving efficiency in computation.

Finally, the column named Parallelization indicates that, though LSTM and GRU suffer from limited parallelization since they process data sequentially, both the Conv1D and Transformer models allow tremendous parallelism. Such characteristics make Conv1D and Transformer significantly more amenable to modern hardware such as GPUs, where parallel processing can significantly lower training time. Overall, this comparison illustrates that, although recurrent models may capture temporal dependencies more effectively, nonsequential models such as Conv1D and Transformer can provide better scalability and efficiency, particularly for large datasets, like the drone and bird detection dataset used in this research. This is a hands-on evaluation of models against predetermined baselines—it measures performance only with noise. The curves show that there are sizable divergences between these algorithms in terms of their behavior under different noise types, thus marking out each algorithm's level of sensitivity to noise in signals.

An analysis of the boxplots (as shown in Figure 6) for classification probability under a white noise condition showed that the algorithms, namely LSTM, GRU, Conv1D, and Transformer, significantly differ in performance.

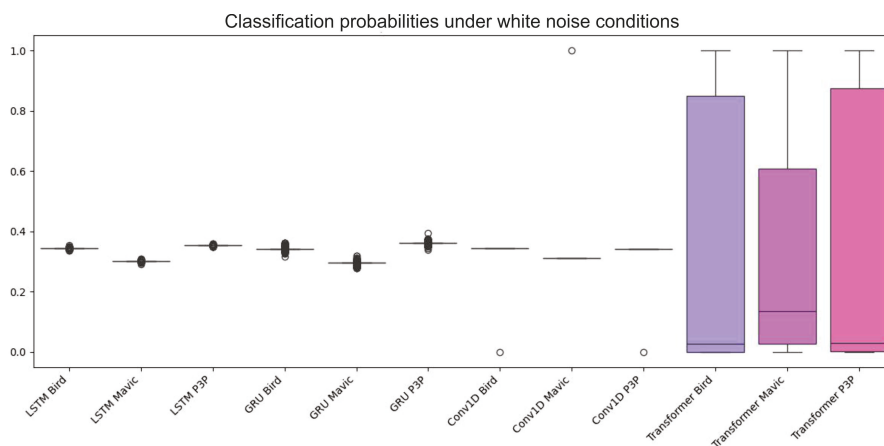


Figure 6. This boxplot illustrates the classification probability outputs of four algorithms—LSTM, GRU, Conv1D, and Transformer—under white noise conditions. The boxplot reveals that LSTM, GRU, and Conv1D exhibited tightly clustered probability distributions with low variance, and their median probabilities remained around or below 0.4 across all classes (Bird, Mavic drone, and P3P drone). This low variability and clustered median values suggests poor classification performance, with predictions lacking high confidence and distinguishing power. In contrast, the Transformer algorithm demonstrated a markedly different behavior, with wider interquartile ranges and higher median probabilities for all classes. The wider spread indicates that Transformer is more resilient to white noise, producing more varied and accurate probability outputs, thus highlighting its superior robustness in handling noisy data compared to the other models.

When looking for the ability of any of the algorithms to classify a given object (Bird, Mavic drone, or P3P drone, as shown in Figure 7), it shows low variability for those algorithms: LSTM, GRU, and Conv1D.

It also shows that their median probability values cluster near 0.4 or lower. These figures highlight the ineffectiveness of the algorithms in classification tasks as these values are not near those levels expected for highly confident predictions. Finally, there were also very close clusters with low variance, which shows that they have equally produced output that is not satisfactory in classification while under the white noise conditions. Transformer, however, displayed completely different behavior, characterized by wider interquartile

ranges and higher probability medians for all three classes. This indicates that it is a significantly more robust and capable engine in dealing with the noise dataset than all the other algorithms; hence, the wider and more accurate range of output probabilities predicted.

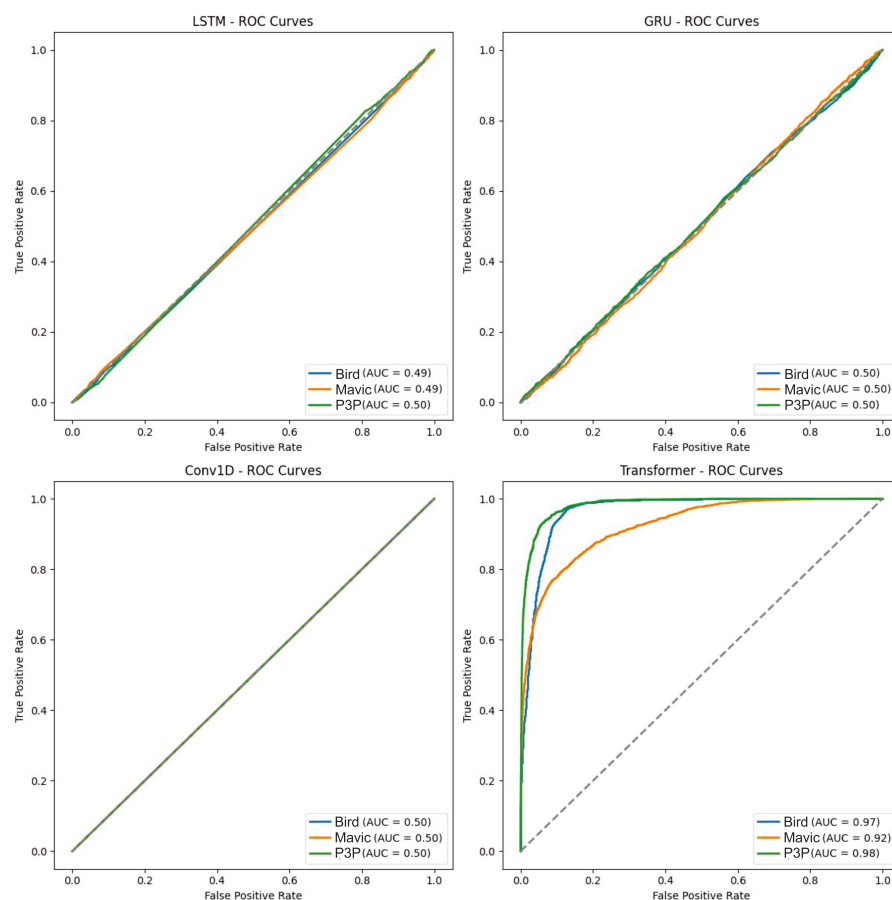


Figure 7. This boxplot presents the classification probabilities of four machine learning models—LSTM, GRU, Conv1D, and Transformer—under white noise conditions for the three target classes: Bird, Mavic drone, and P3P drone. The LSTM, GRU, and Conv1D models display tightly grouped probability distributions with narrow interquartile ranges and median values clustered near or below 0.4 across all classes. This indicates that these models struggle to produce confident predictions in noisy environments as their output probabilities remain low and exhibit limited variability, suggesting a uniform inability to distinguish between the classes under these conditions. In contrast, the Transformer model showed a significantly wider interquartile range and higher median probability values for all classes. This broader distribution highlights Transformer’s superior robustness to white noise, enabling it to generate more confident and diverse predictions across the noise dataset, outperforming the other models in terms of classification reliability under challenging noise conditions.

The ROC curves of LSTM, GRU, and Conv1D appeared almost diagonal and had AUC values of 0.49–0.50 compositively across the three classes. This shows that the algorithms are not capable of distinguishing the classes because their depicting performance is like random guessing under white noise conditions. Transformer, however, had exceptionally high AUC values: 0.97 for Bird, 0.92 for Mavic, and 0.98 for P3P. This resulted in poorly inclined ROC curves for Transformer since their placement was not in the direction toward the top left corner. However, they remained quite steep toward said corner, which is indicative of its superior classification ability even with noise.

The boxplot in Figure 8 indicates a considerable difference in the probabilities assigned to class labels by different models and classes when the noise being considered is Pareto

noise. Both LSTM and GRU seem to have higher median probabilities assigned to the ‘bird’ class, indicating a better performance in this category.

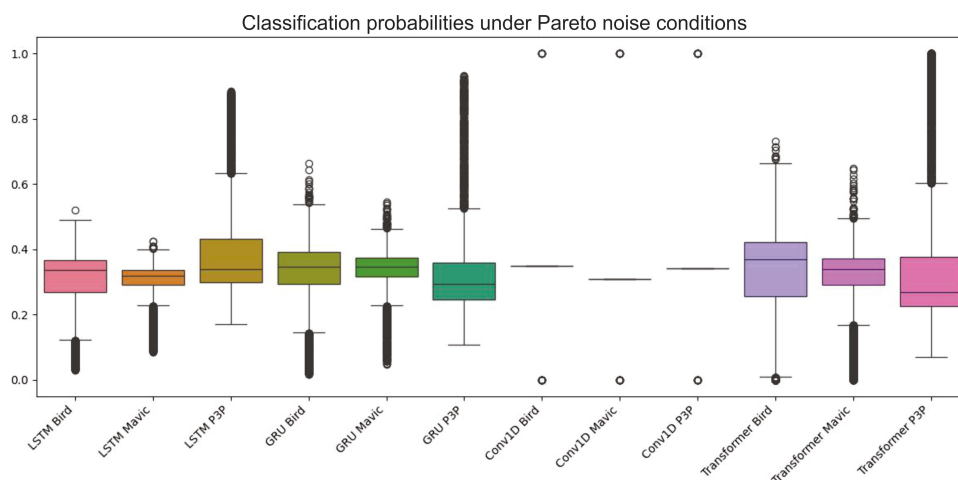


Figure 8. This boxplot illustrates the classification probabilities for different models—LSTM, GRU, Conv1D, and Transformer—under Pareto noise conditions across three target classes: Bird, Mavic drone, and P3P drone. The LSTM and GRU models exhibited higher median probabilities for the “bird” class, suggesting better performance in this specific category compared to other classes. However, the overall performance across all models was negatively impacted by Pareto noise, which introduces frequent extreme values (outliers) and disrupts the models’ ability to confidently assign accurate probabilities, particularly those affecting classification consistency.

Moreover, all of the classes of Transformer models performed fairly well because they showed a reduced number of outliers in the performance scores for all of the outputs.

These findings are further supported by ROC curves, where the Transformer model outperformed the rest of the models in terms of AUC scores by indicating better discriminative power across classes. Such Pareto noise definitely affects the performance of all models as it systematically decreased the overall accuracy and affected the positive side by increasing the number of false positives. This highlights the challenge these algorithms face in generalizing to noisy data. A potential area for future work could involve advancing modeling efforts to design more robust models, as stated in Figure 9.

The boxplot in Figure 10 shows the distribution of the classification probabilities across models and classes under impulsive noise, where the Transformer model consistently had a higher median probability, suggesting a higher confidence in predictions. The ROC curves in Figure 11 support these results.

The fact that the boxplot also shows wider ranges of probabilities for Transformer implies that, in some cases, it also has a higher level of uncertainty. In the results, typically, the Transformer model is better in terms of AUC values, with the major focus on Bird and Mavic classes, which mean there is better discrimination through Transformer among those classes with impulsive noise compared to the rest. The scenarios are similar for LSTM and GRU, although slightly lower in AUC than Transformer. The Conv1D model, however, performed poorly in most cases and particularly for the ‘bird’ class, indicating that it is a convolutional architecture, which is not suitable for the temporal dependencies within the data.

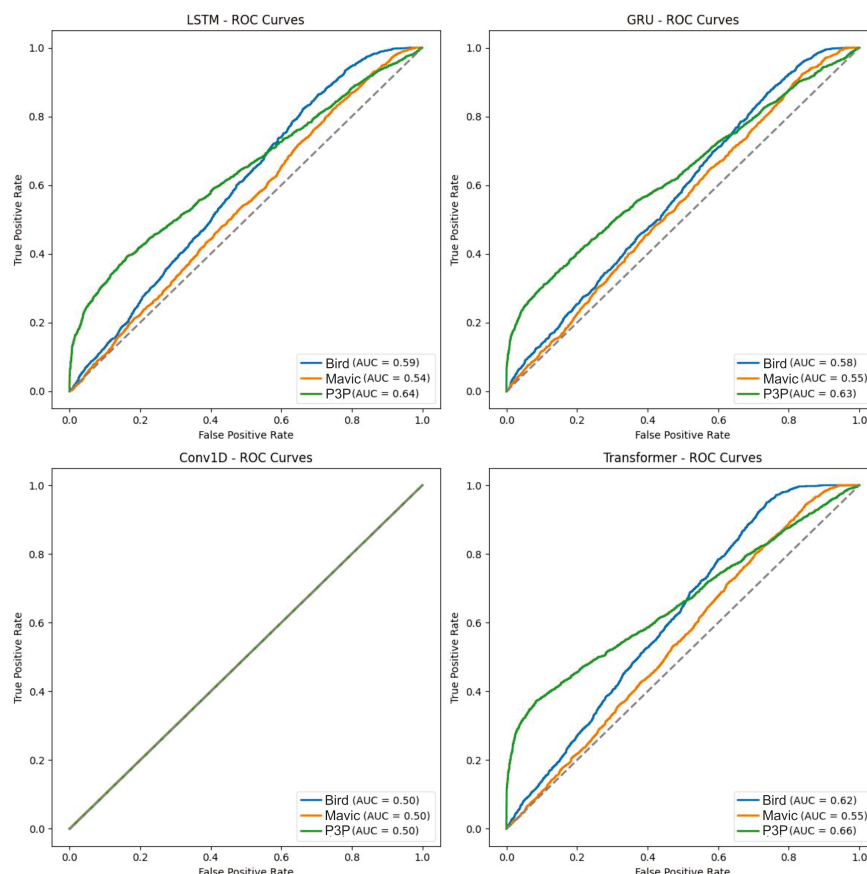


Figure 9. This figure presents the ROC curves and AUC scores for the classification performance of LSTM, GRU, Conv1D, and Transformer models under Pareto noise conditions. The Transformer model demonstrated superior performance, with higher AUC scores across all target classes, indicating better discriminative ability compared to the other models. Additionally, Transformer exhibited fewer outliers in classification scores, highlighting its robustness to Pareto noise. In contrast, the LSTM, GRU, and Conv1D models showed higher false positive rates, suggesting difficulty in generalizing to data with high variability caused by noise. These results emphasize the need for further model optimization to handle noise-induced challenges effectively.

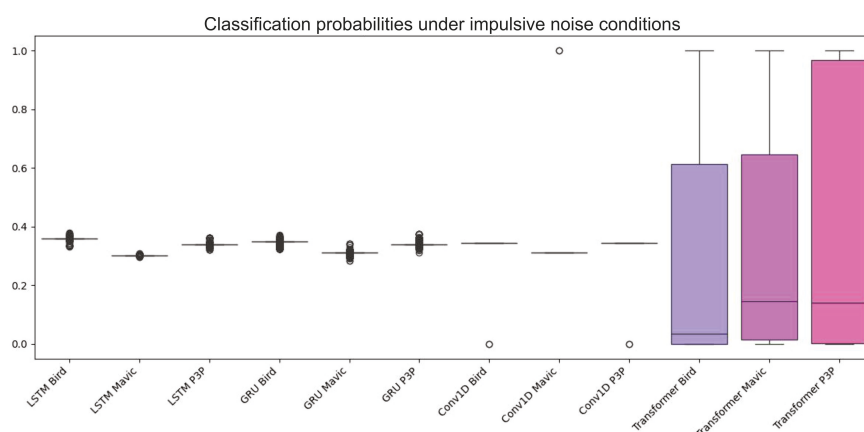


Figure 10. This boxplot displays the distribution of classification probabilities for the LSTM, GRU, Conv1D, and Transformer models under impulsive noise conditions across the Bird, Mavic drone, and P3P drone classes. The Transformer model consistently showed a higher median probability across all classes, indicating more confident predictions. However, its wider interquartile range suggests that it also exhibits greater uncertainty in some predictions. In contrast, the other models—LSTM, GRU, and Conv1D—showed lower median probabilities and tighter ranges, indicating less confidence and lower variability in their predictions under impulsive noise conditions.

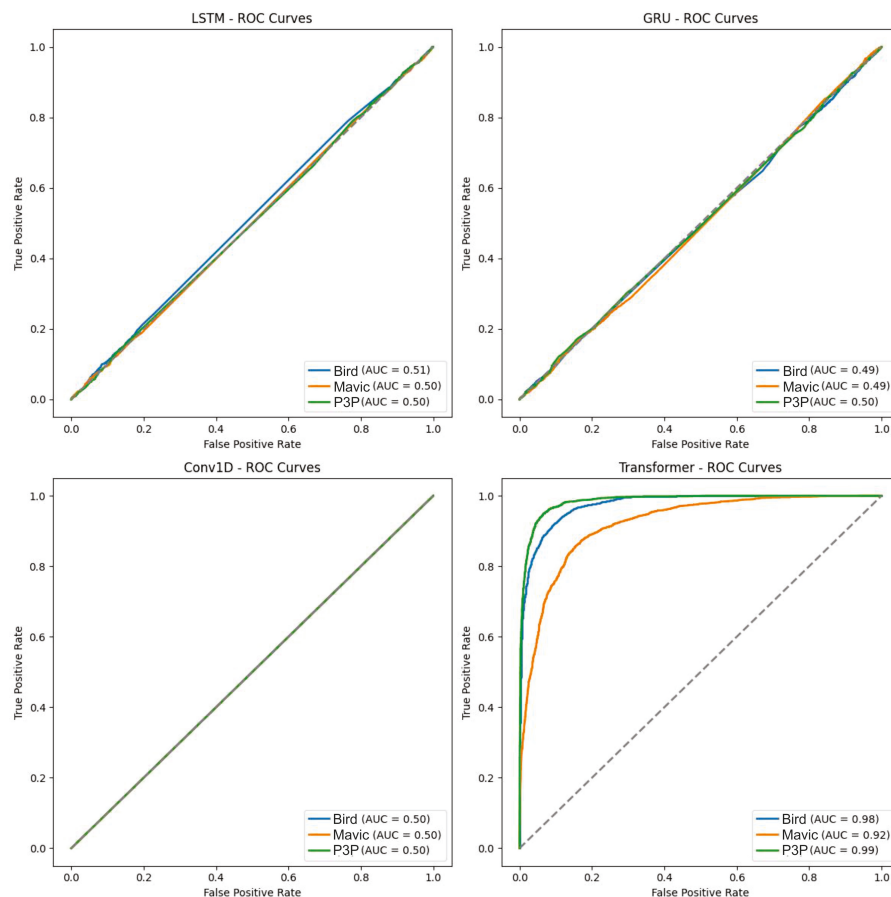


Figure 11. The ROC curves and AUC values demonstrate the classification performance of LSTM, GRU, Conv1D, and Transformer models under impulsive noise conditions. The Transformer model outperforms the other models, particularly for the Bird and Mavic drone classes, with higher AUC values, indicating better discrimination capabilities. The LSTM and GRU models show moderate performance but are slightly less effective than Transformer. The Conv1D model performs poorly across most classes, especially for the Bird class, reflecting its inability to effectively handle temporal dependencies in the presence of impulsive noise.

Classification performance suffers from multipath interference; in particular, this was evident for LSTM, GRU, and Conv1D. The higher variability of classification probabilities, as seen in the boxplot, indicates that these models, while not infallible, display less sensitivity to the noise created from multipath. In comparison, Transformer shows more stability in performance as it is robust under these extreme conditions. All these claims are borne out by the ROC curves, where Transformer's AUC scores were almost consistently superior to all the other models and can be considered the best under noisy conditions.

The boxplot in Figure 12 shows the data distribution of the classification algorithms considering multipath interference.

Figure 13 graphically shows the performance differences of the four machine learning algorithms when working on the classification of objects with multipath interference.

The Transformer model consistently outperformed LSTM, GRU, and Conv1D in terms of probabilities of classification and having a greater area under the ROC curve (AUC) value. This superlative performance is attributed to the extraordinary attention mechanism of Transformer in weighing the different parts of an input sequence and catching long dependencies. Both LSTM and GRU are wonderful in performing other tasks, however, and they have a considerable amount of confusion for multi-patterns created through multipath interference and, therefore, yield lower classification accuracies. Albeit, they

were poorly designed for the temporal dependencies prevailing in radar data while the CNN model is optimal for spatial data.

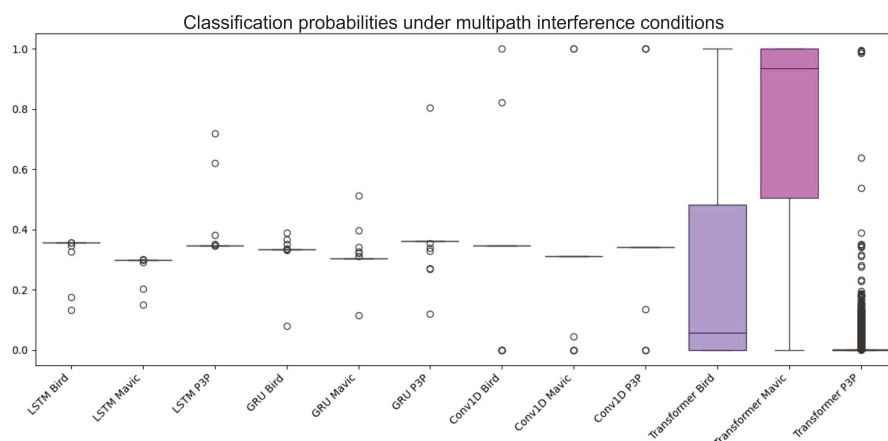


Figure 12. Boxplot illustrating the classification accuracy variability of the four machine learning algorithms under multipath interference. The Transformer model consistently demonstrated higher accuracy and lower variability, indicating superior stability and performance compared to the LSTM, GRU, and Conv1D models.

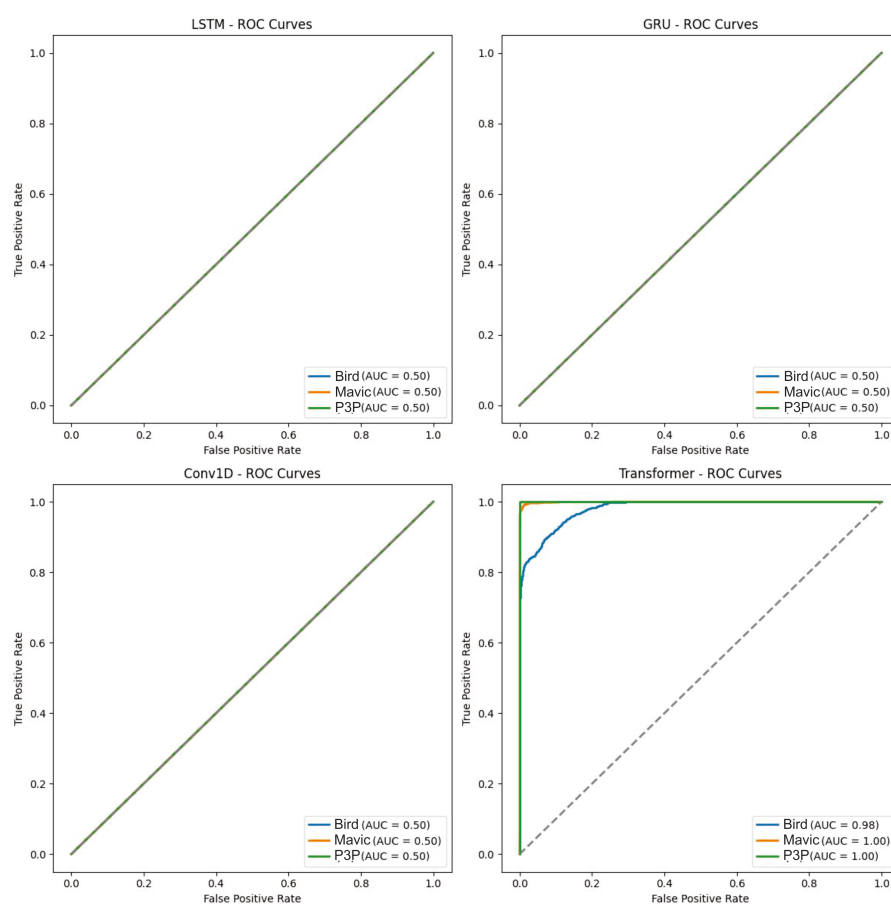


Figure 13. ROC curves showing the performance of the Transformer, LSTM, GRU, and Conv1D models in object classification with multipath interference. The Transformer model consistently outperformed the others, with higher AUC values, highlighting its robustness and superior attention mechanism for handling noise and temporal dependencies.

4.4. Proposed Method for Enhanced Drone Detection Efficiency

Voluminous and complicated data from several millimeter-wave sensors was put through a methodology that functions efficiently to classify the drones at computation and efficiency, as well as accuracy. The noisy and complex radar signal dataset was processed into four distinct features: the amplitude and phase as real physical features of the radar signal that are more direct inferences from the signal's inherent characteristics; and skewness and kurtosis because they are the derived statistical features that can capture higher-order moments of the distribution of signals.

The methods for extracting features from radar data are presented in Table 3.

Table 3. Extracted features from the radar data and their computation methods.

