



Article

A Deep Reinforcement Learning Method with a Low Intercept Probability in a Netted Synthetic Aperture Radar

Longhao Xie , Ziyang Cheng, Ming Li and Huiyong Li *

School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China; xlh@csj.uestc.edu.cn (L.X.); zycheng@uestc.edu.cn (Z.C.); liming1176@uestc.edu.cn (M.L.)

* Correspondence: hyl@uestc.edu.cn

Abstract

A deep reinforcement learning (DRL)-based power allocation method is proposed to achieve a low probability of intercept (LPI) in a netted synthetic aperture radar (SAR). To provide a physically meaningful and intuitive assessment of a netted radar for LPI performance, a netted circular equivalent vulnerable radius (NCEVR) is proposed and adopted. For SAR detection performance, the resolution, signal-to-noise ratio in a single pulse, and signal-to-noise ratio in SAR imaging are integrated at the task level. The LPI performance is achieved by minimizing NCEVR within the constraints of SAR detection performance. The powers in multiple moments are optimized using the DRL proximal policy optimization algorithm with the designed reward and observation. A DRL-based solver is provided for LPI radar, which handles problems that are difficult to optimize using traditional methods. The effectiveness is verified by simulations.

Keywords: deep reinforcement learning; low probability of intercept; resource allocation; synthetic aperture radar; netted radar



Academic Editors: Jingwei Xu, Weijian Liu, Junpeng Shi, Yuan Liu and Haowei Zhang

Received: 26 June 2025

Revised: 4 July 2025

Accepted: 5 July 2025

Published: 8 July 2025

Citation: Xie, L.; Cheng, Z.; Li, M.; Li, H. A Deep Reinforcement Learning Method with a Low Intercept Probability in a Netted Synthetic Aperture Radar. *Remote Sens.* **2025**, *17*, 2341. <https://doi.org/10.3390/rs17142341>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

A synthetic aperture radar (SAR) is an all-weather, high-resolution imaging system, which is widely used in military reconnaissance applications [1,2]. With the increasing complexity of the battlefield environment and the development of electronic technology, the survival of a radar in combat is seriously threatened [3]. To ensure survivability, it is necessary to develop a SAR with low probability of intercept (LPI). The LPI radar technology has been greatly developed since it was proposed [4]. A netted radar system consisting of several multiple input multiple output (MIMO) radars has more robust detection and LPI performance [5–12]. In this paper, a power allocation optimization method of a netted MIMO radar is proposed, aiming to implement LPI in SARs.

The resource of the radar is optimized by LPI techniques, resulting in good detection performance and LPI performance. Currently, the performance metric of LPI for a netted radar is still an interesting issue. Existing LPI metrics for a netted radar mainly include interception factor, interception probability, and radiated power. The interception factor for a netted radar is proposed in [13], which equates the netted radar to a single radar without considering the distribution of the radars, thus calculating the equivalent interception factor. The equivalent interception factor is minimized under the condition of the mutual information for target parameter estimation, achieving LPI by adjusting the powers of the netted radar. The interception factor of the netted radar is also used as the LPI metric

in [14], which is minimized by optimizing the power allocation and target assignment under the constraint of a minimum mean square error. As touching the interception probability for a netted radar, several papers achieve LPI by minimizing it [6,15–17]. For instance, Ref. [6] minimizes the interception probability of the radar system by optimizing the revisit interval, dwell time, and transmit power under the constraint of target tracking performance. Similarly, multiple approaches are proposed in target search and target tracking scenarios [15–17], which minimize the interception probability for a netted radar under the constraints of detection performance, thus enhancing the LPI performance. However, the interception probability for a netted radar is a function of the interceptor, which usually requires the known parameters and locations of the interceptors. It is not easy to obtain in practice. The radiated power for a netted radar is used as the LPI performance metric in several papers [18–22]. The system's radiated energy is considered an objective function to be minimized by optimizing the node selection, dwell time, bandwidth, and power while satisfying the constraint of target tracking accuracy [18,19]. The LPI performance of the radar system is enhanced while maintaining a certain level of target tracking performance. The radiated power of the radar system is considered a constraint in [20–22]. The target tracking performance is improved by optimizing the node selection, dwell time, power, and bandwidth, which realizes the LPI performance by the certain radar system power. The probability of being intercepted by an enemy reconnaissance aircraft reduces with a reduction in the radiated power due to the reduction in the propagation strength of the radar signal in space. But the LPI performance is considered indirectly; it is difficult for radiated power to achieve a quantitative and intuitive assessment of LPI. Circular equivalent vulnerable radius (CEVR) is proposed as an LPI metric for a single radar, which is calculated by the easily intercepted area [23]. It is an intuitive LPI metric, but the complexity of the calculation results in few applications. The LPI metric for a netted radar is still an interesting concern. Hence, the CEVR is extended to netted radars in this paper.

Existing LPI methods can be classified into two kinds: maximizing the current detection performance while satisfying the condition of the current LPI performance, and minimizing the current LPI performance while satisfying the condition of the current detection performance. The long-term performance at the task level has to be considered in some scenarios. In SAR, the result is imaged based on data acquired over the entire aperture time. It is necessary to consider the detection performance over the mission period as a whole. Similarly, for LPI performance, one should only consider the effect of the current action on the current situation, which does not optimize the performance over the entire task period. It can not guarantee the performance of the whole task. Therefore, it is necessary to consider the long-term performance of the whole task to ensure the overall optimization of the entire mission for LPI in SAR. Due to scheduling during the entire task in multiple moments, traditional optimization methods may no longer be applicable, which poses a challenge to the solver.

With the development of artificial intelligence technology, deep reinforcement learning (DRL) has become a novel solver for radar resource scheduling in recent years [24]. The optimal action policy is optimized using a data-driven approach, which utilizes the accumulated interaction data between the agent and the environment [25]. DRL not only considers the scheduling for the current moment but also the scheduling over the entire task period for multiple moments. At the same time, it has the characteristic of adaptability, which avoids the complex solution process of traditional optimization methods. In the single-target tracking scenario, a DRL-based resource allocation method is proposed to achieve better long-term and short-term tracking performance, which optimizes node selection and power allocation using a deep deterministic policy gradient (DDPG) method [26].

For multi-target tracking, the DRL method is also adopted to schedule the beam selection and transmit power, resulting in better multi-target tracking performance [27,28]. In the air combat scenario, to ensure the survival and the detection performance of aircraft, a DQN-based reinforcement learning method is employed to achieve real-time scheduling of sensor resources in a dynamic environment, which considers the threat level of the target, the accumulated power radiation, and the acquisition of detection information [29]. In the multi-target detection scenario, a domain knowledge-assisted DRL framework was proposed to achieve better target detection probability by allocating the power of netted MIMO radar, which utilizes the generated timely reward and final task reward to train the policy network [30]. In the integration of target detection and communication scenario, DRL is used to optimize massive MIMO systems to improve the probability of detecting weak targets in strong clutter environments by controlling the number of targets to be focused [31]. In the LPI radar scenario, Ref. [32] considered the transmit signal power and the cumulative signal-to-noise ratio (SNR) intercepted by the electronic intelligence (ELINT) system by optimizing the transmit signal power and the modulation strategy and realizing the LPI in SAR imaging using the reinforcement learning Q-learning algorithm. The LPI performance is considered by the long-term cumulative SNR intercepted by ELINT, and the SAR imaging performance is considered by the current power at each moment without long-term performance. Reinforcement learning-based resource scheduling methods have been applied in multiple scenarios; however, the research in SAR with LPI is not sufficient.

Deep reinforcement learning-based scheduling methods achieve better long-term performance through data-driven approaches, and they can be applied to problems that are difficult to optimize for traditional methods. In this paper, a reinforcement learning-based resource scheduling method is proposed to reduce the probability of a netted MIMO radar being intercepted in SAR detection. For the LPI performance, based on CEVR [23], we propose and adopt the netted circular equivalent vulnerable radius (NCEVR) to provide a physically meaningful and intuitive assessment of a netted radar. For the SAR detection performance, multiple factors of the resolution, single-pulse SNR, and SNR in SAR imaging are considered at the level of the entire long-term imaging task. The LPI performance is achieved by optimizing the power at multiple moments in the entire synthetic aperture time by minimizing the cumulative NCEVR while maintaining the SAR detection performance. The innovation of this paper is as follows: 1. A reinforcement learning-based LPI method for a netted radar is proposed in SAR detection, which is achieved by the joint control of power at multiple moments in aperture time considered at the entire task level. 2. A netted radar LPI metric, NCEVR, is proposed by considering the interactions of multiple radars in areas susceptible to interception. It achieves a physically meaningful and intuitive assessment of LPI performance for any radar distribution and number. 3. Due to the complex computation of NCEVR and the joint scheduling on multiple moments for the performance, it is difficult to solve for traditional methods. This paper provides a novel solver for the intractable problems in LPI radar scenarios.

The rest of this paper is organized as follows. In Section 2, the performance characterizations of SAR imaging and LPI are introduced, and the optimization model of NCEVR in netted radar SAR imaging is described. In Section 3, the observation, action, reward, and the proposed algorithm are designed, and the applicability of the algorithm to the optimization model is described. The proposed method is tested, and the simulations are presented in Section 4. The conclusions are presented in Section 5.

2. Problem Formulation

A netted MIMO radar is considered, which is used to achieve LPI for SAR in spotlight mode. In this section, the performance characterizations of low interception and SAR

detection are presented separately. And the optimization model of the netted LPI for a SAR is established.

