

## Article

# DSADNet: A Dual-Source Attention Dynamic Neural Network for Precipitation Nowcasting

Jinliang Yao <sup>1,2</sup>, Junwei Ji <sup>1,\*</sup>, Rongbo Wang <sup>1,2</sup>, Xiaoxi Huang <sup>1,2</sup> , Zhiming Kang <sup>3,\*</sup> and Xiaoran Zhuang <sup>3</sup>

<sup>1</sup> School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou 310018, China; yaojinl@hdu.edu.cn (J.Y.); wangrongbo@hdu.edu.cn (R.W.); huangxx@hdu.edu.cn (X.H.)

<sup>2</sup> Shangyu Institute of Science and Engineering, Hangzhou Dianzi University, Shaoxing 312000, China

<sup>3</sup> Jiangsu Meteorological Observatory, Nanjing 210008, China; zrxz3212009@163.com

\* Correspondence: 222050126@hdu.edu.cn (J.J.); kangzm@cma.gov.cn (Z.K.)

**Abstract:** Accurate precipitation nowcasting is of great significance for flood prevention, agricultural production, and public safety. In recent years, spatiotemporal sequence models based on deep learning have been widely used for precipitation nowcasting and have achieved better prediction results than traditional methods. These models commonly use radar echo extrapolation and utilize the Z-R relationship between radar and rainfall to predict rainfall. However, radar echo data can be affected by various noises, and the Z-R correlation linking radar and rainfall encompasses several variables influenced by factors like terrain, climate, and seasonal variations. To solve this problem, we propose a dual-source attention dynamic neural network (DSADNet) for precipitation nowcasting, which is a network model that utilizes a fusion module to extract valid information from radar maps and rainfall maps, together with dynamic convolution and the attention mechanism, to directly predict future rainfall through encoding and decoding structure. We conducted experiments on a real dataset in Jiangsu, China, and the experimental results show that our model had better performance than the other examined models.

**Keywords:** precipitation nowcasting; DSADNet; fusion module; dynamic convolution; attention mechanism



**Citation:** Yao, J.; Ji, J.; Wang, R.; Huang, X.; Kang, Z.; Zhuang, X. DSADNet: A Dual-Source Attention Dynamic Neural Network for Precipitation Nowcasting. *Sustainability* **2024**, *16*, 3696. <https://doi.org/10.3390/su16093696>

Academic Editor: Mike Spiliotis

Received: 27 March 2024

Revised: 19 April 2024

Accepted: 27 April 2024

Published: 28 April 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Precipitation nowcasting is an effective way to make a relatively short-period prediction of the intensity of rainfall at the kilometer level in a local area, which can help local meteorological departments to make timely and accurate decisions about rainfall warnings, and to guide governmental departments in the transportation, agriculture, and electric power industries to make responses in advance, thus reducing potential economic losses [1]. The role of heavy precipitation weather forecasting in today's society cannot be overemphasized. Intense rainfall within a brief timeframe has the potential to trigger natural calamities like floods and landslides, leading to significant societal harm and adverse effects on ecological stability. Accurate rainfall forecasting holds immense significance for pre-emptive flood management, agricultural productivity, and sustainable progress, empowering proactive measures. Currently, there exist two primary methodologies for precipitation nowcasting: numerical weather prediction (NWP) [2] and radar echo extrapolation. NWP relies on numerical simulations derived from a series of mathematical physics equations, which involves the calculation of complex physics equations, but it is difficult to meet the requirements of accuracy and real-time precipitation prediction due to the uncertainty of its initial and boundary conditions, the incompleteness of the physical mechanism, and the fact that the technique involves the calculation of complex physics equations [3]. With the continuous increase of meteorological data and the rapid development of deep learning technology, most of the current research work on precipitation nowcasting mainly adopts radar echo extrapolation.

Radar echo extrapolation-based precipitation nowcasting methods can be broadly classified into two categories: conventional extrapolation algorithms and deep learning extrapolation algorithms. Among the traditional extrapolation algorithms, numerous algorithms based on optical flow techniques [4–6] have been shown to have better radar echo extrapolation performance. Optical flow techniques, based on the assumption of constant brightness, are globally consistent techniques that can predict the motion of the entire cloud layer. However, in reality, clouds tend to produce aggregation and dissipation phenomena, which can affect the accuracy of optical flow techniques.

