atmosphere

MDPI

Article

SPFANet: A Spatial Polarization Feature Aggregation Network for Radar-Based Hydrometeor Classification

Jianlin Guo [1], Xu Zhang [2], Zhiying Lu [1,*] and Gang Wang [1,3]

[1] School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China; guojianlin0317@tju.edu.cn (J.G.); wanggg@tju.edu.cn (G.W.)
[2] State Grid Tianjin Electric Power Company Chengnan Power Supply Branch, Tianjin 300010, China; gelk@tju.edu.cn
[3] CETC Key Laboratory of Aerospace Information Applications, The 54th Research Institute of China Electronics Technology Group Corporation, Shijiazhuang 050081, China
* Correspondence: luzy@tju.edu.cn

Abstract

Hydrometeor classification from dual-polarization weather radar is important for precipitation monitoring, cloud microphysical analysis, and severe weather warnings. However, operational performance remains limited by non-meteorological echoes, incomplete use of the relationships among polarimetric variables, and severe class imbalance. To address these issues, we propose a Spatial Polarization Feature Aggregation Network (SPFANet) for hydrometeor classification using dual-polarization weather radar. The proposed method combines a non-meteorological echo removal procedure and a multi-feature radar image constructed from reflectivity, differential reflectivity, and the correlation coefficient. SPFANet is designed as a semantic segmentation model that integrates multi-level semantic representations, aggregates multiscale spatial features, and enhances latent polarimetric features. Experiments using KTLX (USA) observations show a high level of agreement between SPFANet predictions and the NOAA reference labels across different precipitation scenarios. Across five independent runs, SPFANet achieves mean *Recall* values above 81% for all five categories. It also obtains the highest mean *F*1 and *IoU* for ice crystals, graupel, and hail, as well as the highest mean *Precision* for hail. The overall evaluation metrics further demonstrate the effectiveness of SPFANet in hydrometeor classification. These findings highlight the potential of the proposed model for radar-based precipitation analysis and severe weather monitoring.

Keywords: radar applications; dual-polarization weather radar; hydrometeor classification; semantic segmentation; transformer

Check for updates

Academic Editors: Eiichi Yoshikawa and Sounak Biswas

Received: 15 September 2026
Revised: 4 October 2026
Accepted: 13 October 2026
Published: 16 October 2026

Copyright: © 2026 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.

1. Introduction

Weather radar is an important instrument for atmospheric observation, enabling continuous monitoring of clouds, precipitation, and their evolution [1,2]. Hydrometeor classification is a central task in weather radar applications because it provides direct information on the microphysical composition of precipitating systems [3,4]. Accurate identification of hydrometeor types supports precipitation monitoring, severe weather analysis, aviation safety, agricultural operations, and water-resource management [5–7]. Before dual-polarization weather radar became widely available, hydrometeor identification depended mainly on sounding-assisted analysis [8]. In addition, it relied on bright band or melting layer detection near the freezing level [9]. Although these approaches provided

useful guidance, they offered only limited information on the fine-scale distribution and evolution of hydrometeor types.

Dual-polarization weather radar has substantially strengthened the observational basis for hydrometeor classification [10]. It simultaneously transmits and receives horizontally and vertically polarized signals, thereby providing multiple polarimetric variables that describe the physical characteristics of precipitation particles more effectively than conventional radar measurements [11]. These measurements have made dual-polarization radar an important tool for distinguishing hydrometeor categories and analyzing precipitation microphysical processes.

Despite these advances, accurate hydrometeor classification remains difficult in operational settings. Non-meteorological echoes, including ground clutter and biological echoes, can contaminate radar observations and reduce classification reliability [1]. Existing methods also do not always make full use of the physical relationships among polarimetric variables [12,13]. Radar datasets are often dominated by frequently occurring categories. Operationally important classes such as graupel and hail occupy only a small fraction of labeled pixels. This imbalance can bias model training toward majority classes and increase the risk of missed detections for minority classes. Classification is further complicated in mixed-precipitation and phase-transition regions, where different hydrometeor types often share similar polarimetric signatures, making reliable discrimination even more difficult [10].

To address these challenges, this study proposes a hydrometeor classification method that combines radar data processing with the Spatial Polarization Feature Aggregation Network (SPFANet). The model integrates multi-level semantic representations, aggregates multiscale spatial features, and enhances latent polarimetric features while reducing the influence of class imbalance. The main contributions of this paper are summarized as follows:

- In the dual-polarization radar data processing stage, a cascaded removal method for non-meteorological echoes is proposed to reduce interference from clutter and biological echoes, thereby improving data quality and classification accuracy;
- To address the limited representation of relationships among polarimetric variables, a multi-feature radar image synthesis strategy is proposed to achieve a joint representation of radar physical information;
- A hydrometeor classification model based on SPFANet is proposed. The model integrates multi-level semantic representations and captures multiscale spatial structures while enhancing latent polarimetric features. A class-weighting strategy is further designed to reduce the influence of class imbalance.

This paper is organized as follows: Section 1 introduces the research background and motivation. Section 2 reviews the related work. Section 3 describes the dual-polarization radar data, the non-meteorological echo removal method, and the multi-feature radar image synthesis strategy. Section 4 presents the hydrometeor classification model based on SPFANet. Section 5 presents the experimental setup and analyzes the results. Section 6 discusses the findings and limitations of the proposed framework. Section 7 concludes the paper.

2. Related Work

Hydrometeor classification has been widely studied in weather radar applications. In this section, we briefly review the research most relevant to the methods used in this paper. Existing hydrometeor classification methods can be grouped into two broad categories: fuzzy logic methods and neural network methods.

2.1. Hydrometeor Classification Based on Fuzzy Logic

Fuzzy logic methods [12,13] usually combine radar variables with empirically designed membership functions to distinguish among precipitation particle types. These methods are relatively interpretable and can also handle overlapping class boundaries reasonably well. However, methods of this kind depend heavily on manually designed rules and membership functions [14]. As a result, their generalization ability is limited, and they are less effective at capturing complex nonlinear relationships in high dimensional feature spaces.

Some studies have also attempted to combine fuzzy logic with data-driven methods. For example, Wang [15] proposed a hydrometeor classification method that combines deep learning with a fuzzy-logic algorithm and adopts a two-stage strategy. In the first stage, the model is used to classify precipitation patterns. In the second stage, the classification results are combined with a fuzzy-logic algorithm for hydrometeor classification. These studies indicate that meteorological research has begun to explore data driven solutions, and that the focus has gradually shifted toward deep learning methods.

2.2. Hydrometeor Classification Based on Deep Learning

With the rapid development of deep learning, hydrometeor classification has increasingly been treated as a semantic segmentation problem in which each pixel is assigned a class label. General semantic segmentation models built on Convolutional Neural Networks (CNNs), including the Fully Convolutional Network (FCN) [16], U-Net [17], SegNet [18], and the DeepLab family [19–22], have been widely used in remote sensing image semantic segmentation. For hydrometeor classification, Lu and Kumar [23] investigated deep learning-based hydrometeor classification. They evaluated five CNN models for identifying four common hydrometeor categories under class imbalanced conditions. They used the NEXRAD hydrometeor classification product as the training label, providing a practical training strategy for deep learning-based hydrometeor classification. Even so, conventional CNN semantic segmentation models still struggle to capture global context and multiscale information effectively [24], which limits their performance in complex precipitation scenes. This limitation is particularly evident in cloud and precipitation systems, where spatial structures are complex and hydrometeor phase changes are frequent.

Transformer models are a promising alternative because attention mechanisms can model long-range dependencies and capture global structure and texture patterns more effectively [25–27]. Xu [28] used a Transformer encoder to extract features for the classification of five hydrometeor categories. Although hydrometeor classification with Transformer models is still at an early stage, related studies in meteorology have reported promising results in cloud classification [29,30] and adverse weather classification [31,32]. Given the similarities among these tasks in both data characteristics and physical processes, the application of Transformer models to hydrometeor classification deserves further investigation. From a methodological perspective, this trend reflects a shift away from models that rely mainly on local receptive fields, toward approaches that can capture both global dependencies and multiscale features.

Despite this progress, several issues remain unresolved. Non-meteorological echoes can still degrade the quality of dual-polarization radar data and undermine classification reliability. Existing methods also make limited use of the physical relationships among multiple polarimetric variables. In addition, severe class imbalance continues to hinder the recognition of meteorologically important categories such as graupel and hail. These unresolved issues motivate the work presented in this paper.

3. Data Preparation and Preprocessing

This section describes the data preparation and preprocessing procedure used to construct high-quality input data for hydrometeor classification from dual-polarization weather radar observations. It consists of three parts. Section 3.1 introduces the dual-polarization radar data used in this study and their basic characteristics. Section 3.2 presents a non-meteorological echo removal method to reduce interference from clutter and biological echoes in the subsequent classification task. Section 3.3 then proposes a multi-feature radar image synthesis strategy that combines multiple dual-polarization radar products into a unified representation suitable for network input.

3.1. Dual-Polarization Radar Data

Weather radar transmits electromagnetic waves into the atmosphere and receives the backscattered signals from atmospheric particles. By analyzing these returned signals, the radar can retrieve information about atmospheric targets. Conventional weather radar generally uses a single polarization state, whereas dual-polarization weather radar transmits and receives both horizontally and vertically polarized waves. Different hydrometeors scatter horizontally and vertically polarized waves in different ways. As a result, dual-polarization weather radar can provide conventional variables such as reflectivity, radial velocity, and spectrum width. It can also provide polarimetric variables such as differential reflectivity, specific differential phase, and the correlation coefficient. These variables provide information on hydrometeor shape and size, thereby supporting more accurate precipitation classification and analysis.

We use dual-polarization NEXRAD observations from the KTLX radar in Oklahoma, USA, provided by the National Oceanic and Atmospheric Administration (NOAA). The NOAA NEXRAD Level-III products were exported as gridded radar data using the official NOAA Weather and Climate Toolkit (WCT). The 0.5° elevation angle was selected to maintain spatial consistency among radar variables. The products used in this study were collected between April and September 2021, with a nominal detection radius of 162 nautical miles. Within this coverage, the study area spans 31.92–38.64 degrees N and 93.93–100.65 degrees W. During data export, the spatial size of the gridded radar data was set to 256 × 256 using WCT. This setting provides a practical balance between preserving the main radar echo structures and controlling computational cost. The dataset includes reflectivity (Z_H, dBZ), differential reflectivity (Z_{DR}, dB), specific differential phase (K_{DP}, °/km), the correlation coefficient (ρ_{hv}, dimensionless), and the NOAA hydrometeor classification product.

Based on the polarimetric observations, the NOAA operational Hydrometeor Classification Algorithm (HCA) assigns each radar cell to the dominant hydrometeor category inferred for that cell. This product describes the dominant hydrometeor category aloft within the radar coverage area, rather than the precipitation type measured at the surface. Because direct pixel-level labels for cloud hydrometeor categories are difficult to obtain at the same observation time and at the radar-cell scale, this product was used as the NOAA reference label for model training and testing, following a strategy used in previous hydrometeor classification studies [23]. The NOAA reference labels are not fully independent of the model inputs because the NOAA HCA product is derived from closely related polarimetric radar variables.

However, physically similar hydrometeor categories may exhibit overlapping radar signatures, which can increase uncertainty in the NOAA reference labels. This uncertainty may be greater in transition areas, where multiple hydrometeor types may be present in the same radar cell, making it difficult for a single NOAA reference label to fully represent the hydrometeor composition within that cell. When the NOAA reference labels are used

for evaluation, a correct model prediction may still be counted as a classification error if the corresponding reference label is incorrect. The reported quantitative metrics may therefore reflect both model classification errors and uncertainty in the NOAA reference labels.

3.2. Cascaded Removal of Non-Meteorological Echoes

Removing non-meteorological echoes is a critical step in hydrometeor classification. Radar observations contain not only backscatter from hydrometeors but also non-meteorological signals, such as clutter and biological echoes, which can substantially degrade the performance of downstream classification models. Threshold screening of radar variables is an efficient and commonly used approach for non-meteorological echo removal. To remove such echoes, this study proposes a non-meteorological echo removal method that combines radar data processing with meteorological feature screening. The overall workflow is shown in Figure 1. The default WCT color tables and value intervals used to visualize radar variables are provided in Figure A1 in Appendix A.

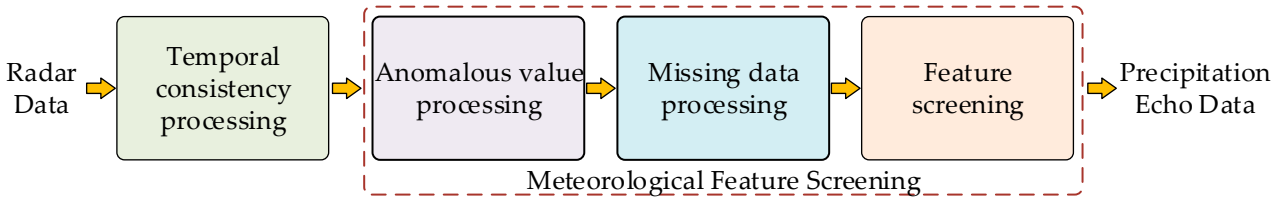


Figure 1. Flowchart of non-meteorological echo removal.

Temporal consistency processing is first performed across the radar variables. Any sample with missing observations in one or more radar variables is discarded to ensure that all variables are available at the same observation time.

After temporal consistency processing, meteorological feature screening is applied based on the radar variables. Pixels with reflectivity lower than 0 dBZ are removed. Pixels with missing values in any product are also removed synchronously from all five radar products. For pixels with reflectivity lower than 5 dBZ, an eight-neighbor connectivity check is performed. If at least one neighboring pixel also has reflectivity lower than 5 dBZ, the pixel is regarded as a non-outlier and retained. Otherwise, it is treated as an isolated outlier and removed synchronously from all radar products. Because specific differential phase is generally positive for horizontally oriented raindrops [33], pixels with specific differential phase lower than 0 °/km are removed from the corresponding locations in all radar products. After screening, the removed pixel locations are masked as background in all five radar products. This cascaded procedure reduces contamination from non-meteorological echoes while preserving the dominant meteorological structures as much as possible. It therefore provides cleaner and more consistent inputs for subsequent model training. To evaluate the proposed method, two representative cases were analyzed: a non-meteorological echo-dominated case and a precipitation-dominated case, as shown in Figures 2 and 3.

The visual results show that the proposed cascaded strategy remains effective under different echo compositions. In the non-meteorological echo-dominated case, the unwanted echoes occupy a relatively large spatial extent and are removed more completely. In the precipitation-dominated case, the non-meteorological echoes are more scattered and embedded in a more complex background, yet the proposed method still achieves reliable suppression.