2.1. Performance Characterization of SAR Detection

In this subsection, the performance characterization of a SAR is introduced, which includes the single-pulse SNR, the SNR in SAR imaging, and the range and azimuth resolution.

2.1.1. Single-Pulse SNR

For a MIMO radar, according to the radar equation [33], the single-pulse SNR can be described as follows:

$$SNR_{Pm,n,t} = \frac{P_{Tm,t} G_T G_R \lambda^2 \sigma G_{RP}}{(4\pi)^3 R_{m,n,t}^4 K T_R F_R B_R}, \quad (1)$$

where m , n , and t are the number of radar, target, and time, respectively; for example, $SNR_{Pm,n,t}$ is the SNR for a single pulse for radar m to target n at time t . $P_{Tm,t}$ is the transmitting pulse power of the radar m at time t . G_T and G_R are the gains of the transmitting and receiving antennas for all radars. λ is the signal wavelength. σ is the radar cross-section (RCS) of the target, which is a reflectivity number that normalizes per unit area [34]. G_{RP} is the pulse compression gain attributed to the range processing. $R_{m,n,t}$ is the distance between the radar m and the target n at time t . K is the Boltzmann constant. T_R , F_R , and B_R represent absolute temperature, noise figure, and bandwidth, respectively, for all radar receivers. The single-pulse SNR is a function of transmit power and distance when the radar system and the target are fixed.

2.1.2. SNR in SAR Imaging

A SAR is a high-resolution imaging radar technology that utilizes the echoes reflected by the target during the aperture time. The SNR in SAR imaging is a key performance measure that is critical for obtaining high-quality images [34,35]. The imaging process can be regarded as a two-dimensional matched filtering operation in both range and azimuth dimensions [36]. A good SNR improves signal coherence and suppresses noise, leading to the features of the targets not being lost in the noise, and thus making detection more accurate and reliable [37]. Conversely, a low SNR leads to blurry shapes of the structures and mixed pixels after filtering [38–40]. Thus, the SNR in SAR imaging is a fundamental performance metric that critically determines image quality. For the single-pulse SNR fixed at each moment, the SNR in SAR imaging of target n for radar m can be described as [34]:

$$\begin{aligned} SNR_{m,n} &= SNR_{Pm,n,t_0} G_A \\ &= SNR_{Pm,n,t_0} N_0 \end{aligned} \quad (2)$$

where SNR_{Pm,n,t_0} is used to represent the SNR over the entire measurement line. G_A is the azimuth process gain due to N_0 coherent pulse integrals. For a more general scenario, when the single-pulse SNR varies, $SNR_{m,n}$ is modified as follows:

$$SNR_{m,n} = \sum_{t=1}^T SNR_{Pm,n,t} N_t \quad (3)$$

where T is the overall number of time steps for SAR imaging. N_t is the pulse number at step t . For the netted radar, the integrated SNR after coherent integration of M radars can be expressed as follows:

$$SNR_n = \sum_{m=1}^M SNR_{m,n} \quad (4)$$

To ensure the SNR requirements are met, the system can be characterized as follows:

$$SNR_n \geq SNR_{th} \quad (5)$$

where SNR_{th} is the threshold in SAR imaging.

2.1.3. Range and Azimuth Resolution

The resolution of SAR is an important parameter of imaging capability, which affects the clarity and detail of SAR imaging. The resolution is usually divided into range resolution and azimuth resolution. The range resolution ρ_R is a function with respect to the radar bandwidth B_R , which can be described as follows [33]:

$$\rho_R = \frac{c}{2B_R} \quad (6)$$

where c is the speed of light. The azimuth resolution of SAR in spotlight mode can be described as follows [41]:

$$\rho_A = \frac{\lambda}{4 \sin(\Delta\theta/2) \cos \theta_c} \quad (7)$$

where ρ_A is the azimuth resolution. $\Delta\theta$ is the beam rotation angle during synthetic aperture time. θ_c is the beam center oblique angle. For a given target, the azimuth resolution depends on the beam angle of the start time and the end time.

In order to obtain clear images, the resolution is usually required to be better than the threshold, which can be described as follows:

$$\rho_R \leq \rho_{Rth} \quad (8)$$

$$\rho_A \leq \rho_{Ath} \quad (9)$$

Usually, the PRF of the SAR should be greater than a threshold to avoid aliasing in the azimuthal direction. The PRF can be obtained as follows [34]:

$$f_P = \frac{2kv}{D} \quad (10)$$

where k is a constant factor; typically, $k \geq 1.5$. v is the platform velocity in the horizontal and orthogonal to the target direction. D is the physical aperture dimension of the antenna.

2.2. Performance Characterization of LPI Radar

With the increasing complex electromagnetic threats, netted radar is being used in low-intercept radar systems to achieve better performance, while at the same time, the characterization of LPI in netted radars is still a topic of interest. In this paper, a new low-intercept performance metric called NCEVR is proposed for netted radars based on CEVR [23]. It achieves the evaluation of LPI performance for any radar distribution and number, which can be described as follows:

$$NCEVR_t = \left(\text{Area} \left[\left(\frac{P_{1,t}}{N_I} > SNR_{Ith} \right) \text{ or } \left(\frac{P_{12,t}}{N_I} > SNR_{Ith} \right) \text{ or } \dots \text{ or } \left(\frac{P_{1M,t}}{N_I} > SNR_{Ith} \right) \right] / \pi \right)^{\frac{1}{2}} \quad (11)$$

where $NCEVR_t$ is the netted circular equivalent vulnerable radius at time t . $P_{1m,t}$ is the power of radar m received by the interceptor at time t . N_I is the noise power of the interceptor. SNR_{Ith} is the required SNR of the interceptor satisfying a certain detection

probability p_D and a certain false alert probability p_F . Function *Area* represents the area that satisfies the condition. It can be seen that the netted radar vulnerability area is calculated, and then the equivalent radius is calculated based on the area. Specifically, according to the radar equation, $P_{Im,t}$ can be described as follows:

$$P_{Im,t} = \frac{P_{Am,t}}{(4\pi)R_{In}^2} = \frac{P_{Tm,t}T_E f_P}{(4\pi)R_{In}^2} \quad (12)$$

where $P_{Am,t} = P_{Tm,t}T_E f_P$ is the average transmitting power of radar m at time t . T_E is the effective pulse width, and f_P is the pulse repetition frequency. R_{Im} is the distance between radar m and the interceptor. N_I can be described as follows:

$$N_I = KT_I F_I B_I \quad (13)$$

where T_I , F_I , and B_I are the absolute temperature, the noise figure, and the bandwidth of the interceptor receiver, respectively. The relationship between SNR_{Ith} , p_D , and p_F can be described as follows [16]:

$$p_D = \frac{1}{2} \operatorname{erfc}\left(\sqrt{-\ln p_F} - \sqrt{SNR_{Ith} + 0.5}\right) \quad (14)$$

It should be noted that NCEVR is not a simple summation of CEVR but considers multiple radars in the vulnerability areas. It takes into account the interactions of multiple radars on the vulnerability areas, providing a physically meaningful and intuitive evaluation of LPI for a netted radar. However, the need to identify and compute the vulnerability areas when calculating NCEVR also poses significant challenges for traditional optimization methods.

2.3. Optimization Model of NCEVR in Netted Radar SAR Imaging

The coherent combining echoes from multiple pulses increased the SNR in the SAR image, but it also increased the probability of being intercepted. Therefore, it is necessary to consider low interception performance and SAR detection performance comprehensively. An optimization method is used to obtain the transmit power at each moment, minimizing the probability of interception at a certain SAR detection performance. In our paper, only signals with a single-pulse SNR greater than the threshold are considered to be echo signals, and the indicator function $I_{m,n,t}$ can be expressed as follows:

$$I_{m,n,t} = \begin{cases} 1, & \text{if } SNR_{Pm,n,t} \geq SNR_{Pth} \\ 0, & \text{else} \end{cases}, \quad (15)$$

where SNR_{Pth} is the single-pulse SNR threshold. According to (3), (4), and (15), the SNR in SAR imaging for target n until aperture time T can be modified as follows:

$$SNR_n = \sum_{m=1}^M \sum_{t=1}^T SNR_{Pm,n,t} I_{m,n,t} N_t \quad (16)$$

In the absence of the azimuthal direction aliasing, the resolution is obtained from the beam rotation angle and the beam center oblique angle at a certain PRF during the synthetic aperture time. However, in order to minimize the $NCEVR_t$, the proposed LPI netted radar system adjusts the transmit strategy at each moment, which can not guarantee the continuity of the aperture time. For example, the radar only transmits to the target at the start position and the end position, whereas it does not transmit in the span. It causes a very high sidelobe due to the lack of acquisition at the positions between the start and

end. Therefore, we calculate the resolution using the beam rotation angle and the beam center oblique angle during continuous acquisition, thus satisfying the resolution without azimuthal aliasing. For a known radar, the specific process is shown in Algorithm 1. In summary, the data that satisfy the single pulse SNR are used for SAR imaging. Then, by judging whether the data acquisition is consecutive or not, the start and end positions of consecutive acquisitions are updated, and the corresponding resolution is calculated and recorded, respectively. Finally, the resolution of each target during the whole task time is calculated.

Algorithm 1 Calculate the azimuth resolution of LPI netted radar for SAR

Input: The positions $Pos_{m,t}$ of the M radars in time T . The positions of the N targets. The single-pulse signal-to-noise ratio $SNR_{Pm,n,t}$ from M radars to N targets in time T .