Feature	Description	Computation Method
Amplitude	Represents the magnitude of the radar signal, indicating the strength of the reflected wave	$ x $ (absolute value of the complex number)
Phase	Captures the angular component of the radar signal, which provides information on the relative position of the target	$\angle x$ (angle of the complex number)
Skewness	Measures the asymmetry of the amplitude distribution, helping to detect anomalies in the signal	$\frac{E[(X-\mu)^3]}{\sigma^3}$ (third standardized moment)
Kurtosis	Quantifies the peakedness of the amplitude distribution, indicating how heavy the tails of the distribution are	$\frac{E[(X-\mu)^4]}{\sigma^4} - 3$ (fourth standardized moment)

Amplitude is obtained from the magnitude of the complex signal, and phase is computed from the angular component. The skewness and kurtosis computed along the temporal axis capture the asymmetry and peakedness of the signal distribution. All temporal and signal channel measurements were incorporated into the derivation of these features.

Then, the algorithm processed the multimodal data with Transformer after feature extraction. Each one of these feature maps was saved in independent arrays in order to keep amplitude, phase, skewness, and kurtosis as distinct representations. Each input had its corresponding sample label mapped to numeric values indicative of the class it fell under (Mavic, Phantom 3 Pro, or bionic bird).

The extracted features were then subject to the Z-score normalization, which normalized all the data to a zero mean and unit variance across all of the samples. Then, all the feature maps were fused along the last dimension to create a coherent input tensor. This final fused tensor was a vector consisting of the whole multidimensional input, which captures complementary aspects of the radar signal as input to the Transformer network for classification.

These features were the input features for the Multimodal Transformer model [50], which was developed for drones and artificial bird classification. Combining the physical and statistical characteristics of the input signals leads to increased classification accuracy and robustness.

The Multimodal Transformer function constructs a deep learning model for classifying time-series data that is produced by radar sensors in order to categorize their data as that of Mavic drones, Phantom 3 Pro drones, and bionic birds. There have been several modifications made to the Transformer architecture to increase its capability in handling and interpreting multimodal radar data such as amplitude, phase, kurtosis, and skewness. The goal was thus to analyze these features and detect and classify small aerial targets based on micro-Doppler signatures, which is difficult because of the similarity characteristics between the two types of targets.

First, this function depicted in Figure 14 defines an input layer with a shape compatible with the dimensions of the dataset. Input consists of radar-derived features capturing

essential aspects of the motion of the target, such as the amplitude and phase variations over time. These features help distinguish various targets because drones and birds have distinct motion patterns. The input data have been normalized using a LayerNormalization layer, which means that, essentially, the data will have a common distribution and keep fat tails from dominating learning. This stabilizes and hastens convergence by reducing internal covariate shift.

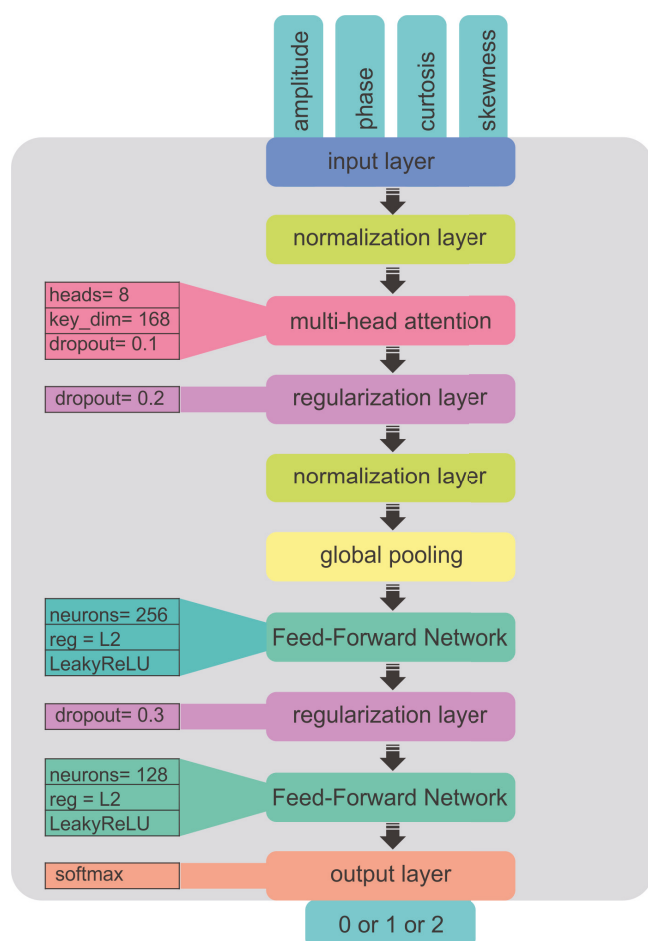


Figure 14. Schematic of the Multimodal Transformer Model (MMT) used for radar-based target classification. The model begins with an input layer processing radar features, followed by LayerNormalization for stable learning. Multi-head attention (8 heads) captures complex temporal dependencies in radar signals. Dropout layers (0.1 and 0.2) prevent overfitting. A GlobalAveragePooling1D layer reduces dimensionality, followed by two dense layers with L2 regularization and LeakyReLU activation. The final dense layer outputs classification probabilities using softmax, where the model's output is 0 for the Mavic drone, 1 for the Phantom 3 Pro drone, or 2 for bionic bird. The model was optimized with Adam and sparse categorical cross-entropy loss for multi-class classification.

The core of multi-head attention lies in its ability to simultaneously focus on different parts of the input sequence. It specifies eight attention heads and a key dimension that determines the size of the query, key, and value vectors used in the attention calculations. The dropout rate of 0.1 adds a regularization effect to prevent overfitting, where some neurons are randomly deactivated during training. Such an attention mechanism helps this model capture more complex temporal dependence relations within those radar signals, which is crucial for distinguishing between targets, such as different drone models and bionic birds.

Subsequently, following the attention layer, and subjecting to a dropout rate of 0.2, the model further continued to prevent overfitting. The out-of-attention mechanism proceeded with a normalized output that meets a GlobalAveragePooling1D layer that summarizes information throughout the entire sequence to a fixed-size vector. This pooling operation guarantees that the global capture of the model in the data occurs while output keeps the dimension down so that easier processing can be performed by the next dense layers.

Next, the model entered a new phase consisting of two dense layers that constituted a feed-forward network through which the first, in total, comprised 256 neurons. The layer incorporates L2 regularization to overhaul the weights and prevent them from being excessively large. Here, the activation function in use is a LeakyReLU as it addresses the problem of dead neurons, which arises with standard ReLU processing, by letting negative values sag slightly to allow a small gradient to move through them. This means it suits radar data well, where models should learn both positive and negative variations as input. To impose further regularization, a dropout layer with a dropout rate of 0.3 was applied subsequent to the first dense layer. The second dense layer was constructed in an identical manner, containing 128 units and LeakyReLU activation, refining the feature representation further.

Thus, the last layer is a dense layer with softmax activation as its classification output probabilities for the respective target class: Mavic drone, Phantom 3 Pro drone, or bionic bird. Thus, the sum of the output values was found to be equal to 1 and they were treated as a multi-class classification problem. The model is compiled with Adam as the optimizer with a learning rate of 0.0001 to allow the model weight rearrangement during training, which has a sparse categorical cross-entropy as the loss function that is suitable for multi-class classification tasks with integer-encoded labels.

The aforementioned modifications were made in the Multimodal Transformer model with regard to optimizing the performance concerning detection of Mavic drones, Phantom 3 Pro drones, and bionic birds in RADAR data. These modifications are described in Table 4. Learning rate schedulers, early stopping, and class weights are some of these measures that come into play concerning certain layers and blocks of the model to ensure accurate and balanced predictions in the device's different multidimensional classes.

First would be the Learning Rate Scheduler, which changes the learning rate dynamically during training, starting from a higher rate, thus allowing quick exploration of the parameter space, and then lowering toward the end for better convergence. Affected by this is the multiple head attention layers that work their magic in guaranteeing that the model receives relevant patterns of the radar signatures without instability. Furthermore, class weights are used to mitigate the class imbalance in the dataset, which is essentially critical for avoiding a biasing of the more represented classes and ensuring that the model trains freely for classifying Mavics, Phantom 3 Pro drones, and bionic birds.

The early stopping mechanism keeps track of the validation loss and interrupts the training after a specific point of not improving for five epochs. This is to avoid overfitting, thus maintaining the generalization capabilities of the models. Thus, dropout layers also benefit from this mechanism by regularizing without needing excessive training. The feed-forward network (FFN) was thus affected by class weights, which makes it strong in learning representations balanced across the classes.

The output layer using softmax for classification was influenced by class weights and early stopping. Hence, it ensures the probability distributions produced by the model are balanced and not biased toward any particular class, thus giving an overall improved performance in object detection across classes in the dataset.

Table 4. Impact of the training adjustments on the Multimodal Transformer model.

Model Block	Adjustment Applied	Expected Benefit
Input Layer	Class Weights	Balanced predictions across all classes
Multi-Head Attention	Learning Rate Scheduler	Stable convergence and improved pattern detection
Regularization Layer	Early Stopping	Reduced overfitting and better generalization
Feed-Forward Network	Class Weights & Learning Rate Scheduler	Improved representation of underrepresented classes
Output Layer	Class Weights & Early Stopping	Balanced predictions and optimal stopping point

These optimizations ensure Transformer effectively handles the complexities of the radar dataset, improving the detection accuracy of drones and birds, even under challenging conditions, such as noisy signals or similar movement patterns between different objects.

The model was trained for 15 epochs with a batch size of 64, and it used 10% of the data for validation. The performance metrics tracked the training and validation accuracy and loss.

The distribution of radar feature data, polluted with white noise (Figure 15), shows a much narrower and more compact distribution for the Mavic and Bird classes, with less dispersion, and it shows a wider data distribution for the drone P3P class.

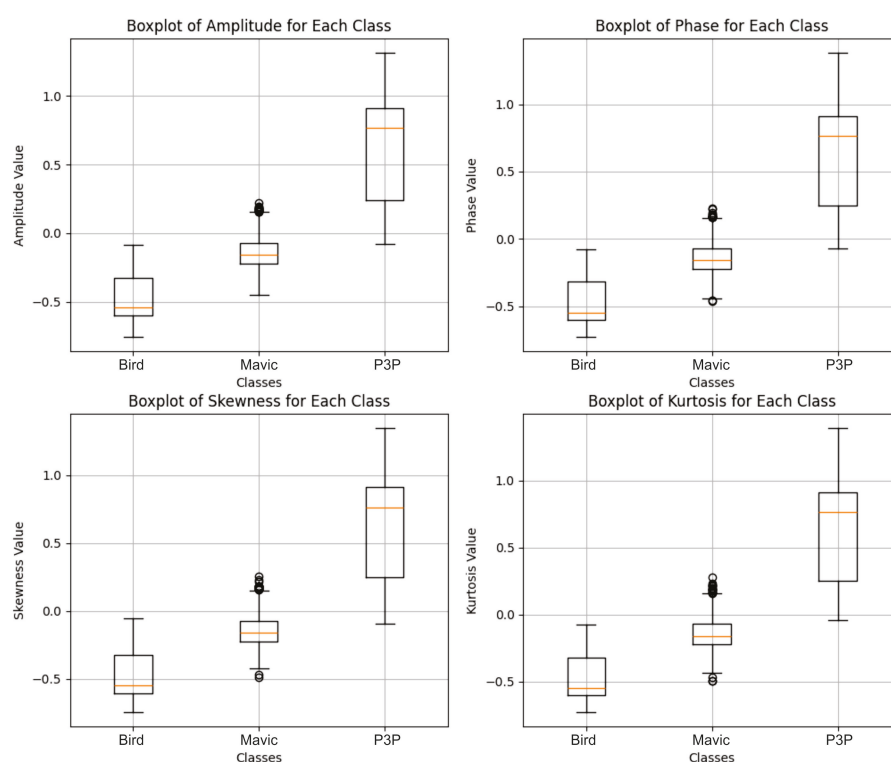


Figure 15. Boxplots showing the amplitude, phase, skewness, and kurtosis values for the Bird, Mavic, and P3P classes, as extracted from radar signals with added white noise. The Bird class showed lower and more stable values across all features. The Mavic class had moderate values with noticeable outliers. The P3P class consistently showed the highest medians and broader ranges, indicating stronger and more variable radar reflections. The differences in these features help to distinguish the classes in the Transformer model's classification process.

The original structure of the signal was maintained, which most likely facilitates the discrimination between the classes; consequently, a better classification is performed by the algorithm, as can be seen in Figure 16.

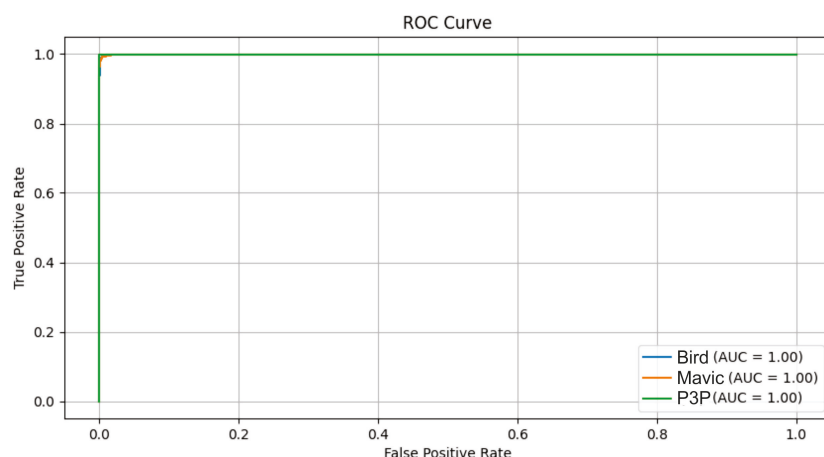


Figure 16. ROC. curves for the classification of the Bird, Mavic, and P3P classes. The model achieved a perfect AUC for all the classes in both noise cases. For white noise, the curves were slightly farther from the vertical axis compared to the Pareto noise, indicating a slightly better robustness to the latter noise type.

For Pareto noise, which was previously our critical case, even with Transformer, there were a couple of more visible outliers, at least in the amplitude and kurtosis measures, but the main bulk of distribution, as represented in the boxplot (Figure 17) through the interquartile ranges and medians, remained very similar across the classes.

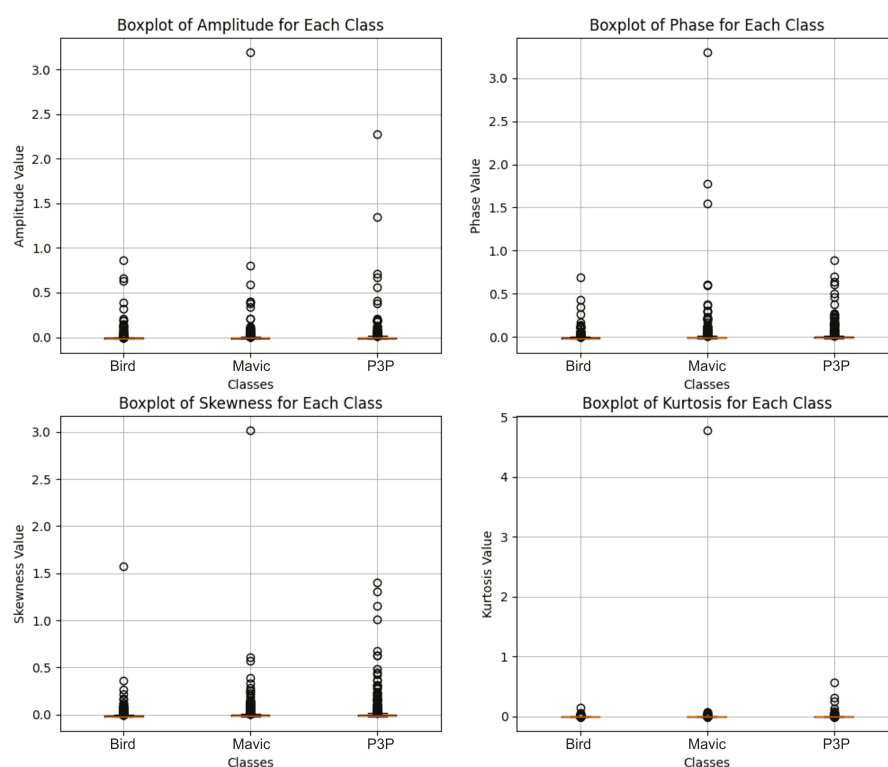


Figure 17. Boxplots showing the amplitude, phase, skewness, and kurtosis values for the Bird, Mavic, and P3P classes, as extracted from radar signals with added Pareto noise. The median values remained close to zero across the features, with fewer outliers and lower variability compared to white noise. The Bird class showed the most stable distribution, while the Mavic and P3P classes exhibited moderate spreads with fewer extreme values. This reduced variability under Pareto noise led to diminished class separability, resulting in lower classification performance compared to white noise.

Pareto noise had slightly positively skewed distribution characteristics when compared to impulsive noise, but such a difference did not tend to produce anything remarkably dissimilar in the overall similarity between classes, as shown in Figure 18.

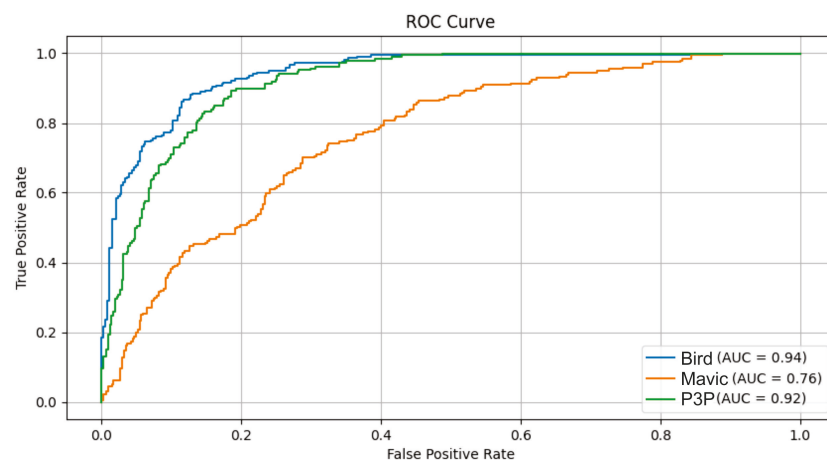


Figure 18. The ROC curves for the classification of Bird, Mavic drone, and P3P drone when considering Pareto noise with a Multimodal Transformer.

Based on the graphs analyzed in Figure 19, the boxplots for impulsive noise revealed similar distributions among the classes Bird, Mavic, and P3P in all the evaluated metrics (amplitude, phase, skewness, and kurtosis). There were no significant differences in the central values, dispersion, or presence of outliers indicating any substantial changes caused by impulsive noise.

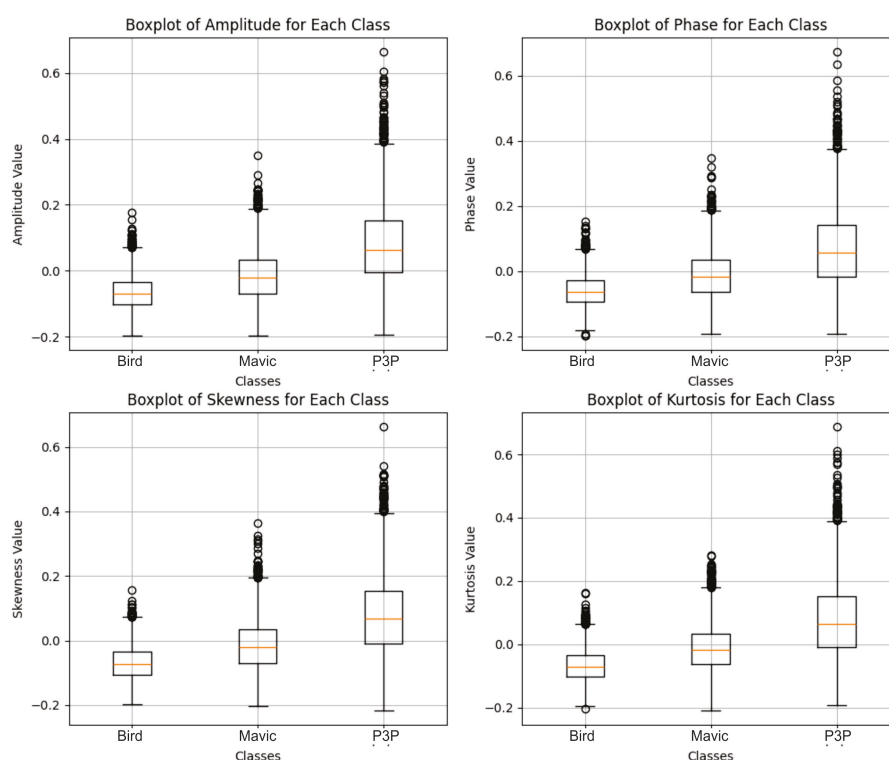


Figure 19. Boxplots showing the amplitude, phase, skewness, and kurtosis values for the Bird, Mavic, and P3P classes under impulsive noise. The Bird class exhibited the most stable and compact distribution across all features, while the P3P class showed the highest median values and widest variability, especially in amplitude and kurtosis. The Mavic class displayed intermediate behavior. Impulsive noise significantly increased outliers, particularly in the P3P class, indicating that larger or more complex targets produce more erratic radar reflections.

This patchiness is not strong enough with respect to the effect of impulsive noise on the data—at least in reference to those metrics—to distinguish patterns that overwhelmingly classify the data into groups (Figure 20).

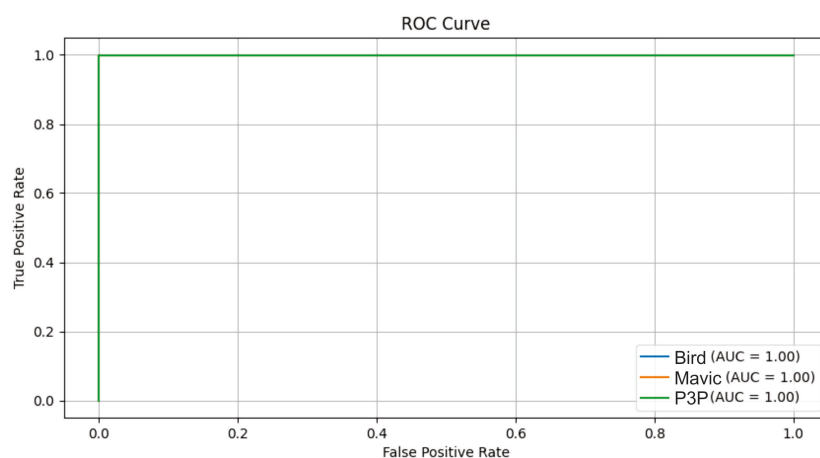


Figure 20. ROC curves for the classification of the Bird, Mavic drone, and P3P drone classes when considering impulsive noise with the Multimodal Transformer.

The result in Figure 21 indicates perfect separability despite the uniform signal distortion caused by multipath effects.

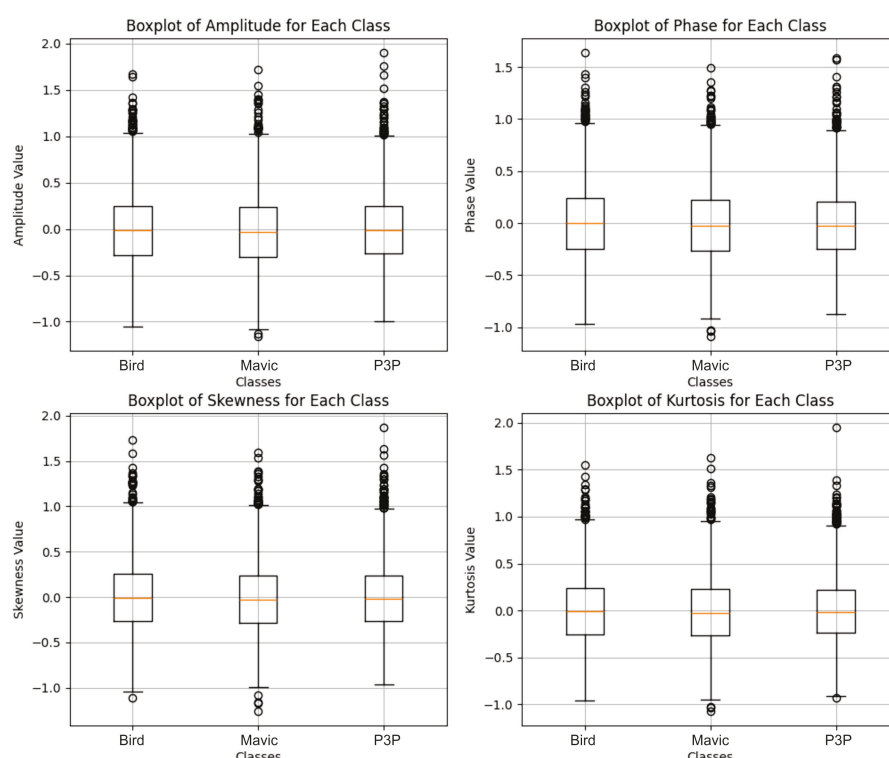


Figure 21. Boxplots showing the amplitude, phase, skewness, and kurtosis values for the Bird, Mavic, and P3P classes under multipath interference. Unlike impulsive noise, the distributions were more uniform across classes, with median values close to zero and consistent interquartile ranges. The P3P class showed a slightly wider spread in amplitude and skewness, suggesting higher susceptibility to multipath effects. Outliers were evenly distributed across classes, indicating random variations in signal reflections that reduce separability between classes.