With the development of deep learning, some deep learning models have been proposed, such as RNN [7], LSTM [8], and GRU [9] to deal with precipitation nowcasting problems. Both satellite data and radar echo data are important tools used to obtain precipitation information. However, satellite data cannot directly reflect the precipitation inside the cloud, which needs to be judged by the cloud location, shape, cloud top temperature, and other characteristics. Radar echo data, on the other hand, can directly reflect the precipitation situation based on the size of the echo value. Therefore, radar echo data are more suitable for precipitation nowcasting. By learning from a large amount of historical radar echo data, these models establish a mapping relationship from previous observations to future prediction data. Shi et al. [10] innovatively regarded the radar echo extrapolation task as a spatiotemporal sequence prediction problem and designed a ConvLSTM model, which substitutes the fully connected architecture of the FC-LSTM model with a convolutional structure, thus enhancing radar echo extrapolation. This addresses the limitation of RNN models, which primarily emphasize temporal aspects of image sequences while neglecting spatial features. However, the memory units of the ConvLSTM model tend to focus on local spatial dependencies and are inadequate for capturing long-range spatiotemporal dependencies; therefore, Lin et al. [11] introduced the self-attention memory (SAM) module, enabling effective feature extraction with extended spatiotemporal dependencies via the self-attention mechanism. This facilitates improved accuracy in forecasting future radar echo states. Nevertheless, as the lead time increases, these methods tend to underestimate areas with high reflectivity, which typically indicate intense precipitation. This underestimation of such regions can have significant repercussions. Furthermore, many previous models solely rely on radar maps to forecast future radar conditions and, subsequently, use the Z-R relationship to derive rainfall maps. While radar echo intensity provides insight into the size and density of precipitation particles to some extent, establishing a relationship between echo intensity and precipitation encounters two common challenges: inconsistency of radar echo data, due to radar's operating principle, and various noises; since the Z-R relationship is empirical, this relationship may change under different meteorological conditions, resulting in inaccurate predicted rainfall data.

Numerous studies have shown that the use of multi-source data is more accurate than the use of single data in precipitation prediction [12]. For this reason, Geng et al. [13] proposed the AF-SRNet model for fusing radar and rainfall data to predict future rainfall data, which extracts the information from radar and rainfall, respectively, through the SRU module, then fuses the extracted radar and rainfall information using the attention fusion module, and, finally, decodes and outputs the future rainfall data. The experimental results show that fusing multi-source data helps to improve the accuracy of prediction. However, the above model uses a late-fusion strategy to extract information from radar and rainfall separately before fusing the information, which does not effectively extract the evolution of the radar data in the two phases and only fuses the radar and rainfall data at the same moment, which fails to capture the aggregation and dissipation of the radar echo intensity in the two phases. Therefore, we propose a dual-source attention dynamic neural network (DSADNet), which utilizes a fusion module (Fusion-Module) to effectively fuse the radar and rainfall information to extract the evolution of the two-phase radar data with higher prediction accuracy than the AF-SRNet model. Our model directly predicts future rainfall using encoding and decoding structure, which overcomes the problem of inaccurate radar echo data due to noise and inaccurate prediction due to the inability to accurately

determine its parameters using the Z-R relationship. Rainfall data can accurately reflect the real rainfall situation through direct measurements such as rain gauges. However, rainfall data may not fully capture climate fluctuations, whereas radar echo data can depict the aggregation and dissipation of echo intensity, and the fusion module designed by us can better fuse the effective information in the radar echo data and the rainfall data, as well as better capture the aggregation and dissipation of the radar echo in two phases, which improves the prediction accuracy. To pay more attention to the high-echo heavy-precipitation region, we replaced static convolution with dynamic convolution [14]. To improve the ability to capture long-range spatiotemporal dependencies, we added the SAM [11] module to each module in DSADNet. Numerous experimental results show that DSADNet can effectively predict future rainfall with better metrics than existing models.

## 2. Related Work

Shi et al. [10] innovatively considered the radar echo extrapolation task as a spatiotemporal sequence prediction problem. The authors proposed ConvLSTM by replacing the input-to-state and state-to-state fully connected structures in FC-LSTM with a convolutional structure, and the experimental results show that this model outperforms both the optical flow techniques and FC-LSTM. Afterward, to solve the problem that the convolutional operation in the ConvLSTM model is position-invariant, while the meteorological variations are usually position-variant, Shi et al. [15] introduced the TrajGRU model, capable of autonomously learning motion position changes and achieving superior prediction performance compared to ConvLSTM. ConvGRU [16] has a similar structure and effect as ConvLSTM, but the number of parameters is decreased, which optimizes the model's prediction time.