Output: Azimuth resolution ρ_{An} of each target for the whole task time.

```

1: Initialize the record of azimuth resolution  $\mathbf{Rec}_{m,n,t}$  with element  $\infty$ .
2: for  $t = 1$  to  $T$  do
3:   for  $m = 1$  to  $M$  do
4:     for  $n = 1$  to  $N$  do
5:       if ( $SNR_{Pm,n,t} \geq SNR_{Pth}$  and  $SNR_{Pm,n,t+1} \leq SNR_{Pth}$ ) or  $t = 1$  then
6:         Obtain the start position during continuous acquisition  $Pos_{start} = Pos_{m,t}$ .
6:         Obtain the end position during continuous acquisition  $Pos_{end} = Pos_{m,t+1}$ .
7:         Calculate the the beam rotation angle  $\Delta\theta$  and beam center oblique angle  $\theta_c$ .
7:         Calculate the azimuth resolution  $\rho_{Am,n,t}$  from radar  $m$  to target  $n$  at time  $t$ , and
7:         store in  $\mathbf{Rec}_{m,n,t}$ .
8:       end if
9:       if  $SNR_{Pm,n,t} \geq SNR_{Pth}$  and  $SNR_{Pm,n,t+1} \geq SNR_{Pth}$  then
10:        Update the end position during continuous acquisition  $Pos_{end} = Pos_{m,t+1}$ .
11:        Calculate the the beam rotation angle  $\Delta\theta$  and beam center oblique angle  $\theta_c$ .
11:        Calculate the azimuth resolution  $\rho_{Am,n,t}$  from radar  $m$  to target  $n$  at time  $t$ , and
11:        store in  $\mathbf{Rec}_{m,n,t}$ .
12:      end if
13:    end for
14:  end for
15: end for
16: Calculate the azimuth resolution of target  $n$  at time  $t$ ;  $\rho_{An,t} = \min_m \mathbf{Rec}_{m,n,t}$ . Calculate
    the azimuth resolution of target  $n$  for the whole task time;  $\rho_{An} = \min_t \rho_{An,t} =$ 
     $\min_{m,t} \mathbf{Rec}_{m,n,t}$ .
```

In this paper, the cumulative NCEVR in the whole task time is minimized under the conditions of a single-pulse SNR, an SNR in SAR imaging, and the range and azimuth resolution, thus achieving LPI for the netted radar. It should be noted that the single-pulse SNR is included in the calculation of SNR_{In} in (16) and the calculation of ρ_{An} in Algorithm 1. The optimization model can be described as follows:

$$\begin{aligned}
 & \underset{P_{Tm,t}}{\text{minimize}} && \sum_{t=1}^T NCEVR_t \\
 & \text{s.t.} && SNR_n \geq SNR_{th}, \forall n \\
 & && \rho_{An} \leq \rho_{Ath}, \forall n, \\
 & && \rho_{Rn} \leq \rho_{Rth}, \forall n \\
 & && P_{Am,t} \leq P_{max}, \forall m, t \\
 & && P_{Am,t} \geq P_{min}, \forall m, t
 \end{aligned} \tag{17}$$

where $P_{\max}$ and $P_{\min}$ are the upper and lower limits of the average transmitting power $P_{Am,t}$, respectively.

3. Data-Driven Resource Allocation Policy Method

In this paper, a deep reinforcement learning method, proximal policy optimization (PPO) [42] is adopted to solve the optimization model. The proposed specific algorithm based on PPO and the applicability to the optimization model are described separately.

3.1. Algorithm

The observation, action, and reward of the agent are presented separately. For a certain SAR, the observation at time step t is defined as follows:

$$\mathbf{obs}_t = [\mathbf{Pos}_t, \mathbf{CR}_{At}, \mathbf{CR}_{Et}] \quad (18)$$

where $\mathbf{Pos}_t$ is the M radar positions at time t ; specifically,

$$\mathbf{Pos}_t = [Pos_{1,t}, \dots, Pos_{M,t}] \quad (19)$$

$\mathbf{CR}_{At}$ is the completion rate of the azimuth resolution for N targets at time t ; specifically,

$$\mathbf{CR}_{At} = \left[\min\left(\frac{\rho_{Ath}}{\rho_{A1,t}}, 1\right), \dots, \min\left(\frac{\rho_{Ath}}{\rho_{AN,t}}, 1\right) \right] \quad (20)$$

The highest completion rate is 1. $\rho_{An,t}$ can be obtained as shown in line 16 of Algorithm 1. $\mathbf{CR}_{St}$ is the completion rate of the SNR in imaging for N targets at time t ; specifically,

$$\mathbf{CR}_{St} = \left[\min\left(\frac{SNR_{1,t}}{SNR_{th}}, 1\right), \dots, \min\left(\frac{SNR_{N,t}}{SNR_{th}}, 1\right) \right] \quad (21)$$

Element $SNR_{n,t}$ is the SNR in SAR imaging of target n until time t for the netted radar. Similar to the SNR until aperture time T as shown in (16), $SNR_{n,t}$ can be described as follows:

$$SNR_{n,t} = \sum_{m=1}^M \sum_{i=1}^t SNR_{Pm,n,i} I_{m,n,i} N_i \quad (22)$$

The observation contains information on the positions and the completion rates. The radar adjusts the transmitting power according to the observation.

The action at time step t is defined as follows:

$$\mathbf{act}_t = [P_{T1,t}, \dots, P_{TM,t}] \quad (23)$$

It is the optimization variables that control the transmitting pulse power of M radars to achieve an LPI netted radar in SAR detection.

The reward is the core component of deep reinforcement learning, which is used to guide the agent to achieve the optimal strategy. In order to achieve a better strategy, the total reward r_t includes the primary reward r_{Pt} and auxiliary reward r_{At} . In our method, it can be described as follows:

$$r_t = r_{Pt} + r_{At} \quad (24)$$

Specifically, r_{Pt} can be described as follows:

$$r_{Pt} = -\alpha_1 NCEVR_t + \alpha_2 r_{cond1t} + \alpha_3 r_{cond2t} \quad (25)$$

where α_n is the coefficient constant of the corresponding reward. $-NCEVR_t$ is the reward to realize the LPI performance by maximizing the cumulative rewards in reinforcement learning, which corresponds to $\sum_{t=1}^T NCEVR_t$. $r_{\text{cond}1t}$ is the reward for the azimuth resolution constraint, which is defined as follows:

$$r_{\text{cond}1t} = \begin{cases} 5, & \text{if } t = T \text{ and } \forall n, \rho_{An} \leq \rho_{A\text{th}} \\ -5, & \text{if } t = T \text{ and } \exists n, \rho_{An} > \rho_{A\text{th}} \\ 0, & \text{else} \end{cases} \quad (26)$$

$r_{\text{cond}2t}$ is the reward for the SNR in SAR imaging, which is defined as follows:

$$r_{\text{cond}2t} = \begin{cases} 5, & \text{if } t = T \text{ and } \forall n, SNR_n \geq SNR_{\text{th}} \\ -5, & \text{if } t = T \text{ and } \exists n, SNR_n < SNR_{\text{th}} \\ 0, & \text{else} \end{cases} \quad (27)$$

To maximize returns, the strategy tends to satisfy the constraints with training.

For the strategy converging better, the auxiliary reward r_{At} is considered, which can be defined as follows:

$$r_{At} = \alpha_4 \Delta_{PA_t} + \alpha_5 \Delta_{PS_t} + \alpha_6 \Delta_{NA_t} + \alpha_7 \Delta_{NS_t} \quad (28)$$

where Δ_{PA_t} and Δ_{PS_t} are the improving completion rate of the azimuth resolution and SNR in imaging for the targets at time t . Δ_{NA_t} and Δ_{NS_t} are the improved satisfied number of the azimuth resolution and SNR in imaging for the targets at time t . Specifically, the improving completion rate of the azimuth resolution Δ_{PA_t} is defined as follows:

$$\Delta_{PA_t} = \sum_{n=1}^N CR_{At,n} - \sum_{n=1}^N CR_{At-1,n} \quad (29)$$

where $CR_{At,n}$ is the n -th element of $\mathbf{CR}_{At}$, which is the completion rate of the azimuth resolution for the n -th target at time t . $\sum_{i=1}^N CR_{At,n}$ is the total completion rate of the azimuth resolution at time t . Thus, Δ_{PA_t} is the improving completion rate compared to the previous moment. Similar to $\mathbf{CR}_{At}$, the improving completion rate of the SNR in imaging Δ_{PS_t} is defined as follows:

$$\Delta_{PS_t} = \sum_{n=1}^N CR_{St,n} - \sum_{n=1}^N CR_{St-1,n} \quad (30)$$

where $CR_{St,n}$ is the n -th element of $\mathbf{CR}_{St}$. The improving satisfied number of the azimuth resolution Δ_{NA_t} is defined as follows:

$$\Delta_{NA_t} = \sum_{n=1}^N 1_{\{CR_{At,n}=1\}} - \sum_{n=1}^N 1_{\{CR_{At-1,n}=1\}} \quad (31)$$

$\sum_{i=1}^N 1_{\{CR_{At,n}=1\}}$ is the total satisfied number of the azimuth resolution at time t , where the completion rate $CR_{At,n} = 1$. Thus, Δ_{NA_t} is the improvement satisfied number compared to the previous moment. Similar to Δ_{NA_t} , the improvement satisfied number of the SNR in imaging Δ_{NS_t} is defined as follows:

$$\Delta_{NS_t} = \sum_{n=1}^N 1_{\{CR_{St,n}=1\}} - \sum_{n=1}^N 1_{\{CR_{St-1,n}=1\}} \quad (32)$$

In our paper, the proximal policy optimization (PPO) method is adopted to realize the LPI netted radar for SAR imaging. The algorithm is shown in Algorithm 2. The policy network π is used for controlling radar action. Value network V is used for estimating the cumulative reward. N_{train} is the training epoch. N_{update} is the number of stochastic gradient descent (SGD) iterations in each epoch. After training, the radar action in time T can be obtained, realizing LPI in SAR imaging. The optimization results are obtained utilizing the interaction of the agent with the environment through a data-driven approach. The agent learns the optimal strategy through a trial-and-error process. Compared with traditional optimization methods, it can effectively handle complex models and avoid the process that is difficult to solve.