The Multimodal Transformer model demonstrated excellent performance in distinguishing between the Bird, Mavic, and P3P classes under multipath interference (Figure 22),

achieving an AUC of 1 for all classes. A similarly high performance was observed under white noise and impulsive noise conditions, confirming the robustness of the model across different noise types. However, the classification accuracy was slightly lower under Pareto noise, highlighting that heavy-tailed noise distributions pose a greater challenge to the model. Overall, Transformer effectively handles various noise interferences, maintaining reliable classification even under complex signal distortions.

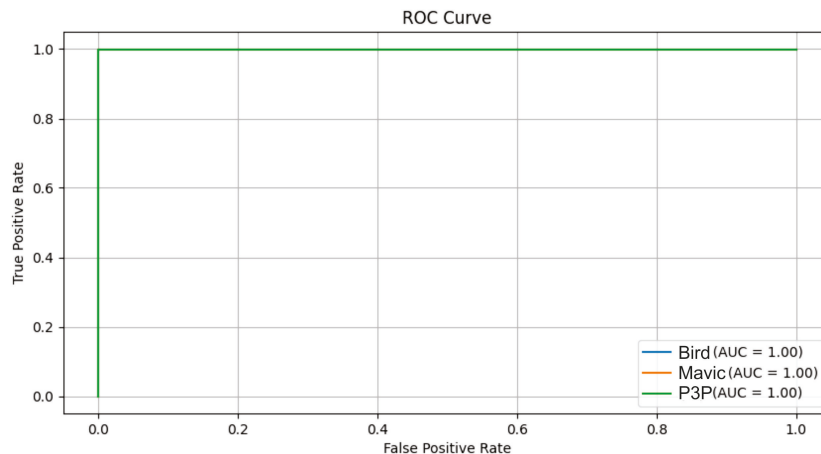


Figure 22. ROC curves for the classification of Bird, Mavic drone, and P3P drone classes when considering multipath interference with the Multimodal Transformer.

The Multimodal Transformer algorithm, indicated by the above complexity forms, followed the same time complexity logic as Transformer, with slight differences. The attention operation, at this stage, was essentially going to incur a time complexity or storage requirements that are quadratic with respect to the sequence length in the self-attention mechanism because it, in actuality, takes part in all dimensions. This implies that the time complexity is $O(n^2cd)$, n in this case, when it really refers to the length c of the sequence and d according to the dimension of the word representations extracted features. This means that the associated space complexity is also about $O(ncd)$ since the intermediate representations need to be eliminated as the model will also have a need for storing attention weights.

However, even if the normalization, pooling, and dense layers with regularization have no impact on the order of magnitude as far as complexity is concerned, they do mold the specific implementation behavior. As far as parallelization is concerned, the implementation would be considered highly parallelizable since the Multimodal Transformer processes nearly all incoming information input in parallel (due to the main usage of the MultiHeadAttention mechanism). Therefore, it can execute this on modern hardware like GPUs.

5. Discussion

The excellent performance exhibited by the Multimodal Transformer model in distinguishing drone and bird classes showed that there are several factors that are crucial in its robustness against noise interference. One important attribute is the use of attention mechanisms as these allow the model to only refer the input data to the relevant parts, thus reducing the effect of impact from noise. Features such as amplitude, phase, skewness of amplitude, as well as the kurtosis of amplitude, are critical in retaining significant signal characteristics. These features still remain informative despite the introduction of white noise, Pareto noise, impulsive noise, or multipath interference, and they enable Transformer to achieve near perfect classification.

The architecture of Transformer is structurally built to capture complex non-linear relationships between features that may not be captured by other models. Unlike traditional methods that may be considered biased toward simple or localized patterns, the mechanisms of Transformer tend to weigh the input features on the entire set. This flexibility allows the system to detect those subtle features embedded in the data, which is important for accurate classification, even under noisy conditions.

In addition, there is a self-embedded feature of the model of extraction and ranking of critical information from various input characteristics so that the irrelevant noise does not overpower the important parts of the signal. This then depicts the noise resilience of Transformer, showing its propensity for intricate feature extraction and analytical interpretation—an invaluable asset for classifying radar signals in challenging real-world environments.

The robustness demonstrated by Transformer also highlights its potential for practical applications in domains such as public safety and military defense. In public safety, the model could be employed to detect unauthorized drones in sensitive areas such as airports, government buildings, or public events. Similarly, in military scenarios, Transformer could be used to identify and classify drones in complex and hostile environments. However, deploying Transformer in real-world operations presents challenges, especially concerning computational requirements for real-time analysis and the need for continuous model updates to adapt to evolving threats and conditions.

6. Conclusions

This study demonstrates the Multimodal Transformer model as performing exceptionally in classifying drone and bird signals with different types of noise such as white noise, Pareto noise, impulsive noise, and multipath interference without any real degradation in the accuracy. This superior performance of the model is mainly because the attention mechanisms are effectively used to focus on the most relevant features in data that are significantly affected by noise. The features that were considered were amplitude, phase, skewness of amplitude, and kurtosis of amplitude, and they were used to thoroughly cover the characteristics of the signals that Transformer takes to class description.

Noise is, of course, a big problem, yet the very high AUCs of all the noise types demonstrate how robust the Multimodal Transformer model is in handling complex and noisy input data. However, the additional capability of the model to include both statistical ones, namely skewness and kurtosis, and physical ones, that is, amplitude and phase, is expected to contribute to considerably capturing subtle but meaningful patterns and improving generalization beyond immediate requirements. This trait becomes more important in real-world practical use cases where radar and sensor data are invariably noisy due to various sources, as seen in aerial object detection.

From these findings, it can be inferred that the architecture of Transformer can most likely fulfill the need for noisy radar signal classification. Further research could test more feature combinations in the model or the model could be optimized further to face other operational issues. The results laid a solid foundation for improving the reliability and efficiency of automatic drone detection systems in environments of different and highly noisy conditions.

Author Contributions: Conceptualization, M.L. and C.S.; methodology, M.L. and C.S.; software, M.L. and C.S.; validation, M.L. and C.S.; formal analysis, M.L. and C.S.; investigation, M.L. and C.S.; resources, M.L. and C.S.; data curation, M.L. and C.S.; writing—original draft preparation, M.L. and C.S.; writing—review and editing, M.L. and C.S.; visualization, M.L. and C.S.; supervision, M.L. and C.S.; project administration, M.L. and C.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data supporting the findings of this study are available upon reasonable request from the corresponding author. Due to the nature of the research, some data may be restricted for confidentiality or ethical reasons.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

DJI	Dà-Jiāng Innovations Science and Technology Co., Ltd.
Conv1D	One-dimensional Convolutional Neural Network
LSTM	Long Short-Term Memory
GRU	Gated Recurrent Unit
CNN	Convolutional Neural Network
UAV	Unmanned Aerial Vehicle
C-UAV	Counter-Unmanned Aerial Vehicle
mmWave	Millimeter Wave
CVMD	Complex-Valued Micro-Doppler
SVD	Singular Value Decomposition
ECSB	Echo Cancellation for Signal-Based Systems
FMCW	Frequency-Modulated Continuous Wave
RCS	Radar Cross Section

References

- Clark, B. United States. In *Above and Beyond: State of the Art of Uncrewed Combat Aerial Systems and Future Perspectives*; Calcagno, E., Marrone, A., Eds.; Istituto Affari Internazionali (IAI): Rome, Italy, 2023; pp. 31–37. Available online: <http://www.jstor.org/stable/resrep55314.6> (accessed on 1 October 2024).
- Pettyjohn, S. Types of drones. In *Evolution Not Revolution: Drone Warfare in Russia's 2022 Invasion of Ukraine*; Center for a New American Security: Washington, DC, USA, 2024; pp. 16–37. Available online: <http://www.jstor.org/stable/resrep57900.7> (accessed on 1 October 2024).
- Mohammed, A.B.; Fourati, L.C.; Fakhrudeen, A.M. Comprehensive systematic review of intelligent approaches in UAV-based intrusion detection, blockchain, and network security. *Comput. Netw.* **2023**, *239*, 110140. [CrossRef]
- Seidaliyeva, U.; Omarbekova, A.; Abdrayev, S.; Mukhamezhanov, Y.; Kassenov, K.; Oyewole, O.; Koziel, S. Advances and Challenges in Drone Detection and Classification Techniques: A state-of-the-art review. *Sensors* **2023**, *24*, 125. [CrossRef] [PubMed]
- Solaiman, S.; Alsuwat, E.; Alharthi, R. Simultaneous Tracking and Recognizing Drone Targets with Millimeter-Wave Radar and Convolutional Neural Network. *Appl. Syst. Innov.* **2023**, *6*, 68. [CrossRef]
- Dumitrescu, C.; Pleșu, G.; Băbuț, M.; Dobrinoiu, C.; Dospinescu, A.; Botezatu, N. Development of an Acoustic System for UAV Detection. *Sensors* **2020**, *20*, 4870. [CrossRef]
- Song, C.; Li, H. An Acoustic Array Sensor Signal Recognition Algorithm for Low-Altitude Targets Using Multiple Five-Element Acoustic Positioning Systems with VMD. *Appl. Sci.* **2024**, *14*, 1075. [CrossRef]
- Li, D.; Li, Z.; He, S.; Du, W.; Wei, Z. Multi-source threatening event recognition scheme targeting drone intrusion in the fiber optic DAS system. *IEEE Sens. J.* **2024**, *24*, 32185–32195. [CrossRef]
- Chen, T.; Yu, J.; Yang, Z. Research on a Sound Source Localization Method for UAV Detection Based on Improved Empirical Mode Decomposition. *Sensors* **2024**, *24*, 2701. [CrossRef] [PubMed]
- Tejera-Berengue, D.; Zhu-Zhou, F.; Utrilla-Manso, M.; Gil-Pita, R.; Rosa-Zurera, M. Acoustic-Based Detection of UAVs Using Machine Learning: Analysis of Distance and Environmental Effects. In *Proceedings of the 2023 IEEE Sensors Applications Symposium (SAS)*, Ottawa, ON, Canada, 18 July 2023; pp. 1–6. [CrossRef]
- Utebayeva, D.; Ilipbayeva, L.; Matson, E.T. Practical study of recurrent neural networks for efficient real-time drone sound detection: A review. *Drones* **2023**, *7*, 26. [CrossRef]

12. He, J.; Fakhreddine, A.; Alexandropoulos, G.C. RIS-Augmented Millimeter-Wave MIMO Systems for Passive Drone Detection. *arXiv* **2024**, arXiv:2402.07259.
13. Zhai, X.; Zhou, Y.; Zhang, L.; Wu, Q.; Gao, W. YOLO-Drone: An Optimized YOLOv8 Network for Tiny UAV Object Detection. *Electronics* **2023**, *12*, 3664. [CrossRef]
14. Sen, C.; Singh, P.; Gupta, K.; Jain, A.K.; Jain, A.; Jain, A. UAV Based YOLOV-8 optimization technique to detect the small size and high speed drone in different light conditions. In Proceedings of the 2024 2nd International Conference on Disruptive Technologies (ICDT), Greater Noida, India, 15–16 March 2024; pp. 1057–1061.
15. Cheng, Q.; Li, X.; Zhu, B.; Shi, Y.; Xie, B. Drone Detection Method Based on MobileViT and CA-PANet. *Electronics* **2023**, *12*, 223. [CrossRef]
16. Muhamad Zamri, F.N.; Gunawan, T.S.; Yusoff, S.H.; Alzahrani, A.A.; Bramantoro, A.; Kartiwi, M. Enhanced Small Drone Detection Using Optimized YOLOv8 With Attention Mechanisms. *IEEE Access* **2024**, *12*, 90629–90643. [CrossRef]
17. Zeng, Y.; Wang, G.; Hong, T.; Wu, H.; Yao, R.; Song, C.; Zou, X. UAVData: A Dataset for Unmanned Aerial Vehicle Detection. *Soft Comput.* **2021**, *25*, 5385–5393. [CrossRef]
18. El-Latif, E.I.A. Detection and Identification of Drones Using Long Short-Term Memory and Bayesian Optimization. *Multimed. Tools Appl.* **2024**, *83*, 4465–4517.
19. Xiao, J.; Pisutsin, P.; Tsao, C.W.; Feroskhan, M. Clustering-based Learning for UAV Tracking and Pose Estimation. *arXiv* **2024**, arXiv:2405.16867.
20. Johnstone, C.B.; Cook, C.R.; Aldisert, A.; Klaas, L.; Michienzi, C.; Rubinstein, G.; Sanders, G.; Szechenyi, N. Case study two: Electro-optical sensors. In *Building a Mutually Complementary Supply Chain Between Japan and the United States: Pathways to Deepening Japan-U.S. Defense Equipment and Technology Cooperation*; Center for Strategic and International Studies (CSIS): Washington, DC, USA, 2024; pp. 16–20. Available online: <http://www.jstor.org/stable/resrep62403.6> (accessed on 1 October 2024).
21. Yang, D.; Xie, X.; Zhang, Y.; Li, D.; Liu, S. A Multi-Rotor Drone Micro-Motion Parameter Estimation Method Based on CVMD and SVD. *Remote Sens.* **2022**, *14*, 3326. [CrossRef]
22. Hanif, A.; Ahmad, F.; Zhang, G.; Wang, X. Micro-Doppler Based Target Recognition With Radars: A review. *IEEE Sens. J.* **2022**, *22*, 2948–2961.
23. Soumya, A.; Mohan, C.K.; Cenkeramaddi, L.R. Recent advances in mm Wave-radar-based sensing, its applications, and machine learning techniques: A review. *Sensors* **2023**, *23*, 8901. [CrossRef] [PubMed]
24. Kumawat, H.C.; Chakraborty, M.; Raj, A.A.B. DIAT-RadSATNet—A novel lightweight DCNN architecture for micro-Doppler-based small unmanned aerial vehicle (SUAV) targets' detection and classification. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 1–11.
25. Kumawat, H.C.; Chakraborty, M.; Arockia Basil Raj, A.; Dhavale, S.V. DIAT-μSAT: Micro-Doppler Signature Dataset of Small Unmanned Aerial Vehicle (SUAV). *IEEE Dataport* **2022**, 1–5. [CrossRef]
26. Texas Instruments. The Fundamentals of Millimeter Wave Radar Sensors. Available online: <https://www.ti.com/lit/spyy005> (accessed on 10 September 2024).
27. Lu, G.; Bu, Y. Mini-UAV Movement Classification Based on Sparse Decomposition of Micro-Doppler Signature. *IEEE Geosci. Remote Sens. Lett.* **2024**, *21*, 3506405.
28. Kang, S.; Forsten, H.; Semkin, V.; Rangan, S. Millimeter Wave 60 GHz Radar Measurements: UAS and Birds. *IEEE Dataport*, March 2024. Available online: <https://ieee-dataport.org/open-access/millimeter-wave-60-ghz-radar-measurements-uas-and-birds> (accessed on 1 October 2024)
29. Narayanan, R.M.; Tsang, B.; Bharadwaj, R. Classification and Discrimination of Birds and Small Drones Using Radar Micro-Doppler Spectrogram Images. *Signals* **2023**, *4*, 337–358. [CrossRef]
30. Yan, J.; Hu, H.; Gong, J.; Kong, D.; Li, D. Exploring Radar Micro-Doppler Signatures for Recognition of Drone Types. *Drones* **2023**, *7*, 280. [CrossRef]
31. Fan, S.; Wu, Z.; Xu, W.; Zhu, J.; Tu, G. Micro-Doppler Signature Detection and Recognition of UAVs Based on OMP Algorithm. *Sensors* **2023**, *23*, 7922. [CrossRef] [PubMed]
32. Ye, J.; Gu, F.; Wu, H.; Han, X.; Yuan, X. A new Frequency Hopping Signal Detection of Civil UAV Based on Improved k-means Clustering Algorithm. *IEEE Access* **2021**, *9*, 53190–53204.
33. Xu, C.; Tang, S.; Zhou, H.; Zhao, J.; Wang, X. Adaptive RF fingerprint decomposition in micro UAV detection based on machine learning. In Proceedings of the ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, ON, Canada, 6–11 June 2021.
34. Inani, K.N.; Sangwan, K.S.; Dhiraj. Machine learning based framework for drone detection and identification using RF signals. In Proceedings of the 2023 4th International Conference on Innovative Trends in Information Technology (ICITIIT), Kottayam, India, 16–17 February 2023; pp. 1–8. [CrossRef]
35. Mandal, S.; Satija, U. Time-Frequency Multiscale Convolutional Neural Network for RF-Based Drone Detection and Identification. *IEEE Sens. Lett.* **2023**, *7*, 7003304. [CrossRef]

36. Garvanov, I.; Kanev, D.; Garvanova, M.; Ivanov, V. Drone Detection Approach Based on Radio Frequency Detector. In Proceedings of the 2023 International Conference Automatics and Informatics (ICAI), Varna, Bulgaria, 5–7 October 2023; pp. 230–234. [CrossRef]
37. Wang, J.; Gao, Y.; Li, D.; Yu, W. Multi-Domain Features Guided Supervised Contrastive Learning for Radar Target Detection. *arXiv* **2024**, arXiv:2412.12620.
38. Shi, Y.; Du, L.; Li, C.; Guo, Y.; Du, Y. Unsupervised Domain Adaptation for SAR Target Classification Based on Domain-and Class-Level Alignment: From Simulated to Real Data. *ISPRS J. Photogramm. Remote Sens.* **2024**, *207*, 1–13. Available online: <https://www.sciencedirect.com/science/article/abs/pii/S0924271623003155> (accessed on 1 October 2024).
39. Chen, Y.; Jia, Y. Target detection method based on unsupervised domain adaptation for through-the-wall radar imaging. In Proceedings of the 2023 6th International Conference on Electronics Technology (ICET), Chengdu, China, 12–15 May 2023; pp. 156–160. Available online: <https://ieeexplore.ieee.org/document/10137672> (accessed on 1 October 2024).
40. Guo, Y.; Du, L.; Lyu, G. SAR Target Detection Based on Domain Adaptive Faster R-CNN with Small Training Data Size. *Remote Sens.* **2021**, *13*, 4202. [CrossRef]
41. Ai, J.; Mao, Y.; Wang, Y.; Xing, M. SAR Target Classification Using the Multikernel-Size Feature Fusion-Based Convolutional Neural Network. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–15. Available online: <https://ieeexplore.ieee.org/document/9528903> (accessed on 1 October 2024).
42. Liu, Y.; Ding, Z.; Cao, Y.; Chang, M. Multi-scale Feature Fusion UAV Image Object Detection Method Based on Dilated Convolution and Attention Mechanism. In Proceedings of the 2020 8th International Conference on Information Technology: IoT and Smart City (ICIT 2020), Xi'an, China, 25–27 December 2020; pp. 1–12. [CrossRef]
43. Tan, S.; Duan, Z.; Pu, L. Multi-scale object detection in UAV images based on adaptive feature fusion. *PLoS ONE* **2024**, *19*, e0300120. [CrossRef]
44. Othman, N.A.; Aydin, I. Development of a Novel Lightweight CNN Model for Classification of Human Actions in UAV-Captured Videos. *Drones* **2023**, *7*, 148. [CrossRef]
45. Dewangan, O.; Vij, P. CNN-LSTM framework to automatically detect anomalies in farmland using aerial images from UAVs. *BIO Web Conf.* **2024**, *82*, 5015. [CrossRef]
46. Utebayeva, D.; Almagambetov, A.; Alduraibi, M.; Temirgaliyev, Y.; Ilipbayeva, L.; Marxuly, S. Multi-label UAV sound classification using Stacked Bidirectional LSTM. In Proceedings of the 2020 International Conference on Robotics and Control (IRC), Taichung, Taiwan, 9–11 November 2020; pp. 453–458. [CrossRef]
47. Rodriguez, Y. Drone Preto e Branco Voando Sob o céu Azul Durante o dia, Unsplash, 28 January 2021. Available online: <https://unsplash.com/pt-br/fotografias/drone-preto-e-branco-voando-sob-o-ceu-azul-durante-o-dia-mVI7sD0nTIA> (accessed on 1 October 2024). [CrossRef]
48. Spratt, A. Quadricóptero Branco Voando Durante o Dia, Unsplash, 2 March 2017. Available online: https://unsplash.com/pt-br/fotografias/quadricoptero-branco-voando-durante-o-dia-I9TaTsU_VnE (accessed on 1 October 2024).
49. Amazon. Bionic Bird—Smartphone Controlled and Remotely Operated. Amazon. Available online: <https://www.amazon.in/Bionic-Bird-Smartphones-Remotely-Controlled/dp/B00TFCS4KC> (accessed on 1 October 2024).
50. Xu, P.; Zhu, X.; Clifton, D. A. Multimodal learning with transformers: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *45*, 12113–12132.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



Article

Two-Dimensional Scattering Center Estimation for Radar Target Recognition Based on Multiple High-Resolution Range Profiles

Kang-In Lee ¹, Jin-Hyeok Kim ² and Young-Seek Chung ^{3,*}

¹ Defense Rapid Acquisition Technology Research Institute, Agency for Defense Development, Seoul 07062, Republic of Korea; leekangin@add.re.kr

² Group for Quantum Electricity & Magnetism Metrology, Korea Research Institute of Standards and Science, Daejeon 34113, Republic of Korea; jinhykim@kriss.re.kr

³ Department of Electronic Convergence Engineering, Kwangju University, Seoul 01897, Republic of Korea

* Correspondence: yschung@kw.ac.kr

Abstract: A new estimation strategy on locations of two-dimensional target scattering centers for radar target recognition is developed by using multiple high-resolution range profiles (HRRPs). Based on the range information contained in multiple HRRPs obtained from various observation angles, the estimated target scattering centers can be successfully located at the intersection points of the lines passing through the multiple HRRP points. This geometry-based algorithm can significantly reduce the computational complexity while ensuring the ability to estimate the two-dimensional target scattering centers. The computational complexity is formulated and compared to that of the conventional methods based on the synthetic aperture radar (SAR) images and HRRP sequences. In order to verify the performance of the proposed algorithm, the numerical and experimental results for three different types of aircraft were compared to those from SAR images. At the end of this article, the estimated radar scattering centers are used as the target features for the conventional classifier machine to confirm its target classification performance.

Keywords: high resolution range profile; microwave propagation; radar target recognition; scattering center estimation; target signature

1. Introduction

Target recognition technology has been used in the military field to improve the early assessment of potential threats, such as missiles and hostile aircraft, and to respond with appropriate means. Reliable target recognition capabilities in coastal surveillance and air traffic control systems can also monitor and confront intrusions of illegal immigrants, smugglers, and terrorists by land or sea [1]. In addition to the military field, target recognition algorithms that are implemented with radar systems to detect pedestrians or unexpected oncoming threats have become a breakthrough technology in the autonomous vehicle industry.

Target signatures commonly used for target recognition include wideband radar cross-section (RCS) frequency response, high-resolution range profile (HRRP), and synthetic aperture radar (SAR) images. Multiple wideband RCS frequency responses for different angles are generally used to recognize small targets compared to the system range resolution. Huang and Lee [2] proposed a target identification method using the RCS data for three different targets. In this method, the similarity between the unknown and known targets was defined to perform a statistical classification, where the dataset was based on the RCS data that are observed in directions of different elevation and azimuth angles.

If a range profile provides sufficient information to identify a target, either the HRRP or the SAR image can be used in the target recognition process. The HRRP is a one-dimensional representation of the target characteristics obtained by expressing the distribution of target scattering points projected onto the observation axis. Despite the low

computational complexity, accurate target discrimination cannot be achieved due to the possibility that similar HRRPs can be obtained from different targets [3]. On the contrary, in the SAR images, which are expressed in a two-dimensional target scattering point distribution, the probability of having similar characteristics from different targets is relatively low, but higher computational complexity is required to obtain high-dimensional target characteristics [4].

In terms of using the HRRPs and SAR techniques for real-time recognition, the former, which can be obtained by performing the one-dimensional inverse Fourier transform of the received signals, has an advantage over the latter, which requires complex signal processing, such as two-dimensional Fourier transforms and back projection of the HRRPs obtained from multiple angles [5,6]. However, as Novak evaluated in [7], the classification results from two-dimensional algorithms using SAR images show better performance than those from one-dimensional HRRPs.

Over the past few decades, numerous studies have been conducted to secure both performance and computational efficiency, which can be divided into two main categories according to their approaches. Studies in the first category are focused on developing effective feature extraction methods as preprocessing techniques prior to classification. Numerous feature extraction methods have been introduced since the early 2000s, including principal component analysis (PCA) [8–12] and linear discriminant analysis (LDA) [13–16], and their common purpose was mainly to reduce redundancy and dimensionality in extracting a feature vector for a given target. Studies in the second category are focused on designing robust classifiers, such as the Bayesian decision-based classifier, pattern recognition-based classifier and neural network-based classifier [17,18]. With advances in computing and information technologies, object classification capabilities of machine learning (ML) methods, such as the support vector machine (SVM) [19,20] and the extreme learning machine (ELM) [21], have been widely utilized in the field of target recognition.