Wang et al. [17] improved the previous model by proposing a new recurrent neural network framework called predictive recurrent neural network (PredRNN), which allows the memory states between different LSTM units to interact from layer to layer. The authors devised an innovative spatiotemporal LSTM (ST-LSTM) unit, which integrates spatial and temporal relationships within a unified memory unit. This unit facilitates the horizontal and vertical transmission of memories, thereby considering both temporal and spatial states. The experimental results indicated the favorable performance of the model. Since the accuracy of ConvRNN-based precipitation nowcasting methods is greatly affected by the gradient vanishing problem [18,19], Wang et al. [20] improved PredRNN and proposed PredRNN++. This model utilizes a novel recursive structure called Causal LSTM, which has a cascading double memory, thus making the network structure deeper in time. In addition, the authors proposed a gradient highway unit that provides alternative quick routes for the gradient flows from outputs back to long-range previous inputs. Causal LSTM works seamlessly with the gradient highway to solve the problem of gradient vanishing. Following this, Wang et al. [21] introduced the PredRNN-V2 version, which introduced a novel loss function derived from PredRNN and implemented a new sampling strategy, leading to improved prediction outcomes. Wu et al. [22] presented ISA-PredRNN, integrating a self-attention mechanism and long-term memory to enhance capability in handling global and long-term dependencies. Tuyen et al. [23] proposed the RainPredRNN model, which combines PredRNN-V2 and U-Net [24], which not only reduces the amount of computation for model prediction but also improves performance.

Highly non-stationary states such as accumulation, deformation, and dissipation of radar echo in precipitation prediction have a crucial impact on the accuracy of the prediction; therefore, Wang et al. [25] proposed MIM, which utilizes the difference signals between neighboring cyclic states to model non-stationary and approximately stationary states in spatiotemporal dynamics and improves the model's ability to deal with higher-order non-stationary states.

In recent years, the attention mechanism has received much attention, so some ConvRNN models based on the attention mechanism have been proposed for precipitation nowcasting. Lin et al. [11] proposed the self-attention memory (SAM) module and embedded SAM into ConvLSTM. The model can improve the ability to capture long-range

spatiotemporal dependencies through the self-attention mechanism. Since CongGRU has fewer parameters compared to ConvLSTM, but the prediction effect is similar, Zhou et al. [26] embedded the SAM module into ConvGRU and obtained a model with better precipitation nowcasting capability compared to the benchmark model. To amplify the effectiveness of the attention mechanism, Luo et al. [27] introduced a novel interactive dual-attention LSTM (IDA-LSTM) model. This model integrates both channel attention and spatiotemporal attention to retrieve long-term forgotten information.

Previous models have used radar echo extrapolation for precipitation nowcasting tasks. Radar maps are extrapolated to predict future radar conditions, which are subsequently converted into rainfall maps using the Z-R relationship. However, the acquisition of radar echo data is affected by various noises, such as wind speed, humidity in the atmosphere, and the obstruction of mountains and buildings, resulting in inaccurate extrapolated radar maps. Moreover, the Z-R relationship linking radar to rainfall involves numerous parameters that fluctuate based on factors such as terrain, climate, and season. Failure to precisely ascertain these parameters can result in decreased prediction accuracy. Precipitation nowcasting requires a high degree of ability to capture long-range spatiotemporal dependencies as well as to accurately predict areas of heavy precipitation, but ConvRNN models rarely consider both of these aspects simultaneously.

### 3. Methods

#### 3.1. Problem Definition

The precipitation nowcasting model described in our method utilizes past radar echo and rainfall data to forecast future precipitation patterns. Given a historical radar echo sequence  $R_{t-k+1:t} = \{R_{t-k+1}, R_{t-k+2}, \dots, R_t\}$  and a historical rainfall sequence  $P_{t-k+1:t} = \{P_{t-k+1}, P_{t-k+2}, \dots, P_t\}$ , we forecast the upcoming rainfall sequence  $\hat{P}_{t+1:t+m} = \{\hat{P}_{t+1}, \hat{P}_{t+2}, \dots, \hat{P}_{t+m}\}$ . This can be defined as follows:

$$\hat{P}_{t+1}, \dots, \hat{P}_{t+m} = \arg \max(P_{t+1}, \dots, P_{t+m} | R_{t-k+1}, \dots, R_t; P_{t-k+1}, \dots, P_t) \quad (1)$$

where  $k$  denotes the input sequence length and  $m$  denotes the output sequence length.