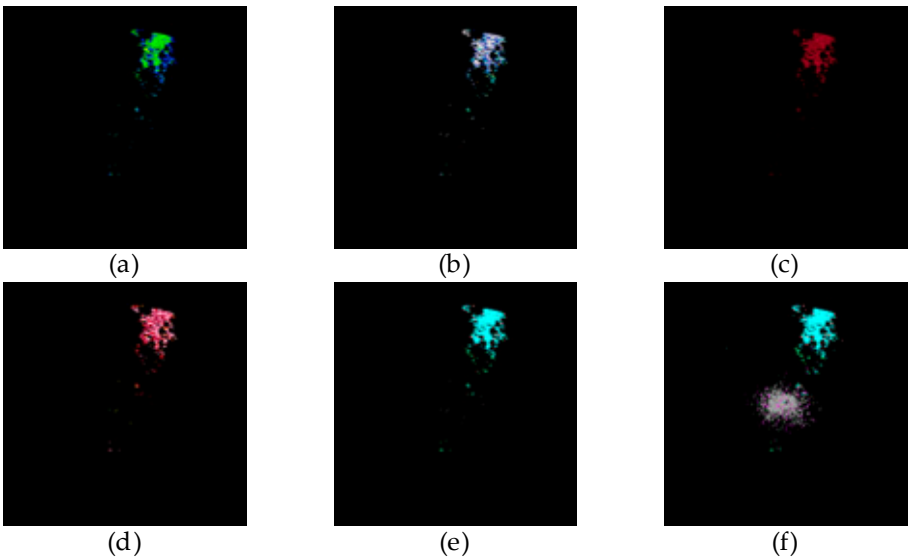


Figure 2. Example of non-meteorological echo removal in a non-meteorological echo-dominated scene at 17:22 UTC on 6 April 2021: (**a**) reflectivity after removal; (**b**) differential reflectivity after removal; (**c**) specific differential phase after removal; (**d**) correlation coefficient after removal; (**e**) HCA product after removal; (**f**) original HCA product.

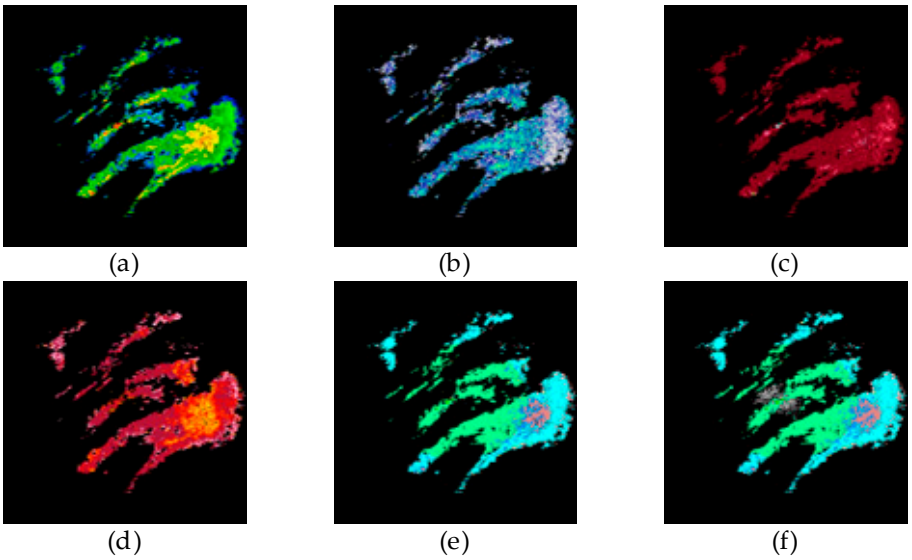


Figure 3. Example of non-meteorological echo removal in a precipitation-dominated scene at 18:25 UTC on 28 April 2021: (**a**) reflectivity after removal; (**b**) differential reflectivity after removal; (**c**) specific differential phase after removal; (**d**) correlation coefficient after removal; (**e**) HCA product after removal; (**f**) original HCA product.

For each category, pixels labeled as that category in the NOAA HCA product were used as reference pixels. *Precision* was calculated as the proportion of correctly removed non-meteorological reference pixels among all pixels removed after preprocessing. *Recall* was calculated as the proportion of correctly removed non-meteorological reference pixels among all non-meteorological reference pixels. The *False-removal rate* was calculated as the proportion of incorrectly removed meteorological reference pixels among all meteorological reference pixels. The quantitative evaluation was conducted using 21,431 KTLX samples collected from 1 April to 30 September 2021. These samples were divided into non-meteorological echo-dominated and precipitation-dominated cases. The *Precision*, *Recall*, and *False-removal rates* for the two scenarios are summarized in Table 1.

Table 1. Evaluation metrics of non-meteorological echo removal.

Metric	Non-Meteorological Echo-Dominated	Precipitation-Dominated
Precision	99.4%	97.6%
Recall	98.2%	94.9%
False-removal rate	0.13%	0.20%

The results show that the proposed method achieves high *Precision* and *Recall* while maintaining a consistently low *False-removal rate* in both scenarios. The incorrectly removed meteorological pixels are mainly located in light/moderate rain and dry snow regions adjacent to non-meteorological echoes, while the main meteorological structures are generally preserved. The role of non-meteorological echo removal in radar input processing is further evaluated in Section 5.3.

3.3. Multi-Feature Radar Image Synthesis

Dual-polarization radar provides multiple variables for hydrometeor classification, but these variables differ substantially in numerical range, statistical distribution, and classification relevance. Directly using multiple radar variables as network inputs may weaken their complementary information because they have different scales and physical meanings. To preserve the complementary information among radar variables while maintaining physical interpretability, this study constructs a unified input representation in the form of a multi-feature radar image.

Different radar variables characterize distinct physical properties of hydrometeors. Reflectivity represents radar echo intensity. Differential reflectivity provides information on particle shape and size. Specific differential phase is related to liquid water content and rainfall intensity. The correlation coefficient indicates the uniformity of hydrometeor populations.

Reflectivity, differential reflectivity, specific differential phase, and the correlation coefficient were considered as candidate variables for multi-feature radar image synthesis. Experimental comparisons were conducted to select the final variable combination, which consisted of reflectivity, differential reflectivity, and the correlation coefficient. The combination of these three variables provides complementary information on radar echo intensity, particle shape and size, and the uniformity of hydrometeor populations. This complementary information helps distinguish rain from ice-phase hydrometeors and differentiate among ice crystals, snow, graupel, and hail. The three selected variables are mapped to the R, G, and B channels, respectively, so that their combined characteristics can be incorporated into a single network input.

Direct fusion of the raw values can degrade model performance because the three selected variables differ markedly in numerical scale. To reduce these scale differences, a non-uniform piecewise mapping scheme is applied before image synthesis. This mapping transforms radar variables with different value ranges into a common intensity domain, thereby providing a unified image representation for network input while retaining the principal physical intervals relevant to hydrometeor classification.

For each selected radar variable, the valid value range after non-meteorological echo removal is divided into several intervals according to its physical meaning and the NOAA interval definitions. These intervals preserve the main meteorological meaning of the selected variables. For reflectivity, the intervals correspond to different echo intensity levels. For differential reflectivity, the intervals capture variations associated with particle shape and size. For the correlation coefficient, the intervals indicate the uniformity of hydrometeor populations. Because the physical attributes and statistical distributions differ among variables, both the value range and the interval partitioning are designed separately

for each variable and remain intentionally non-uniform. According to the NOAA product interval definitions, each interval is equally divided into 11 sub-intervals. The mapped values are then assigned to the range of 0–255 by Equation (1).

$$value = \left\lceil 255 \times [\frac{1}{n}(a-1) + \frac{1}{11n}b] \right\rceil \quad (1)$$

where a denotes the index of the interval containing the input value, b denotes the index of the corresponding sub-interval and ranges from 0 to 10, n denotes the total number of intervals, and *value* denotes the mapped output. The sub-intervals follow the same open/closed endpoint rule as the intervals in Tables A1–A3 in Appendix A, and $\lceil\ \rceil$ denotes the ceiling operation.

This transformation reduces scale differences while preserving the physical interpretability of the original radar measurements. The mapped reflectivity, differential reflectivity, and correlation coefficient are assigned to the R, G, and B channels, respectively, to form a three-channel multi-feature radar image for network input. An example of the resulting image is shown in Figure 4, and the mapping intervals for reflectivity, differential reflectivity, and the correlation coefficient are summarized in Tables A1–A3 in Appendix A. The roles of the multi-feature radar image synthesis and polarimetric variable selection are evaluated in Section 5.3, where the non-uniform piecewise mapping scheme is further compared with linear normalization.

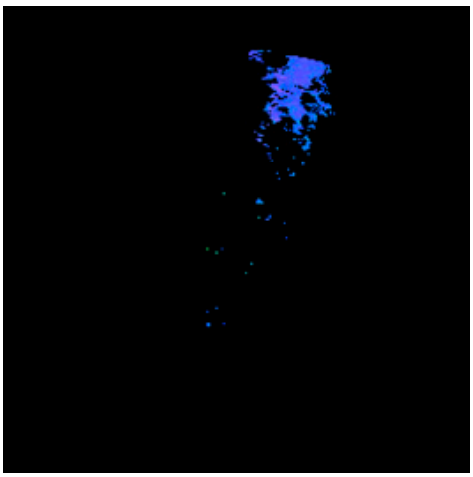


Figure 4. Multi-feature radar image at 17:22 UTC on 6 April 2021.

The proposed multi-feature radar image offers two main advantages over conventional input schemes:

- From the perspective of feature representation, the multi-feature radar image preserves the complementary information among different radar products. This allows the network to capture spatial structures more effectively. It also facilitates the learning of coupling relationships among different radar variables, thereby reducing reliance on any single feature;
- In terms of input consistency, this strategy reduces scale differences among radar variables through non-uniform piecewise mapping. As a result, the selected variables are integrated into a unified image representation while preserving physical interpretability.

4. Hydrometeor Classification Model with SPFANet

4.1. Overall Architecture

Hydrometeor classification remains limited by weak contextual modeling, insufficient multiscale feature fusion [23], and severe class imbalance [22]. To address these issues, this study proposes SPFANet, a semantic segmentation model for hydrometeor classification. The main design of SPFANet is the coordinated use of two complementary branches: a frozen Contrastive Language-Image Pre-Training (CLIP) branch based on ViT-B/16 and a Spatial Polarization Feature Aggregation (SPFA) branch. The frozen CLIP branch provides multi-level semantic representations, whereas the SPFA branch aggregates multiscale spatial features and enhances latent polarimetric features. The overall architecture of SPFANet is shown in Figure 5a.

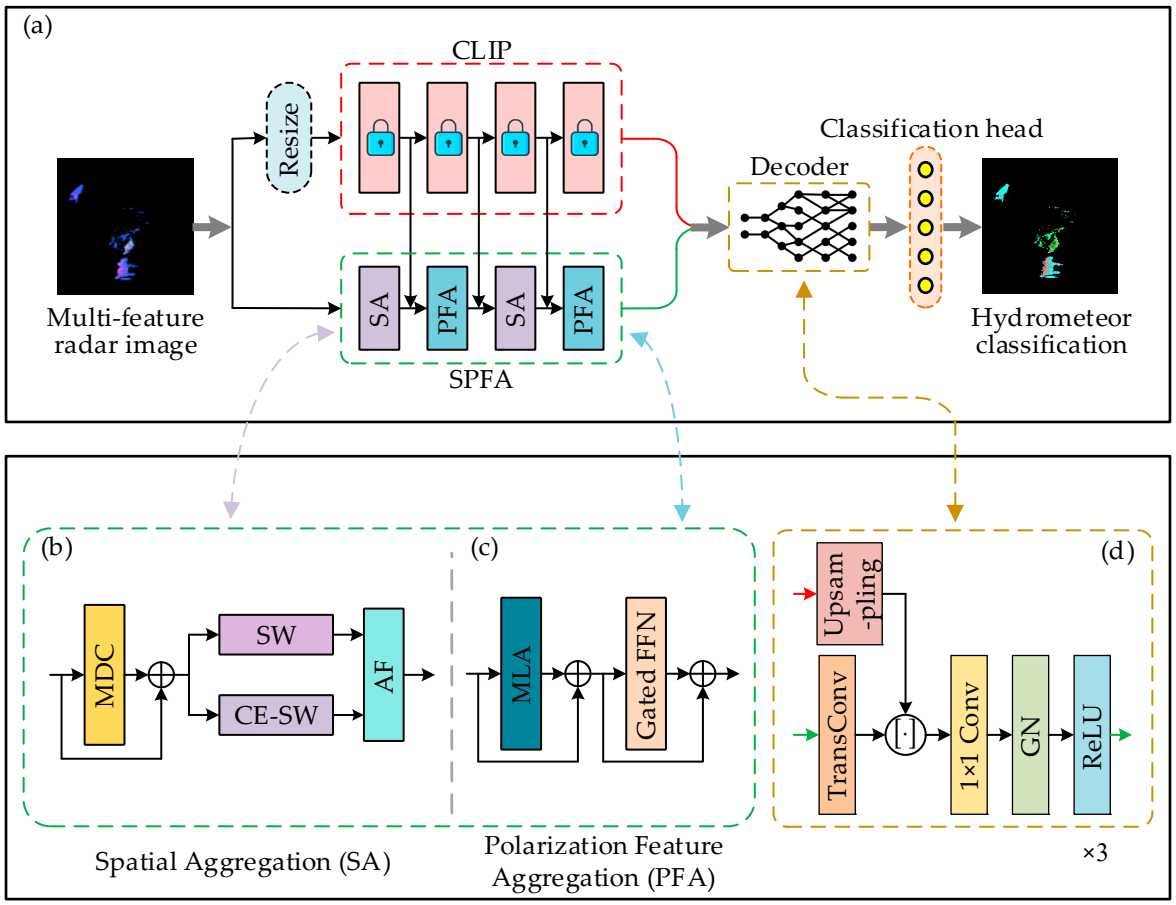


Figure 5. Architecture of SPFANet and its main components: (**a**) overall architecture; (**b**) SA module; (**c**) PFA module; (**d**) decoder block.

To meet the different requirements of the two branches, an asymmetric input-resolution strategy is adopted. The CLIP branch receives images upsampled to 320 × 320 to obtain higher-resolution semantic features. In contrast, the SPFA branch retains the original 256 × 256 resolution to preserve local details. Within the SPFA branch, the 256 × 256 × 3 input is first projected into non-overlapping 8 × 8 patches before entering the Spatial Aggregation (SA) and Polarization Feature Aggregation (PFA) modules. The SPFA branch contains two SA modules and two PFA modules arranged in an alternating cascade.

Meanwhile, a three-stage decoder is designed to reconstruct SPFA features progressively from low to high resolution. At each stage, transposed convolution performs twofold upsampling, and the upsampled features are concatenated with aligned CLIP features. The fused features are refined by convolution, group normalization (GN), and Rectified

Linear Unit (ReLU) activation to recover local texture and adjust channel dimensions. The decoder is followed by a 1 × 1 convolutional classification head for pixel-level hydrometeor classification. The structure of the decoder is shown in Figure 5d.