However, since both approaches are biased towards the processing of the acquired data, they cannot be freed from the inherent limitations of the target signature itself, which are mentioned in the fourth paragraph of this article. To mediate the trade-off between classification accuracy and computational load in target recognition, some studies on developing the estimation method of target scattering centers have been conducted since the early 2000s.

Kyung-Tae Kim and Hyo-Tae Kim [22] proposed a new method that estimates two-dimensional locations of target scattering centers using a set of one-dimensional down-range locations and associated amplitudes over several aspect angles. In this method, the one-dimensional total least squares estimation of signal parameters was applied via rotational invariance techniques, and an intersection point equation based on geometric techniques, such as projection and triangulation, was derived to provide input training data for multiple elastic modules (MEMs). As a result, the two-dimensional locations of target scattering centers were estimated using an iterative scheme for MEM network optimization. The authors showed better estimation accuracy and computational complexity compared to the well-known two-dimensional matrix enhancement and matrix pencil technique. However, it is also stated that the algorithm tends to rely on the initial guess of module center locations due to its iterative nature, and the computational complexity can be significantly increased when the number of aspect angles increases.

Jianxiong et al. [23] proposed a reconstruction method, which estimates three-dimensional scattering centers from multiple HRRPs without knowledge of aspect angles or relative radar-target movement based on the signal space estimation method. In this method, the singular value decomposition technique was used to estimate signal subspace, and the performance was evaluated by Cramer–Rao Lower Bounds of the reconstruction problem. However, the uniqueness of the solution in the algorithm could not be ensured due to an arbitrary rotation matrix between the reconstructed and real scattering center coordinates. To overcome the ambiguity in [23], Jun et al. [24] developed a parameter matrix based on the geometric projection of the scattering center on the line of radar sight in multi-

aspect, which is unique for any rotational translation. The effectiveness of the method was investigated for a flat-bottomed cone target with five scattering points, and the authors emphasized the advantages in resolution, stability, data requirement, and computational complexity. Recently, Su et al. [25] pointed out the ambiguity in [23] and the degradation of performance with increasing numbers of scattering centers and pulses. In this article, the authors divided the spatial target scattering centers with micro-motion properties into two types: stable and sliding scattering centers. Based on the factorization method used [24], the coordinates of stable and sliding scattering centers were reconstructed in three-dimensional and two-dimensional space, respectively. However, this method still cannot ensure the uniqueness of the solution, and at least six aspect angles are required to perform the scattering center estimation.

Ji-Hoon Bae and Kyung-Tae Kim [26] proposed a compressive sensing (CS)-based two-dimensional scattering center extraction (SCE) technique to extract scattering centers from incomplete radar cross-section (RCS) data. Unlike traditional methods, this approach is designed to accurately extract scattering centers when some RCS data are missing. However, due to the nature of the compressive sensing technique used in the proposed method for estimating scattering centers, the method inherently has high computational complexity.

Jianxiong et al. [27] proposed a method for estimating a global 3D scattering center model using multiple HRRPs. In this method, the HRRPs obtained from various angles are visualized on an OTSM (One-Two/Three-Dimensional Scatter Map), and a Hough Transform is then applied to automatically convert the 1D projection positions into 3D scattering centers. In [28], the 3D scattering centers of the target obtained using the method from [27] were pre-acquired and then used to build a database for SAR image-based automatic target recognition.

In summary, the methods described in [27,28] involve observing the target model of interest from numerous aspect angles and investing significant computational resources to estimate the global 3D scattering center model before new observations and identification. This model is then used to build a database necessary for target recognition using SAR images.

However, the method proposed in this paper can acquire target signatures with less computational complexity than the SAR images in environments where moving targets are being tracked, serving as an alternative to the SAR images. Therefore, it differs in purpose from the methods in [27,28]. Additionally, due to the computational complexity of the dataset required in [27] and the Hough Transform, the proposed method has significantly lower computational requirements compared to [27].

There is a certain trade-off between the estimation accuracy and computational load, both proportional to the number of dimensions to be processed. Therefore, we propose a method for constructing target signatures that provides dimensional qualities similar to those of a SAR technique with greatly reduced computational load. In the proposed method, the geometric relationship between multiple HRRPs and the location of scattering centers was used to define an error function based on the least-squares method (LSM), and two-dimensional locations of target scattering centers were estimated to represent the points, such as edges, corners, and planes, where the RCSs were expected to be large. Compared to the previous methods in [22–25], the estimation performance of the algorithm is fully proportional to the quality of HRRPs obtained by the radar system, and only three aspect angles of HRRP are required to ensure the uniqueness of the result. In addition, the estimation process is applicable even in an environment in which the number of visible scattering centers varies depending on the observation angle. Due to the non-iterative nature based on the geometric approach, the computational complexity of the proposed algorithm can be significantly reduced.

The remainder of this article is organized as follows. In Section 2, the scattering center estimation method for a point target model is developed using multiple HRRPs and expanded to be applicable for an actual target model. The computational complexity of the method is compared to that of the conventional techniques. In Section 3, the numerical and

experimental results of scattering center estimation are provided using measured HRRPs from three different types of aircraft and applied to the conventional classifier to confirm its target classification performance. Finally, conclusions are given in Section 4.

2. Scattering Center Estimation

In this section, a strategy for the two-dimensional estimation of radar target scattering centers is developed based on multiple HRRPs and their geometric relations. The multiple HRRPs are assumed to be obtained in the same manner as a spotlight-mode-SAR observation model that collects radar signals for various aspect angles by continuously steering the radar beam at a target on a rotating platform [29].

2.1. Point Target Model

The HRRP contains the range information of target scattering centers projected onto the observation axis. A basic geometric concept of the HRRP projection is shown in Figure 1. Since (u_m^k, v_m^k) is the point at which the scattering center (x_m, y_m) is projected onto the observation axis, the range profile can be expressed as

$$h_m^k = x_m \cos(\alpha_k) + y_m \sin(\alpha_k) \quad (1)$$

where h_m^k is the profile which contains range information of the m -th scattering center with the k -th observation angle α_k . The range profile in (1) can be obtained from the peak location in the HRRP, which is the inverse discrete Fourier transform of the received signal in the frequency domain. Therefore, the resulting magnitude of h_m^k in (1) is equal to the distance of a point (u_m^k, v_m^k) from the origin.

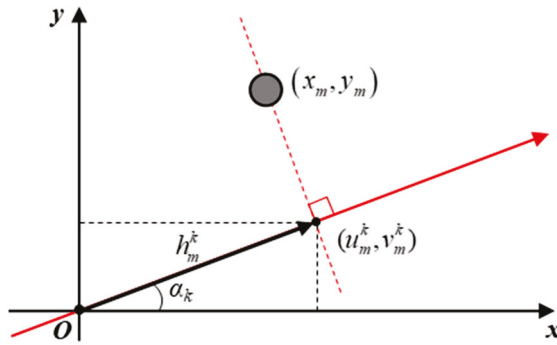


Figure 1. Concept of HRRP projection onto the observation axis.

Using the obtained range profile with known value of α_k , the projected point (u_m^k, v_m^k) can be defined as

$$(u_m^k, v_m^k) = (h_m^k \cos(\alpha_k), h_m^k \sin(\alpha_k)) \quad (2)$$

The imaginary line passing through (2) and perpendicular to the observation axis can be formulated as

$$y = -\frac{\cos(\alpha_k)}{\sin(\alpha_k)}(x - u_m^k) + v_m^k \quad (3)$$

Substituting (2) in (3), another expression for the range profile h_m^k can be obtained by

$$h_m^k = x \cos(\alpha_k) + y \sin(\alpha_k) \quad (4)$$

By the definition of (4), the same range profile can be obtained from any point on the imaginary line passing through the projected point and perpendicular to the observation axis. Clearly, the real scattering center (x_m, y_m) is located on the line defined by (4), which yields (1), but it is impossible to inversely specify the exact location from h_m^k alone. Therefore, in the case shown in Figure 1, more range profiles are needed to decide the exact coordinate of the scattering center. With the help of several range profiles obtained from

different observation angles, a set of simultaneous equations from (4) can be derived, which can be finally solved for the unknown coordinates of scattering centers.

Once the multiple HRRPs are obtained for M scattering centers with K observation angles, the total dataset can be expressed in matrix form as

$$\mathbf{h} = \mathbf{x}\mathbf{P} \quad (5)$$

where

$$\mathbf{h} = \begin{bmatrix} h_1^1 & h_1^2 & \cdots & h_1^K \\ h_2^1 & h_2^2 & \cdots & h_2^K \\ \vdots & \vdots & \ddots & \vdots \\ h_M^1 & h_M^2 & \cdots & h_M^K \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_M & y_M \end{bmatrix}, \mathbf{P} = \begin{bmatrix} \cos(\alpha_1) & \cos(\alpha_2) & \cdots & \cos(\alpha_K) \\ \sin(\alpha_1) & \sin(\alpha_2) & \cdots & \sin(\alpha_K) \end{bmatrix}.$$

The rows and columns of $\mathbf{h}$ represent the obtained HRRPs for different angles and scattering centers, respectively. The $\mathbf{x}$ matrix contains two-dimensional location of scattering centers, and $\mathbf{P}$ is the projection matrix, whose column vectors are $(\cos(\alpha_k), \sin(\alpha_k))^T$.

Considering K observations, the simultaneous equations for the m -th scattering center can be expressed in the matrix form as

$$\mathbf{h}_m = \mathbf{x}_m \mathbf{P} \quad (6)$$

where $\mathbf{h}_m$ is the 1-by- K size row vector whose elements are $[h_m^1, h_m^2, \dots, h_m^K]$ and $\mathbf{x}_m$ is the row vector that contains the location of scattering center (x_m, y_m) .

Based on the LSM, the location of m -th scattering center can be estimated by matrix inversion of (6) as follows:

$$\mathbf{x}_m = \mathbf{h}_m \mathbf{P}^\dagger \quad (7)$$

where $(\cdot)^\dagger$ is the pseudo-inverse operation. As stated above, the two-dimensional location of a single scattering center is able to be specified by finding the solution of simultaneous equations in (7), which can be described as the intersection point of the vertical lines from two different HRRPs. However, there are two issues to be simultaneously considered under the general circumstances: authenticity of the solution and the matching problem.

As shown in Figure 2, the targets are observed in the form of peaks on the HRRP axes, but they appear in different order depending on the observation angle, so the target scattering center index m for each peak is actually unknown. If the components of matrix $\mathbf{h}$ in (5) are 'mismatched' to the order of actual targets, the unique solution cannot be obtained.

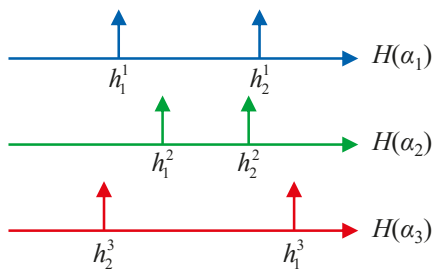


Figure 2. Target information shown as peaks on HRRP axes for three observation angles.

Figure 3 shows an example of two scattering centers observed from three different angles. The dotted lines represent the vertical lines passing through the projected scattering centers on each HRRPs, derived as in (4). In Figure 3, two observation axes of H_1 and H_2 generate four intersections of (h_1^1, h_1^2) , (h_1^1, h_2^2) , (h_2^1, h_1^2) , and (h_2^1, h_2^2) . Of these four points, only two intersections, created by (h_1^1, h_1^2) and (h_2^1, h_2^2) , are the authentic scattering centers, while the other two are not. Therefore, to sort out the real solution, one more range profile H_3 is needed. As shown in Figure 3, the authentic scattering centers are located only when

the three vertical lines intersect, which means the unique solution of (7) can be obtained when $K = 3$.

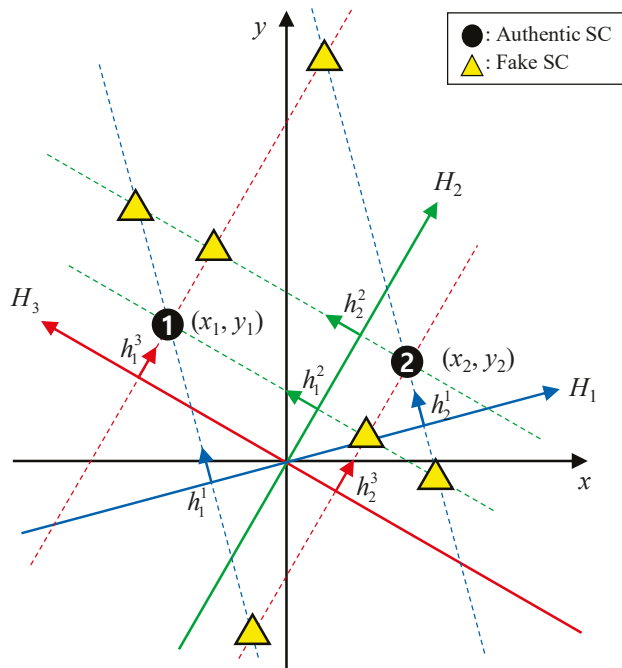


Figure 3. HRRPs for three observation angles when two scattering centers exist.

However, even with the three range profiles, the three vertical lines may not intersect, as the components of $\mathbf{h}$ may still be mismatched, resulting in the obtaining of an incorrect solution of (7). In this case, the LSM places the solution point at the center of gravity of the triangle formed by the three lines. Therefore, to overcome the matching problem, an authentication algorithm is proposed based on the error defined by the geometric property of all possible combinations of HRRP peaks.

The number of combinations can be calculated as

$$Q = \prod_{k=1}^K q_k \quad (8)$$

where q_k is the number of peaks on k -th observation axis. In most cases, q_k is equal to the number of scattering centers M , but it may be less depending on the location of the scattering centers and observation angle.

To sort out the matched combinations, each of them is used as a HRRP vector in (7) to estimate the scattering center $\mathbf{x}'$. If the incorrect combination was used, the unique solution cannot be obtained, resulting in $\mathbf{x}'$ being located at the center of gravity of the triangle formed by the mismatched combination of HRRP vertical lines. Therefore, an error function can be defined based on the geometric relationship between each combination of HRRP peaks and the HRRP recalculated from the point estimated by the corresponding combination. The error function is defined as

$$\varepsilon_i = \|\mathbf{h}'_i - \mathbf{x}'_i \mathbf{P}\|_2, \quad i = 1, 2, \dots, Q \quad (9)$$

where $\mathbf{x}'_i$ denotes the estimated location of the scattering center from $\mathbf{h}_i$, the i -th combination of HRRP peaks, and $\|\cdot\|_2$ denotes the L_2 norm.

Theoretically, the number of zeros out of the Q errors obtained by (9) should be equal to the number of authentic scattering centers. However, in practice, even the authentic scattering centers cannot have exactly zero errors due to the limitations of the radar measurement system. Therefore, we should introduce a threshold γ to the errors to determine the authenticity. The threshold γ must be less than or equal to $\Delta R/2$ to ensure

that the target exists within the same range-bin as the measured HRRP, where ΔR is the range resolution of the measurement system. The flowchart of the proposed algorithm is shown in Figure 4. The overall procedure can be summarized as follows:

Step 1: Obtain HRRPs from K observations

Step 2: Make Q combinations of HRRPs by considering the number of peaks on the k -th observation axis using Equation (8)

Step 3: Estimate the scattering centers using Equation (7) for each of the Q combinations of HRRPs. Note that the scattering centers estimated in this step include fake SCs.

Step 4: Calculate the HRRPs for all estimated scattering centers using the K observations and Equation (6).

Step 5: Calculate the error between the combinations of HRRPs obtained in Step 2 and the HRRPs estimated through Step 3 and Step 4 using Equation (9).

Step 6: If the calculated error in Step 5 is smaller than the threshold, the scattering centers are classified as authentic.

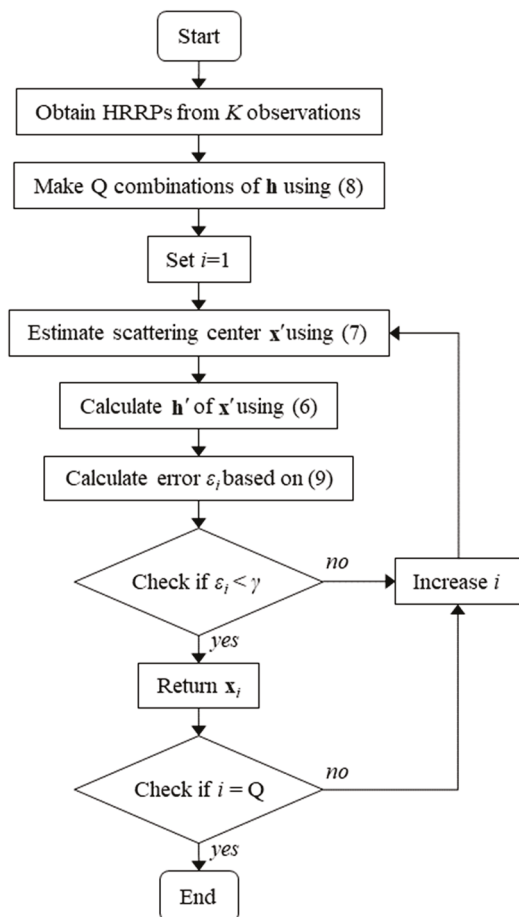


Figure 4. Flowchart of the proposed algorithm for two-dimensional radar target scattering center estimation.

2.2. Actual Target Model

The point target model assumes that all scattering centers can be observed by the radar. That is, each HRRP contains the range information for all scattering centers. However, in the actual target model, there are cases where a specific region is not visible depending on the observation angle. Fortunately, multiple HRRPs obtained from adjacent angles have the information of the scattering centers within the same area, so the proposed algorithm can be applied to estimate the entire scattering centers on the target.

Figure 5 shows the concept of choosing K neighboring HRRPs out of the entire HRRP dataset of size N to estimate all scattering centers around the target. In this

figure, $\mathbf{h}(n)$ is the n -th observed HRRP, δ is the interval of each HRRP selection, and $[\mathbf{h}(n), \mathbf{h}(n + \delta), \dots, \mathbf{h}(n + (K - 1)\delta)]$ are the K HRRPs used for the estimation procedure. Therefore, the proposed algorithm derived for the point target model can be applied to each of the $N - \delta(K - 1)$ neighboring sets. The resulting scattering points estimated through the above process are used for further processing to find the dominant scattering centers located on an object's surface that are likely to be observed from multiple viewpoints, such as edges, corners, and planes.

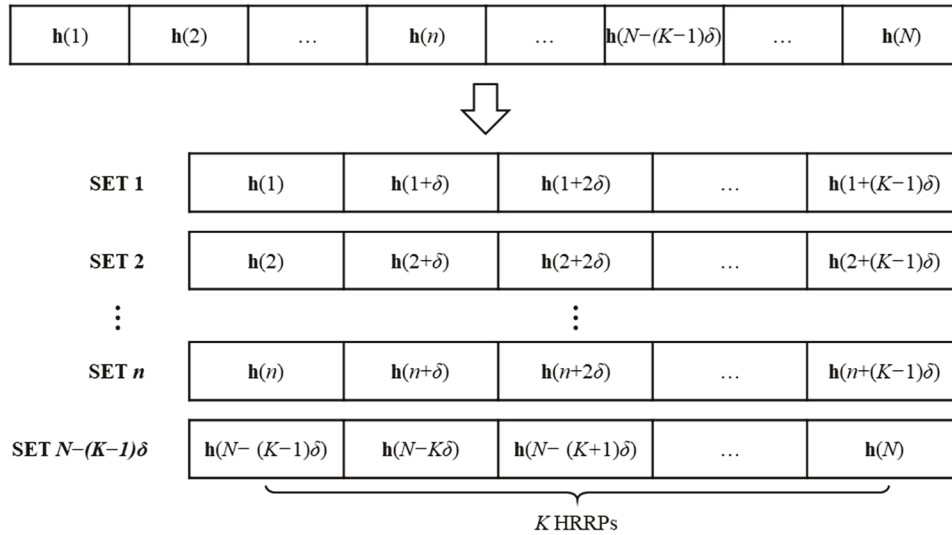


Figure 5. Generation of $N - (K - 1)\delta$ sub-datasets by choosing K neighboring HRRPs out of the entire HRRP dataset of size N .

Based on the property of the dominant scattering centers, we introduce an exposure factor that defines the frequency at which the scattering centers estimated using each of the neighboring HRRPs are observed in the entire dataset. The HRRPs for all observation angles of the i -th scattering center can be calculated as follows:

$$h'_i(n) = \mathbf{x}'_i \mathbf{p}(n), \quad n = 1, 2, \dots, N \quad (10)$$

where $\mathbf{x}'_i \equiv [x_i, y_i]$ and $\mathbf{p}(n) \equiv [\cos(\alpha_n), \sin(\alpha_n)]^T$.

The error between the estimated HRRP $h'_i(n)$ and the measured HRRP can be expressed as follows:

$$\beta_i(n) = \min(|\mathbf{h}(n) - h'_i(n)|) \quad (11)$$

where $\min(\cdot)$ stands for the minimum value of the vector elements. To verify the existence of the scattering center at a specific observation angle, the error obtained in (11) is applied to

$$D_i(n) = \begin{cases} 1, & \beta_i(n) < \gamma \\ 0, & \beta_i(n) > \gamma \end{cases} \quad (12)$$

where γ is the threshold used for the point target model.

Using (12), the exposure factor of the i -th scattering center ρ_i is given by

$$\rho_i = \frac{1}{N} \sum_{n=1}^N D_i(n) \quad (13)$$

The locations of the dominant scattering centers of the actual target model can be finally estimated by comparing (13) to the exposure threshold γ_ρ .

The overall performance of the proposed algorithm depends on the values of K and δ . In the theoretical environment, $K = 3$ is a sufficient condition for two-dimensional

estimation, but if the range resolution of the radar system is considered, the performance can be clearly improved in proportion to K . However, as the value of K increases, the computation complexity also increases as well, which is discussed in the following section. Smaller values of δ allow more scatter centers to be estimated, but the pseudo-inverse matrix inversion of $\mathbf{P}$ in (2) becomes more singular because the column vectors of the matrix $\mathbf{P}^T \mathbf{P}$ are not orthogonal. Therefore, the parameter δ should be determined considering the condition number of $\mathbf{P}$ [30].

The overall estimation process for the actual target model can be summarized as follows:

Step 1: Estimate the location of the scattering centers using (7) and (9) from observed HRRP data.

Step 2: Estimate HRRP $\hat{h}_i(n)$ using (10).

Step 3: Calculate the error between the estimated HRRP and measured HRRP $\beta_i(n)$ using (11).

Step 4: Compare (11) to threshold γ to determine whether the i -th scattering center is exposed from the observation angles.

Step 5: Process all estimated scattering points through Step 2 to Step 4.

Step 6: Compare (13) to the exposure threshold to decide whether each estimated scattering center becomes the main scattering center.

2.3. Computational Complexity

In this section, the computational complexity of the proposed algorithm is calculated and compared to that of the conventional methods. For comparison under the same conditions, only the process after obtaining HRRP through S frequency samples received through N observations was considered.

The proposed algorithm selects K out of N HRRPs to construct $\mathbf{h}$ in (7). Since the pseudo-inverse operation $\mathbf{P}^\dagger$ can be pre-calculated with the knowledge of observation angles, it is excluded from the complexity derivation. Considering the matrix multiplication for Q combinations of HRRPs, the complexity can be express as $O(2KQ)$ using the Big-O notation. The error function in (9) uses Q scattering centers to estimate their authenticities; it also has the complexity of $O(2KQ)$. For the actual target model, this process repeats approximately N times for estimating total scattering centers, which is assumed to be $M_{avg} \times N$. Finally, the estimated total scattering centers are used in (10) to calculate the HRRPs for N observation angles in order to determine their dominance.

Therefore, the complexity becomes

$$C_{prop} \approx O(4KQN + 2M_{avg}N^2), \quad (14)$$

And with approximated Q as M_{avg}^K ,

$$C_{prop} \approx O(4KM_{avg}^K N + 2M_{avg}N^2) \quad (15)$$

The computational complexity of the conventional SAR image algorithm can be expressed as the following equation [31,32]:

$$C_{SAR} \approx O(S^2N) \quad (16)$$

For example, in the case of $K = 4$, $M_{avg} = 4$, $N = 121$, and $S = 1024$, the complexities of the proposed method and SAR algorithm are $O(6.13 \times 10^5)$ and $O(1.27 \times 10^8)$, respectively. It can be seen that the proposed method has a complexity of about 0.5% of that of the SAR algorithm.

In [22], MEM network optimization was used in the scattering center estimation, and the associated computational complexity can be expressed as

$$C_{MEM} \approx O(z(J + L \log L)) \quad (17)$$

where z is the number of iterations of the MEM operation and J is the number of neurons over the entire network, which is defined as $P \times_N C_2$. The number of modules P should be larger than the number of scattering points M , and N is the number of observations. Finally, L is the number of neurons in one module defined as $_N C_2$.