Algorithm 2 LPI netted radar for SAR imaging in PPO

```

1: Initialize the policy network  $\pi$  with parameter  $\psi$ , initialize the value network  $V$  with
   parameter  $\phi$ , and initialize the hyperparameter of PPO.
2: for  $n_{\text{train}} = 1$  to  $N_{\text{train}}$  do
3:   Initialize the replay buffer  $D$ . Initialize the positions of the  $N$  targets and the positions
      $Pos_{m,t}$  of the  $M$  radars in time  $T$ .
4:   for  $t$  to  $T$  do
5:     Calculate the observation  $\mathbf{obs}_t$  according to (18). Calculate action  $\mathbf{act}_t =$ 
        $\pi(\mathbf{obs}_t | \psi)$ . Calculate the reward  $r_t$  according to (24).
6:     if  $t < T$  then
7:       Calculate  $\mathbf{obs}_{t+1}$ .
8:     else
9:       Go to step 3.
10:    end if
11:    Store  $(\mathbf{obs}_t, \mathbf{act}_t, r_t, \mathbf{obs}_{t+1})$  in replay buffer  $D$ .
12:  end for
13:  for  $n_{\text{update}} = 1$  to  $N_{\text{update}}$  do
14:    Sample mini-batch of transitions from  $D$ . Update parameters  $\psi$  and parameters  $\phi$ 
      based on PPO.
15:  end for
16: end for

```

3.2. Applicability to the Optimization Model

The original optimization problem as shown in (17) can be relaxed to the following unconstrained problem by the Lagrangian relaxation method [43], which can be described as follows:

$$\begin{aligned} \underset{P_{Am,t}}{\text{minimize}} \left[\sum_{t=1}^T NCEVR_t + \sum_{n=1}^N \mu_n (SNR_{\text{th}} - SNR_n) \right. \\ \left. + \sum_{n=1}^N \beta_n (\rho_{An} - \rho_{Ath}) \right] \end{aligned} \quad (33)$$

where μ_n and β_n are the Lagrange multipliers corresponding to inequality constraints.

Based on the DRL theory, the agent learns the optimal strategy by maximizing the objective function. The objective function $J(\pi_\theta)$ is the expectation of cumulative discount reward, which can be defined as [44]:

$$J(\pi_\psi) = E_{\tau \sim \rho_\psi(\tau)} \left[\sum_{t=0}^T \gamma^t r_t \right] \quad (34)$$

where ψ is the parameters of the policy network π . τ is the trajectory observed by the agent. ρ_ψ is the probability distribution of the trajectory. $\gamma \in [0, 1]$ is the reward discount factor, which is used to denote the importance of the current reward for the future. The total reward r_t can be divided into the primary reward r_{Pt} and the auxiliary reward r_{At} .

The potential-based reward shaping is used to design the auxiliary reward r_{At} , which can accelerate the convergence of the algorithm under the optimal policy invariance [45]. Thus, the objective function of the optimal policy can be simplified to

$$J(\pi_\psi) = E_{\tau \sim \rho_\psi(\tau)} \left[\sum_{t=0}^T \gamma^t r_{Pt} \right] \quad (35)$$

According to (25) and (35) can be described as follows:

$$\begin{aligned} J(\pi_\psi) &= E_{\tau \sim \rho_\psi(\tau)} \left[\sum_{t=0}^T (-\gamma^t \alpha_1 NCEVR_t \right. \\ &\quad \left. + \gamma^t \alpha_2 r_{\text{cond}1t} + \gamma^t \alpha_3 r_{\text{cond}2t}) \right] \\ &= E_{\tau \sim \rho_\psi(\tau)} \left[\left(\sum_{t=0}^T -\gamma^t \alpha_1 NCEVR_t \right) \right. \\ &\quad \left. + \gamma^T \alpha_2 r_{\text{cond}1t} + \gamma^T \alpha_3 r_{\text{cond}2t} \right] \end{aligned} \quad (36)$$

When $\gamma = 1$, maximizing $J(\pi_\psi)$ is equivalent to minimizing the original problem with Lagrangian relaxation as shown in (33). μ_n , β_n , α_n , and the value of $r_{\text{cond}1t}$ and $r_{\text{cond}2t}$ are coefficients that correspond to each other. For example, (33) can correspond to $\alpha_1 = 1$ in (36). When $SNR_1 < SNR_{\text{th}}$, $r_{\text{cond}2t} = -5\alpha_3$; this corresponds to μ_1 being equal to a value such that $\mu_1(SNR_{\text{th}} - SNR_1) = 5\alpha_3$ and $\mu_{n \neq 1} = 0$. In summary, the optimal solution of the optimization model as shown in (17) can be solved by the proposed method. However, the algorithm is not easy to converge when $\gamma = 1$ [46], which leads to bad results. In practice, γ can be set to converge to 1, thus obtaining a suboptimal solution to the original optimization model.

4. Simulation

In this section, the proposed method's performance is verified by evaluating the simulations performed. The optimization model in (17) can be solved not only by the proposed algorithm based on PPO but also by other intelligent optimization algorithms. The particle swarm optimization algorithm [47] is used for comparison, which optimizes the transmitting power of the netted radar at all times, referred to as PSO-A. Furthermore, a conventional situation with fixed transmitting power across all times is considered and optimized using particle swarm optimization, referred to as PSO-S. The simulation setting and simulation results are introduced in turn.

4.1. Simulation Setting

As shown in Figure 1, an LPI netted radar consisting of 4 movement radars was considered, which realized SAR imaging with a high resolution. The total time of the task was 198 s with 99 time steps. The initial positions of the radar were [10 km, 10 km], [10 km, 18 km], [10 km, 22 km], and [10 km, 30 km], respectively, and the velocities were the same [100 m/s, 0 m/s]. The positions of the 10 targets were [20 km, 0 km], [25 km, 5 km], [15 km, 12 km], [20 km, 14 km], [25 km, 16 km], [20 km, 26 km], [12 km, 40 km], [15 km, 40 km], [27 km, 40 km], and [28 km, 40 km], respectively. The RCS of each target was 1. For the radars, the transmitting antenna gain G_T was 25 dB, and the receiving antenna gain G_R was 25 dB. The signal wavelength λ was 0.03 m. The pulse repetition frequency f_P was 2000 Hz, and the pulse width T_E was about 2.1×10^{-6} s. The pulse compression gain G_{RP} was 30 dB. The absolute temperature of the receiver T_r was 290 K, and the noise figure F_r was 3 dB. The single-pulse SNR threshold SNR_{Pth} was set to -30 dB. The SNR threshold in SAR imaging SNR_{th} was set to 40 dB. The azimuth resolution threshold ρ_{Ath} and range

resolution threshold ρ_{Rth} were both 0.1 m. In order to satisfy the range resolution $\rho_R \leq 0.1$ m, the bandwidth B_R was set to 1.5 GHz according to (6). The upper limit P_{max} and lower limit P_{min} of the average transmitting power were set to 1 and 0, respectively. Regarding the interceptor, the receiver temperature, noise figure, and bandwidth were $T_I = 290$ K, $F_I = 3$ dB, and $B_I = 1.5$ GHz. The required SNR of the interceptor SNR_{Ith} was set to 15 dB, whose detection probability p_D was about 0.99, with the false alert probability $p_F = 1 \times 10^{-7}$, according to (14).

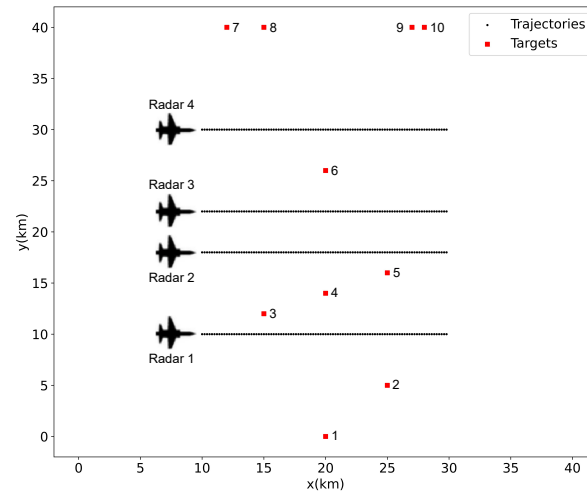


Figure 1. Target positions and radar trajectories in scenario one.