4.2. CLIP Branch and Feature Fusion

CLIP [34] learns transferable visual representations through large-scale image-text pretraining. Although it was originally developed as a dual-encoder architecture with both image and text encoders, its image encoder has demonstrated strong feature extraction ability and robust generalization across a wide range of visual tasks. For this reason, this study introduces the CLIP image encoder as an auxiliary branch to strengthen the representation of complex radar echo structures.

A pretrained ViT-B/16 model is used as the backbone of the CLIP branch, and the input to this branch is resized to 320 × 320. The ViT-B/16 backbone contains 12 Transformer blocks. For the 320 × 320 input, the patch size of 16 × 16 produces 400 patch tokens. This setting provides a suitable balance between local structural detail and global semantic information. To avoid relying solely on the final layer output, which may lead to the loss of local image detail, features from the third, sixth, and ninth Transformer blocks are injected into the SPFA branch.

To fuse the features from the two branches, this study designs a feature fusion unit, as shown in Figure 6. A 1 × 1 convolution is first used to project the 20 × 20 × 768 CLIP feature into 20 × 20 × 192, followed by upsampling to adjust its spatial resolution to 32 × 32. The aligned CLIP and SPFA features are then concatenated along the channel dimension. Finally, another 1 × 1 convolution reduces the channel dimension to 192, followed by layer normalization (LN).

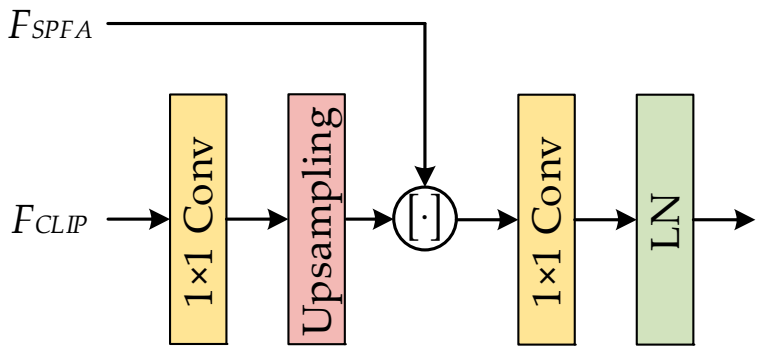


Figure 6. Feature fusion between the CLIP and SPFA branches.

The CLIP branch remains frozen during training and is not updated by backpropagation. This design helps preserve the general visual knowledge acquired during large-scale pretraining while reducing the number of trainable parameters and the overall computational cost. It also allows the optimization process to focus more directly on the SPFA branch and the decoder.

4.3. SA Module

Each channel of the multi-feature radar image corresponds to a polarimetric variable with a distinct physical meaning and therefore carries useful microphysical information about hydrometeors. However, real cloud and precipitation systems exhibit strong spatial continuity together with complex multiscale organization. Local regions may reflect fine-scale phase transitions, whereas larger regions are jointly controlled by broader meteorological systems. As a result, relying only on single-pixel features or limited local receptive fields is often insufficient to describe the spatial distribution of meteorological targets.

To enhance the spatial modeling capability of SPFANet, this study designs an SA module to aggregate multiscale spatial features in radar inputs. Unlike conventional convolutional feature extraction with fixed sampling locations, the SA module combines adaptive sampling and window-based attention to capture local structures and exchange information across window boundaries. To balance spatial modeling capability and computational efficiency, the input and output channel dimensions of the SA module are both set to 192. The input feature map is first processed by a Modulated Deformable Convolution (MDC) with a residual connection. MDC learns sampling offsets and modulation weights to adaptively capture the spatial structures of meteorological targets. According to the attention-head allocation, the 192-channel feature map is divided into two 96-channel inputs for the parallel Swin (SW) and Context-Enhanced Swin (CE-SW) branches. The SW branch captures local spatial dependencies, whereas the CE-SW branch incorporates information from a larger context region. Finally, an Adaptive Fusion (AF) module is designed to assign adaptive weights to the outputs of the two branches and fuse them. The overall structure of the SA module is shown in Figure 5b.

The SW branch follows the standard Swin Transformer structure [35], in which Window-based Multi-head Self-Attention (W-MSA) and Shifted Window-based Multi-head Self-Attention (SW-MSA) are applied consecutively. This branch uses a window size of 4×4 to focus on local spatial variations. In contrast, the CE-SW branch replaces W-MSA with the proposed Context-Enhanced Multi-head Self-Attention (CE-MSA), while retaining standard SW-MSA with an 8×8 window in the second block. This design allows the CE-SW branch to incorporate additional contextual information while preserving cross-window information exchange. The structures of the SW and CE-SW branches are shown in Figure 7.

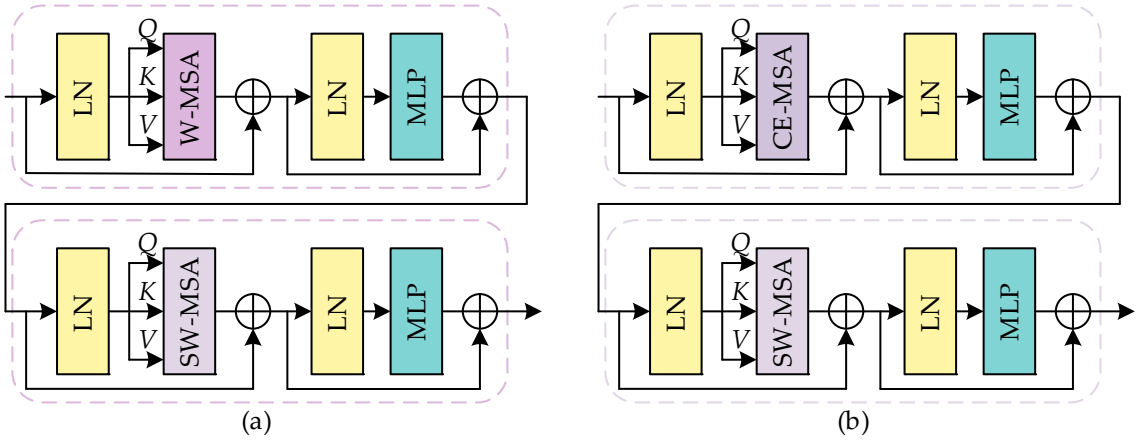


Figure 7. Structures of the window-attention branches: (**a**) SW branch; (**b**) CE-SW branch.

In the CE-MSA module, each 8×8 local window generates queries, whereas keys and values are constructed from both the local window and its corresponding 16×16 context region. For local windows near the feature map boundaries, the corresponding context regions are shifted inward to maintain a fixed size. The context features are downsampled to 4×4 by average pooling and projected into context keys and values, which are concatenated with the local keys and values along the token dimension. The resulting queries, keys, and values are then processed using the standard MSA formulation. The structure of the CE-MSA module is shown in Figure 8.

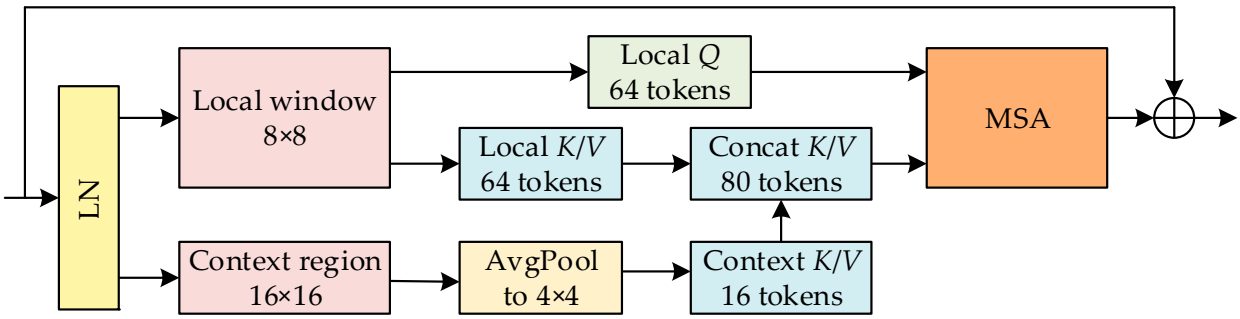


Figure 8. Structure of the CE-MSA module.

To combine the complementary spatial information extracted by the two branches, the SW and CE-SW branch features are passed to the AF module. The structure of the AF module is shown in Figure 9.

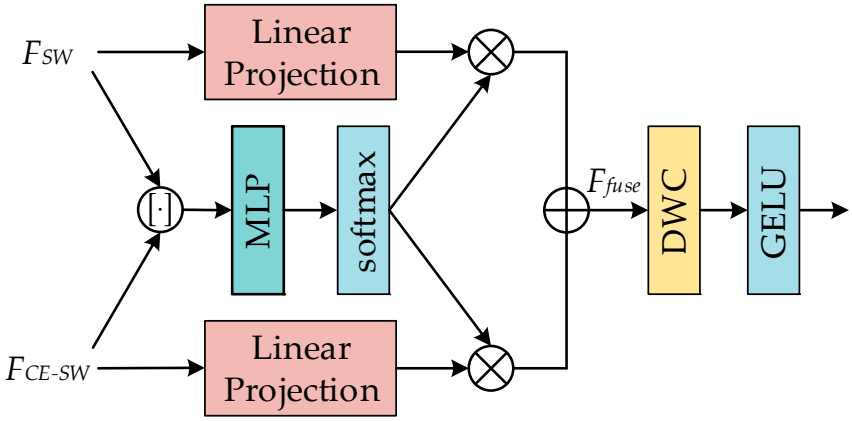


Figure 9. Structure of the AF module.

Within the AF module, the two branch features are concatenated and processed by a Multilayer Perceptron (MLP) followed by softmax to generate adaptive weights. Meanwhile, the feature from each branch is projected from 96 channels to 192 channels. The projected features are multiplied by their corresponding weights and added to obtain a fused feature. Finally, the fused feature is refined by Depthwise Convolution (DWC) and Gaussian Error Linear Unit (GELU).

4.4. PFA Module

In hydrometeor classification, a single polarimetric variable is usually insufficient for determining the underlying hydrometeor category. Reliable classification depends more strongly on the coordinated behavior of multiple polarimetric variables. Through patch embedding, the three polarimetric variables are jointly encoded into latent polarimetric features. To enhance the representation of these features, this study designs a Polarization Feature Aggregation (PFA) module.

To model long-range dependencies among spatial tokens, this study designs the PFA module based on the Linear Transformer architecture [36]. The module retains Multi-head Linear Attention (MLA) while replacing the original MLP with a Gated Feed-Forward Network (Gated FFN). The channel dimension is set to 192, with pre-normalization and residual connections applied to both the MLA layer and the Gated FFN. The structure of the PFA module is shown in Figure 10.

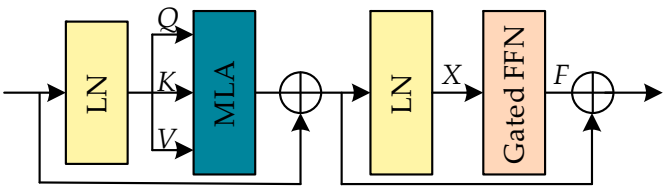


Figure 10. Structure of the PFA module.

The MLA layer applies linear attention to aggregate information across spatial tokens. The core idea of linear attention is to replace the softmax function in traditional scaled dot-product attention (SDPA) with a feature mapping function so that the similarity function can be reformulated in a feature-mapped form. The similarity function is given in Equation (2).

$$\mathrm{sim}(Q_i, K_j) = \phi(Q_i)^T \phi(K_j) \tag{2}$$

where Q_i and K_j denote the query vector of the i-th token and the key vector of the j-th token, sim denotes the similarity function, and ϕ denotes the feature mapping function, which is defined in Equation (3).

$$\phi(x) = \mathrm{ELU}(x) + 1 = \begin{cases} x + 1 & x > 0 \\ \alpha(e^x - 1) + 1 & x \leq 0 \end{cases} \tag{3}$$

where α denotes the scaling parameter for the negative part of the exponential linear unit (ELU) activation function and is set to 1 in this study.

Following the MLA layer, the Gated FFN is used to model nonlinear interactions across feature channels. Compared with the original MLP, the Gated FFN divides the projected feature into two branches and combines them through a gating mechanism. The normalized input feature X is first projected from 192 to 768 channels by a linear layer and then evenly split into two 384-channel branches:

$$[U_1, U_2] = \mathrm{Split}(XW_1) \tag{4}$$

The first branch is activated by GELU and then combined with the second branch through element-wise multiplication:

$$Z = \mathrm{GELU}(U_1) \odot U_2 \tag{5}$$

Finally, a second linear layer projects the gated feature from 384 to 192 channels:

$$F = ZW_2 \tag{6}$$

where W_1 and W_2 denote the learnable projection matrices, Split$(\cdot)$ denotes an equal split along the channel dimension, and $\odot$ denotes element-wise multiplication.

4.5. Loss Function

The hydrometeor classification dataset exhibits class imbalance. Hazardous categories such as hail and heavy rain are relatively scarce, whereas background and common hydrometeor classes dominate the data distribution. As a result, the model tends to favor majority classes, which reduces its performance on meteorologically important categories. To alleviate this problem, this study designs a class-weighting strategy for the *W*eighted *C*ross *E*ntropy (*WCE*) loss. The class weights are determined according to the class proportions and capped at a predefined threshold.

The *WCE* formulation is given in Equation (7).

$$WCE = -\sum_{i=1}^{N}\sum_{k=1}^{K} w_k \times y_{(i,k)} \times \log p_{(i,k)} \quad (7)$$

where i denotes the pixel index, k denotes the category index, N denotes the number of pixels, K denotes the number of categories, y denotes the one-hot NOAA reference label, and p denotes the predicted probability. The weight w_k is determined by the proposed class-weighting strategy in Equations (8)–(10).

$$f_k = n_k / n_{all} \quad (8)$$

$$w_k^{raw} = \text{median}(f) / f_k \quad (9)$$

$$w_k = \min(w_k^{raw}, w_{threshold}) \quad (10)$$

where n_{all} denotes the total number of pixels in the training set, n_k denotes the number of pixels belonging to class k, f_k denotes the proportion of class k, median(f) denotes the median of all class proportions, and $w_{threshold}$ denotes the clipping threshold. In this study, $w_{threshold}$ is set to 10 to prevent excessively large class weights. The total number of labeled pixels and the corresponding class weight for each category in the training set are shown in Table A4 in Appendix A.

5. Experimental Results and Analysis

5.1. Experimental Settings

The dataset used in this study was constructed from KTLX radar observations collected between April and September 2021. Each sample corresponds to one observation time and consists of a 256 × 256 multi-feature radar image and its corresponding hydrometeor classification label. A total of 21,431 samples were obtained and randomly divided by observation date into 15,002 training samples, 2143 validation samples, and 4286 test samples. The observation dates assigned to the training, validation, and test sets are provided in Table A5 in Appendix A. This date-based split helps keep adjacent scans from the same precipitation process in the same subset. The NOAA reference labels were obtained from the HCA product. Accordingly, the quantitative results presented in Section 5 quantify the agreement between model predictions and the NOAA reference labels. The original NOAA hydrometeor labels include 12 echo categories, of which two correspond to non-meteorological echoes and ten correspond to precipitation-related echoes. For evaluation, the precipitation-related categories were merged into five broader groups according to operational meteorological practice and physical similarity: ice crystals, snow, rain, graupel, and hail. The category merging strategy and the corresponding physical rationale for each group are provided in Table 2.