In the case of $N = 10$, $M = 4$, $P = 8$, and $z = 300$, the complexity of the MEM-based method is $O(1.59 \times 10^5)$, while the proposed method has the complexity of $O(4.18 \times 10^4)$ with $K = 4$. Moreover, under the same assumption as in [22], the method for the point target model can be used for the comparison, which means the complexity of $O(4.10 \times 10^3)$ can be reached for the given environment. It can be seen that the proposed method has a complexity of 26.3% and 2.6% compared to the algorithm proposed in [22].

3. Results and Discussion

3.1. Simulation on Numerical Model

The proposed algorithm was applied to the reduced scale numerical models for three different types of aircraft: the airbus A380, the Eurofighter, and the F-15. The HRRP datasets were obtained from the CAD models at elevation angle $\theta = 0^\circ$ and azimuth angle ϕ in a range of -90° to 90° with an interval of 1° and observation angles N of 181 using the physical optics technique at the X-band, 8 GHz to 12 GHz. Furthermore, the material of the CAD model was assumed to be a PEC (Perfect Electric Conductor). The number of frequency steps, denoted as frequency samples S , was 1024, and the frequency step size was 3.9 MHz, with a range resolution ΔR of 3.75 cm. Moreover, a noise-free condition was assumed for effective comparison with the SAR technique. The specifications of the airborne targets and a list of the parameters used in the proposed algorithm are listed in Tables 1 and 2, respectively. The proposed algorithm was applied to the HRRPs obtained for each numerical model, and the peak detection results are shown in Figure 6. The red dots in Figure 6a,c,e represent the detected peaks within the measured HRRPs for each observation angle. In addition, the number of peaks obtained for each observation angle is shown in Figure 6b,d,f, where the A380, Eurofighter, and F-15 had an average of 2.97, 2.38, and 4.1 HRRP peaks, respectively, regarding all observation angles. The different number of peaks confirms that the actual target model has different visible scattering points depending on the observation angle. The red dots in Figure 7a,c,e represent the estimated scattering centers in two-dimensional space using the angular HRRP peaks detected in Figure 6. The SAR images for corresponding models are shown in Figure 7b,d,f as well. From the results of SAR imaging shown in Figure 7, we can see that the scattering centers of the A380 arise at the nose, the jet engines, and the tail wing. The Eurofighter model has the scattering centers at its canopy, intake, and tail wing, while the F-15 model has the scattering centers at the intake, missiles, wing edges, and tail wing. We can see that the scattering centers estimated by the proposed algorithm are distributed similarly to those of the SAR images.

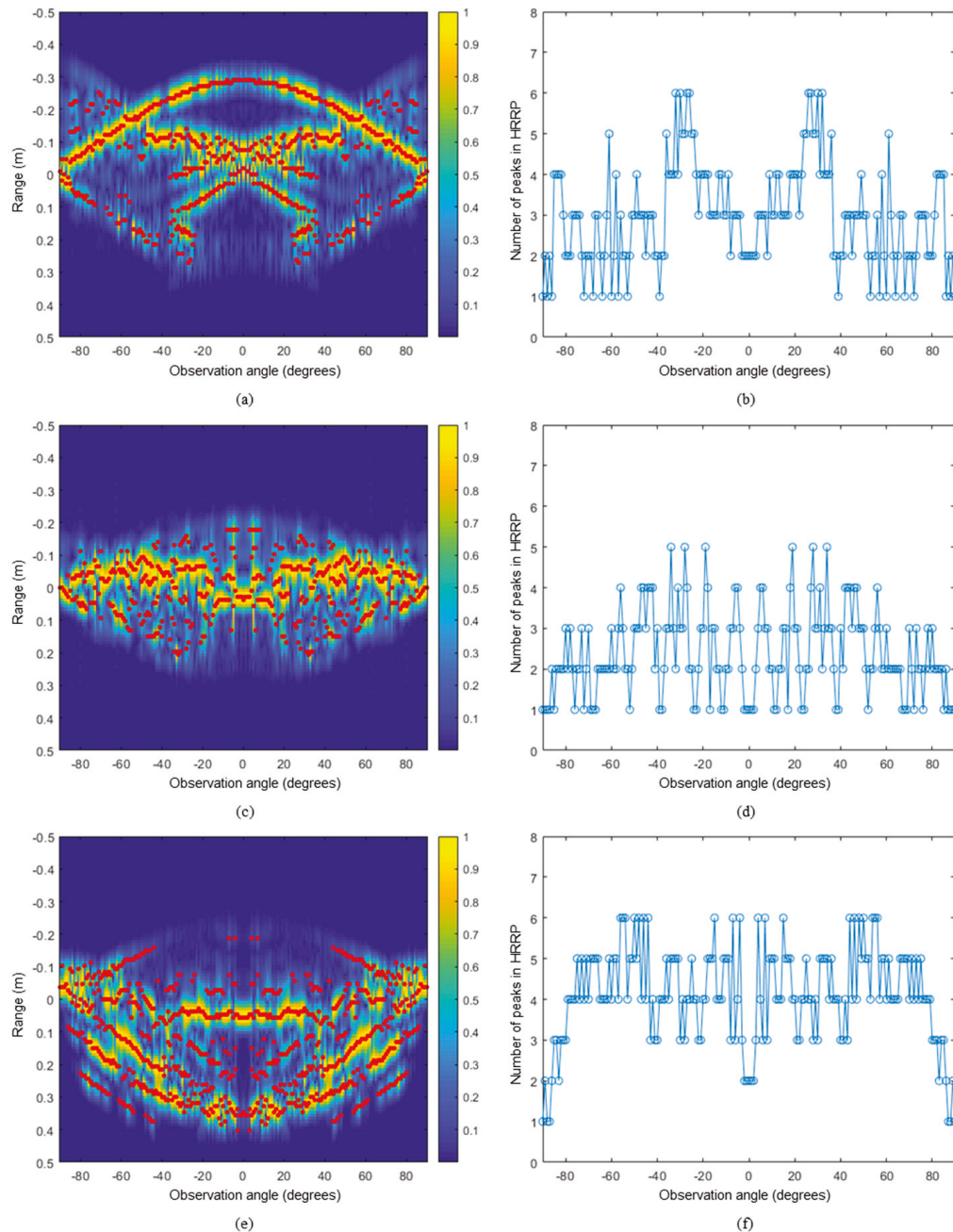
Finally, after 1000 trials on an Intel Core i7-4930 CPU @ 3.4 GHz, the average computation time of the proposed algorithm was 2.2 ms, which is only 2% of the 150 ms required for the SAR algorithm.

Table 1. Specifications of target models.

Model (Scale)	Length (mm)	Wingspan (mm)
A380 (1:120)	650	580
Eurofighter (1:32)	500	350
F-15 (1:32)	600	400

Table 2. Algorithm parameters for numerical evaluation.

Parameter	Value
Number of angle selections K	4
Angle interval δ	3°
Matching threshold γ	$\Delta R/2$
Exposure threshold γ_ρ	70%

**Figure 6.** Peak detection for numerical models with the number of peaks in HRRPs with respect to the observation angle. A380: (a,b), Eurofighter: (c,d), F-15: (e,f).

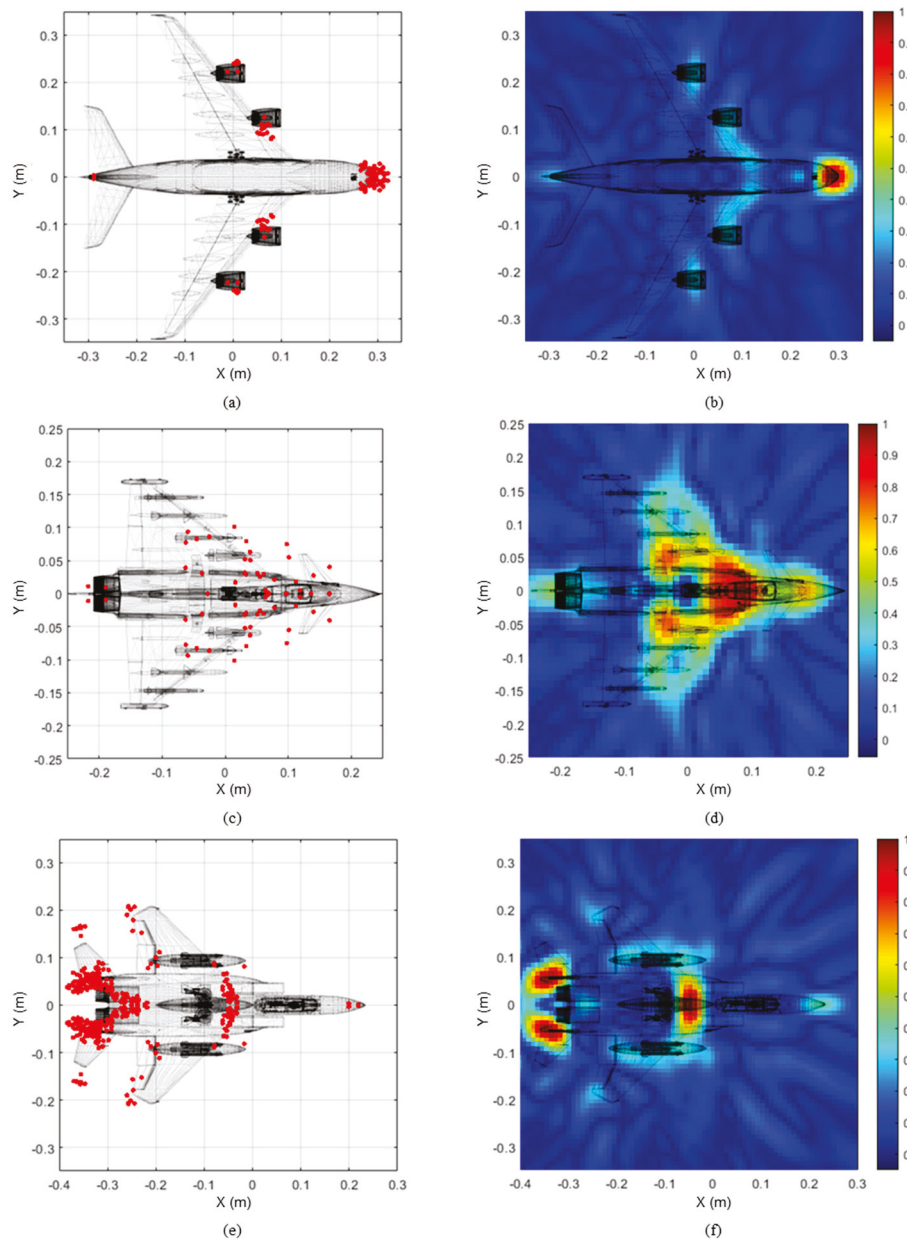


Figure 7. Two-dimensional scattering center estimation and corresponding SAR images of the A380: (a,b), Eurofighter: (c,d), and F-15: (e,f).

3.2. Experiment on Actual Target

To validate the proposed algorithm, a radar measurement system was designed consisting of Agilent's 8720ES vector network analyzer (VNA) and two X-band standard gain horn antennas. The target is placed on a rotating platform for angular variation.

The VNA was used at the X-band, 8 GHz to 12 GHz, with 1024 frequency steps, denoted as frequency samples S , and the frequency step size was 3.9 MHz with a range resolution ΔR of 3.75 cm. Also, the desired transmit power is 5 dBm. The gain of the horn antenna was 20 dBi, and the measured data from the VNA was sent to a storage system via Agilent's 82357B GPIB connection, which is manufactured by Keysight Technologies, located in Santa Clara, CA, USA. The azimuth and elevation angles of the rotating platform were controlled by an Autronics A200K-M599-G10 stepper motor which is manufactured by Autronics, located in Busan, South Korea and a MG995 geared servomotor which is manufactured by TowerPro, located in Taiwan, respectively, through serial communication.

A block diagram of the entire radar measurement system is shown in Figure 8, and the experimental setup is shown in Figure 9.

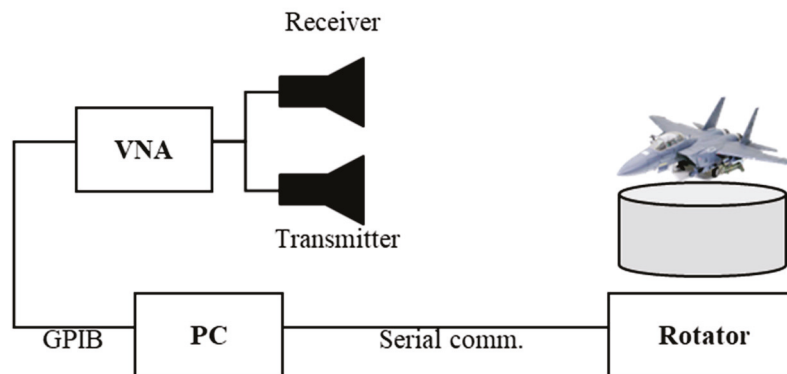


Figure 8. Block diagram of the entire measurement system.

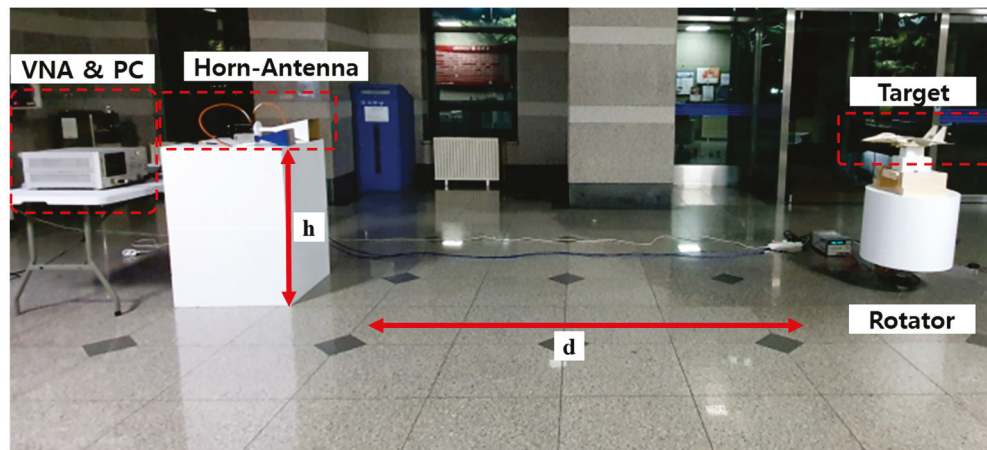


Figure 9. Experimental environment with $d = 4.2$ m and $h = 1$ m.

As performed in the numerical simulation, the scattering centers of the plastic die-cast models of the airbus A-380, Eurofighter, and F-15 were estimated by applying the proposed algorithm. A conductive silver paste was applied to the plastic target surface to improve the signal response. Since the dimensions of each numerical model were designed to be the same as the die-cast model, they follow the specifications listed in Table 1.

The HRRP datasets were obtained for each model at an elevation of $\theta = 0^\circ$, and azimuth angle ϕ in a range of -60° to 60° with an interval of 1° using the proposed measurement system, and the number of observation angles N was 121. The antenna was positioned at a height of 1 m above the ground at a distance of 4.2 m from the target. The same set of algorithm parameters was used as in the numerical simulation and is listed in Table 2.

The scattering point estimation results are shown in Figures 10–12 as red dots on the SAR image obtained for each model, and the corresponding dominant scattering centers of the actual target are highlighted in red circles. As we have already seen from the simulation results, the dominant scattering centers are distributed at the main points of the SAR images, such as the engines, intakes, and the wings.

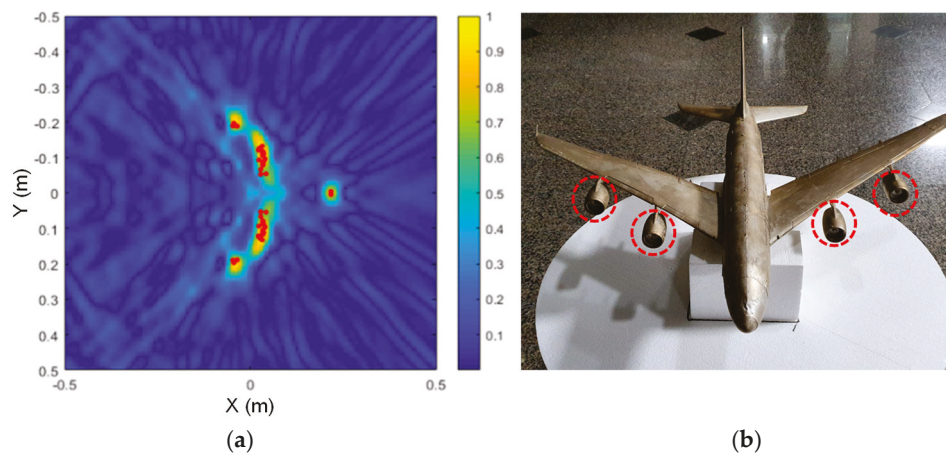


Figure 10. Estimated scattering centers of the A380 model; (a) drawn on a SAR image, and (b) the corresponding dominant scattering centers.

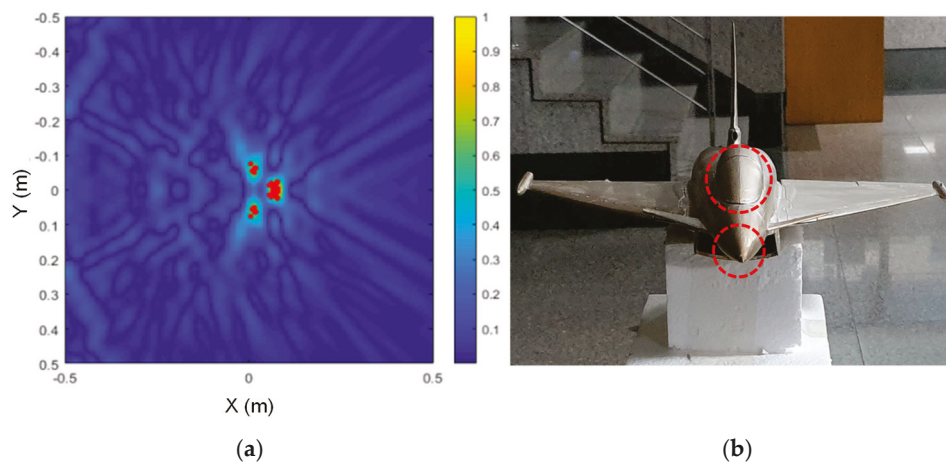


Figure 11. Estimated scattering centers of the Eurofighter model; (a) drawn on a SAR image, and (b) the corresponding dominant scattering centers.

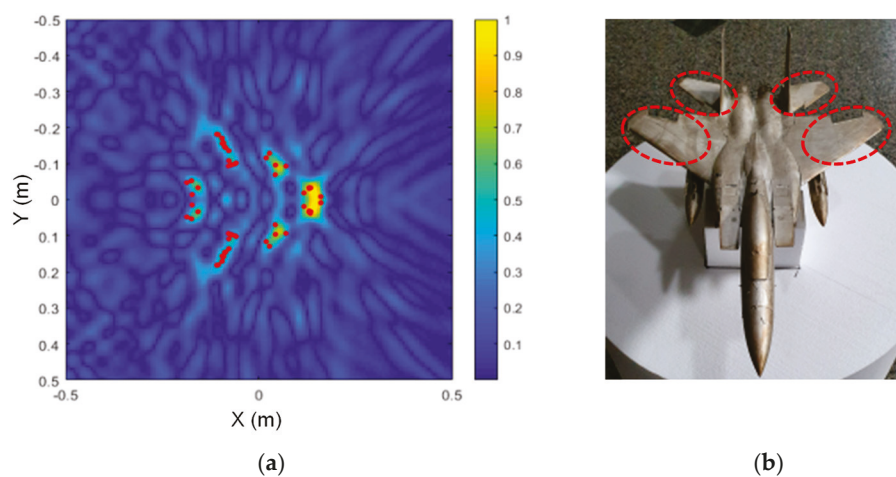


Figure 12. Estimated scattering centers of the F-15 model; (a) drawn on a SAR image, and (b) the corresponding dominant scattering centers.

3.3. Target Classification

The scattering points estimated by the proposed algorithm are subject to being expressed in the standardized data dimension to be applied to the classifier machine. In order

to standardize the positions of the estimated scattering centers, a scattering center image needs to be generated. In this case, the position and value of each pixel are required to consider the two-dimensional location of the scattering centers. Therefore, when the estimated scattering centers exist within the predefined two-dimensional pixel, it is determined that the image pixel exists, and the value of each pixel is defined as the maximum value of the amplitude of the scattering points within the pixel, which is defined as the sum of the amplitudes of K HRRPs used for scattering center estimation. In the classification process, the SVM is used, which has the advantage of classifying feature points that are difficult to linearly classify through high-dimensional mapping using a kernel function.

The HRRP datasets for each model were obtained at elevation θ and azimuth angles ϕ in ranges of -30° to 30° and -60° to 60° , respectively, with intervals of 2° and 1° , under the same conditions as in Experiment 3.2. The image was generated using HRRP in the azimuth section defined at a specific elevation angle, and the same pixel size and image size were considered for comparison of its classification performance to the SAR images. The size of the image used for training and validation is $-0.5\text{ m} \times 0.5\text{ m}$, where the size of one pixel is $0.05\text{ m} \times 0.05\text{ m}$, so it has 21 pixels in width and depth respectively.

The images generated from the proposed method and SAR with three different elevation angles are shown in Figures 13–15 for the given target models. In these figures, it can be seen that the pixels of the scattering center image are distributed at the main points of the SAR image, and it can also be confirmed that the characteristics of each target are clearly distinguishable.

The verification was performed by dividing the training set and validation set, as shown in Table 3. The azimuth section for image generation was defined as three sections, and the elevation section was defined to be different so that the images for training and verification did not overlap. For training and validation, a total of 31 images were generated for each model, where 16 images were used for training and 15 images were applied for validation.

To train and validate the classifier, feature extraction was performed using PCA (Principal Component Analysis) and LDA (Linear Discriminant Analysis). PCA reduces the dimensionality of high-dimensional input data by projecting it onto the dimensions with maximum variance. LDA, on the other hand, considers the data classes and finds a coordinate space that minimizes the variance within classes while maximizing the variance between classes. Using these techniques, distinctive feature points were extracted.

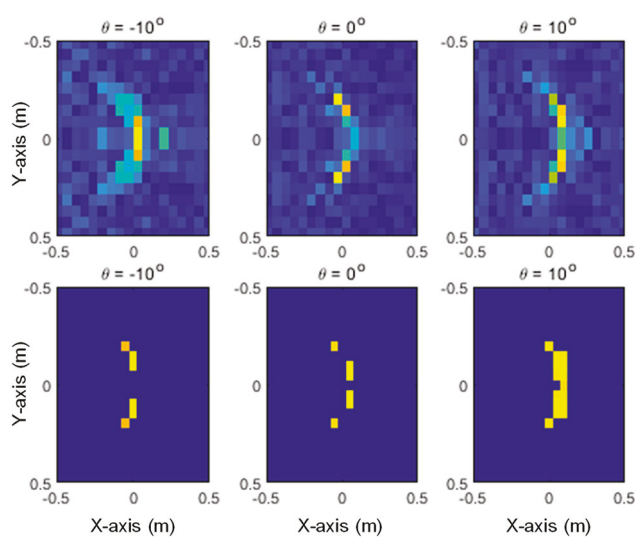


Figure 13. Generated images for the A380: SAR (upper row) and the proposed method (lower row).

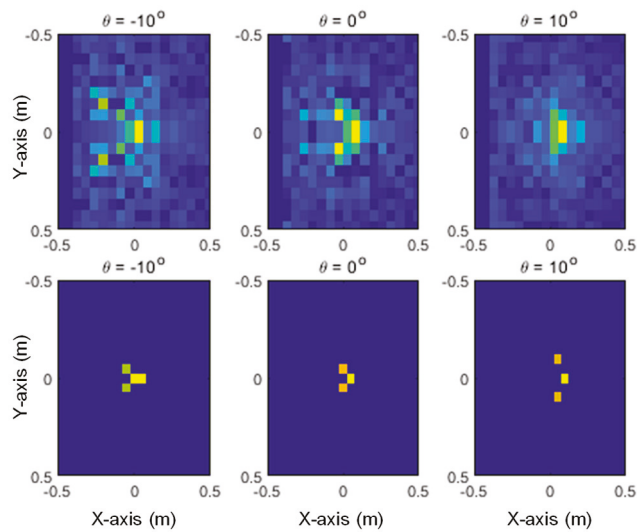


Figure 14. Generated images for the Eurofighter: SAR (**upper row**) and the proposed method (**lower row**).