In this study, the policy network ψ and the value network V were constructed as fully connected neural networks with two hidden layers, each containing 256 neurons. The coefficient constants of the reward were set to $\alpha_1 = 2 \times 10^{-5}$, $\alpha_2 = 1$, $\alpha_3 = 1$, $\alpha_4 = 2$, $\alpha_5 = 2$, $\alpha_6 = 0.2$, and $\alpha_7 = 0.2$. The reward discount factor γ was set to 0.99. The replay buffer size D_m was 10,000. The learning rate decreased gradually from 1×10^{-5} to 1×10^{-7} . The training epoch N_{train} was 800, and the number of SGD iterations N_{update} in each epoch was 30. As shown in the reward in Figure 2, the algorithm gradually converged with the network training.

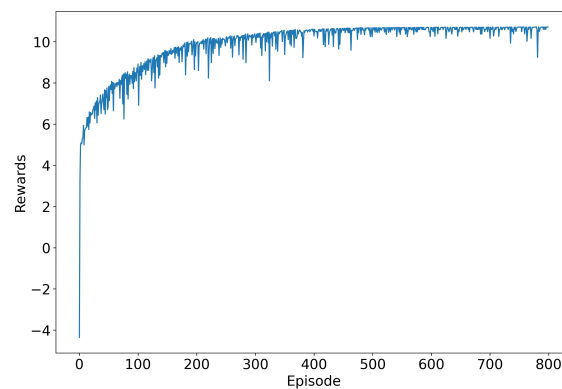


Figure 2. Reward for the training process in scenario one.

4.2. Simulation Results

In scenario one, the average transmitting power $P_{Am,t}$ of 4 radars was obtained after the policy network ψ was trained, as shown in Figure 3. The results were obtained by minimizing the total NCEVR in task time subject to constraints on detection performance. It can be seen that, compared with Radar 2 and Radar 3, the average transmitting power of Radar 1 and Radar 4 is higher, which is because the distance of the target at the top

(Target 1) and the bottom (Targets 7, 8, 9, 10) in scenario one is farther. In particular, Radar 4 has to take on more SAR imaging and transmit more power since multiple long-range targets are at the top.

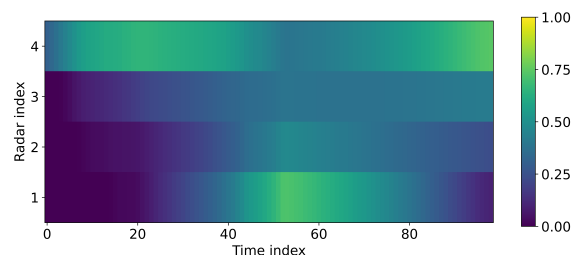


Figure 3. The average transmitting power $P_{Am,t}$ in scenario one.

The easily intercepted area corresponding to the transmitting power is shown in Figure 4. Due to space limitations, only a few moments are shown. It should be noted that the easily intercepted area is not the sum of the respective radars' areas, which should take into account the radar's positional distribution due to the overlap. The easily intercepted areas for PPO and PSO-A are presented in the first and second rows of Figure 4, respectively. While both methods optimize the transmitting power across all time steps, their computational architectures differ fundamentally, with PPO's deep reinforcement learning framework contrasting with PSO's intelligent optimization algorithms. The easily intercepted area for PSO-S is presented in the third row of Figure 4, resulting from the fixed transmitting power configuration in the conventional scenario, where transmitting powers are optimized using PSO. Since each radar's transmitting power remains fixed, the easily intercepted area remains constant across all time steps.

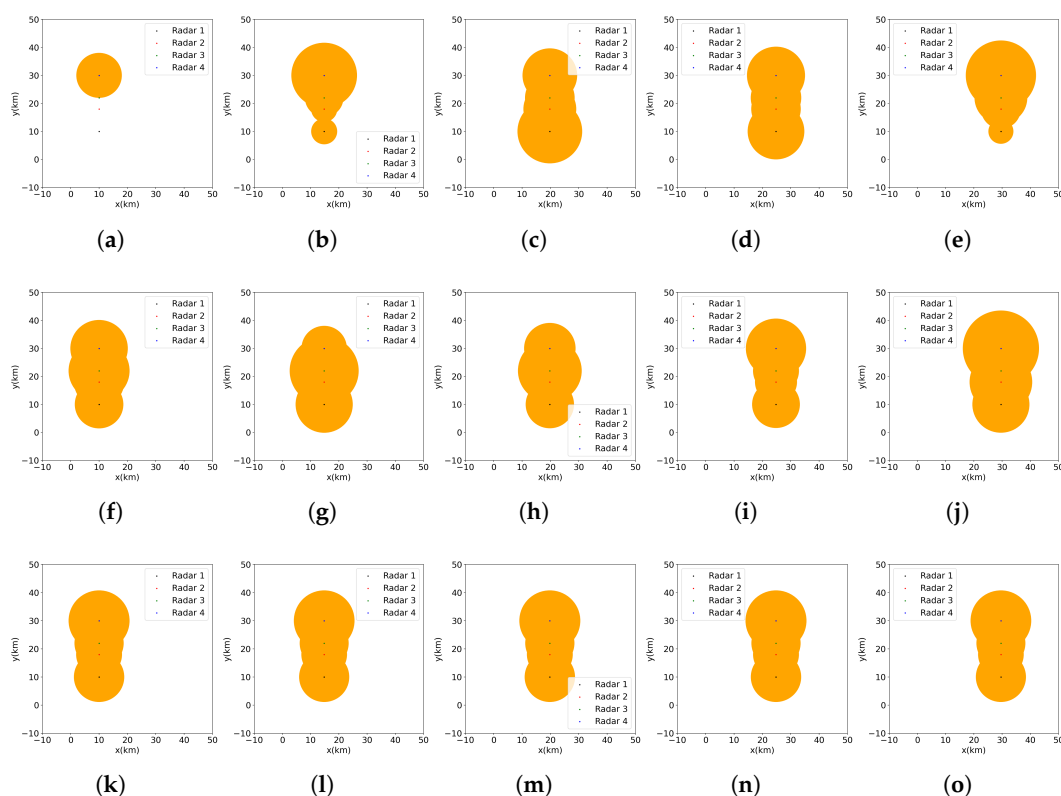


Figure 4. The easily intercepted area in scenario one. (a) Time index 1 for PPO. (b) Time index 25 for PPO. (c) Time index 50 for PPO. (d) Time index 75 for PPO. (e) Time index 99 for PPO. (f) Time index 1 for PSO-A. (g) Time index 25 for PSO-A. (h) Time index 50 for PSO-A. (i) Time index 75 for PSO-A. (j) Time index 99 for PSO-A. (k) Time index 1 for PSO-S. (l) Time index 25 for PSO-S. (m) Time index 50 for PSO-S. (n) Time index 75 for PSO-S. (o) Time index 99 for PSO-S.

1 for PSO-A. (g) Time index 25 for PSO-A. (h) Time index 50 for PSO-A. (i) Time index 75 for PSO-A. (j) Time index 99 for PSO-A. (k) Time index 1 for PSO-S. (l) Time index 25 for PSO-S. (m) Time index 50 for PSO-S. (n) Time index 75 for PSO-S. (o) Time index 99 for PSO-S.

The proposed PPO method enhances LPI performance while maintaining imaging performance. To verify its LPI performance, we compare the results of PPO, PSO-A, and PSO-S, where NCEVR computed from the easily intercepted area for a netted radar is evaluated at different time, as shown in Figure 5.

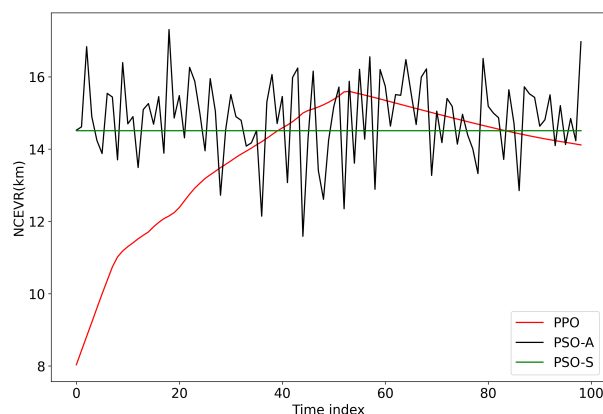


Figure 5. The NCEVR of PPO, PSO-A, and PSO-S at each time in scenario one.

For the proposed PPO-based method, the mid-mission NCEVR is larger than that of PSO-A and PSO-S, but the NCEVR of PPO is smaller at most of the other moments. The objective function is to optimize the cumulative NCEVR over the entire task time; the results are shown in Figure 6, scenario one. The average NCEVR in the task time for PPO is 13.77, which is less than 14.85 for PSO-A and 14.50 for PSO-S. It should be noted that the PSO-S method can be viewed as a special case of the more general PSO-A approach. In theory, if both algorithms converge to their global optima, PSO-A should achieve performance that is either superior or at least equivalent to that of PSO-S. However, the long-term power allocation problem, which involves optimizing transmit power for multiple radars over time, poses a significant challenge for PSO-A. Due to the high dimensionality and non-convexity of the problem, PSO-A tends to fall into local optima. Therefore, a reinforcement learning-based method, PPO, was adopted to solve the long-term optimization problem in this study. By leveraging its ability to explore and exploit high-dimensional state-action spaces effectively, the proposed PPO avoids local optima and achieves better LPI performance.