The training settings were specified as follows. All models used the same data split and the *WCE* loss. AdamW was used as the optimizer, with an initial learning rate of 1×10^{-4}. The batch size was set to 16, and each model was trained for 50 epochs. The model-specific architectural settings of the comparison models followed their original designs or public implementations. Each experiment was independently repeated five times. The evaluation metrics are reported as mean ± standard deviation. SPFANet contains approximately 4.1 million parameters, and its average inference time was 1.7 s per image on a workstation equipped with an NVIDIA GeForce RTX 3090 GPU.

Table 2. Hydrometeor category grouping and physical rationale.

Merged Category	Original Categories	Physical Rationale
Ice crystals	Ice Crystals	Ice crystal characteristics
Snow	Dry Snow; Wet Snow	Melting state
Rain	Light/Moderate Rain; Heavy Rain; Big-Drops Rain	Rainfall intensity and drop size
Graupel	Graupel	Degree of riming
Hail	Small Hail; Large Hail; Giant Hail	Hail size

5.2. Quantitative Evaluation of Hydrometeor Classification

To evaluate model performance for the five hydrometeor categories, a confusion matrix was used. The NOAA hydrometeor classification product provided the reference labels. Therefore, the metrics reported in this section measure the agreement between each model and the NOAA reference labels. The columns and rows of the confusion matrix correspond to the predicted labels and the NOAA reference labels, respectively. Table 3 lists the confusion-matrix definitions used in this study, where *TP*, *TN*, *FP*, and *FN* denote true positive, true negative, false positive, and false negative, respectively.

Table 3. Confusion-matrix definitions.

	Predicted Positive	**Predicted Negative**
Label positive	*TP*	*FN*
Label negative	*FP*	*TN*

5.2.1. Class-Wise Evaluation Metrics

Based on the confusion matrix, *Precision*, *Recall*, *F*1, and Intersection over Union (*IoU*) were selected as class-wise evaluation metrics commonly used in hydrometeor classification. *Precision* measures the reliability of positive predictions, *Recall* reflects the completeness of target detection, *F*1 represents their harmonic mean and reflects the balance between false alarms and missed detections, and *IoU* measures the overlap between predicted and target regions. In meteorology, *IoU* is equivalent to the Critical Success Index (*CSI*). Their formulations are given in Equations (11)–(14).

$$Precision = \frac{TP}{TP + FP} \tag{11}$$

$$Recall = \frac{TP}{TP + FN} \tag{12}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{13}$$

$$IoU = \frac{TP}{TP + FP + FN} \tag{14}$$

Representative CNN semantic segmentation models were selected as comparison methods, including FCN [16], U-Net [17], DeepLabv3+ [22], and SegNet [18]. In addition, SegFormer [37] and DINOv2 [38], which are based on Transformer architectures, were introduced as supplementary benchmarks. The class-wise evaluation results of different models are summarized in Table 4.

Table 4. Class-wise evaluation metrics for different models.

Metric	Model	Ice Crystals	Snow	Rain	Graupel	Hail
Precision (%)	FCN [16]	41.1 ± 0.73	77.6 ± 0.68	89.4 ± 0.59	55.3 ± 0.47	4.4 ± 0.81
	U-Net [17]	45.5 ± 0.61	78.2 ± 0.47	92.3 ± 0.39	57.5 ± 0.32	14.2 ± 0.65
	DeepLabv3+ [22]	**51.7 ± 0.51**	79.4 ± 0.36	**94.5 ± 0.33**	61.4 ± 0.45	51.4 ± 0.71
	SegNet [18]	37.3 ± 1.03	**81.4 ± 0.58**	87.5 ± 0.43	57.3 ± 0.44	41.0 ± 0.74
	SegFormer [37]	46.3 ± 0.42	78.4 ± 0.37	91.7 ± 0.36	66.4 ± 0.33	57.1 ± 0.57
	DINOv2 [38]	24.5 ± 0.87	45.4 ± 0.38	82.6 ± 0.26	**91.0 ± 0.19**	52.4 ± 0.47
	SPFANet	30.5 ± 0.43	69.8 ± 0.46	89.7 ± 0.30	90.7 ± 0.53	**64.3 ± 0.51**
Recall (%)	FCN [16]	0.3 ± 0.02	89.4 ± 0.89	89.5 ± 0.35	44.6 ± 0.45	0.1 ± 0.01
	U-Net [17]	3.5 ± 0.60	90.5 ± 0.61	88.6 ± 0.36	59.6 ± 0.32	11.3 ± 0.44
	DeepLabv3+ [22]	8.4 ± 0.50	**91.6 ± 0.65**	88.7 ± 0.22	64.5 ± 0.36	19.8 ± 0.33
	SegNet [18]	6.0 ± 0.77	91.5 ± 0.72	88.6 ± 0.33	54.7 ± 0.41	3.4 ± 0.41
	SegFormer [37]	6.6 ± 0.31	89.4 ± 0.50	**92.7 ± 0.20**	56.9 ± 0.31	12.4 ± 0.41
	DINOv2 [38]	85.4 ± 0.53	87.5 ± 0.33	82.4 ± 0.39	70.3 ± 0.55	74.2 ± 0.57
	SPFANet	**86.3 ± 0.47**	89.5 ± 0.63	90.6 ± 0.31	**82.6 ± 0.27**	**81.3 ± 0.46**
*F*1 (%)	FCN [16]	0.5 ± 0.05	83.1 ± 0.69	89.4 ± 0.43	49.4 ± 0.44	0.2 ± 0.02
	U-Net [17]	6.5 ± 1.03	83.9 ± 0.45	90.4 ± 0.35	58.5 ± 0.28	12.6 ± 0.51
	DeepLabv3+ [22]	14.4 ± 0.74	85.1 ± 0.42	91.5 ± 0.26	62.9 ± 0.38	28.6 ± 0.44
	SegNet [18]	10.4 ± 1.16	**86.2 ± 0.60**	88.0 ± 0.30	56.0 ± 0.39	6.3 ± 0.70
	SegFormer [37]	11.6 ± 0.48	83.5 ± 0.42	**92.2 ± 0.26**	61.3 ± 0.26	20.4 ± 0.59
	DINOv2 [38]	38.1 ± 1.07	59.8 ± 0.38	82.5 ± 0.32	79.3 ± 0.37	61.4 ± 0.49
	SPFANet	**45.1 ± 0.53**	78.4 ± 0.43	90.1 ± 0.27	**86.5 ± 0.35**	**71.8 ± 0.50**
IoU (%)	FCN [16]	0.3 ± 0.02	71.0 ± 1.01	80.9 ± 0.70	32.8 ± 0.39	0.1 ± 0.01
	U-Net [17]	3.3 ± 0.55	72.2 ± 0.67	82.5 ± 0.58	41.4 ± 0.28	6.7 ± 0.29
	DeepLabv3+ [22]	7.7 ± 0.43	74.0 ± 0.64	84.4 ± 0.45	45.9 ± 0.41	16.7 ± 0.30
	SegNet [18]	5.5 ± 0.65	**75.7 ± 0.92**	78.6 ± 0.48	38.9 ± 0.38	3.3 ± 0.37
	SegFormer [37]	6.2 ± 0.27	71.7 ± 0.62	**85.5 ± 0.44**	44.2 ± 0.27	11.4 ± 0.37
	DINOv2 [38]	23.6 ± 0.81	42.6 ± 0.39	70.2 ± 0.46	65.7 ± 0.50	44.3 ± 0.51
	SPFANet	**29.1 ± 0.44**	64.5 ± 0.58	82.1 ± 0.44	**76.1 ± 0.54**	**56.0 ± 0.61**

The results reveal substantial differences among the models in recognizing ice crystals, graupel, and hail. For most comparison models other than DINOv2, the mean *Recall* remains below 65% for graupel and below 20% for both ice crystals and hail. These results indicate severe missed detections and suggest that these models have difficulty capturing the microphysical characteristics of these three particle types. By contrast, SPFANet achieves mean *Recall* values above 81% for all three categories. Notably, these values are also higher than those of DINOv2, which achieves the highest mean *Recall* among the comparison models for each of the three categories. This improvement helps reduce missed detections of meteorologically important categories, which is particularly valuable for operational forecasting. The relatively small standard deviations indicate that the class-wise metrics remain stable across the five runs.

The improvement in *Recall* for ice crystals is achieved partly at the expense of *Precision*. Even so, SPFANet still achieves a mean *Precision* value of 90.7% for graupel, close to the highest mean value of 91.0%. For hail, it achieves the highest mean *Precision* of 64.3%. It also achieves the highest mean *F*1 for ice crystals, graupel, and hail, indicating a more favorable balance between *Precision* and *Recall*. In addition, SPFANet achieves the highest mean *IoU* for these categories, indicating better spatial agreement.

Although SPFANet does not achieve the best results for rain and snow, it maintains high mean *Recall* values of 90.6% and 89.5%, respectively. In addition, the pixel distributions of rain and snow are highly imbalanced, with light/moderate rain and dry snow dominating their respective categories. Therefore, the metrics for rain and snow do not separately reflect the classification performance of their original NOAA hydrometeor

categories. To further analyze the fine-grained hydrometeor classification results, these important subcategories are examined through visual cases in Section 5.5.

5.2.2. Overall Evaluation Metrics

In addition to the class-wise evaluation metrics, four overall evaluation metrics were further used to evaluate model performance more comprehensively. *Macro-F*1 and mean Intersection over Union (*mIoU*) are calculated by averaging *F*1 and *IoU* across categories. Overall accuracy (*OA*) measures the proportion of correctly classified pixels among all pixels. Balanced accuracy (*BA*) is obtained by averaging the correctly classified pixel ratio of each category. These metrics provide an overall assessment of model performance under class imbalance and serve as a complement to the class-wise results. Their formulations are given in Equations (15)–(18).

$$Macro - F1 = \frac{1}{C}\sum_{c=1}^{C} F1_c \tag{15}$$

$$mIoU = \frac{1}{C}\sum_{c=1}^{C} IoU_c \tag{16}$$

$$OA = \sum_{c=1}^{C} TP_c / \sum_{c=1}^{C} N_c \tag{17}$$

$$BA = \frac{1}{C}\sum_{c=1}^{C} (TP_c / N_c) \tag{18}$$

where C denotes the number of categories, and N_c denotes the number of reference pixels in category c.

Based on the above definitions, the overall evaluation results of different models are summarized in Table 5.

Table 5. Overall evaluation metrics for different models.

Model	*Macro-F*1 (%)	*mIoU* (%)	*OA* (%)	*BA* (%)
FCN [16]	44.5 ± 0.20	37.0 ± 0.26	59.7 ± 0.29	44.8 ± 0.19
U-Net [17]	50.4 ± 0.45	41.2 ± 0.40	67.9 ± 0.27	50.7 ± 0.33
DeepLabv3+ [22]	56.5 ± 0.38	45.7 ± 0.36	71.2 ± 0.26	54.6 ± 0.28
SegNet [18]	49.4 ± 0.50	40.4 ± 0.44	64.9 ± 0.25	48.8 ± 0.35
SegFormer [37]	53.8 ± 0.32	43.8 ± 0.33	67.9 ± 0.24	51.6 ± 0.26
DINOv2 [38]	64.2 ± 0.43	49.3 ± 0.44	76.1 ± 0.44	80.0 ± 0.39
SPFANet	**74.4 ± 0.31**	**61.6 ± 0.40**	**85.8 ± 0.30**	**86.1 ± 0.37**

The results show that SPFANet achieves the best overall performance among all models. It obtains the highest *Macro-F*1, *mIoU*, *OA*, and *BA*, with mean values of 74.4%, 61.6%, 85.8%, and 86.1%, respectively. This indicates that SPFANet provides better overall classification performance while maintaining more balanced performance across different hydrometeor categories.

5.3. Evaluation of Radar Input Processing

This section evaluates the effect of radar input processing on hydrometeor classification performance. Three aspects of radar input processing are considered: non-meteorological echo removal, multi-feature radar image synthesis, and the selection of polarimetric variables. These experiments provide a quantitative assessment of the input processing steps described in Sections 3.2 and 3.3. To ensure a comparable evaluation, all experiments

used the same model architecture and training settings, so that the comparison focused on differences in radar input processing. Five radar input settings were compared. The class-wise and overall evaluation results are summarized in Tables 6 and 7, respectively.

Table 6. Class-wise evaluation metrics for different inputs.

Metric	Input	Ice Crystals	Snow	Rain	Graupel	Hail
Precision (%)	No echo removal	27.5 ± 0.55	64.7 ± 0.45	85.8 ± 0.32	87.8 ± 0.45	58.7 ± 0.50
	Direct input	26.6 ± 0.60	60.7 ± 0.50	83.8 ± 0.35	84.8 ± 0.55	51.8 ± 0.60
	K_{DP} input	28.6 ± 0.55	66.7 ± 0.48	87.8 ± 0.32	88.7 ± 0.48	61.7 ± 0.55
	Normalized input	29.6 ± 0.50	67.7 ± 0.48	88.8 ± 0.30	**91.7 ± 0.45**	62.8 ± 0.50
	Proposed input	**30.5 ± 0.43**	**69.8 ± 0.46**	**89.7 ± 0.30**	90.7 ± 0.53	**64.3 ± 0.51**
Recall (%)	No echo removal	81.7 ± 0.60	85.7 ± 0.45	84.8 ± 0.30	78.8 ± 0.44	75.6 ± 0.52
	Direct input	75.7 ± 0.70	78.7 ± 0.55	82.8 ± 0.35	75.8 ± 0.66	69.6 ± 0.65
	K_{DP} input	83.8 ± 0.62	86.7 ± 0.45	87.7 ± 0.32	79.7 ± 0.56	77.6 ± 0.55
	Normalized input	84.8 ± 0.55	87.7 ± 0.45	89.9 ± 0.30	81.8 ± 0.64	78.7 ± 0.50
	Proposed input	**86.3 ± 0.47**	**89.5 ± 0.63**	**90.6 ± 0.31**	**82.6 ± 0.27**	**81.3 ± 0.46**
*F*1 (%)	No echo removal	41.2 ± 0.58	73.7 ± 0.44	85.3 ± 0.17	83.0 ± 0.40	66.1 ± 0.47
	Direct input	39.4 ± 0.65	68.5 ± 0.29	83.3 ± 0.32	80.0 ± 0.59	59.4 ± 0.53
	K_{DP} input	42.7 ± 0.69	75.4 ± 0.26	87.8 ± 0.20	84.0 ± 0.50	68.8 ± 0.52
	Normalized input	43.8 ± 0.57	76.4 ± 0.40	89.3 ± 0.18	**86.5 ± 0.54**	69.8 ± 0.46
	Proposed input	**45.1 ± 0.53**	**78.4 ± 0.43**	**90.1 ± 0.27**	**86.5 ± 0.35**	**71.8 ± 0.50**
IoU (%)	No echo removal	25.9 ± 0.46	58.4 ± 0.55	74.3 ± 0.25	71.0 ± 0.59	49.4 ± 0.52
	Direct input	24.5 ± 0.50	52.1 ± 0.33	71.4 ± 0.47	66.7 ± 0.82	42.2 ± 0.54
	K_{DP} input	27.1 ± 0.55	60.5 ± 0.33	78.2 ± 0.32	72.4 ± 0.74	52.4 ± 0.60
	Normalized input	28.1 ± 0.47	61.8 ± 0.52	80.7 ± 0.30	**76.1 ± 0.84**	53.6 ± 0.54
	Proposed input	**29.1 ± 0.44**	**64.5 ± 0.58**	**82.1 ± 0.44**	**76.1 ± 0.54**	**56.0 ± 0.61**

Table 7. Overall evaluation metrics for different inputs.