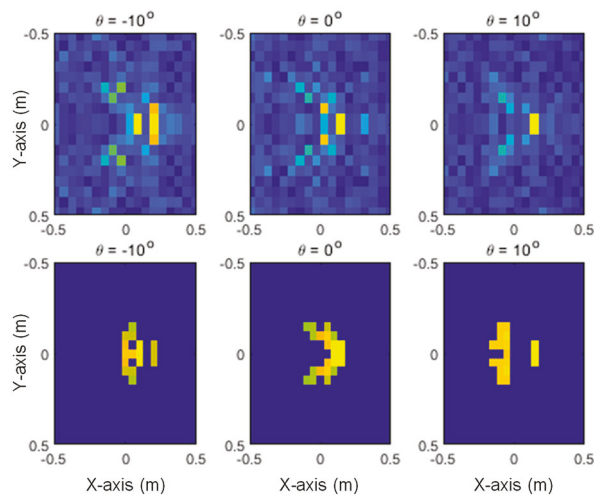


Figure 15. Generated images for the F-15: SAR (**upper row**) and the proposed method (**lower row**).

Table 3. HRRP acquisition angle sections for training and testing.

Item	Azimuth (deg)	Step	Elevation (deg)	Step
Training	$-40^{\circ} \sim 40^{\circ}$	1°	$-30^{\circ} \sim 30^{\circ}$	4°
	$-50^{\circ} \sim 50^{\circ}$			
	$-60^{\circ} \sim 60^{\circ}$			
Validation	$-40^{\circ} \sim 40^{\circ}$	1°	$-28^{\circ} \sim 28^{\circ}$	4°
	$-50^{\circ} \sim 50^{\circ}$			
	$-60^{\circ} \sim 60^{\circ}$			

The feature points obtained from the training images were used to train an SVM (Support Vector Machine) with an RBF kernel. The feature points extracted from the validation images were then input into the trained SVM for classification.

The classification results using the datasets described in Table 3 are shown as the confusion matrix form in Figures 16 and 17. In the case of using a SAR image, it can be confirmed that all targets were classified in the right place, while the classification rate of the proposed algorithm is 98.52%, which is somewhat inferior to that of the SAR images. This can be seen as a performance loss due to the trade-off for reducing the computation

load confirmed in the previous section. Based on the results, it can be judged that the proposed algorithm can be applied to situations requiring real-time target recognition or two-dimensional target information.

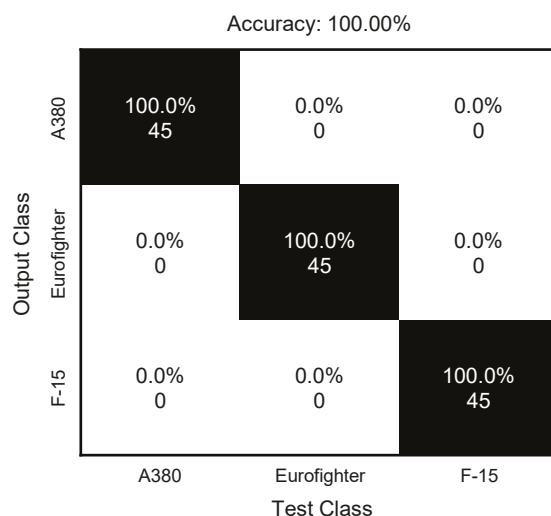


Figure 16. Classification results using SAR images.

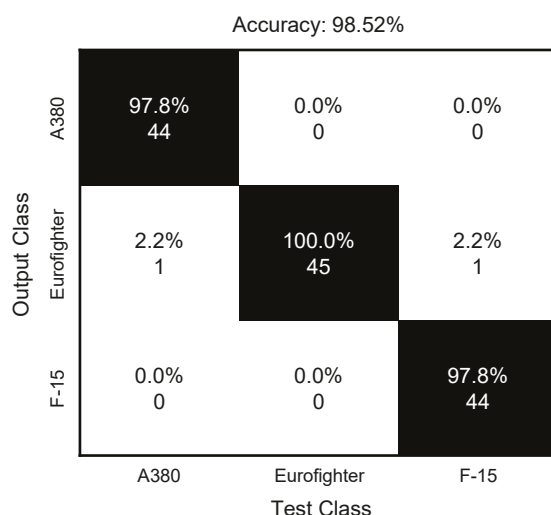


Figure 17. Classification results using the proposed method.

4. Conclusions

In this paper, we proposed a new method for estimating target scattering centers from multiple HRRP datasets. These scattering centers can help obtain a high resolution image of the target and thus can also be used for the classification of the target. Based on the range information contained in multiple HRRPs obtained from various observation angles, the estimated target scattering centers could be located at the intersection points of the lines perpendicular to each of the observation axes. The computational complexity of the proposed algorithm was proven to be significantly less than that of the conventional SAR method or other methods in the previous literature.

The performance of the proposed algorithm was verified by applying the multiple HRRPs measured in the experiment using the X-band radar measurement system. The system was designed using the VNA, a standard gain horn antenna, and a rotator for angle rotation, and the HRRP dataset was obtained using the received signal. As a result of estimating the scattering centers by applying the proposed algorithm to the obtained

HRRPs, it was confirmed that the scattering centers can be successfully estimated at the main points of the SAR images.

Finally, a strategy for generating images was introduced to be used in the classifier from the estimated scattering centers. The classification performance using the scattering center image had a failure rate of less than 2% compared to the discrimination performance using the SAR image.

As demonstrated in this article, the proposed algorithm shows similar target classification capabilities but much less necessary computation than the SAR-based method. Therefore, the proposed algorithm is advantageous in real-time image acquisition and the reduction of the required storage capacity of the classification database, so it is expected to be used in various fields, from defense fields, such as air defense, maritime, and ground surveillance, to private industries, such as vehicle radar and drone radar.

Author Contributions: Conceptualization and methodology, K.-I.L. and Y.-S.C.; software, K.-I.L.; validation and formal analysis, K.-I.L., Y.-S.C. and J.-H.K.; investigation, K.-I.L. and J.-H.K.; resources, Y.-S.C.; data curation, K.-I.L.; writing—original draft preparation, K.-I.L. and J.-H.K.; writing—review and editing, K.-I.L. and Y.-S.C.; visualization, K.-I.L. and J.-H.K.; supervision, Y.-S.C.; project administration, Y.-S.C.; funding acquisition, Y.-S.C. All authors have read and agreed to the published version of the manuscript.

Funding: The present research has been conducted by the Excellent researcher support project of Kwangwoon University in 2024.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Tait, P. *Introduction to Radar Target Recognition*; IET: Stevenage, UK, 2005; Volume 18.
2. Huang, C.W.; Lee, K.C. Application of Ica Technique to PCA Based Radar Target Recognition. *Prog. Electromagn. Res.* **2010**, *105*, 157–170. [CrossRef]
3. Zyweck, A.; Bogner, R.E. Radar target classification of commercial aircraft. *IEEE Trans. Aerosp. Electron. Syst.* **1996**, *32*, 598–606. [CrossRef]
4. Chen, S.; Wang, H.; Xu, F.; Jin, Y. Target Classification Using the Deep Convolutional Networks for SAR Images. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 4806–4817. [CrossRef]
5. Li, H.-J.; Yang, S.-H. Using range profiles as feature vectors to identify aerospace objects. *IEEE Trans. Antennas Propag.* **1993**, *41*, 261–268. [CrossRef]
6. Wong, S.K. High Range Resolution Profiles as Motion-Invariant Features for Moving Ground Targets Identification in SAR-Based Automatic Target Recognition. *IEEE Trans. Aerosp. Electron. Syst.* **2009**, *45*, 1017–1039. [CrossRef]
7. Novak, L. A comparison of 1D and 2D algorithms for radar target classification. In Proceedings of the IEEE 1991 International Conference on Systems Engineering, Dayton, OH, USA, 1–3 August 1991; pp. 6–12.
8. Turk, M.; Pentland, A. Eigenfaces For Recognition. *J. Cogn. Neurosci.* **1991**, *3*, 71–86. [CrossRef] [PubMed]
9. Chen, B.; Liu, H.; Bao, Z. PCA and kernel PCA for radar high range resolution profiles recognition. In Proceedings of the IEEE International Radar Conference, Arlington, VA, USA, 9–12 May 2005; pp. 528–533.
10. Turhan-Sayan, G. Real time electromagnetic target classification using a novel feature extraction technique with PCA-based fusion. *IEEE Trans. Antennas Propag.* **2005**, *53*, 766–776. [CrossRef]
11. Mishra, A.K. Validation of PCA and LDA for SAR ATR. In Proceedings of the TENCON 2008—2008 IEEE Region 10 Conference, Hyderabad, India, 19–21 November 2008; pp. 1–6.
12. Wu, H.; Dai, D.; Wang, X. A novel radar HRRP recognition method with accelerated T-distributed stochastic neighbor embedding and density-based clustering. *Sensors* **2019**, *19*, 5112. [CrossRef] [PubMed]
13. Balakrishnama, S.; Ganapathiraju, A. *Linear Discriminant Analysis—A Brief Tutorial*; Institute for Signal and Information Processing: Simrall, MS, USA, 1998; pp. 1–8.
14. Noushath, S.; Kumar, G.H.; Shivakumara, P. (2D)2LDA: An efficient approach for face recognition. *Pattern Recognit.* **2006**, *39*, 1396–1400. [CrossRef]

15. Liu, J.; Zhang, J.; Zhao, F. High Resolution Radar Automatic Target Recognition Based on an Improved LDA Method. In Proceedings of the 2008 Fifth International Conference on Fuzzy Systems and Knowledge Discovery, Jinan, China, 18–20 October 2008; pp. 301–305.
16. Fu, J.S.; Liao, K.; Yang, W.L. Radar HRRP target recognition using multi-KFD-based LDA Algorithm. *Prog. Electromagn. Res. C* **2012**, *30*, 15–26. [CrossRef]
17. Kishore, T.R.; Rao, K.D. Automatic Intrapulse Modulation Classification of Advanced LPI Radar Waveforms. *IEEE Trans. Aerosp. Electron. Syst.* **2017**, *53*, 901–914. [CrossRef]
18. Fu, Z.; Li, S.; Li, X.; Dan, B.; Wang, X. A neural network with convolutional module and residual structure for radar target recognition based on high-resolution range profile. *Sensors* **2020**, *20*, 586. [CrossRef] [PubMed]
19. Kent, S.; Kasapoglu, N.G.; Kartal, M. Radar target classification based on support vector machines and High Resolution Range Profiles. In Proceedings of the 2008 IEEE Radar Conference, Rome, Italy, 26–30 May 2008; pp. 1–6.
20. Jasinski, T.; Brooker, G.; Antipov, I. W-band multi-aspect high resolution range profile radar target classification using support vector machines. *Sensors* **2021**, *21*, 2385. [CrossRef] [PubMed]
21. Huang, G.; Zhou, H.; Ding, X.; Zhang, R. Extreme Learning Machine for Regression and Multiclass Classification. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* **2012**, *42*, 513–529. [CrossRef] [PubMed]
22. Kim, K.-T.; Kim, H.-T. Two-dimensional scattering center extraction based on multiple elastic modules network. *IEEE Trans. Antennas Propag.* **2003**, *51*, 848–861.
23. Zhou, J.; Zhao, H.; Fu, Q. A Novel Method for Reconstructing 3D Scattering Centers Based on Multiple HRR Profiles and Its Performance Bounds. In Proceedings of the 2006 CIE International Conference on Radar, Shanghai, China, 16–19 October 2006; pp. 1–4.
24. Jun, W.; Dong, Y.; Yuxi, Z.; Shaoming, W. 3-D imaging algorithm for targets with micro-motion based on HRRP sequence. In Proceedings of the 2010 2nd International Conference on Signal Processing Systems, Dalian, China, 5–7 July 2010; pp. V2-283–V2-288.
25. Su, N.; Dai, F.; Liu, H.; Zhang, B. Three-Dimensional Absolute Attitude Reconstruction of a Rigid Body Based on Multi-Station HRRP Sequences. *IEEE Access* **2020**, *8*, 27793–27806. [CrossRef]
26. Bae, J.-H.; Kim, K.-T. Compressive sensing-based two-dimensional scattering-center extraction for incomplete RCS data. *Etri J.* **2020**, *42*, 815–826. [CrossRef]
27. Zhou, J.; Zhao, H.; Shi, Z.; Fu, Q. Extracting global 3D scattering center model of radar target from multiple HRR profiles. In Proceedings of the 2007 IEEE Radar Conference, Waltham, MA, USA, 17–20 April 2007; IEEE: Piscataway, NJ, USA, 2007.
28. Zhou, J.; Shi, Z.; Cheng, X.; Fu, Q. Automatic target recognition of SAR images based on global scattering center model. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 3713–3729.
29. Jakowatz, C.V.; Wahl, D.E.; Eichel, P.H.; Ghiglia, D.C.; Thompson, P.A. *Spotlight-Mode Synthetic Aperture Radar: A Signal Processing Approach*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012.
30. Van Loan, C.F.; Golub, G.H. *Matrix Computations*; Johns Hopkins University Press: Baltimore, MD, USA, 1983.
31. Pipatsrisawat, T.; Gacic, A.; Franchetti, F.; Puschel, M.; Moura, J.M.F. Performance analysis of the filtered backprojection image reconstruction algorithms. In Proceedings of the Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing. 2005, Philadelphia, PA, USA, 23–23 March 2005; Volume 5, pp. 153–156.
32. Toft, P. The radon transform—Theory and Implementation. Ph.D. Dissertation, Technical University of Denmark, Copenhagen, Denmark, 1996.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Motion Clutter Suppression for Non-Cooperative Target Identification Based on Frequency Correlation Dual-SVD Reconstruction

Weikun He *, Yichuan Luo and Xiaoxiao Shang

Tianjin Key Lab for Advanced Signal Processing, Civil Aviation University of China, Tianjin 300300, China; luoyichuan1234567@163.com (Y.L.); xxshang512@163.com (X.S.)

* Correspondence: hwkcauc@126.com; Tel.: +86-022-24092437

Abstract: Non-cooperative targets, such as birds and unmanned aerial vehicles (UAVs), are typical low-altitude, slow, and small (LSS) targets with low observability. Radar observations in such scenarios are often complicated by strong motion clutter originating from sources like airplanes and cars. Hence, distinguishing between birds and UAVs in environments with strong motion clutter is crucial for improving target monitoring performance and ensuring flight safety. To address the impact of strong motion clutter on discriminating between UAVs and birds, we propose a frequency correlation dual-SVD (singular value decomposition) reconstruction method. This method exploits the strong power and spectral correlation characteristics of motion clutter, contrasted with the weak scattering characteristics of bird and UAV targets, to effectively suppress clutter. Unlike traditional clutter suppression methods based on SVD, our method avoids residual clutter or target loss while preserving the micro-motion characteristics of the targets. Based on the distinct micro-motion characteristics of birds and UAVs, we extract two key features: the sum of normalized large eigenvalues of the target's micro-motion component and the energy entropy of the time–frequency spectrum of the radar echoes. Subsequently, the kernel fuzzy c-means algorithm is applied to classify bird and UAV targets. The effectiveness of our proposed method is validated through results using both simulation and experimental data.

Keywords: target identification of birds and UAVs; strong motion clutter; frequency correlation dual-SVD reconstruction; clutter suppression; micro-motion characteristics

1. Introduction

The detection of low-altitude targets by radar is often compromised by strong motion clutter from sources such as airplanes (in apron environments) and cars (on a highway). This interference makes it challenging to distinguish between birds and UAVs, which possess weak scattering characteristics and can easily be obscured by clutter. Therefore, developing effective clutter suppression methods for the identification of birds and UAVs in environments with strong motion clutter is crucial for enhancing the performance of radar detection systems [1,2].

Regarding strong clutter suppression, scholars have conducted extensive research focusing on methods such as adaptive moving target indication (AMTI) [3], the CLEAN algorithm [4], wavelet transform techniques [5–10], and singular value decomposition (SVD) [11–15]. Among these, SVD-based motion clutter suppression methods have garnered significant attention. Zheng Lin et al. [11] proposed a clutter suppression method that combines phase encoding with SVD. This method first whitens the clutter through phase decoding and then separates the clutter based on eigenvalue differences in the autocorrelation matrix between the whitened clutter and the target. Additionally, the wavelet-SVD algorithm involves performing a two-dimensional wavelet transform and SVD on the transformed wavelet coefficient matrix, followed by the removal of larger singular values

to suppress clutter [12]. M. Garcia Fernandez et al. [13] introduced a clutter suppression method for synthetic aperture radar (SAR) images. In this method, clutter and target subspaces are constructed through the SVD of SAR images, and the bases corresponding to clutter are selected. Radar echoes are then projected onto the orthogonal complement of the clutter subspaces to achieve clutter suppression. Reference [14] utilizes three statistical features: singular value spectral distribution, singular vector space correlation (where ground clutter in different range resolution cells exhibits strong correlation), and average Doppler frequency. These features are used to adaptively determine the singular vectors corresponding to the clutter subspace via the K-means clustering algorithm, and clutter components are subsequently suppressed through orthogonal subspace projection. Wu Linyong et al. [15] proposed a first-order differential spectral SVD reconstruction method. This technique employs the first-order differential spectrum of singular values to classify and reconstruct radar echoes, effectively distinguishing between clutter and targets. Cui Weicheng et al. [16] proposed a method for selecting singular values based on the principle of minimizing the fitting error.

In the realm of target identification methods for UAVs and birds, there are two primary approaches: those based on feature extraction and those that combine feature extraction with machine learning. Key extracted features include spectral symmetry pairs of radar echoes [17], motion model conversion frequency [18], the first left singular vector corresponding to singular value decomposition, and the cadence velocity diagram (CVD) [19], as well as target velocity, spectral period, and width [20]. Additionally, features such as the mean and variance of singular values are also utilized [21]. Identification methods that integrate feature extraction with machine learning include various combinations of feature vectors and classifiers. For instance, methods include using a feature vector corresponding to the first two largest eigenvalues combined with a support vector machine (SVM) [22], a micro-Doppler bandwidth combined with a K-nearest neighbor (KNN) classifier [23], Doppler images derived from micro-Doppler signatures (MDS) and rhythmic velocimetry maps combined with a convolutional neural network (CNN) [24], and synthetic micro-Doppler spectrograms combined with a CNN [25]. Reference [26] compares the classification effectiveness of CNNs and autoencoders. Beom Seok Oh et al. [27] designed an automatic recognition and classification system. This system employs empirical mode decomposition (EMD) to decompose and reconstruct radar echoes from UAVs, birds, humans, vehicles, and other objects. Signal energy, logarithmic energy entropy, and spectral energy density are then extracted to classify UAVs, birds, pedestrians, and vehicles using SVM. Additionally, wavelet decomposition is applied to radar echoes of UAVs and robotic birds, allowing for the estimation of bird flapping frequencies and drone rotor speeds for target identification [28].

The previously discussed methods for identifying birds and UAVs do not address the interference caused by strong motion clutter in real-world observation environments (as shown in Figure 1, the target is obscured by strong clutter). As a result, detecting the target within this clutter becomes challenging.

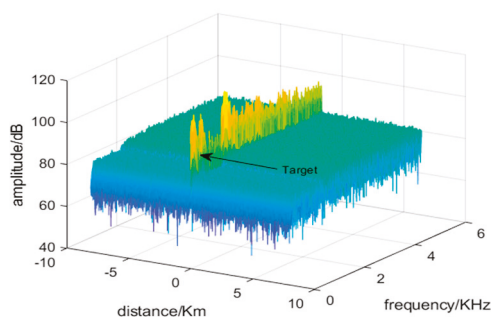


Figure 1. Range-Doppler spectrum with strong motion clutter interference.

Moreover, SVD-based clutter suppression methods encounter several challenges. First, selecting the appropriate singular values is difficult. Second, in practical scenarios, the intermingling of clutter and target information can cause the singular vectors corresponding to these singular values to contain both, potentially leading to the loss of target signals.

The frequency correlation dual-SVD reconstruction method is proposed to address the impact of strong motion clutter on the discrimination performance of birds and UAVs. This method aims to remove motion clutter while preserving the micro-motion information of targets. Compared to first-order differential spectrum SVD reconstruction (FODS-SVD) in reference [15] and fitting error minimum principle SVD reconstruction (FEMP-SVD) in reference [16], the main differences in the proposed method lie in the selection of singular values and the handling of clutter intertwined with target information. The FODS-SVD and FEMP-SVD method only analyze how to select singular values and do not consider the intermingling of target and clutter information. Additionally, they do not account for the possibility that, after singular value decomposition, the singular values representing clutter components may be smaller than those representing target information (leading to residual clutter). The proposed method has already solved the two issues mentioned above. The target identification section is primarily used to verify the effectiveness of this method.

The content of this paper is arranged as follows. The second section primarily discusses the innovative clutter suppression method proposed in this study. The third section focuses on extracting the features of birds and UAVs and implementing their identification using the kernel fuzzy c-means (KFCM) method. The fourth section analyzes the results based on both simulated and experimental data. The fifth section provides a summary of the content of this paper.

2. Frequency Correlation Dual-SVD Reconstruction Based Clutter Suppression

2.1. General Approach to Clutter Suppression

Before discriminating between non-cooperative targets, ground clutter is suppressed using moving target detection (MTD) [29]. Initially, the constant false alarm rate (CFAR) [30] technique is employed to detect strong motion clutter and suppress it using the proposed method. Next, the micro-motion characteristics of the targets are assessed. Only range resolution cells exhibiting these characteristics are retained, while cells containing only clutter are removed. Consequently, subsequent processing focuses on the range resolution cells where both targets and clutter coexist.

In the proposed method, the strong power and spectral correlation of the motion clutter are utilized, and the received signal is processed using the corresponding Hankel matrix [16]. The spectral correlation of the strong motion clutter is transformed into the spectral correlation of the singular vectors obtained through singular value decomposition (SVD) of the Hankel matrix corresponding to the radar echoes. Since the power of the motion clutter significantly exceeds that of the target [31], singular vectors mainly containing clutter components can be identified based on the correlation between the spectrum of each singular vector and that of the singular vector with the maximum singular value.

To preserve target micro-motion information, spectral waveform entropy is used as a quantitative evaluation index for extracting singular vectors containing both clutter and target information. This spectral correlation characteristic is then utilized again to suppress the strong motion clutter component in singular vectors containing both clutter and target information. Then, new singular vectors containing only target information can be obtained. Finally, by selecting different singular values and singular vectors corresponding to the target and clutter, the new Hankel matrix is reconstructed. The inverse Hankel transform is subsequently applied to obtain the target components. A block diagram of the proposed method is shown in Figure 2.

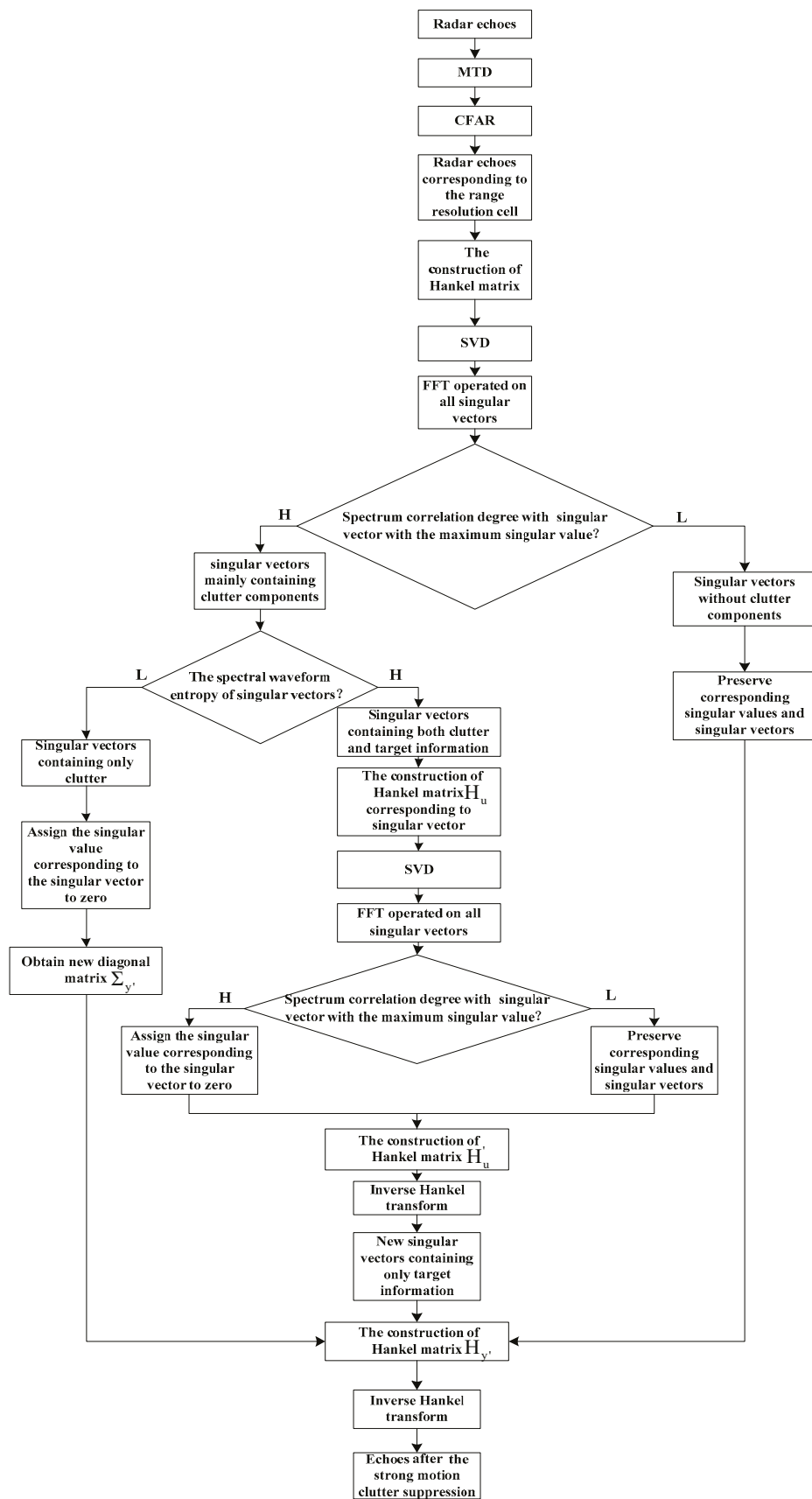


Figure 2. The block diagram of the method based on frequency correlation dual SVD reconstruction.