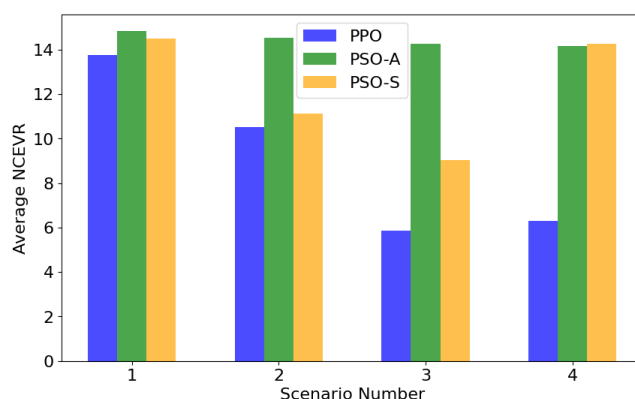


Figure 6. The average NCEVR in the task time for PPO, PSO-A, and PSO-S.

To verify the imaging performance, the SNR in SAR imaging and the azimuth resolution of 10 targets are shown in Figure 7, where the red dashed line is the threshold. The SNR in the SAR imaging of the targets SNR_n is greater than the threshold, and the azimuth resolution of the targets ρ_{An} is less than the threshold, satisfying the constraints. As for the range resolution ρ_R , it is ensured by the radar bandwidth B_R in (6).

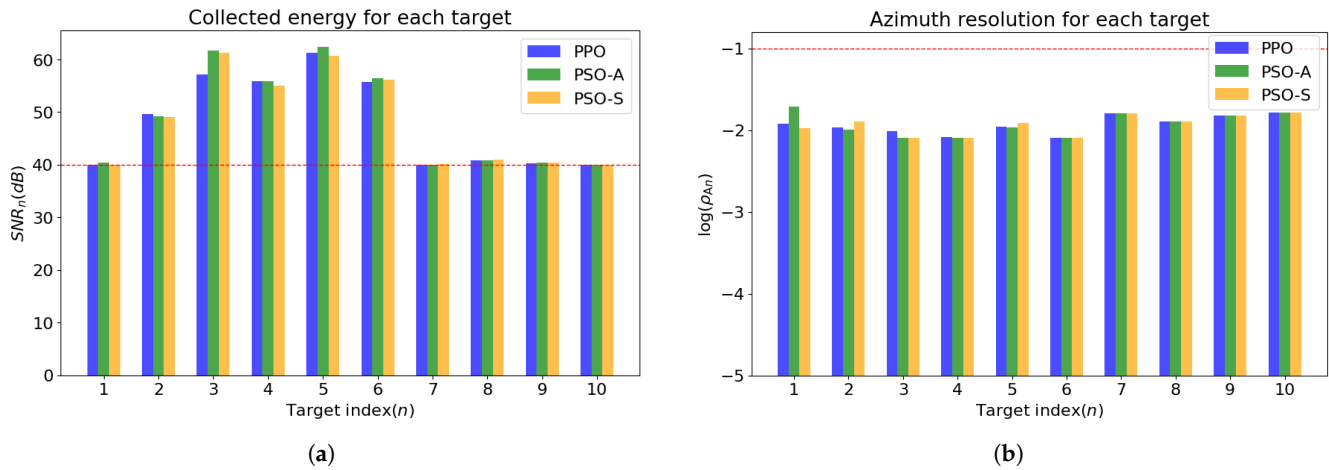


Figure 7. (a) The SNR in the SAR imaging of each target SNR_n in scenario one. (b) The azimuth resolution of each target ρ_{An} in scenario one.

To further verify the SAR imaging performance, the back projection (BP) algorithm was adopted [48]. In the BP algorithm, the target signal energy distributed across multiple echoes is concentrated at its true position through phase compensation and coherent integration, thereby reconstructing the image. Meanwhile, the noises in the echoes do not accumulate due to incoherence. Therefore, the imaging performance is represented by the sum of the SNR of multiple radars at multiple times, that is, the sum SNR after coherent accumulation in (16). The mean square error (MSE) of 10 targets in scenario one after the BP algorithm imaging is shown in Figure 8. Targets with a higher SNR generally exhibit a lower MSE.

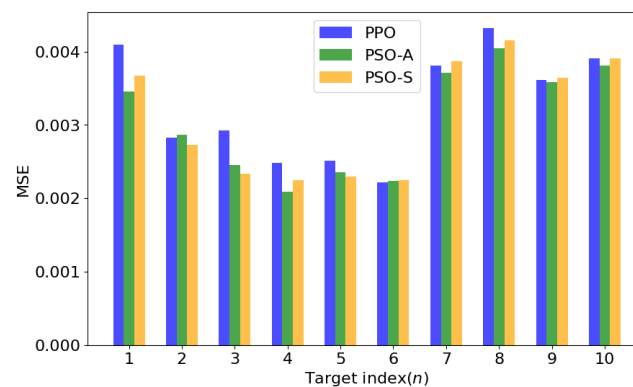


Figure 8. MSE of 10 targets for PPO, PSO-A, and PSO-S in scenario one.

Due to space limitations, part of the imaging results are shown in Figure 9, corresponding to the maximum MSE = 0.0043 (target 8) and minimum MSE = 0.0022 (target 6) in Figure 8 for the proposed method. As shown in Figure 9a, the ground truths of all targets are identical, featuring a $0.3 \text{ m} \times 0.3 \text{ m}$ square with a $0.1 \text{ m} \times 0.1 \text{ m}$ hollow centered at its midpoint. Target 8, which produced the worst imaging result (highest MSE) under the proposed PPO method, PPO, PSO-A, and PSO-S, achieved similar imaging results as shown in the second row of Figure 9, meeting the required SAR imaging performance.

Target 6, which achieved the minimum MSE among all targets under the proposed PPO method, exhibited imaging results with significantly reduced noise as shown in Figure 9e. In general, the proposed PPO method achieved satisfactory imaging of the targets while providing better LPI performance compared with PSO-A and PSO-S.

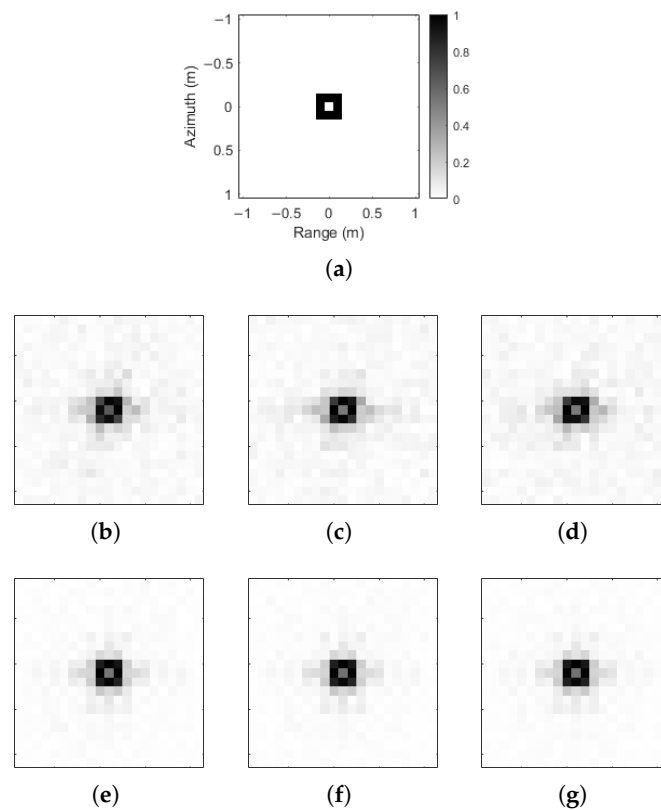


Figure 9. Imaging results in scenario one. (a) Ground truth for all targets. (b) Target 8 for PPO. (c) Target 8 for PSO-A. (d) Target 8 for PSO-S. (e) Target 6 for PPO. (f) Target 6 for PSO-A. (g) Target 6 for PSO-S.

The proposed PPO method's LPI performance improvement compared with PSO-A and PSO-S is related to the scenarios. A specific case in scenario two is considered as shown in Figure 10a.

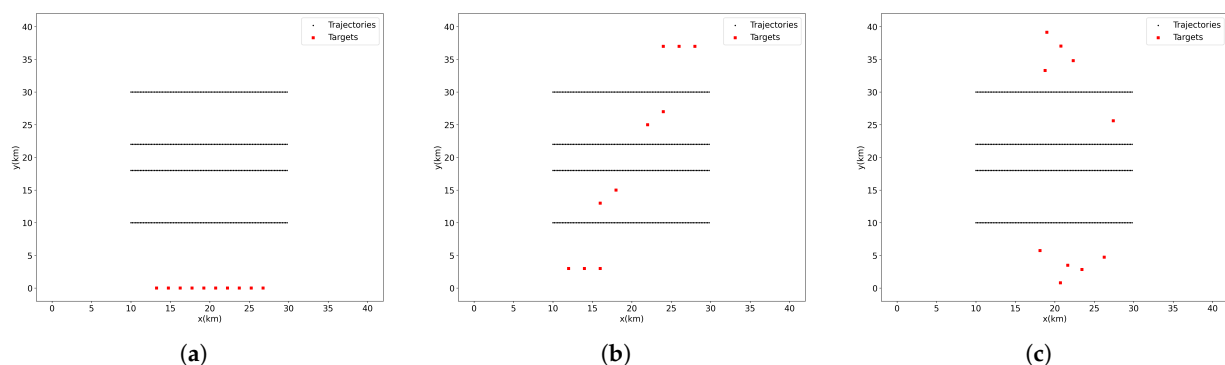


Figure 10. Target positions and radar trajectories in different scenarios. (a) Scenario two. (b) Scenario three. (c) Scenario four.