Input	*Macro-F*1 (%)	*mIoU* (%)	*OA* (%)	*BA* (%)
No echo removal	69.9 ± 0.28	55.8 ± 0.33	81.1 ± 0.26	81.3 ± 0.28
Direct input	66.1 ± 0.29	51.4 ± 0.33	77.9 ± 0.35	76.5 ± 0.27
K_{DP} input	71.7 ± 0.27	58.1 ± 0.31	82.8 ± 0.27	83.1 ± 0.25
Normalized input	73.2 ± 0.29	60.1 ± 0.34	84.7 ± 0.28	84.6 ± 0.29
Proposed input	**74.4 ± 0.31**	**61.6 ± 0.40**	**85.8 ± 0.30**	**86.1 ± 0.37**

For each comparison, only the specified component differs from Proposed input, while all other processing steps remain unchanged. No echo removal denotes the setting without non-meteorological echo removal. Direct input denotes the setting in which the selected radar variables are provided directly as three input channels after non-meteorological echo removal, without applying the non-uniform piecewise mapping scheme. K_{DP} input denotes the setting in which specific differential phase is used instead of the correlation coefficient as one of the input variables. Normalized input denotes the setting in which each selected radar variable is independently linearly mapped from its valid range to 0–255 after non-meteorological echo removal to form a three-channel input. Proposed input denotes the radar input processing setting adopted in this study. By comparing these settings, the effects of non-meteorological echo removal, multi-feature radar image synthesis, and polarimetric variable selection can be evaluated separately.

The effect of non-meteorological echo removal is reflected by the comparison between No echo removal and Proposed input. With non-meteorological echo removal, the evaluation metrics improve for all five hydrometeor categories. This result indicates that removing non-meteorological echoes reduces the interference of clutter and biological echoes in hydrometeor classification. The resulting input is more effective for the model while preserving the main meteorological structures.

The role of multi-feature radar image synthesis is evaluated by comparing Direct input, Normalized input, and Proposed input. Direct input shows the lowest classification performance among the input settings. Directly using the selected radar variables as three input channels ignores their differences in numerical scale and value distribution. In this input setting, the input representation is more strongly influenced by reflectivity, while the contributions of differential reflectivity and the correlation coefficient are not fully represented.

Normalized input shows higher values for all metrics than Direct input. This result indicates that mapping the radar variables to a common intensity domain improves the input representation. Proposed input achieves better results than Normalized input for most evaluation metrics. This indicates that the non-uniform piecewise mapping scheme provides a more effective input representation by accounting for differences in the physical attributes and statistical distributions of the radar variables.

The comparison between K_{DP} input and Proposed input shows the effect of polarimetric variable selection. K_{DP} input performs better than Direct input and No echo removal, further supporting the roles of non-meteorological echo removal and multi-feature radar image synthesis. In contrast, Proposed input achieves better class-wise and overall classification performance than K_{DP} input. This result indicates that the correlation coefficient is preferable to specific differential phase when combined with reflectivity and differential reflectivity for multi-feature radar image synthesis. The performance advantage may result from the information on the uniformity of hydrometeor populations provided by the correlation coefficient.

5.4. Ablation Study

To assess the contributions of the key network modules to model performance, ablation experiments were conducted. The class-wise and overall evaluation results are summarized in Tables 8 and 9, where C denotes the frozen CLIP branch, L denotes the *WCE* loss function, S denotes the SA module, and P denotes the PFA module.

Table 8. Class-wise evaluation metrics for the ablation study.

Metric	Model	Ice Crystals	Snow	Rain	Graupel	Hail
Precision (%)	SPFANet with no C	32.6 ± 0.61	73.7 ± 0.47	88.6 ± 0.34	86.6 ± 0.57	42.2 ± 0.53
	SPFANet with no L	**39.5 ± 0.73**	**81.6 ± 0.60**	86.6 ± 0.38	81.5 ± 0.59	31.0 ± 0.72
	SPFANet with no S	22.7 ± 0.57	59.7 ± 0.43	85.6 ± 0.32	84.6 ± 0.57	56.2 ± 0.61
	SPFANet with no P	27.7 ± 0.53	65.7 ± 0.43	88.6 ± 0.32	89.7 ± 0.51	63.3 ± 0.54
	SPFANet	30.5 ± 0.43	69.8 ± 0.46	**89.7 ± 0.30**	**90.7 ± 0.53**	**64.3 ± 0.51**
Recall (%)	SPFANet with no C	75.1 ± 0.62	88.5 ± 0.62	89.6 ± 0.30	73.5 ± 0.66	69.2 ± 0.53
	SPFANet with no L	59.0 ± 0.74	81.3 ± 0.55	88.5 ± 0.38	71.6 ± 0.75	51.0 ± 0.73
	SPFANet with no S	76.0 ± 0.75	81.5 ± 0.62	84.5 ± 0.40	74.8 ± 0.50	74.2 ± 0.56
	SPFANet with no P	82.3 ± 0.54	86.5 ± 0.59	88.5 ± 0.35	80.8 ± 0.51	78.2 ± 0.54
	SPFANet	**86.3 ± 0.47**	**89.5 ± 0.63**	**90.6 ± 0.31**	**82.6 ± 0.27**	**81.3 ± 0.46**
*F*1 (%)	SPFANet with no C	45.5 ± 0.68	80.4 ± 0.47	89.1 ± 0.29	79.5 ± 0.62	52.5 ± 0.56
	SPFANet with no L	**47.3 ± 0.68**	**81.5 ± 0.56**	87.5 ± 0.38	76.2 ± 0.67	38.6 ± 0.77
	SPFANet with no S	34.9 ± 0.73	68.9 ± 0.48	85.0 ± 0.35	79.4 ± 0.53	63.9 ± 0.60
	SPFANet with no P	41.5 ± 0.64	74.7 ± 0.46	88.6 ± 0.33	85.0 ± 0.51	70.0 ± 0.54
	SPFANet	45.1 ± 0.53	78.4 ± 0.43	**90.1 ± 0.27**	**86.5 ± 0.35**	**71.8 ± 0.50**
IoU (%)	SPFANet with no C	29.5 ± 0.57	67.3 ± 0.66	80.3 ± 0.47	66.0 ± 0.85	35.6 ± 0.51
	SPFANet with no L	**31.0 ± 0.59**	**68.8 ± 0.80**	77.8 ± 0.60	61.6 ± 0.88	23.9 ± 0.59
	SPFANet with no S	21.1 ± 0.53	52.6 ± 0.56	74.0 ± 0.53	65.8 ± 0.73	47.0 ± 0.65
	SPFANet with no P	26.2 ± 0.51	59.6 ± 0.58	79.5 ± 0.53	73.9 ± 0.77	53.8 ± 0.64
	SPFANet	29.1 ± 0.44	64.5 ± 0.58	**82.1 ± 0.44**	**76.1 ± 0.54**	**56.0 ± 0.61**

Table 9. Overall evaluation metrics for the ablation study.

Model	*Macro-F*1 (%)	*mIoU* (%)	*OA* (%)	*BA* (%)
SPFANet with no C	69.4 ± 0.36	55.7 ± 0.44	79.8 ± 0.45	79.2 ± 0.44
SPFANet with no L	66.2 ± 0.48	52.6 ± 0.55	76.4 ± 0.57	70.3 ± 0.54
SPFANet with no S	66.4 ± 0.36	52.1 ± 0.43	78.5 ± 0.42	78.2 ± 0.47
SPFANet with no P	71.9 ± 0.32	58.6 ± 0.42	83.6 ± 0.39	83.3 ± 0.41
SPFANet	**74.4 ± 0.31**	**61.6 ± 0.40**	**85.8 ± 0.30**	**86.1 ± 0.37**

The class-wise evaluation results in Table 8 reflect the performance of SPFANet across hydrometeor categories. Although SPFANet does not achieve the highest mean *Precision* in every category, it achieves the highest mean *Recall* for all five hydrometeor categories. For rain, graupel, and hail, it also obtains the highest mean values of *Precision*, *F*1, and *IoU*. These results indicate that SPFANet improves *Recall* across hydrometeor categories while maintaining reliable performance for rain, graupel, and hail.

Removing the *WCE* loss function increases *Precision* for ice crystals and snow. However, it decreases *Recall* for all five categories. It also reduces *F*1 and *IoU* for rain, graupel, and hail. These results indicate that the *WCE* loss function contributes to improving model performance under class imbalance.

Removing the CLIP branch decreases *Recall* for all five categories and reduces *Precision*, *F*1, and *IoU* for rain, graupel, and hail. The mean values of the four metrics decrease across all categories after removing either the SA or PFA module, with larger decreases following SA removal. Together, these results indicate that the CLIP branch, SA module, and PFA module contribute to hydrometeor classification by providing multi-level semantic representations, aggregating multiscale spatial features, and enhancing latent polarimetric features, respectively. The overall evaluation results in Table 9 further demonstrate the effectiveness of SPFANet. The proposed model achieves the highest *Macro-F*1, *mIoU*, *OA*, and *BA*, with mean values of 74.4%, 61.6%, 85.8%, and 86.1%, respectively.

5.5. Visual Analysis of Typical Cases

Three representative hydrometeor cases were selected for visual analysis: a snow-related case, a hail-related case, and a complex case. To support the case analysis, three types of observational information independent of the NOAA reference labels were used: Global Precipitation Measurement (GPM) satellite precipitation data, NOAA ground-station precipitation records, and Storm Prediction Center (SPC) storm reports. GPM satellite precipitation data provide spatial information on precipitation intensity at a 30-min temporal resolution. NOAA ground-station precipitation records provide 1-h surface precipitation measurements near the selected radar observation times. SPC storm reports provide independent records of severe convective weather, including hail reports. These observations do not provide pixel-level labels for hydrometeor categories aloft. Instead, they provide independent information on precipitation patterns and severe convective weather in the selected cases.

5.5.1. Snow-Related Hydrometeor Case

For the snow-related hydrometeor case at 17:10 UTC on 28 April 2021, NOAA ground-station precipitation records and GPM satellite precipitation data from 17:00 to 17:30 UTC were used as physically independent observational information. Both datasets were processed using the same study domain and 256 × 256 grid as the radar data to enable spatial comparison. Figure 11a shows the ground stations that recorded precipitation during this time interval. Red crosses mark the station locations. According to the ground-station records, light precipitation occurred near the central part of the study area. Figure 11b

shows the GPM precipitation field. A semicircular precipitation region appears in the southeastern part of the study area. This precipitation region provides supporting evidence for interpreting the snow-related hydrometeor distribution observed by radar.

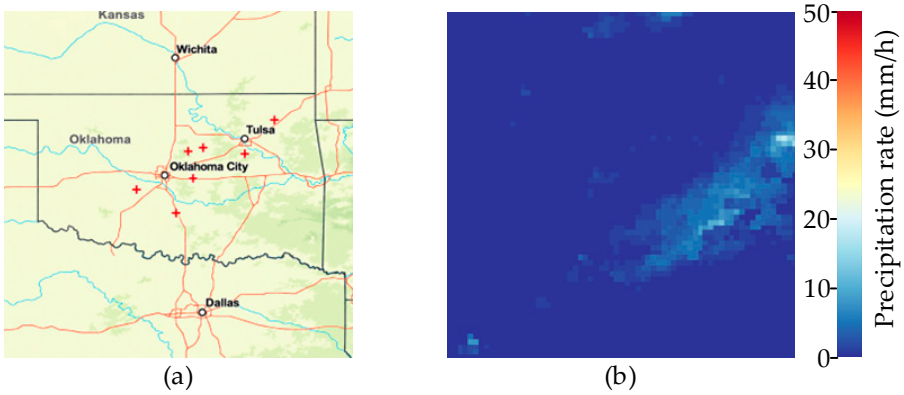


Figure 11. Independent observational information for the snow-related hydrometeor case at 17:10 UTC on 28 April 2021: (**a**) NOAA ground-station precipitation records; (**b**) GPM satellite precipitation data.

The corresponding hydrometeor classification results are shown in Figure 12. The snow-related case contains classified regions of light/moderate rain (light green) and dry snow (light blue). A band of wet snow (dark blue) appears along the boundary between these two regions. This area is also accompanied by a large number of graupel regions (brown). Compared with the other models, SPFANet delineates the transition areas among rain, dry snow, wet snow, and graupel more clearly, showing better boundary definition and improved preservation of local details.

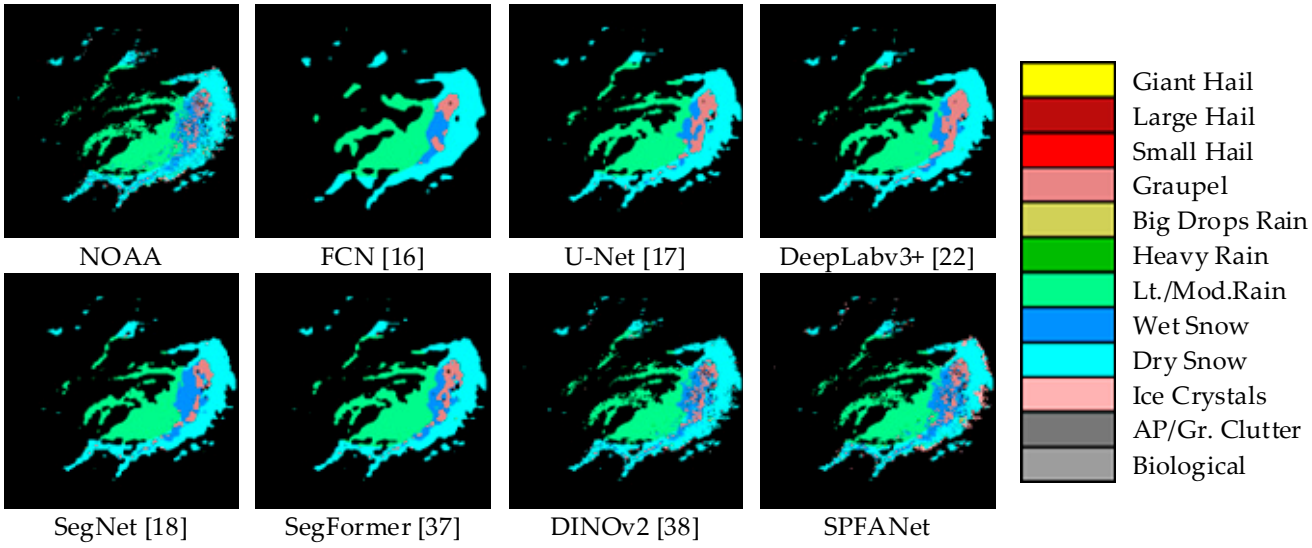


Figure 12. Hydrometeor classification results for the snow-related case at 17:10 UTC on 28 April 2021. The subplots show classification outputs from NOAA, FCN [16], U-Net [17], DeepLabv3+ [22], SegNet [18], SegFormer [37], DINOv2 [38], and the proposed SPFANet model, respectively. The color bar on the right indicates different hydrometeor categories.