2.2. Determination of the Singular Values and Singular Vectors Corresponding to the Clutters

The received echoes corresponding to the range resolution cell of interest are the sequence of sampled instantaneous amplitudes of the received electromagnetic signal over the sampling periods, which can be represented as

$$X = s + c + n \quad (1)$$

where s is the bird or UAV signal, c is strong motion clutter such as a car or plane, and n is noise. The clutter components exhibit spectral correlation due to the similarity in the spectral shape and center frequency of the motion clutter components. To utilize the spectral correlation of strong motion clutter components, the received signal $X = [x(1), x(2), \dots, x(N)]$ is first converted into a two-dimensional Hankel matrix H_x , which can be written as follows:

$$H_x = \begin{bmatrix} x(1) & x(2) & \cdots & x(J) \\ x(2) & x(3) & \cdots & x(J+1) \\ \vdots & \vdots & & \vdots \\ x(M) & x(M+1) & \cdots & x(N) \end{bmatrix} \quad (2)$$

where $N = M + J - 1$ is the received signal length, $M = (N + 1)/2$ when N is an odd number, and $M = N/2$ when N is even. J can be obtained by $J = N + 1 - M$. It can be observed that each row of the Hankel matrix H_x has $J - 1$ elements that are the same as the previous row, indicating a correlation between the rows. H_x is decomposed by the SVD as follows:

$$H_x = U_x \Sigma_x V_x^H \quad (3)$$

The left singular matrix $U_x = [u_1, u_2, \dots, u_M] \in \mathbb{C}^{M \times M}$ and the right singular matrix $V_x = [v_1^H, v_2^H, \dots, v_J^H] \in \mathbb{C}^{J \times J}$ are the orthogonal matrices, $\Sigma_x \in \mathbb{C}^{M \times J}$ is the diagonal matrix composed of the corresponding singular values in descending order, namely

$$\Sigma_x = \begin{cases} \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r) & M = J \\ [\text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r), O]^T & M < J \end{cases} \quad (4)$$

where $r = \min(M, J)$ represents the number of singular values, and O represents the zero matrix. The received signal information is contained in the obtained singular values and singular vectors. Therefore, the essence of the proposed clutter suppression method is to eliminate strong motion clutter components in all singular vectors.

The left and right singular vectors in Equation (3) represent the time-domain characteristics of the received signal. To utilize the spectral correlation of the motion clutter components, the fast Fourier transform (FFT) is applied to each of the left singular vectors, as shown below:

$$\bar{U}_i(k) = \text{FFT}[u_i(n)] = \sum_{n=0}^{M-1} u_i(n) W_M^{kn} = \sum_{n=0}^{M-1} u_i(n) e^{-j2\pi kn/M}, \quad k = 0, 1, 2, \dots, M-1 \quad (5)$$

Since each element on the anti-diagonals of the Hankel matrix H_x is equal and the left and right singular vectors are orthogonal to each other, the spectra of the left and right singular vectors are symmetric about zero frequency [32], and thus only the correlation between the spectra of the left singular vectors needs to be analyzed. Considering that the clutter power significantly exceeds that of the target, the left singular vector corresponding to the maximum singular value is regarded as the clutter component. Therefore, it is necessary to calculate the Pearson correlation coefficients [33] between the spectra of all singular vectors and the spectrum of the left singular vector corresponding to the maximum singular value. Singular vectors with higher Pearson correlation coefficients are extracted because they predominantly contain clutter components. This approach prevents the

issue where singular values representing clutter are smaller than those representing target information, which could lead to residual clutter.

To preserve the micro-motion features of the target while suppressing clutter, the singular vectors mainly containing clutter components are categorized into those containing only clutter and those containing both clutter and target information. Given the complexity and randomness of the spectral distribution of singular vectors, spectral waveform entropy [34] is introduced. Taking any spectral sequence $\bar{U}(k)$ as an example, the probability distribution $P(k)$ of the frequency components is calculated, namely

$$P(k) = |\bar{U}(k)|^2 / \sum_{k=0}^{M-1} (|\bar{U}(k)|^2) \quad (6)$$

The corresponding spectral waveform entropy can be expressed as follows:

$$Q = - \sum_{k=0}^{M-1} P(k) \log(P(k)) \quad (7)$$

That is, the larger the spectral waveform entropy value, the more frequency components are included. A higher entropy value of the spectral waveform is indicative of a singular vector containing both target and clutter information, while a lower entropy value suggests a singular vector containing only clutter. The singular values corresponding to the singular vectors containing only clutter are set to zero, resulting in a new diagonal matrix $\Sigma_{y'}$. However, the left or right singular vectors (u_{c+t} or v_{c+t}) containing both clutter and target information require further processing.

SVD reconstruction is performed again to suppress the motion clutter for the singular vectors containing both clutter and target information. Unlike the initial SVD reconstruction, the Hankel matrix H_u is constructed from each left singular vector containing clutter and target information., specifically:

$$H_u = \begin{bmatrix} u_{c+t}(1) & u_{c+t}(2) & \cdots & u_{c+t}(L) \\ u_{c+t}(2) & u_{c+t}(3) & \cdots & u_{c+t}(L+1) \\ \vdots & \vdots & \ddots & \vdots \\ u_{c+t}(Z) & u_{c+t}(Z+1) & \cdots & u_{c+t}(M) \end{bmatrix} \quad (8)$$

$Z = (M+1)/2$ when M is an odd number, and $Z = M/2$ when M is even, and L can be obtained by $L = M+1-Z$. H_u is subjected to SVD, and U_u , Σ_u , and V_u^H are obtained. Then, the Pearson correlation coefficients between the spectrum of the left singular vector and that of the left singular vector corresponding to the maximum singular value are calculated. The singular values corresponding to singular vectors with relatively larger Pearson correlation coefficient values are set to zero. The new diagonal matrix Σ'_u is obtained. The Hankel matrix H'_u is then reconstructed using U_u , Σ'_u , and V_u^H , specifically:

$$H'_u = U_u \Sigma'_u V_u^H = \begin{bmatrix} u_{tar}(1) & u_{tar}(2) & \cdots & u_{tar}(L) \\ u_{tar}(2) & u_{tar}(3) & \cdots & u_{tar}(L+1) \\ \vdots & \vdots & \ddots & \vdots \\ u_{tar}(Z) & u_{tar}(Z+1) & \cdots & u_{tar}(M) \end{bmatrix} \quad (9)$$

The inverse Hankel transform on H'_u can also be used to obtain new singular vectors containing only target information, which can then replace the original singular vector containing both clutter and target information. The new left and right singular matrices, $U_{y'}$ and $V_{y'}^H$, are obtained.

On this basis, the diagonal matrices $\Sigma_{y'}$, along with the left and right singular matrices $U_{y'}$ and $V_{y'}^H$, are then used to reconstruct the new Hankel matrix $H_{y'}$, specifically:

$$H_{y'} = U_{y'} \Sigma_{y'} V_{y'}^H = \begin{bmatrix} y'(1,1) & y'(1,2) & \cdots & y'(1,j) & \cdots & y'(1,J) \\ y'(2,1) & y'(2,2) & \cdots & y'(2,j) & \cdots & y'(2,J) \\ \vdots & \vdots & & \vdots & & \vdots \\ y'(m,1) & y'(m,2) & \cdots & y'(m,j) & \cdots & y'(m,J) \\ \vdots & \vdots & & \vdots & & \vdots \\ y'(M,1) & y'(M,2) & \cdots & y'(M,j) & \cdots & y'(M,J) \end{bmatrix} \quad (10)$$

where $y'(m,j)$ represents the element in the reconstructed Hankel matrix $H_{y'}$. Then, $H_{y'}$ is subjected to an inverse Hankel transformation, which is shown as follows

$$\begin{cases} y(1) = y'(1,1) \\ y(2) = [y'(1,2) + y'(2,1)]/2 \\ \vdots \\ y(N-1) = [y'(M-1,J) + y'(M,J-1)]/2 \\ y(N) = y'(M,J) \end{cases} \quad (11)$$

The signal after clutter suppression can be written as follows:

$$Y = [y(1), y(2), \dots, y(N)] \quad (12)$$

3. The Kernel Fuzzy C-Means-Based Identification of Birds and UAVs

After suppressing strong moving clutter, considering that energy entropy is easily affected by noise, the wavelet threshold denoising (WTD) method is used to denoise the radar echoes. In this method, the wavelet coefficients corresponding to the signal components are larger, while those corresponding to the noise components are much smaller. By selecting an appropriate threshold, the wavelet coefficients below this threshold are set to zero, and wavelet reconstruction is performed to obtain the denoised signal Y_{denoised} . Feature extraction is then carried out, and the identification of birds and UAVs is achieved using the KFCM method. The flowchart for identifying birds and UAVs is shown below. See Figure 3.

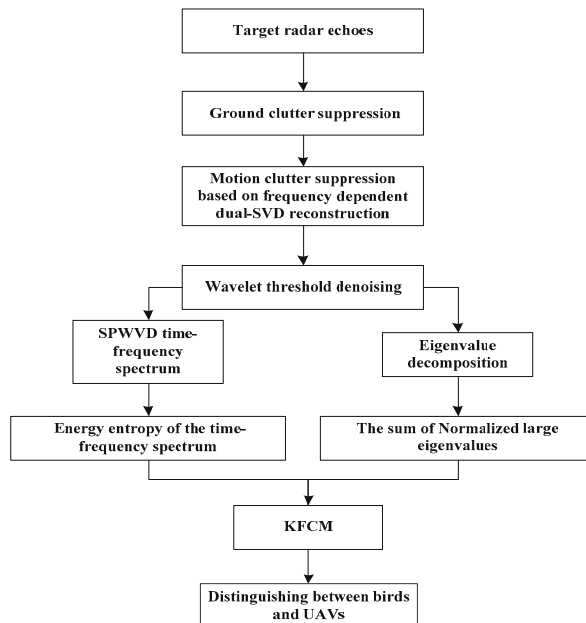


Figure 3. Block diagram of the bird and UAV targets discrimination against the background of strong motion clutter.

3.1. Feature Extraction

3.1.1. Energy Entropy of the Time–Frequency Spectrum

A time–frequency spectrum is used to depict the variance of the frequency distribution with time for a signal. The time–frequency spectrum can simultaneously display the variations of a signal in both the time and frequency domains, making it particularly suitable for analyzing non-stationary signals that change over time. Due to the flapping wings of birds, the time–frequency spectrum is less orderly compared to that of UAVs, which exhibit strong periodicity. According to the physical interpretation of entropy, a larger entropy value indicates a more disordered state. Therefore, the energy entropy of the time–frequency spectrum of the received radar echoes can be extracted as a micro-motion feature to distinguish between birds and UAVs.

The smoothed pseudo-Wigner–Ville distribution (SPWVD) effectively reduces cross-term interference and exhibits better energy concentration through windowing and smoothing in both the time and frequency domains. Therefore, SPWVD is selected for time–frequency analysis, as shown in Equation (13).

$$G(t, f) = \sum_{u=-\infty}^{+\infty} \sum_{\tau=-\infty}^{+\infty} h(\tau)g(u)Y_{\text{denoised}}(t - u + \frac{\tau}{2})Y_{\text{denoised}}^*(t - u - \frac{\tau}{2})e^{j2\pi f\tau} \quad (13)$$

By performing singular value decomposition on $G(t, f)$, the resulting left singular matrix U_g represents the frequency domain information. The time–frequency spectrum energy entropy η is then defined as follows:

$$\eta = -\sum_{n=1}^N q_n \ln q_n \quad (14)$$

where $q_n = E_n / (\sum_{n=1}^N E_n)$, $E_n = |(U_g)_{(n,1)}|^2 + |(U_g)_{(n,2)}|^2$, $|(U_g)_{(n,1)}|^2$ and $|(U_g)_{(n,2)}|^2$, respectively, represent the energy of the first and second left singular vectors for the n -th sample, and σ'_n represents the n -th eigenvalue. The aforementioned entropy can be used to preliminarily distinguish between birds and UAVs.

3.1.2. Sum of the Normalized Large Eigenvalues

Due to the higher rotational speed of a UAV's rotor, the spectrum of the UAV exhibits a wider bandwidth and smaller spacing between spectral lines compared to that of birds. Therefore, the sum of the normalized larger eigenvalues for the autocorrelation matrix of the target echoes is used to characterize the spectral width and the number of spectral lines for the target.

By exploiting the distinct sparse characteristics of the micro-motion components (flapping wing components for birds and rotor components for UAVs) and the body, the micro-motion components y_w of the signals Y_{denoised} are first extracted using the augmented Lagrangian method. The autocorrelation matrix of y_w can be represented as follows:

$$R = \frac{1}{N} y_w y_w^H \quad (15)$$

where $(\bullet)^H$ represents a complex conjugate. The matrix R is subjected to eigenvalue decomposition (EVD), specifically:

$$R = U_r \Sigma_r U_r^H \quad (16)$$

$\Sigma_r = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_N]$. $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$ represents the diagonal matrix composed of the corresponding eigenvalues. The sum of normalized larger eigenvalues is defined as follows:

$$f_Z = \sum_{i=1}^{N_c} \lambda_i / \lambda_1 \quad (17)$$

where N_c is the number of the larger eigenvalues, which are defined as those greater than one percent of the maximum eigenvalue [35]. Compared to the feature spectrum of birds, the spectrum of the UAVs has a greater number of larger eigenvalues and exhibits relatively stable changes. Therefore, the sum of normalized large eigenvalues for birds is generally smaller than that for UAV targets.

3.2. The Kernel Fuzzy C-Means Method for Distinguishing Birds and UAVs

Traditional hard clustering algorithms exhibit poor classification performance due to their strict membership degree of either zero or one. In contrast, fuzzy clustering algorithms extend the membership degree to any value between 0 and 1. This flexibility is not limited by the shape of the samples and can greatly enhance the identification performance of fuzzy datasets. The kernel fuzzy c-means (KFCM) clustering algorithm is an improved version of the traditional fuzzy c-means clustering algorithm. By introducing kernel functions, data can be projected into higher-dimensional feature spaces, facilitating easier separation of the data in these spaces. This approach improves both the discrimination accuracy and the stability of target identification.

For the i -th ($i = 1, 2, \dots, N_s$, where N_s represents the number of samples) given target echo sequence of length N , we calculate a single value η_i of the time–frequency spectrum and a single value f_{Z_i} representing the sum of normalized largest eigenvalues, and then combine them into a point a_i as follows:

$$a_i = (\eta_i \ f_{Z_i}) \quad (18)$$

In the experiment, N target echoes will generate N feature points. Then, by clustering these feature points using the KFCM method, the distinction between birds and drones can be achieved.

4. Results and Analysis

4.1. Results and Analysis of the Simulation Data

In the simulation experiment, an S-band linear frequency modulation pulse radar system was adopted. The signal-to-clutter ratio (SCR) corresponding to the motion clutter was -25 dB. The parameters of the radar, UAVs, and birds are shown in Tables 1–3. The coherent processing interval (CPI) was 256 milliseconds, so the signal length N was 3200. The spectral correlation threshold was set to 0.8 and the spectral entropy threshold was set to 7 bits in the clutter suppression method.

Table 1. Radar parameters.

Parameters	Value
Frequency	5.5 GHz
PRT	80 μ s
Sampling Rate	80 MHz
Bandwidth	20 MHz

Table 2. Parameters of UAVs.

Parameters	Value
Number of rotors	4
Number of blades	2
Blade Length	120 mm
Rotating Rate	50 r/s, 64 r/s

Based on the above parameters, by using the model that combines the movement characteristics of birds and the transmission characteristics of radar electromagnetic waves described in reference [36] and the scattering point superposition model described in reference [37], the

spectrum of the received data without clutter is shown in Figures 4a and 5a. The spectrum of the received data with clutter is shown in Figures 4b and 5b. In the strong motion clutter environment (that is, the plane echoes), the target is essentially covered and cannot be detected. The results after clutter suppression are shown in Figures 4c–e and 5c–e for the FODS-SVD method, the FEMP-SVD method, and the proposed method, respectively.

Table 3. Parameters of birds.

Parameters	Value
Upper arm length	0.15 m
Forearm length	0.2 m
Bird body length	0.3 m
Tapping frequency	8 Hz

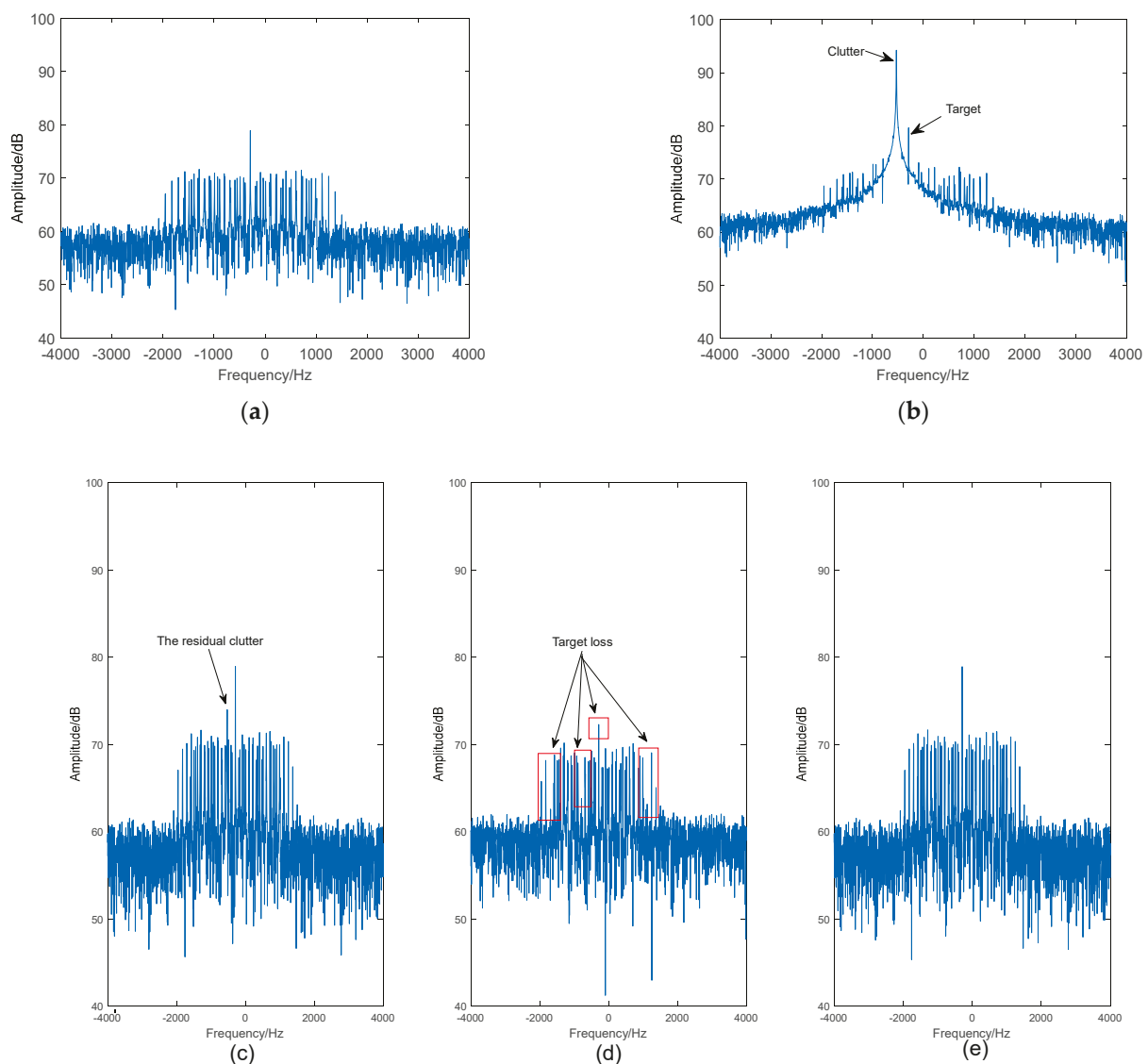


Figure 4. Clutter suppression results of a UAV target. (a) The spectrum of the received signal without clutter. (b) The spectrum of the received signal with clutter. (c) The spectrum after clutter suppression (FODS-SVD method). (d) The spectrum after clutter suppression (FEMP-SVD method). (e) Spectrum after clutter suppression (the proposed method).

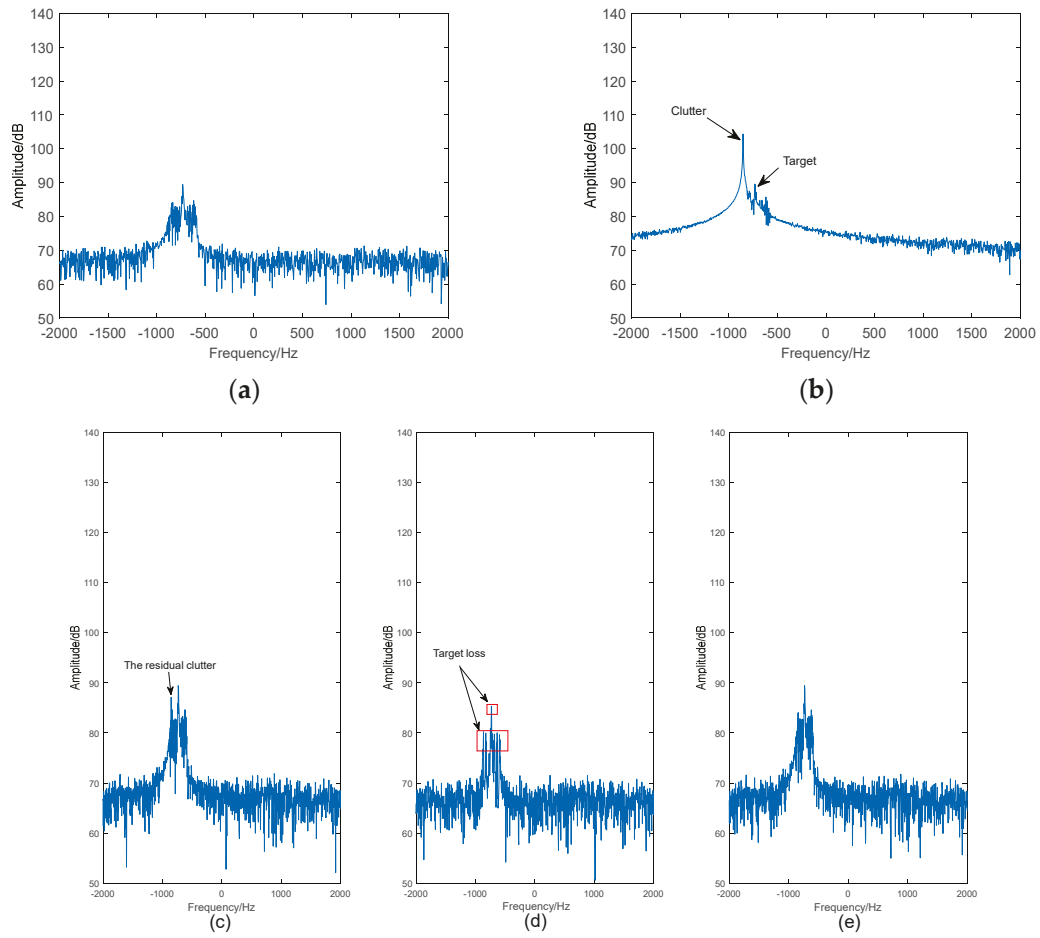


Figure 5. Clutter suppression results of a bird target. (a) The spectrum of the received signal without clutter. (b) The spectrum of the received signal with clutter. (c) The spectrum after clutter suppression (FODS-SVD method). (d) The spectrum after clutter suppression (FEMP-SVD method). (e) Spectrum after clutter suppression (the proposed method).

Compared to the results obtained by the FODS-SVD method, the proposed method successfully suppresses motion clutter, avoiding the existence of residual clutter caused by the situation where the singular values of clutter components are smaller than those of the target information. Compared to the results obtained by the FEMP-SVD method, the proposed method preserves relatively complete target information.