The transmitting power of the PPO method is shown in Figure 11a. Radar 1 continuously emits power toward the targets to ensure LPI and imaging performance, as it operates at a shorter distance to the targets. Since the transmitting power of PPO is similar

to the fixed power of PSO-S in conventional scenarios, PPO achieves an average NCEVR of 10.58, which is comparable to PSO-S's 11.12 as shown in Figure 6, scenario two.

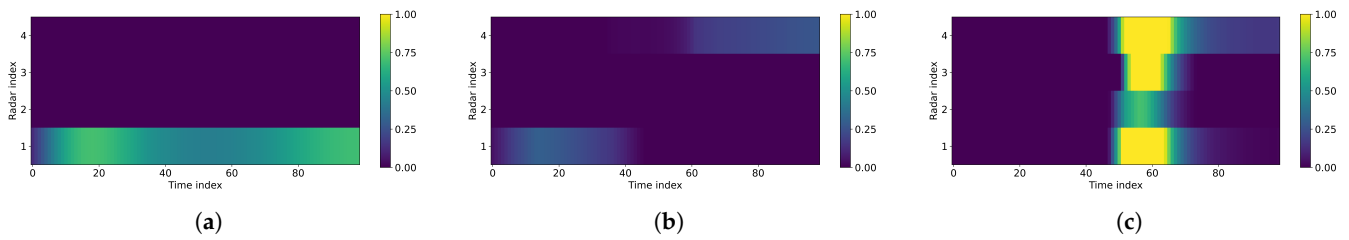


Figure 11. The average transmitting power $P_{Am,t}$ in different scenarios. (a) Scenario two. (b) Scenario three. (c) Scenario four.

As for the imaging results, due to space limitations, the targets with maximum MSE = 0.0043 for PPO are shown in the first row of Figure 12. The results are intentionally chosen to reflect the worst imaging performance of PPO. Nevertheless, PPO achieves satisfactory imaging results compared to PSO-A and PSO-S.

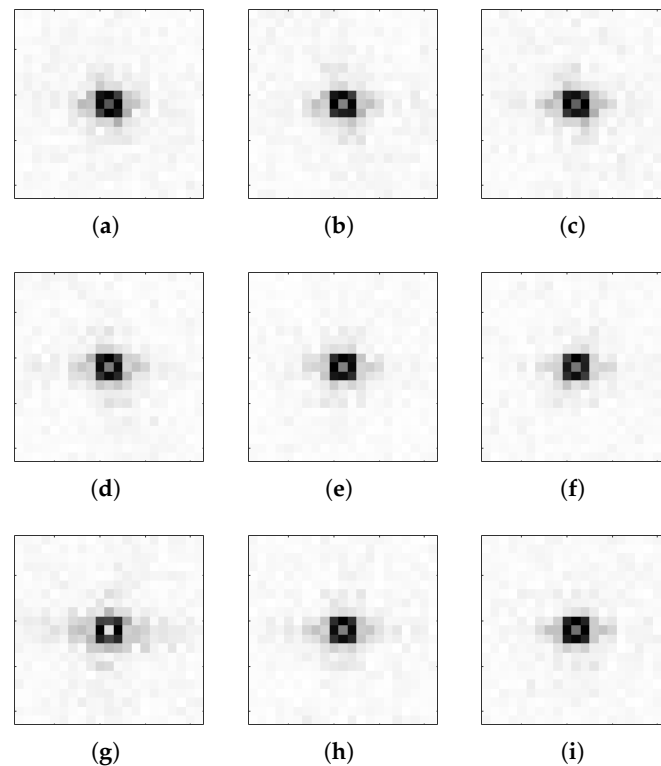


Figure 12. Imaging results of the target with a maximum MSE for PPO. (a) PPO in scenario two. (b) PSO-A in scenario two. (c) PSO-S in scenario two. (d) PPO in scenario three. (e) PSO-A in scenario three. (f) PSO-S in scenario three. (g) PPO in scenario four. (h) PSO-A in scenario four. (i) PSO-S in scenario four.

In scenario three of Figure 10b, the transmitting power of the PPO method is presented in Figure 11b. Since the five targets are located at the starting position of Radar 1, it tends to transmit power at this time. Meanwhile, since three of the five targets [12 km, 3 km], [14 km, 3 km], and [16 km, 3 km] are relatively far away from Radar 1, targets [16 km, 13 km] and [18 km, 15 km] are also adequately detected during the scanning of these three targets, eliminating the need for other radars to operate. When they move away from these targets, Radar 1 stops transmitting, ensuring that the NCEVR is minimized for the entire task. In this scenario, the conventional fixed transmitting power approach in PSO-S

is no longer suitable. As a result, PPO achieves an average NCEVR of 5.88, significantly lower than PSO-S's 9.05 as shown in Figure 6, scenario three. And PSO-A converged to a local optimum solution, resulting in 14.25. As shown in the second row of Figure 12, the target with maximum MSE = 0.0037 for PPO is selected to show its worst-case imaging performance. The results show that PPO achieves the best LPI performance with similar imaging performance.

In scenario four of Figure 10c, a randomly generated scenario is also considered to verify the proposed PPO-based method. The transmitting power is shown in Figure 11c. The imaging results of the target with maximum MSE=0.0047 for PPO are shown in the third row of Figure 12. Within SAR imaging performance constraints, PPO minimizes LPI through optimal transmit power allocation over the whole mission. As shown in Figure 6 scenario four, PPO achieves an average NCEVR of 6.30, which is less than 14.18 for PSO-A and 11.75 for PSO-S. In summary, the proposed PPO algorithm explores the optimal action using a deep reinforcement learning method. It is illustrated and validated in different scenarios, which achieves better performance compared with PSO-A and PSO-S.

5. Conclusions

In this paper, a PPO-based reinforcement learning method is proposed to achieve LPI for netted radar in SAR imaging. For the LPI performance, this paper proposes NCEVR based on CEVR, which provides a physically meaningful characterization of the LPI performance for netted radar. For SAR imaging performance, the single-pulse SNR, the SNR in SAR imaging, and the resolution ratio are considered and verified using the BP algorithm. It should be noted that the LPI performance and SAR imaging performance in this paper are the comprehensive results of radar resource scheduling at different moments during the entire time. A PPO-based method is employed to optimize the cumulative NCEVR for the whole mission, with the constraints of the SAR imaging performance. The proposed method was verified through simulation experiments, and its LPI performance was better than that of the comparison algorithm.

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

Funding: This work was supported in part by the National Natural Science Foundation of China under Grant 62231006 and in part by the Municipal Government of Quzhou under Grant 2023D040.

Data Availability Statement: Certain portions of the data are being used in ongoing related research and projects; premature disclosure may disrupt the progression of these studies or affect the outcomes of associated experiments. To ensure research integrity and orderly project advancement, the feasibility of data disclosure will be evaluated after all related research and projects are completed. For interested researchers, reasonable data requests will be assessed by the corresponding author, and the necessary data support will be provided within permissible boundaries.

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

References

1. Brenner, A.R.; Ender, J.H. Demonstration of advanced reconnaissance techniques with the airborne SAR/GMTI sensor PAMIR. *IEE Proc.-Radar Sonar Navig.* **2006**, *153*, 152–162. [[CrossRef](#)]
2. Ren, H.; Zhou, R.; Zou, L.; Tang, H. Hierarchical Distribution-Based Exemplar Replay for Incremental SAR Automatic Target Recognition. *IEEE Trans. Aerosp. Electron. Syst.* **2025**, *61*, 6576–6588. [[CrossRef](#)]