5.5.2. Hail-Related Hydrometeor Case

For the hail-related hydrometeor case at 01:30 UTC on 11 July 2021, SPC storm reports were used as physically independent observational information. As shown in Figure 13, a hail report was recorded near the central part of the study area at 01:31 UTC, close to the selected radar observation time. The reported hail location is marked by the blue point.

This report provides independent surface hail evidence for interpreting the hail-related hydrometeor distribution observed by radar in this case.



Figure 13. SPC storm report for the hail-related hydrometeor case at 01:30 UTC on 11 July 2021.

The corresponding hydrometeor classification results are shown in Figure 14. The hail-related case contains classified regions of light/moderate rain (light green), dry snow (light blue), graupel (brown), and hail (red). DINOv2 and SPFANet both identify hail regions, broadly consistent with the SPC storm report shown in Figure 13. Compared with the other models, these two models preserve more detailed spatial structures and show clearer local boundaries among the classified hydrometeor regions.

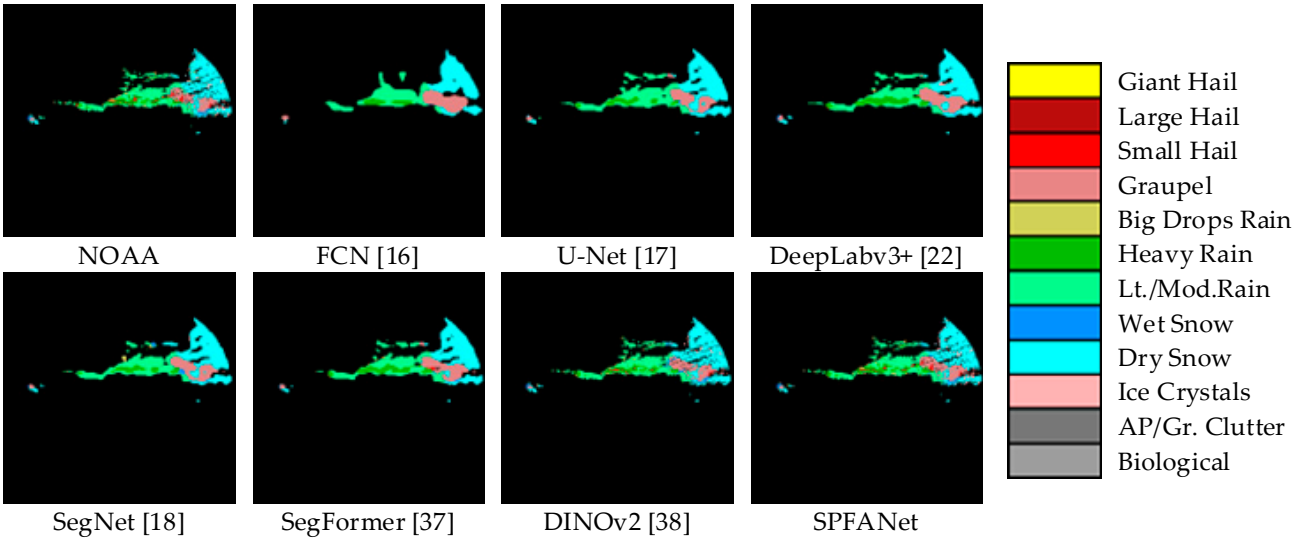


Figure 14. Hydrometeor classification results for the hail-related case at 01:30 UTC on 11 July 2021. The subplots show classification outputs from NOAA, FCN [16], U-Net [17], DeepLabv3+ [22], SegNet [18], SegFormer [37], DINOv2 [38], and the proposed SPFANet model, respectively. The color bar on the right indicates different hydrometeor categories.

To further examine the local differences near the reported hail location, Figure 15 presents an enlarged view centered on the SPC storm report at 01:31 UTC. In the complex mixed-hydrometeor region on the right side of the enlarged view, SPFANet identifies some hail regions that are not detected by DINOv2. In addition, SPFANet shows a more continuous representation of the surrounding rain and dry snow regions. This local comparison indicates that SPFANet provides a clearer representation of hail-related hydrometeor categories aloft near the reported hail location.

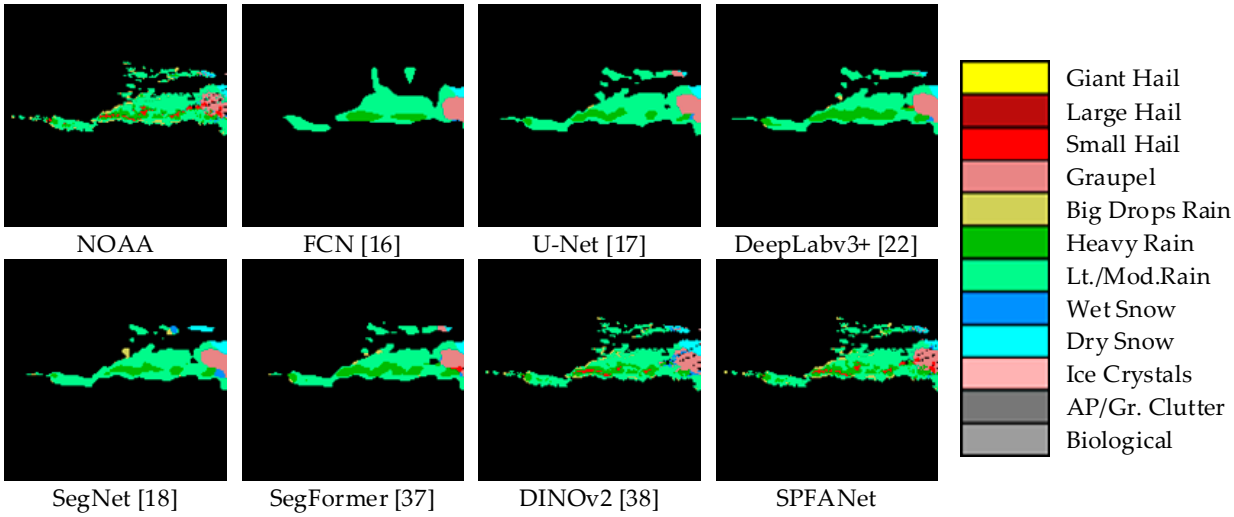


Figure 15. Local enlargement centered on the SPC storm report. The subplots show classification outputs from NOAA, FCN [16], U-Net [17], DeepLabv3+ [22], SegNet [18], SegFormer [37], DINOv2 [38], and the proposed SPFANet model, respectively. The color bar on the right indicates different hydrometeor categories.

5.5.3. Complex Hydrometeor Case

For the complex hydrometeor case at 01:16 UTC on 28 April 2021, NOAA ground-station precipitation records and GPM satellite precipitation data from 01:00 to 01:30 UTC were used as physically independent observational information. Both datasets were processed using the same study domain and 256 × 256 grid as the radar data to enable spatial comparison. Figure 16a shows the ground stations that recorded precipitation during this time interval. Red crosses mark the station locations. According to the ground-station records, precipitation occurred near the central and eastern parts of the study area. Figure 16b shows the GPM precipitation field. The precipitation field exhibits a southwest-to-northeast band across the study area, with higher precipitation intensities concentrated in the southwestern part. This precipitation pattern provides supporting evidence for interpreting the complex hydrometeor distribution observed by radar.

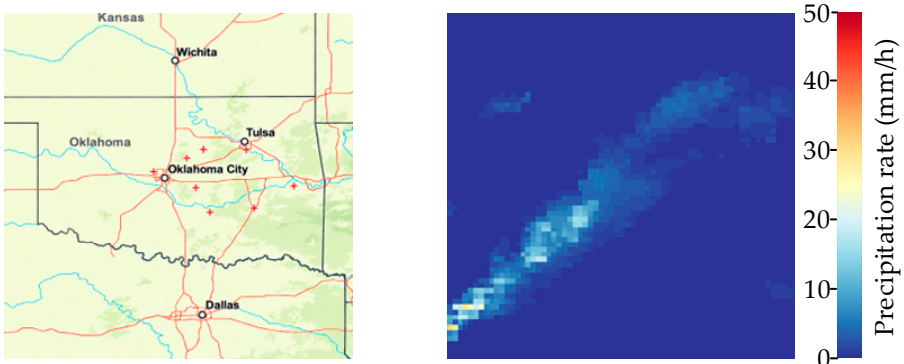


Figure 16. Independent observational information for the complex hydrometeor case at 01:16 UTC on 28 April 2021: (**a**) NOAA ground-station precipitation records; (**b**) GPM satellite precipitation data.

The corresponding hydrometeor classification results are shown in Figure 17. The complex case contains classified regions of light/moderate rain (light green), heavy rain (dark green), hail (red), and scattered big-drops rain (gold).

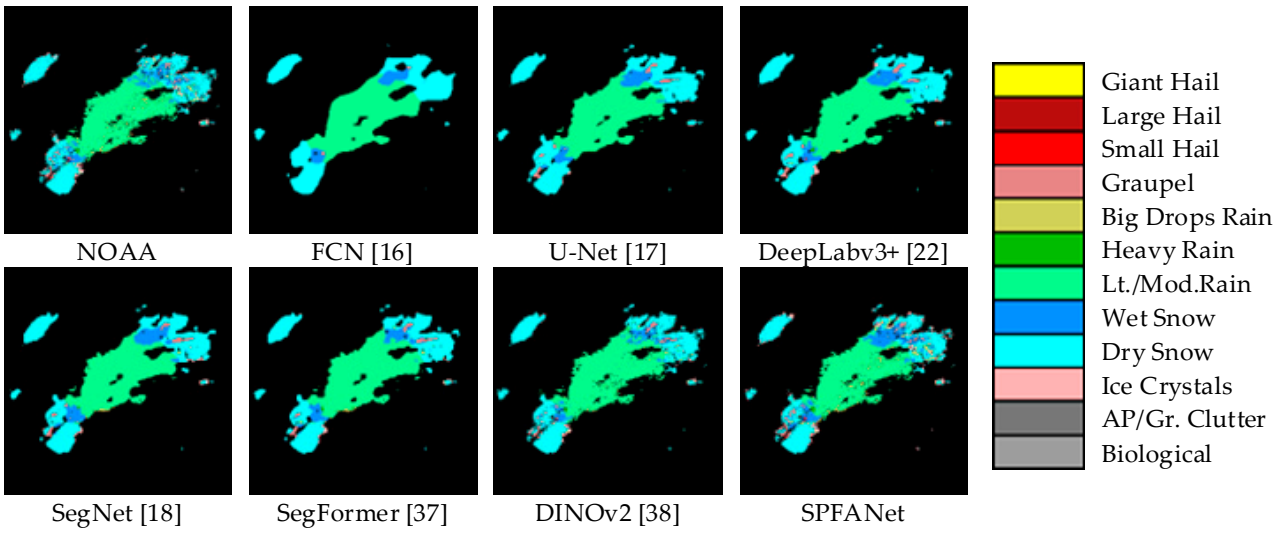


Figure 17. Hydrometeor classification results for the complex case at 01:16 UTC on 28 April 2021. The subplots show classification outputs from NOAA, FCN [16], U-Net [17], DeepLabv3+ [22], SegNet [18], SegFormer [37], DINOv2 [38], and the proposed SPFANet model, respectively. The color bar on the right indicates different hydrometeor categories.

In this case, SPFANet performs better in identifying small-scale regions of heavy rain, big-drops rain, and hail embedded within broad rain and snow areas. To further examine the local classification results of the proposed method, Figures 18 and 19 present enlarged views of two local areas in Figure 17. In the southwestern area shown in Figure 18, an unstable vertical atmospheric structure promotes frequent phase transitions of hydrometeors, leading to the coexistence of heavy rain, big-drops rain, and hail within a limited area. SPFANet provides a clearer representation of light/moderate rain, heavy rain, big-drops rain, dry snow, wet snow, and hail in this area.

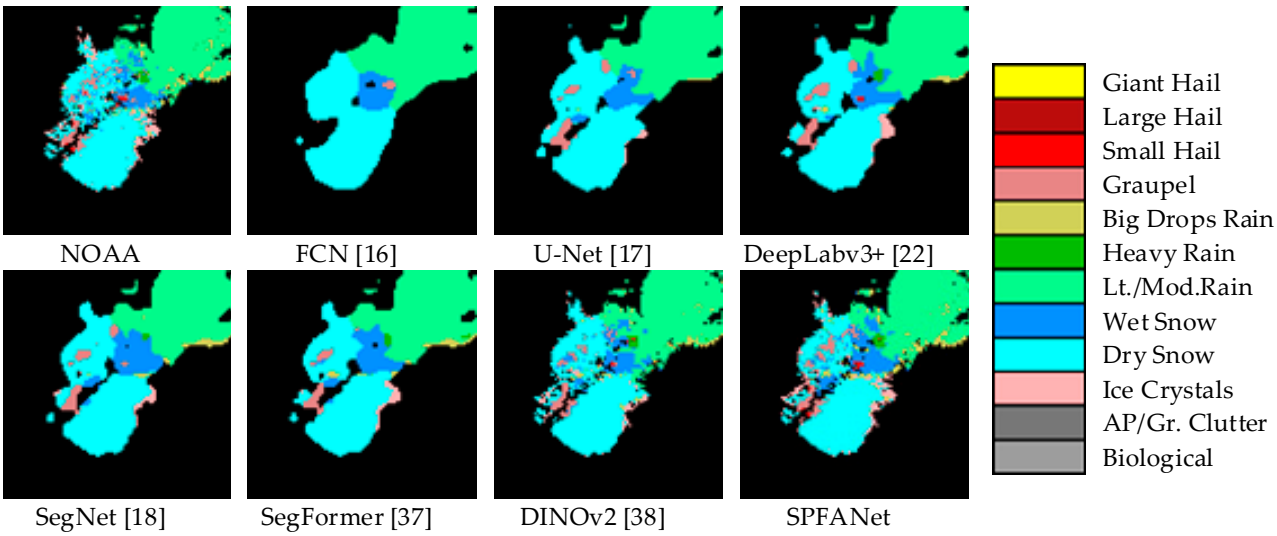


Figure 18. Local enlargement of the hail and heavy rain region. The subplots show classification outputs from NOAA, FCN [16], U-Net [17], DeepLabv3+ [22], SegNet [18], SegFormer [37], DINOv2 [38], and the proposed SPFANet model, respectively. The color bar on the right indicates different hydrometeor categories.