To quantitatively evaluate the performance of the proposed method, the relative variance between the signal power after clutter suppression and the power of the received signal without clutter is defined [38] as follows:

$$\text{err} = \frac{\|y - y_r\|_2^2}{\|y\|_2^2} \quad (19)$$

where y_r is the amplitude of the bird or UAV target without clutter and y is the signal amplitude of the bird or UAV after clutter suppression. The smaller the variance, the better the performance of the clutter suppression method. The relative power variations for the proposed method, the FODS-SVD method, and the FEMP-SVD method are shown in Figure 6. Using 200 sets of Monte Carlo experiments for each signal-to-clutter ratio, with each set containing 200 bird samples and 200 UAV samples, it can be observed that the relative power variance decreases and then tends to stabilize as the SCR increases.

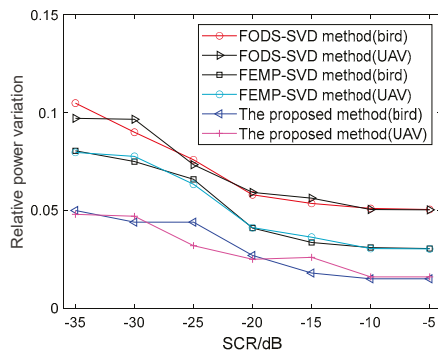


Figure 6. Relative power variation comparison after the clutter suppression.

The computational complexity of the proposed method is analyzed and compared with that of the FODS-SVD and FEMP-SVD methods, as shown in Table 4. The computer CPU model is Intel(R) Core (TM) i7-12700, and the main frequency is 2.10 GHz. It can be seen from Table 4 that the computation efficiency is lower compared with the FODS-SVD and FEMP-SVD methods. The proposed method performs two SVDs to ensure that the target is not lost and to suppress clutter as much as possible. However, this improved performance comes at the cost of increased computational complexity.

Table 4. The analysis of computational complexity.

Clutter Suppression Methods	Calculation Time
The FODS-SVD method	1.42 s
The FEMP-SVD method	1.88 s
The proposed method	2.98 s

After clutter suppression, the accuracy of target discrimination between birds and UAVs is depicted in Figure 7. As the signal-to-clutter ratio increases, the discrimination accuracy also improves. Compared to the FODS-SVD and FEMP-SVD methods, the proposed method significantly enhances discrimination accuracy in environments with strong motion clutter.

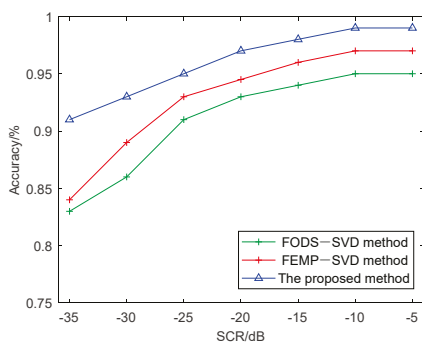


Figure 7. Performance comparison of the three methods.

4.2. Results and Analysis of the Experimental Data

The radar operates in the C-band frequency range in the experimental data, with the clutter being primarily caused by car echoes from a nearby highway (as shown in Figure 8b). The actual radar data acquisition environment and radar beam pattern are shown in Figure 8a,b.

The radar parameters include a pulse repetition time (PRT) of 62 μ s, a fast sampling rate of 16 MHz, and a bandwidth of 8 MHz. The coherent processing interval (CPI) is 127 milliseconds, so the signal length N is 2048. The spectral correlation threshold

is set to 0.8 and the spectral entropy threshold is set to 7 bits in the clutter suppression method. The received radar echoes of pigeons and DJI Phantom 4 UAVs, after ground clutter suppression, are shown in Figures 9a and 10a, respectively. It is evident that strong motion clutter largely obscures the target echoes. The results of clutter suppression using the FODS-SVD, FEMP-SVD, and proposed methods are presented in Figures 9b–d and 10b–d, respectively.

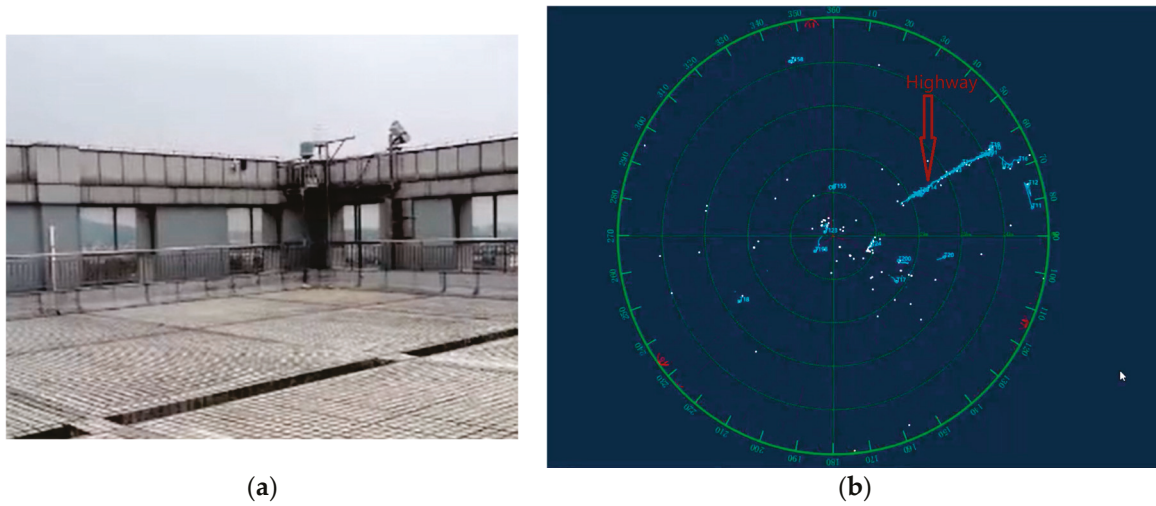


Figure 8. (a) The actual radar environment. (b) Radar beam pattern.

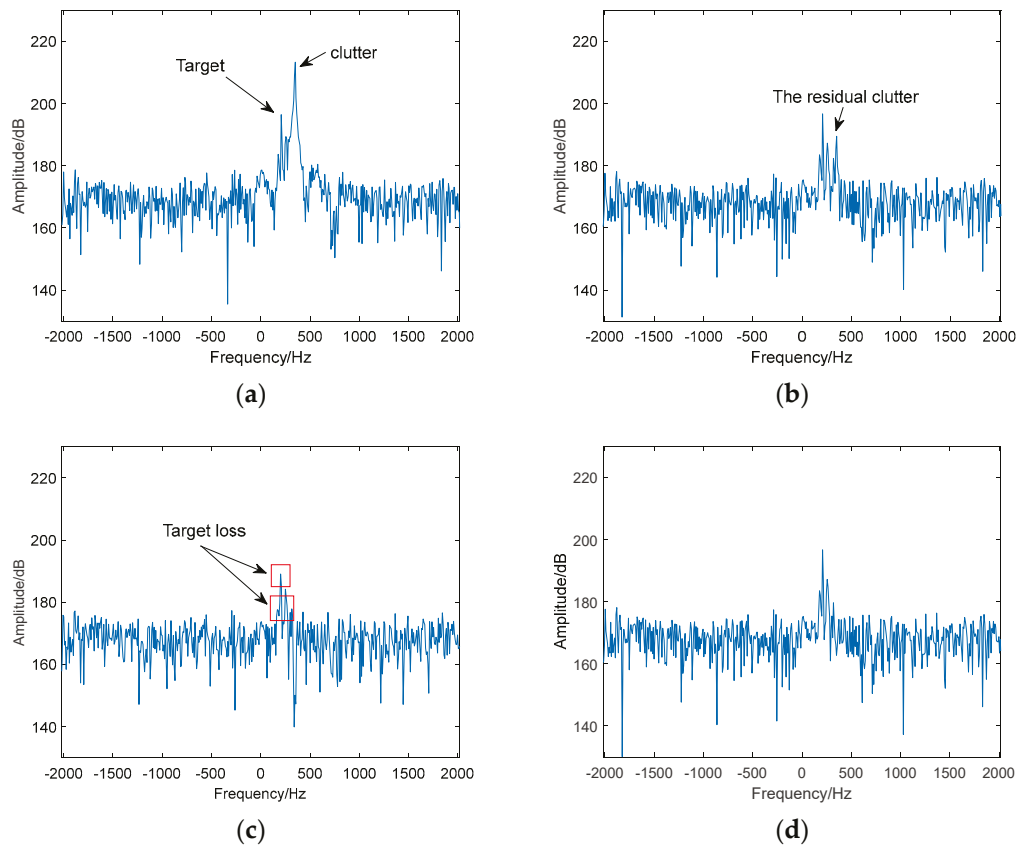


Figure 9. Clutter suppression results of the bird target. (a) The spectrum of the received signal. (b) Spectrum after clutter suppression (FODS-SVD method). (c) Spectrum after clutter suppression (FEMP-SVD method). (d) Spectrum after clutter suppression (the proposed method).

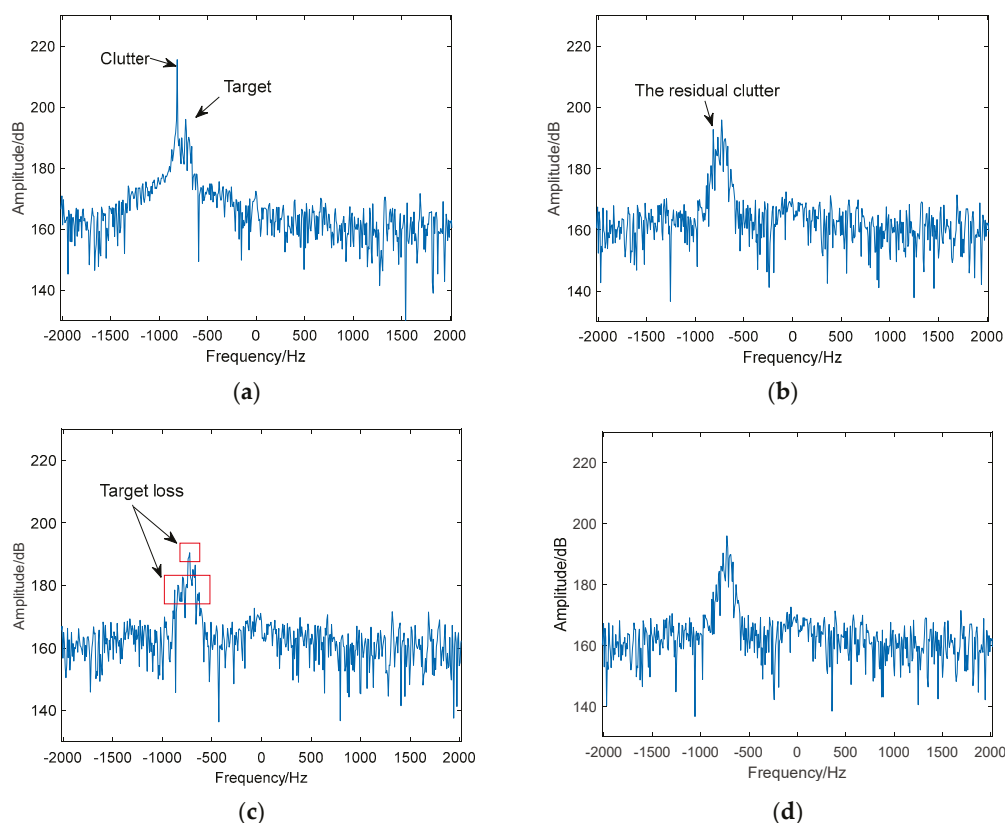


Figure 10. Clutter suppression results of the UAV target. (a) The spectrum of the received signal. (b) Spectrum after clutter suppression (FODS-SVD method). (c) Spectrum after clutter suppression (FEMP-SVD method). (d) Spectrum after clutter suppression (the proposed method).

From Figures 9b–d and 10b–d, compared to the results obtained by the FODS-SVD method, the proposed method successfully suppresses motion clutter, avoiding the existence of residual clutter caused by the situation where the singular values of clutter components are smaller than those of the target information. Compared to the results obtained by the FEMP-SVD method, the proposed method preserves relatively complete target information.

The discrimination results between birds and UAVs are presented in Figure 11. Figure 11a,b display the extracted characteristic spectrum energy entropy and the sum of normalized large eigenvalues, respectively. After fusing these two features, the kernel fuzzy c-means method is employed to identify birds and UAVs, as illustrated in Figure 10c. The discrimination accuracy using the proposed clutter suppression methods is summarized in Table 5. Misjudged samples are indicated with circles in Figure 11c. For experimental data, our method enhances discrimination accuracy by 11.68% compared to the FODS-SVD method and by 6.28% compared to the FEMP-SVD method.

Table 5. Identification accuracy of birds and UAVs.

Clutter Suppression Methods	Number of False Alarms	Total Sample Size	Identification Accuracy
The FODS-SVD method	44	240	81.7%
The FEMP-SVD method	31	240	87.1%
The proposed method	16	240	93.38%

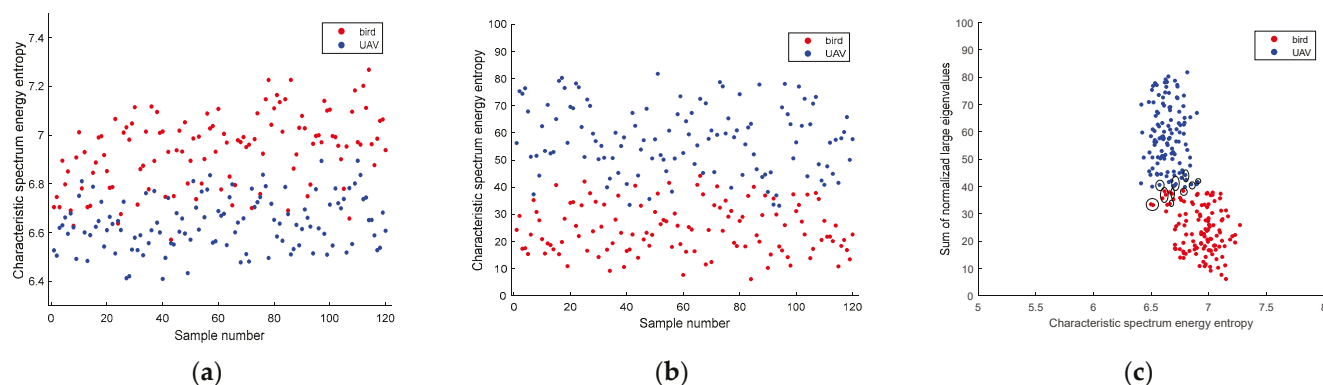


Figure 11. Identification results of birds and UAVs. (a) Characteristic spectral energy entropy. (b) Sum of normalized large eigenvalues. (c) Results obtained by kernel fuzzy c-means clustering.

5. Conclusions

In this paper, we propose a frequency correlation dual-SVD reconstruction method to mitigate the impact of strong motion clutter on distinguishing between UAVs and birds. The proposed method preserves target micro-motion information while effectively suppressing clutter. We validate its effectiveness using both simulated and experimental data. Analysis of the simulation data reveals that the proposed method exhibits smaller relative power variation between the clutter-suppressed signal and the target signal (bird or UAV) without clutter, compared to the FODS-SVD method and the FEMP-SVD method. Our approach significantly improves the discrimination accuracy of birds and UAVs under low signal-to-clutter ratios. Specifically, for experimental data, the proposed method enhances discrimination accuracy by 11.68% compared to the FODS-SVD method and by 6.28% compared to the FEMP-SVD method.

Author Contributions: Conceptualization, Y.L.; Methodology, Y.L.; Software, Y.L.; Validation, Y.L.; Formal analysis, W.H.; Investigation, X.S.; Data curation, X.S.; Writing—original draft, Y.L.; Writing—review & editing, W.H.; Supervision, W.H. All authors have read and agreed to the published version of the manuscript.

Funding: Key Research Projects of Tianjin Municipal Education Commission in China (2022ZD005).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Alhaji Musa, S.; Raja Abdullah, R.S.A.; Sali, A.; Ismail, A.; Abdul Rashid, N.E. Low-slow-small (LSS) target detection based on micro Doppler analysis in forward scattering radar geometry. *Sensors* **2019**, *19*, 3332. [CrossRef]
2. Chen, X.L.; Chen, W.S.; Rao, Y.H. Progress and prospects of radar target detection and recognition technology for flying birds and unmanned aerial vehicles. *J. Radars* **2020**, *9*, 803–827.
3. Xie, W.C.; Zhao, G.L.; Shao, Y.B. Ground moving target detection technique for airborne fire-control radar. In Proceedings of the International Congress on Image and Signal Processing, Hangzhou, China, 16–18 December 2013; pp. 1588–1592.
4. Chai, X.Y.; Zhao, Z.G.; Yue, M.J.; Liu, Y. Adaptive side-lobe cancellation technique based on CLEAN algorithm. *Radar Sci. Technol.* **2023**, *21*, 526–530.
5. Zhang, B.; Sun, L.X.; Yu, H.B. A method for improving wavelet threshold denoising in laser-induced breakdown spectroscopy. *Spectrochim. Acta Part B At. Spectrosc.* **2015**, *107*, 32–44. [CrossRef]
6. Li, H.Y.; Zhou, Y.L.; Tian, F. Wavelet-based vibration signal denoising algorithm with a new adaptive threshold function. *Chin. J. Sci. Instrum.* **2015**, *36*, 2200–2206.
7. Dutt, R.; Baloria, A.; Prasad V., R.C.; ESMP, R.; Acharyya, A. Discrete wavelet transform based unsupervised underdetermined blind source separation methodology for radar pulse deinterleaving using antenna scan pattern. *IET Radar Sonar Navig.* **2019**, *13*, 1350–1358. [CrossRef]

8. Ramakanth, S.; Soundarresan, S.; Arumuganainar, D.; Mamgain, R.; Kumaraswamy, H.V. Extraction of jet engine parameters from micro-Doppler signatures using continuous wavelet transform. *IET Radar Sonar Navig.* **2021**, *15*, 194–207. [CrossRef]
9. Feng, M. Correction of ionosphere phase contamination of high-frequency hybrid sky-surface wave radar using wavelet transform. *IET Radar Sonar Navig.* **2023**, *17*, 1017–1022. [CrossRef]
10. Liang, X.; Deng, J.; Zhang, S.; Jia, D. Improved radiation detection algorithm using wavelet packet decomposition. *IET Radar Sonar Navig.* **2020**, *14*, 547–555. [CrossRef]
11. Lin, Z.; Weiwei, Y.; Chao, Y.; Hongbing, Q. Clutter suppression method for short range slow moving target detection. *J. Electron. Inf. Technol.* **2018**, *40*, 2506–2512.
12. Zheng, J.; Liu, W.P.; Ma, N. Clutter reduction for ground penetrating radar signal. *Electron. Meas. Technol.* **2018**, *41*, 63–66.
13. Garcia-Fernandez, M. SVD-based clutter removal technique for gpr. In Proceedings of the 2017 IEEE International Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting, San Diego, CA, USA, 9–14 July 2017; pp. 2369–2370.
14. Huang, F.Q.; Zheng, L.; Yang, C.; Liu, Z.H.; Deng, X.F.; Fu, M. SVD clutter suppression method based on k-means clustering. *Radar Sci. Technol.* **2020**, *18*, 611–617.
15. Wu, L.Y.; Mao, J.; Bai, W.X. Radar small target detection based on singular value decomposition method. *J. Electron. Sci. Technol. China* **2019**, *48*, 326–330.
16. Cui, W.C.; Xu, A.Q.; Li, W.; Meng, F.L. A new method for determining effective rank order of singular value decomposition based on fitting error minimum principle. *J. Vib. Shock* **2017**, *36*, 132–137.
17. Bennett, C.; Jahangir, M.; Fioranelli, F. Use of symmetrical peak extraction in drone micro-Doppler classification for staring radar. In Proceedings of the 2020 IEEE Radar Conference, Florence, Italy, 21–25 September 2020; pp. 1–6.
18. Chen, W.S.; Liu, J.; Chen, X.L. Non-cooperative UAV target recognition in low-altitude airspace based on motion model. *J. Beijing Univ. Aeronaut. Astronaut.* **2019**, *45*, 687–694.
19. Fuhrmann, L.; Biallawons, O.; Klare, J. Micro-Doppler analysis and classification of UAVs at ka band. In Proceedings of the 2017 18th International Radar Symposium (IRS), Prague, Czech Republic, 28–30 June 2017; pp. 1–9.
20. Harmanny, R.I.A.; de Wit, J.J.M.; Premel-Cabic, G. Radar micro-Doppler mini-UAV classification using spectrograms and cepstrograms. *Int. J. Microw. Wirel. Technol.* **2015**, *7*, 469–477. [CrossRef]
21. Palama, R.; Fioranelli, F.; Ritchie, M.A. Measurements and discrimination of drones and birds with a multi-frequency multistatic radar system. *IET Radar Sonar Navig.* **2021**, *15*, 841–852. [CrossRef]
22. Molchanov, P.; Harmanny, R.; Wit, J.D. Classification of small UAVs and birds by micro-Doppler signatures. *Int. J. Microw. Wirel. Technol.* **2014**, *6*, 435–444. [CrossRef]
23. Park, S.H.; Jung, J.H.; Cha, S.B. In-depth analysis of the micro-Doppler features to discriminate drones and birds. In Proceedings of the 2020 International Conference on Electronics, Information, and Communication (ICEIC), Barcelona, Spain, 19–22 January 2020; pp. 1–3.
24. Kim, B.K.; Kang, H.S.; Park, S.O. Drone classification using convolutional neural networks with merged Doppler images. *IEEE Geosci. Remote Sens. Lett.* **2016**, *14*, 38–42. [CrossRef]
25. White, D.; Jahangir, M.; Baker, C.J.; Antoniou, M. Urban bird-drone classification with synthetic micro-Doppler spectrograms. *IEEE Trans. Radar Syst.* **2024**, *2*, 167–179. [CrossRef]
26. White, D.; Jahangir, M.; Baker, C.; Antoniou, M.; Thiyaalingam, J. Comparing optical transfer learning and autoencoder pretraining for radar bird-drone discrimination. In Proceedings of the 2024 IEEE Radar Conference (RadarConf24), Denver, CO, USA, 6–10 May 2024; pp. 1–6.
27. Oh, B.S.; Guo, X.; Lin, Z. A UAV classification system based on FMCW radar micro-Doppler signature analysis. *Expert Syst. Appl.* **2019**, *132*, 239–255. [CrossRef]
28. Rahman, S.; Robertson, D. Time-frequency analysis of millimeter-wave radar micro-Doppler data from small UAVs. In Proceedings of the IEEE 2017 Sensor Signal Processing for Defence Conference (SSPD), London, UK, 6–7 December 2017.
29. Jia, K.; Zhang, R. A second-order cone programming-based design algorithm for MTD filter banks. *Radar Electron. Warf.* **2019**, *39*, 6–9+45. [CrossRef]
30. Sim, Y.; Heo, J.; Jung, Y.; Lee, S.; Jung, Y. FPGA Implementation of Efficient CFAR Algorithm for Radar Systems. *Sensors* **2023**, *23*, 954. [CrossRef] [PubMed]
31. Wu, R.B.; Wang, J.G.; Wang, X.L.; He, W.K. A detection method for small UAVs in the strong moving clutter environment. *J. Civ. Aviat. Univ. China* **2023**, *41*, 22–29.
32. Mei, W.; Wang, J.Q. Decomposition and generalized inverse for quasi-row(column) symmetric matrix. *Int. Res. J. Pure Algebra* **2018**, *8*, 25–30.
33. Zhao, Y.S.; Lin, W.F. Research on typical scenarios based on fusion density peak value and entropy weight method of pearson's correlation coefficient. *Electr. Power* **2023**, *56*, 193–202.
34. Yu, X.S. Research on low-resolution radar target classification based on micro-motion characteristics. *J. Xidian Univ.* **2019**, *2*, 78. [CrossRef]
35. Gao, L.; Xia, X.; Min, F. An adaptive SVD fusion denoising algorithm in time and frequency domain. *J. Zhengzhou Univ.* **2023**, *55*, 48–54. [CrossRef]
36. He, W.K.; Zhang, X.; Wang, X.L.; Liu, Z.M. Refined modeling and feature analysis of birds echoes for array radar. *J. Instrum.* **2021**, *42*, 262–270.

37. Wang, X.L.; Wang, C.X.; He, W.K. Radar echoes simulation of the quadrotor drone utilizing flight control characteristics. *J. Signal Process.* **2021**, *37*, 292–303. (In Chinese)
38. Li, W.J.; Li, Z.L.; Yuan, Z.T. Sea clutter suppression and target extraction algorithm based on sparse reconstruction. *Syst. Eng. Electron.* **2022**, *44*, 777–785.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

MDPI AG
Grosspeteranlage 5
4052 Basel
Switzerland
Tel.: +41 61 683 77 34

Sensors Editorial Office
E-mail: sensors@mdpi.com
www.mdpi.com/journal/sensors



Disclaimer/Publisher's Note: The title and front matter of this reprint are at the discretion of the Guest Editors. The publisher is not responsible for their content or any associated concerns. The statements, opinions and data contained in all individual articles are solely those of the individual Editors and contributors and not of MDPI. MDPI disclaims responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



Academic Open
Access Publishing

mdpi.com

ISBN 978-3-7258-8580-0