3. Butt, F.A.; Jalil, M. An overview of electronic warfare in radar systems. In Proceedings of the 2013 The International Conference on Technological Advances in Electrical, Electronics and Computer Engineering (TAECE), Konya, Turkey, 9–11 May 2013; pp. 213–217.
4. Schleher, D.C. Low probability of intercept radar. In Proceedings of the International Radar Conference, Arlington, VA, USA, 6–9 May 1985; pp. 346–349.
5. Li, J.; Stoica, P. MIMO radar with colocated antennas. *IEEE Signal Process. Mag.* **2007**, *24*, 106–114. [\[CrossRef\]](#)
6. Shi, C.; Zhou, J.; Wang, F. LPI based resource management for target tracking in distributed radar network. In Proceedings of the 2016 IEEE Radar Conference (RadarConf), Philadelphia, PA, USA, 2–6 May 2016; pp. 1–5.
7. Yi, W.; Yuan, Y.; Hoseinnezhad, R.; Kong, L. Resource scheduling for distributed multi-target tracking in netted colocated MIMO radar systems. *IEEE Trans. Signal Process.* **2020**, *68*, 1602–1617. [\[CrossRef\]](#)
8. Zhang, H.; Xie, J.; Zong, B. Bi-objective particle swarm optimization algorithm for the search and track tasks in the distributed multiple-input and multiple-output radar. *Appl. Soft Comput.* **2021**, *101*, 107000. [\[CrossRef\]](#)
9. Sun, H.; Li, M.; Zuo, L.; Zhang, P. Joint radar scheduling and beampattern design for multitarget tracking in netted colocated MIMO radar systems. *IEEE Signal Process. Lett.* **2021**, *28*, 1863–1867. [\[CrossRef\]](#)
10. Zheng, W.; Shi, J.; Li, Y.; Huang, Z.; Zhang, Z.; Li, Z. D² AF-Net: A Dual-Domain Adaptive Fusion Method for Radar Deception Jamming Recognition. *IEEE Trans. Aerosp. Electron. Syst.* **2025**, 1–15. [\[CrossRef\]](#)
11. Zhang, H.; Liu, W.; Zhang, Q.; Zhang, L.; Liu, B.; Xu, H.X. Joint Power, Bandwidth, and Subchannel Allocation in a UAV-Assisted DFRC Network. *IEEE Internet Things J.* **2025**, *12*, 11633–11651. [\[CrossRef\]](#)
12. Zhang, H.; Weijian, L.; Zhang, Q.; Taiyong, F. A robust joint frequency spectrum and power allocation strategy in a coexisting radar and communication system. *Chin. J. Aeronaut.* **2024**, *37*, 393–409. [\[CrossRef\]](#)
13. Shi, C.; Zhou, J.; Wang, F. Low probability of intercept optimization for radar network based on mutual information. In Proceedings of the 2014 IEEE China Summit & International Conference on Signal and Information Processing (ChinaSIP), Xi'an, China, 9–13 July 2014; pp. 683–687.
14. She, J.; Zhou, J.; Wang, F.; Li, H. LPI optimization framework for radar network based on minimum mean-square error estimation. *Entropy* **2017**, *19*, 397. [\[CrossRef\]](#)
15. Lu, X.; Yi, W.; Lai, Y.; Su, Y. LPI-based joint node selection and power allocation strategy for target tracking in distributed MIMO radar. In Proceedings of the 2022 25th International Conference on Information Fusion (FUSION), Linköping, Sweden, 4–7 July 2022; pp. 1–7.
16. Shi, C.; Zhou, J.; Wang, F. Adaptive resource management algorithm for target tracking in radar network based on low probability of intercept. *Multidimens. Syst. Signal Process.* **2018**, *29*, 1203–1226. [\[CrossRef\]](#)
17. Liu, D.; Wang, F.; Shi, C.; Zhang, J. LPI based optimal power and dwell time allocation for radar network system. In Proceedings of the 2016 CIE International Conference on Radar (RADAR), Guangzhou, China, 10–13 October 2016; pp. 1–5.
18. Shi, C.; Ding, L.; Wang, F.; Salous, S.; Zhou, J. Low probability of intercept-based collaborative power and bandwidth allocation strategy for multi-target tracking in distributed radar network system. *IEEE Sens. J.* **2020**, *20*, 6367–6377. [\[CrossRef\]](#)
19. Zhang, W.; Shi, C.; Zhou, J. LPI-based joint node selection and power allocation for target localization in non-coherent distributed MIMO radar with low complexity. In Proceedings of the IET International Radar Conference (IET IRC 2020), Online, 4–6 November 2020; Volume 2020, pp. 138–143.
20. Hu, J.; Zuo, L.; Varshney, P.K.; Gao, Y. Resource Allocation for Distributed Multi-Target Tracking in Radar Networks with Missing Data. *IEEE Trans. Signal Process.* **2024**, *72*, 718–734. [\[CrossRef\]](#)
21. Xie, M.; Yi, W.; Kirubarajan, T.; Kong, L. Joint node selection and power allocation strategy for multitarget tracking in decentralized radar networks. *IEEE Trans. Signal Process.* **2017**, *66*, 729–743. [\[CrossRef\]](#)
22. Su, Y.; He, Z.; Cheng, T.; Wang, J. Joint Node and Resource Scheduling Strategy for the Distributed MIMO Radar Network Target Tracking via Convex Programming. In Proceedings of the IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, 17–22 July 2022; pp. 4098–4101.
23. Wu, P.H. On sensitivity analysis of low probability of intercept (LPI) capability. In Proceedings of the MILCOM 2005-2005 IEEE Military Communications Conference, Atlantic City, NJ, USA, 17–20 October 2005; pp. 2889–2895.
24. Shyalika, C.; Silva, T.; Karunananda, A. Reinforcement learning in dynamic task scheduling: A review. *SN Comput. Sci.* **2020**, *1*, 306. [\[CrossRef\]](#)
25. Li, Y. Deep Reinforcement Learning: An Overview. *arXiv* **2017**, arXiv:1701.07274.
26. Shi, Y.; Jiu, B.; Yan, J.; Liu, H. Data-driven radar selection and power allocation method for target tracking in multiple radar system. *IEEE Sens. J.* **2021**, *21*, 19296–19306. [\[CrossRef\]](#)
27. Shi, Y.; Jiu, B.; Yan, J.; Liu, H.; Li, K. Data-driven simultaneous multibeam power allocation: When multiple targets tracking meets deep reinforcement learning. *IEEE Syst. J.* **2020**, *15*, 1264–1274. [\[CrossRef\]](#)
28. Shi, Y.; Zheng, H.; Li, K. Data-Driven Joint Beam Selection and Power Allocation for Multiple Target Tracking. *Remote Sens.* **2022**, *14*, 1674. [\[CrossRef\]](#)

29. He, J.; Wang, Y.; Liang, Y.; Hu, J.; Yan, S. Learning-based airborne sensor task assignment in unknown dynamic environments. *Eng. Appl. Artif. Intell.* **2022**, *111*, 104747. [[CrossRef](#)]
30. Wang, Y.; Liang, Y.; Zhang, H.; Gu, Y. Domain knowledge-assisted deep reinforcement learning power allocation for MIMO radar detection. *IEEE Sens. J.* **2022**, *22*, 23117–23128. [[CrossRef](#)]
31. Zhai, W.; Wang, X.; Cao, X.; Greco, M.S.; Gini, F. Reinforcement learning based dual-functional massive MIMO systems for multi-target detection and communications. *IEEE Trans. Signal Process.* **2023**, *71*, 741–755. [[CrossRef](#)]
32. Yuan, Y.; Liu, X.; Zhang, T.; Cui, G.; Kong, L. Reinforcement Learning-Enhanced Adaption of Signal Power and Modulation for LPI Radar System. *IEEE Trans. Aerosp. Electron. Syst.* **2024**, *60*, 8555–8568. [[CrossRef](#)]
33. Richards, M.A. *Fundamentals of Radar Signal Processing*; McGraw-Hill: New York, NY, USA, 2005; Volume 1.
34. Doerry, A.W. *Performance Limits for Synthetic Aperture Radar*; Technical Report; Sandia National Laboratories (SNL): Albuquerque, NM, USA; Livermore, CA, USA, 2006.
35. Van Zyl, J.J. *Synthetic Aperture Radar Polarimetry*; John Wiley & Sons: Hoboken, NJ, USA, 2011.
36. Wei, J.; Li, Y.; Yang, R.; Li, L.; Guo, L. Method of high signal-to-noise ratio and wide swath SAR imaging based on continuous pulse coding. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 2185–2196. [[CrossRef](#)]
37. Makas, E.; Aslan, A.R. Spaceborne SAR System Design Considerations: Minimizing Satellite Size and Mass, System Parameter Trade-Offs, and Optimization. *Appl. Sci.* **2024**, *14*, 9661. [[CrossRef](#)]
38. Jiang, M.; Guarnieri, A.M. Distributed scatterer interferometry with the refinement of spatiotemporal coherence. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 3977–3987. [[CrossRef](#)]
39. Jiang, M.; Hooper, A.; Tian, X.; Xu, J.; Chen, S.N.; Ma, Z.F.; Cheng, X. Delineation of built-up land change from SAR stack by analysing the coefficient of variation. *ISPRS J. Photogramm. Remote Sens.* **2020**, *169*, 93–108. [[CrossRef](#)]
40. Zhao, Y.; Jiang, M. Integration of optical and SAR imagery for dual PolSAR features optimization and land cover mapping. *IEEE J. Miniaturization Air Space Syst.* **2022**, *3*, 67–76. [[CrossRef](#)]
41. Liu, H.-Y.; Song, H.-J.; Cheng, Z.-J. Comparative study on stripmap mode, spotlight mode, and sliding spotlight mode. *J. Univ. Chin. Acad. Sci.* **2011**, *28*, 410.
42. Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; Klimov, O. Proximal policy optimization algorithms. *arXiv* **2017**, arXiv:1707.06347.
43. Fisher, M.L. The Lagrangian relaxation method for solving integer programming problems. *Manag. Sci.* **1981**, *27*, 1–18. [[CrossRef](#)]
44. Bilgin, E. *Mastering Reinforcement Learning with Python: Build Next-Generation, Self-Learning Models Using Reinforcement Learning Techniques and Best Practices*; Packt Publishing Ltd.: Birmingham, UK, 2020.
45. Ng, A.Y.; Harada, D.; Russell, S. Policy invariance under reward transformations: Theory and application to reward shaping. In Proceedings of the ICML, Bled, Slovenia, 27–30 June 1999; Volume 99, pp. 278–287.
46. Kim, M.; Kim, J.S.; Choi, M.S.; Park, J.H. Adaptive discount factor for deep reinforcement learning in continuing tasks with uncertainty. *Sensors* **2022**, *22*, 7266. [[CrossRef](#)] [[PubMed](#)]
47. Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceedings of the ICNN'95-International Conference on Neural Networks, Perth, WA, Australia, 27 November–1 December 1995; Volume 4, pp. 1942–1948.
48. Desai, M.D.; Jenkins, W.K. Convolution backprojection image reconstruction for spotlight mode synthetic aperture radar. *IEEE Trans. Image Process.* **1992**, *1*, 505–517. [[CrossRef](#)] [[PubMed](#)]

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.