As shown in Figure 19, the northeastern area exhibits another complex mixed-hydrometeor pattern. In this area, collisions and melting among precipitation particles cause big-drops rain to appear sporadically among dry snow, wet snow, and light/moderate

rain. SPFANet improves the spatial representation of big-drops rain while reducing the missed detections observed in the comparison models.

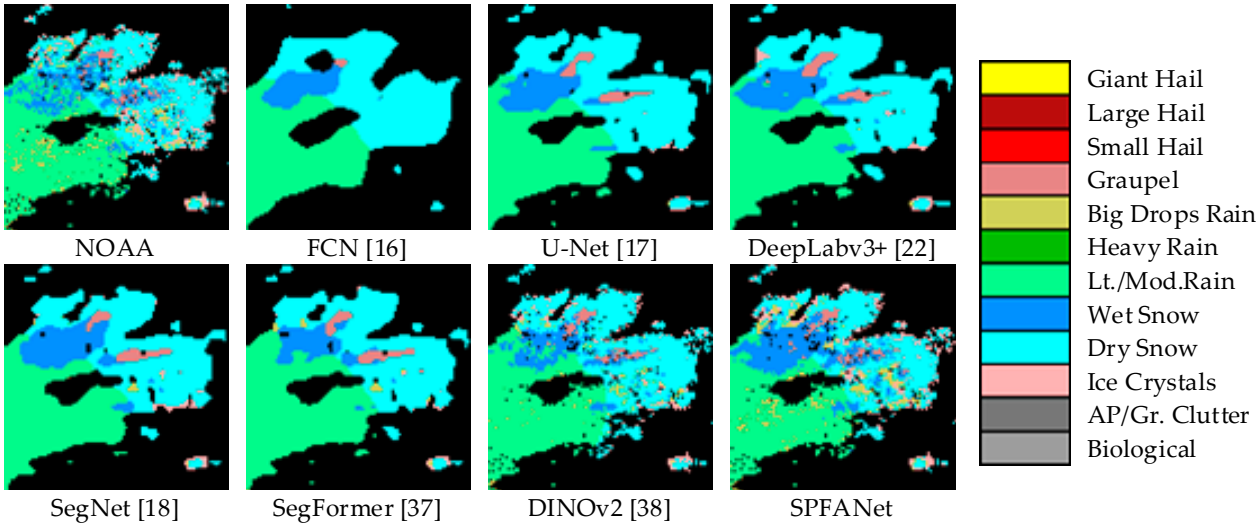


Figure 19. Local enlargement of the big-drops rain region. The subplots show classification outputs from NOAA, FCN [16], U-Net [17], DeepLabv3+ [22], SegNet [18], SegFormer [37], DINOv2 [38], and the proposed SPFANet model, respectively. The color bar on the right indicates different hydrometeor categories.

Overall, the visual results show that SPFANet provides clearer spatial representations of hydrometeor categories across different precipitation scenarios.

5.6. Generalization Across Sites and Seasons

To further evaluate the generalization ability of SPFANet across sites and seasons, an additional test set was constructed using observations from KVNX. KVNX is also located in Oklahoma and is approximately 108 mi (174 km) northwest of KTLX. Its convective-weather background is generally similar to that of KTLX. Compared with the warm-season KTLX dataset, the KVNX test set provides an evaluation period covering the full year from 1 January to 31 December 2025 and contains 37,624 samples. It includes the same radar variables as the original dataset. This test set was used only for evaluation and was not used for training.

The KVNX test set was evaluated using the same trained models and evaluation metrics used for the KTLX test set. The class-wise and overall evaluation results on the KVNX test set are summarized in Tables 10 and 11, respectively.

The class-wise evaluation results on the KVNX test set are generally consistent with those obtained on the KTLX test set. SPFANet achieves the highest mean *Recall*, *F*1, and *IoU* among all models for ice crystals, graupel, and hail. For graupel and hail, it also obtains the highest mean *Precision*. These results show that SPFANet maintains its classification advantage for ice crystals, graupel, and hail on the KVNX test set.

Compared with the KTLX results in Table 4, most models perform better for ice crystals, snow, and rain on the KVNX test set. By contrast, they generally perform worse for graupel and hail. These differences may be related to the broader seasonal coverage of the KVNX test set. In particular, the proportions of rain and snow pixels are higher in the full-year KVNX test set than in the warm-season KTLX test set, further increasing the imbalance among hydrometeor categories. Consistent with this pattern, SPFANet shows increases of 6.0 and 5.7 percentage points in mean *Precision* for ice crystals and snow, respectively, indicating improved prediction reliability for these two categories.

Table 10. Class-wise evaluation metrics on the KVNX test set.

Metric	Model	Ice Crystals	Snow	Rain	Graupel	Hail
Precision (%)	FCN [16]	42.2 ± 0.76	78.4 ± 0.73	90.5 ± 0.70	44.2 ± 0.63	2.4 ± 0.97
	U-Net [17]	47.6 ± 0.61	81.4 ± 0.49	93.5 ± 0.40	53.5 ± 0.39	12.6 ± 0.84
	DeepLabv3+ [22]	**55.7 ± 0.54**	80.8 ± 0.38	**93.8 ± 0.30**	54.3 ± 0.55	49.6 ± 0.89
	SegNet [18]	38.8 ± 1.06	**85.6 ± 0.57**	89.8 ± 0.42	47.4 ± 0.62	34.4 ± 0.93
	SegFormer [37]	50.7 ± 0.45	82.7 ± 0.36	92.7 ± 0.27	56.1 ± 0.65	50.6 ± 0.68
	DINOv2 [38]	32.2 ± 0.92	52.3 ± 0.71	87.5 ± 0.53	84.5 ± 0.32	48.5 ± 0.68
	SPFANet	36.5 ± 0.47	75.5 ± 0.54	92.8 ± 0.25	**88.6 ± 0.46**	**61.6 ± 0.65**
Recall (%)	FCN [16]	1.9 ± 0.17	91.3 ± 0.96	90.4 ± 0.45	43.4 ± 0.64	0.1 ± 0.02
	U-Net [17]	5.4 ± 0.68	91.5 ± 0.68	89.4 ± 0.48	57.6 ± 0.39	7.6 ± 0.51
	DeepLabv3+ [22]	10.3 ± 0.59	**94.5 ± 0.61**	89.7 ± 0.21	60.3 ± 0.56	16.6 ± 0.47
	SegNet [18]	10.3 ± 0.85	92.1 ± 0.78	90.5 ± 0.42	53.6 ± 0.51	1.7 ± 0.46
	SegFormer [37]	7.7 ± 0.35	92.2 ± 0.66	**95.7 ± 0.28**	54.6 ± 0.37	11.2 ± 0.62
	DINOv2 [38]	89.6 ± 0.58	91.6 ± 0.37	84.6 ± 0.46	68.4 ± 0.70	72.0 ± 0.86
	SPFANet	**90.5 ± 0.42**	92.4 ± 0.68	91.6 ± 0.37	**81.6 ± 0.42**	**78.2 ± 0.64**
F1 (%)	FCN [16]	3.6 ± 0.31	84.4 ± 0.61	90.5 ± 0.53	43.8 ± 0.61	0.2 ± 0.03
	U-Net [17]	9.8 ± 1.09	86.2 ± 0.54	91.4 ± 0.44	55.5 ± 0.39	9.5 ± 0.63
	DeepLabv3+ [22]	17.4 ± 0.85	87.1 ± 0.48	91.7 ± 0.17	57.1 ± 0.55	24.8 ± 0.51
	SegNet [18]	16.2 ± 1.09	**88.7 ± 0.49**	90.1 ± 0.25	50.3 ± 0.56	3.1 ± 0.84
	SegFormer [37]	13.3 ± 0.53	87.2 ± 0.48	**94.2 ± 0.21**	55.4 ± 0.51	18.4 ± 0.86
	DINOv2 [38]	47.4 ± 1.03	66.6 ± 0.62	86.0 ± 0.47	75.6 ± 0.52	57.9 ± 0.72
	SPFANet	**52.0 ± 0.51**	83.1 ± 0.47	92.2 ± 0.20	**85.0 ± 0.39**	**68.9 ± 0.52**
IoU (%)	FCN [16]	1.8 ± 0.16	73.0 ± 0.91	82.6 ± 0.88	28.0 ± 0.50	0.1 ± 0.02
	U-Net [17]	5.1 ± 0.60	75.7 ± 0.84	84.2 ± 0.75	38.4 ± 0.37	5.0 ± 0.35
	DeepLabv3+ [22]	9.5 ± 0.51	77.1 ± 0.75	84.7 ± 0.30	40.0 ± 0.54	14.2 ± 0.33
	SegNet [18]	8.8 ± 0.65	**79.7 ± 0.79**	82.0 ± 0.42	33.6 ± 0.50	1.6 ± 0.43
	SegFormer [37]	7.1 ± 0.30	77.3 ± 0.75	**89.0 ± 0.38**	38.3 ± 0.48	10.1 ± 0.52
	DINOv2 [38]	31.0 ± 0.88	49.9 ± 0.69	75.5 ± 0.72	60.8 ± 0.67	40.8 ± 0.72
	SPFANet	**35.1 ± 0.47**	71.0 ± 0.69	85.6 ± 0.35	**73.9 ± 0.60**	**52.5 ± 0.60**

Table 11. Overall evaluation metrics on the KVNX test set.

Model	*Macro-F1* (%)	*mIoU* (%)	*OA* (%)	*BA* (%)
FCN [16]	44.5 ± 0.27	37.1 ± 0.33	59.6 ± 0.43	45.4 ± 0.25
U-Net [17]	50.5 ± 0.44	41.7 ± 0.42	67.0 ± 0.34	50.3 ± 0.35
DeepLabv3+ [22]	55.6 ± 0.40	45.1 ± 0.34	69.4 ± 0.35	54.3 ± 0.35
SegNet [18]	49.7 ± 0.49	41.1 ± 0.42	64.9 ± 0.42	49.6 ± 0.43
SegFormer [37]	53.7 ± 0.44	44.4 ± 0.43	67.9 ± 0.34	52.3 ± 0.39
DINOv2 [38]	66.7 ± 0.48	51.6 ± 0.55	76.1 ± 0.47	81.2 ± 0.40
SPFANet	**76.2 ± 0.36**	**63.6 ± 0.47**	**85.6 ± 0.35**	**86.9 ± 0.39**

The overall evaluation results in Table 11 further indicate the performance advantage of SPFANet on the KVNX test set. SPFANet obtains the highest *Macro-F1*, *mIoU*, *OA*, and *BA*, with mean values of 76.2%, 63.6%, 85.6%, and 86.9%, respectively. Compared with the KTLX test set, SPFANet maintains similar overall performance on the KVNX test set. This result indicates that SPFANet has good generalization ability across sites and seasons under a generally similar meteorological background. However, the present evaluation does not establish the transferability of SPFANet to different radar networks, climates, or radar configurations.

6. Discussion

Evaluation against the NOAA reference labels shows that the proposed pipeline is effective for hydrometeor classification. This performance improvement reflects the combined effects of radar input processing and model design. Hydrometeor classification performance depends on data quality, the physical consistency of radar inputs, and network design.

The proposed non-meteorological echo removal method improves data reliability by reducing contamination from non-meteorological echoes. In addition, the multi-feature radar image reduces scale differences among radar variables and integrates their complementary information into a unified input representation. Radar input processing experiments show that these processing steps improve both class-wise and overall classification performance.

Another advantage of SPFANet is its strong performance on minority categories such as graupel and hail. These categories usually occupy small spatial areas and often exhibit radar signatures similar to those of other hydrometeors, which makes accurate classification difficult. Although the CLIP visual encoder was pretrained on image-text pairs rather than radar data, it can still capture spatial structures in the synthesized radar images. These structures reflect differences in the spatial distributions of hydrometeor categories. By combining multi-level semantic representations from the frozen CLIP branch with multiscale spatial feature aggregation and latent polarimetric feature enhancement in the SPFA branch, SPFANet improves the recognition of these challenging categories. The *WCE* loss function also contributes by mitigating class imbalance and encouraging the model to focus more strongly on meteorologically important categories. The ablation experiments further quantify the contributions of the CLIP branch, SA module, PFA module, and *WCE* loss function. These results demonstrate that the model design provides additional improvements beyond radar input processing.

Despite these advantages, several limitations remain. First, some categories are merged for evaluation. Although this simplifies result interpretation, it may obscure finer distinctions among the original NOAA classes. Second, the NOAA hydrometeor classification product is used as the reference label, so the quantitative metrics mainly reflect agreement with the NOAA reference labels rather than independent pixel-level validation. Finally, confusion among physically similar categories still persists in some transition areas. Future work should therefore examine stronger physical constraints and the integration of additional atmospheric information.

7. Conclusions

This study proposes a hydrometeor classification method based on SPFANet for dual-polarization weather radar observations. The method combines a cascaded procedure for removing non-meteorological echoes with a multi-feature radar image synthesis strategy. SPFANet is designed as a semantic segmentation network that integrates multi-level semantic representations, aggregates multiscale spatial features, and enhances latent polarimetric features. A class-weighting strategy is further designed for the *WCE* loss function to reduce the influence of class imbalance. Because the quantitative evaluation used the NOAA reference labels, the reported results reflect agreement with these labels rather than independent pixel-level validation. Experimental results show that SPFANet performs robustly across different precipitation scenarios. Across five independent runs, only SPFANet achieves mean *Recall* values above 81% for all five hydrometeor categories. It also obtains the highest mean $F1$ and *IoU* for ice crystals, graupel, and hail, as well as the highest mean *Precision* for hail. The overall evaluation metrics further show that SPFANet achieves the highest mean *Macro-F*1, *mIoU*, *OA*, and *BA*. Overall, SPFANet provides a useful basis for precipitation analysis and severe weather monitoring.

Author Contributions: Conceptualization, Z.L. and J.G.; methodology, J.G.; software, J.G.; validation, Z.L. and J.G.; formal analysis, J.G.; investigation, X.Z.; resources, Z.L.; data curation, X.Z.; writing—original draft preparation, J.G.; writing—review and editing, Z.L. and G.W.; project administration, G.W.; funding acquisition, Z.L. 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: For this study, we used an open-source and governmental dataset from the address: National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Information (https://www.ncdc.noaa.gov/nexradinv/, accessed on 12 October 2026). The dataset was constructed from KTLX NEXRAD Level-III radar products collected from 1 April to 30 September 2021. Training, validation, and test samples were randomly divided by observation date. An additional test set was constructed from KVNX NEXRAD Level-III radar products collected from 1 January to 31 December 2025 and was used only for generalization evaluation. For radar product export, the NOAA Weather and Climate Toolkit (WCT) is publicly available from NOAA/NCEI (https://www.ncei.noaa.gov/products/weather-climate-toolkit, accessed on 12 October 2026). The code used in this study is available from the first author upon reasonable request.

Acknowledgments: During the preparation of this manuscript, the authors used ChatGPT (OpenAI, GPT-5.6) for grammar checking and spelling correction.

Conflicts of Interest: Xu Zhang is employed by State Grid Tianjin Electric Power Company Chengnan Power Supply Branch. The remaining authors declare no competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A

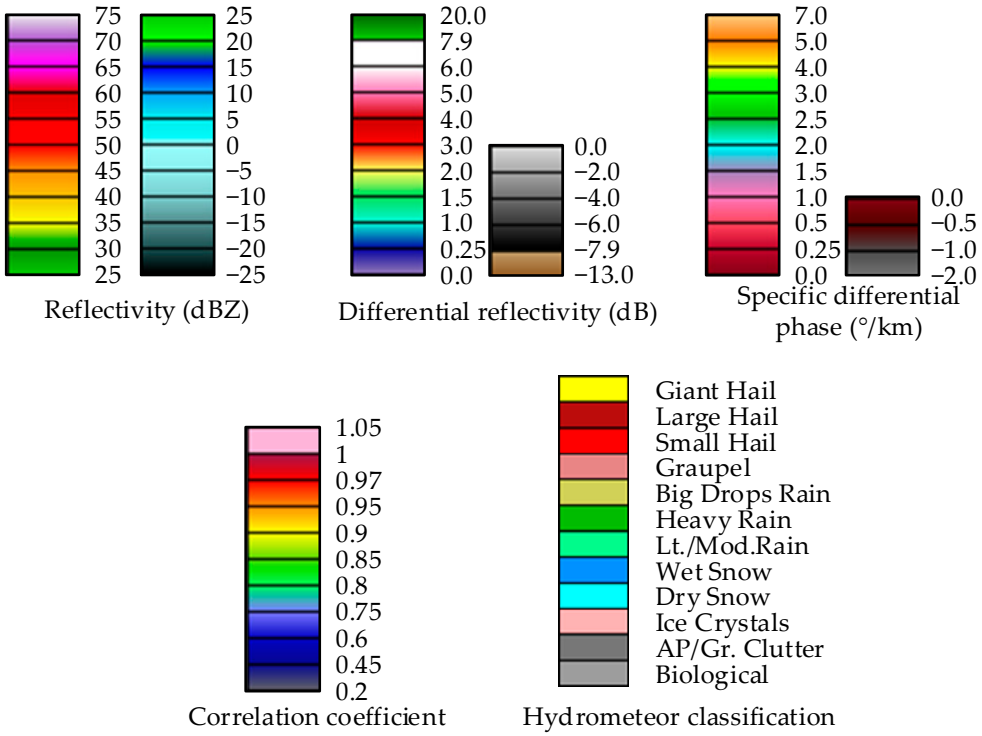


Figure A1. Default WCT color tables used for radar-variable visualization. Abbreviations: Lt./Mod. Rain, light/moderate rain; AP/Gr. Clutter, anomalous propagation/ground clutter.

Table A1. Mapping intervals for reflectivity (Z_H, dBZ).

Original Interval	Mapped Interval	Original Interval	Mapped Interval	Original Interval	Mapped Interval
[0, 5]	[0, 16]	(5, 10]	(16, 33]	(10, 15]	(33, 50]
(15, 20]	(50, 67]	(20, 25]	(67, 84]	(25, 30]	(84, 101]
(30, 35]	(101, 118]	(35, 40]	(118, 135]	(40, 45]	(135, 152]
(45, 50]	(152, 169]	(50, 55]	(169, 187]	(55, 60]	(187, 204]
(60, 65]	(204, 221]	(65, 70]	(221, 238]	(70, 75]	(238, 255]

Table A2. Mapping intervals for differential reflectivity (Z_{DR}, dB).

Original Interval	Mapped Interval	Original Interval	Mapped Interval	Original Interval	Mapped Interval
[−13, −7.9]	[0, 16]	(−7.9, −6]	(16, 33]	(−6, −4]	(33, 50]
(−4, −2]	(50, 67]	(−2, 0]	(67, 84]	(0, 0.25]	(84, 101]
(0.25, 1]	(101, 118]	(1, 1.5]	(118, 135]	(1.5, 2]	(135, 152]
(2, 3]	(152, 169]	(3, 4]	(169, 187]	(4, 5]	(187, 204]
(5, 6]	(204, 221]	(6, 7.9]	(221, 238]	(7.9, 20]	(238, 255]

Table A3. Mapping intervals for the correlation coefficient (ρ_{hv}, dimensionless).

Original Interval	Mapped Interval	Original Interval	Mapped Interval	Original Interval	Mapped Interval
[0.2, 0.45]	[0, 29]	(0.45, 0.6]	(29, 57]	(0.6, 0.75]	(57, 86]
(0.75, 0.8]	(86, 114]	(0.8, 0.85]	(114, 143]	(0.85, 0.9]	(143, 172]
(0.9, 0.95]	(172, 200]	(0.95, 0.97]	(200, 229]	(0.97, 1]	(229, 255]
(1, 1.05]	255				

Table A4. Pixel counts and class weights for the original hydrometeor categories.

Category	Pixel Count	Weight	Category	Pixel Count	Weight
Background	923,660,774	0.00355	Giant hail	2585	10
Large hail	5383	10	Small hail	176,259	10
Graupel	4,857,521	0.675	Big-drops rain	2,257,568	1.45
Heavy rain	382,155	8.58	Light/moderate rain	17,262,720	0.19
Wet snow	3,277,226	1	Dry snow	26,632,073	0.123
Ice crystals	4,656,808	0.704			

Table A5. Observation dates assigned to the training, validation, and test sets in 2021.

Month	Training Dates	Validation Dates	Test Dates
April	All other dates	9	12–16, 20, 28
May	All other dates	15–16, 21, 31	4, 7, 12, 26, 28–29
June	All other dates	1, 13, 16, 20	11, 18, 21, 28
July	All other dates	3, 13, 25, 27	5, 10–12, 21, 24, 31
August	All other dates	5, 16	11, 14, 17, 20, 23–24
September	All other dates	13, 19, 27	5–6, 11, 22, 24, 26

References

1. Keem, M.; Seo, B.-C.; Krajewski, W.F.; Kim, S. Data-Driven Non-Precipitation Echo Removal of NEXRAD Radars Based on a Random Forest Classifier Using Polarimetric Observations and GOES-16 Data. *Remote Sens.* **2026**, *18*, 827. [CrossRef]
2. Li, X.; Mecikalski, J.R.; Otkin, J.A.; Henderson, D.S.; Srikishen, J. A Polarimetric Radar Operator and Application for Convective Storm Simulation. *Atmosphere* **2022**, *13*, 645. [CrossRef]

3. Lukach, M.; Dufton, D.; Crosier, J.; Hampton, J.M.; Bennett, L.; Neely, R.R. Hydrometeor classification of quasi-vertical profiles of polarimetric radar measurements using a top-down iterative hierarchical clustering method. *Atmos. Meas. Tech.* **2021**, *14*, 1075–1098. [CrossRef]
4. Feng, N.; Shu, Z.; Qiu, Y. Classification of Hydrometeors During a Stratiform Precipitation Event in the Rainy Season of Liupanshan. *Atmosphere* **2025**, *16*, 132. [CrossRef]
5. Li, N.; Zhang, J.; Wang, D.; Wang, P. Research on Hail Mechanism Features Based on Dual-Polarization Radar Data. *Atmosphere* **2023**, *14*, 1827. [CrossRef]
6. Ryzhkov, A.; Zhang, P.; Bukovcic, P.; Zhang, J.; Cocks, S. Polarimetric radar quantitative precipitation estimation. *Remote Sens.* **2022**, *14*, 1695. [CrossRef]
7. Kim, Y.; Ye, B.-Y.; Suk, M.-K. AIDER: Aircraft icing potential area detection in real-time using 3-dimensional radar and atmospheric variables. *Remote Sens.* **2024**, *16*, 1468. [CrossRef]
8. Matsuo, T.; Sasyo, Y. Melting of snowflakes below freezing level in the atmosphere. *J. Meteorol. Soc. Jpn. Ser. II* **1981**, *59*, 10–25. [CrossRef] [PubMed]
9. Mroz, K.; Battaglia, A.; Fridlind, A.M. Enhancing consistency of microphysical properties of precipitation across the melting layer in dual-frequency precipitation radar data. *Atmos. Meas. Tech.* **2024**, *17*, 1577–1597. [CrossRef]
10. Wang, L.; Wang, Y.; Liu, M.; Chen, W.; Li, C. Analysis of Dual-Polarimetric Radar Observations of Precipitation Phase during Snowstorm Events in Jiangsu Province, China. *Atmosphere* **2024**, *15*, 321. [CrossRef]
11. Dai, S.; Li, X.; Bu, Z.; Chen, Y.; He, J.; Li, M.; Xiong, M. Signal Simulation of Dual-Polarization Weather Radar and Its Application in Range Ambiguity Mitigation. *Atmosphere* **2022**, *13*, 432. [CrossRef]
12. Lim, S.; Chandrasekar, V.; Bringi, V.N. Hydrometeor classification system using dual-polarization radar measurements: Model improvements and in situ verification. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 792–801. [CrossRef]
13. Al-Sakka, H.; Boumahmoud, A.-A.; Fradon, B.; Frasier, S.J.; Tabary, P. A new fuzzy logic hydrometeor classification scheme applied to the French X-, C-, and S-band polarimetric radars. *J. Appl. Meteorol. Climatol.* **2013**, *52*, 2328–2344. [CrossRef]
14. Besic, N.; Figueras i Ventura, J.; Grazioli, J.; Gabella, M.; Germann, U.; Berne, A. Hydrometeor classification through statistical clustering of polarimetric radar measurements: A semi-supervised approach. *Atmos. Meas. Tech.* **2016**, *9*, 4425–4445. [CrossRef]
15. Wang, H.; Ran, Y.; Deng, Y.; Wang, X. Study on deep-learning-based identification of hydrometeors observed by dual polarization Doppler weather radars. *EURASIP J. Wirel. Commun. Netw.* **2017**, *2017*, 173. [CrossRef]
16. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
17. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; pp. 234–241.
18. Badrinarayanan, V.; Kendall, A.; Cipolla, R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [CrossRef] [PubMed]
19. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. Semantic image segmentation with deep convolutional nets and fully connected CRFs. *arXiv* **2014**, arXiv:1412.7062.
20. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 834–848. [CrossRef] [PubMed]
21. Chen, L.C.; Papandreou, G.; Schroff, F.; Adam, H. Rethinking atrous convolution for semantic image segmentation. *arXiv* **2017**, arXiv:1706.05587.
22. Chen, L.C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European Conference on Computer Vision, Munich, Germany, 8–14 September 2018; pp. 833–851.
23. Lu, Y.; Kumar, J. Convolutional neural networks for hydrometeor classification using dual polarization Doppler radars. In Proceedings of the IEEE International Conference on Data Mining Workshops, Beijing, China, 8–11 November 2019; pp. 288–295.
24. Chen, X.; Li, D.; Liu, M.; Jia, J. CNN and transformer fusion for remote sensing image semantic segmentation. *Remote Sens.* **2023**, *15*, 4455. [CrossRef]
25. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention is all you need. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; pp. 5998–6008.
26. Zheng, S.; Lu, J.; Zhao, H.; Zhu, X.; Luo, Z.; Wang, Y.; Fu, Y.; Feng, J.; Xiang, T.; Torr, P.H.S.; et al. Rethinking semantic segmentation from a sequence-to-sequence perspective with transformers. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 6881–6890.

27. Gu, J.; Kwon, H.; Wang, D.; Ye, W.; Li, M.; Chen, Y.H.; Lai, V.; Chandra, V.; Pan, D.Z. Multi-scale high-resolution vision transformer for semantic segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 19–24 June 2022; pp. 12094–12103.
28. Xu, W. Study on Cloud Particle Phase Identification Algorithms Based on Machine Learning. Master's Thesis, Nanjing University of Information Science and Technology, Nanjing, China, 2023. (In Chinese)
29. Guzel, M.; Kalkan, M.; Bostanci, E.; Celik, M.; Alsharif, M.H. Cloud type classification using deep learning with cloud images. *PeerJ Comput. Sci.* **2024**, *10*, e1779. [CrossRef] [PubMed]
30. Kopec, S.; Duniec, G.; Bochenek, B.; Netzel, P.; Markowicz, K.M. Artificial neural networks in automatic image classification of clouds from ground-based observations using deep learning models. *Q. J. R. Meteorol. Soc.* **2024**, *150*, 5206–5224. [CrossRef]
31. Yang, X.; Yan, W.; Yuan, Y.; Wang, Q.; Zhang, H.; Li, X. Semantic segmentation in multiple adverse weather conditions with domain knowledge retention. In Proceedings of the AAAI Conference on Artificial Intelligence, Vancouver, BC, Canada, 20–27 February 2024; pp. 6558–6566.
32. Kerim, A.; Chamone, F.; Ramos, W.; Franchi, G.; Yamasaki, T.; de Charette, R. Semantic segmentation under adverse conditions: A weather- and nighttime-aware synthetic data-based approach. In Proceedings of the British Machine Vision Conference, London, UK, 21–24 November 2022; pp. 977–991.
33. Shen, F.; Shu, A.; Min, J.; Wu, Z.; Wang, Y.; Xu, D.; Chen, J.; Wan, S. Assimilation of dual-pol radar K_{DP} observations with the GSI ensemble Kalman filter for the analysis and prediction of a squall line. *J. Geophys. Res. Atmos.* **2025**, *130*, e2024JD041933. [CrossRef]
34. Radford, A.; Kim, J.W.; Hallacy, C.; Ramesh, A.; Goh, G.; Agarwal, S.; Sastry, G.; Askell, A.; Mishkin, P.; Clark, J.; et al. Learning transferable visual models from natural language supervision. In Proceedings of the 38th International Conference on Machine Learning, Virtual, 18–24 July 2021; pp. 8748–8763.
35. Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 10–17 October 2021; pp. 10012–10022.
36. Katharopoulos, A.; Vyas, A.; Pappas, N.; Fleuret, F. Transformers are RNNs: Fast autoregressive transformers with linear attention. In Proceedings of the 37th International Conference on Machine Learning, Virtual, 13–18 July 2020; pp. 5156–5165.
37. Xie, E.; Wang, W.; Yu, Z.; Anandkumar, A.; Alvarez, J.M.; Luo, P. SegFormer: Simple and efficient design for semantic segmentation with transformers. In Proceedings of the Advances in Neural Information Processing Systems, Virtual, 6–14 December 2021.
38. Oquab, M.; Darcet, T.; Moutakanni, T.; Vo, H.; Szafraniec, M.; Khalidov, V.; Fernandez, P.; Haziza, D.; Massa, F.; El-Nouby, A.; et al. DINOv2: Learning robust visual features without supervision. *arXiv* **2023**, arXiv:2304.07193.

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.