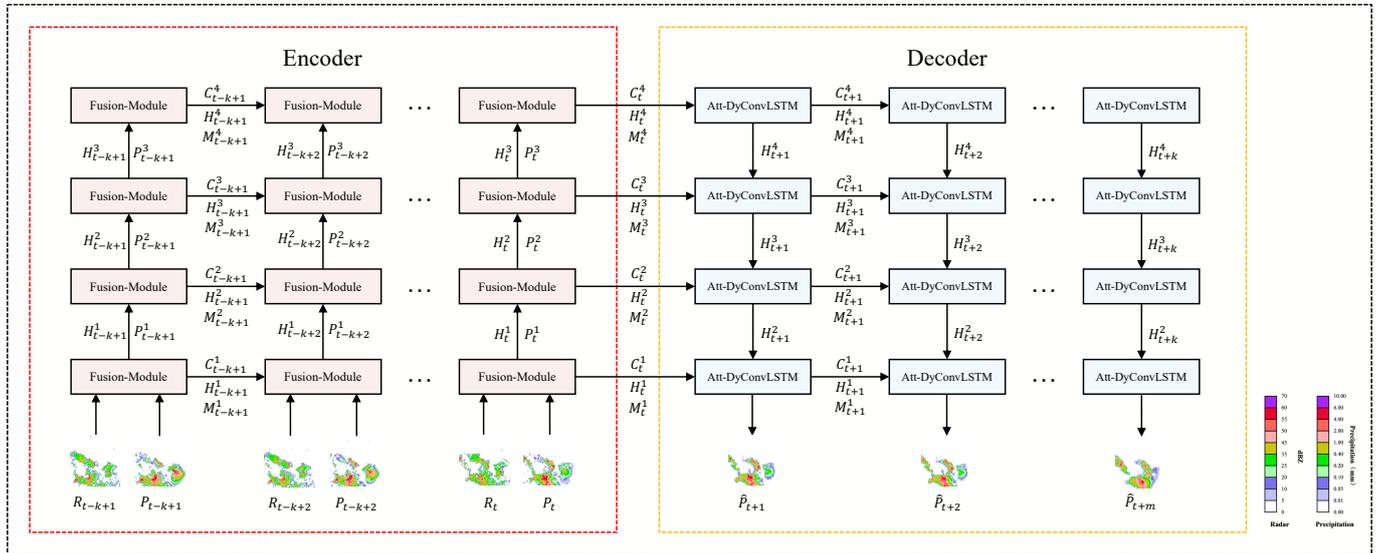
#### 3.2. Model

For precipitation nowcasting, most researchers use the radar echo extrapolation algorithm [28–30], but this does not fully utilize the existing historical data. The collection of radar echo data will be affected by a variety of noises, such as topography, electromagnetic interference, thunderstorms, strong winds, and other meteorological conditions, which greatly affects the accuracy of the extrapolation of radar echoes. Rainfall data can be directly measured using rain gauges and other devices, providing a more accurate representation of actual precipitation conditions, but by using only rainfall data it may not be possible to respond to climate changes such as cloud aggregation or dissipation, which may cause data inconsistency problems. Therefore, in our method in this paper, we designed a new fusion module to improve the accuracy of prediction by fusing radar echo data and rainfall data and effectively extracting the aggregation and dissipation of the radar echo intensity in both phases. Our proposed precipitation nowcasting model adopts the structure of encoding and decoding with dynamic convolution and the attention mechanism to predict rainfall directly, which overcomes the problem of inaccurate prediction due to the need to map the radar map to the rainfall map by using the Z-R relationship in the extrapolation method of radar echoes.

##### 3.2.1. DSADNet

The overall network architecture of DSADNet is illustrated in Figure 1. DSADNet comprises two main components: the encoder, which includes Fusion-Module on the left, and the decoder, which includes Att-DyConvLSTM on the right. Typically, the encoder is constructed with four layers of Fusion-Module to enhance the extraction of deep spatiotemporal features, while the decoder is composed of four layers of Att-DyConvLSTM.

Fusion-Module integrates radar echo and rainfall data, capturing variations in radar echo intensity by encoding data from two phases, and then it merges the rainfall data to consolidate the fused information, which is saved in the cell state  $C_t^l$ , hidden state  $H_t^l$ , and spatiotemporal memory unit  $M_t^l$  and is passed to the decoder; finally, the predicted rainfall data is decoded and output.



**Figure 1.** DSADNet model framework. It comprises both the encoder and decoder structures. The encoder contains Fusion-Module modules and the decoder contains Att-DyConvLSTM modules. The inputs are radar echo maps and rainfall maps, and the outputs are rainfall prediction maps generated by the DSADNet model.

### 3.2.2. Fusion-Module

Radar echo intensity reflects the scale and density of precipitation particles to a certain extent, and the relationship between echo intensity and precipitation can be established, so a radar map can reflect the rainfall situation, but the radar echo data is affected by various kinds of noise, which leads to inaccurate reflection of the rainfall data. While rainfall data provide an accurate representation of actual rainfall conditions, their utilization may overlook factors such as cloud dissipation and atmospheric dynamics, which contribute to climate fluctuations, while radar echo data can reflect climate change, so the joint use of radar and rainfall data can compensate for their respective shortcomings. Therefore, to extract effective information from radar echo data and rainfall data, we designed the Fusion-Module module to effectively fuse radar echo data and rainfall data and to encode and extract the effective information.

The configuration of the Fusion-Module module is depicted in Figure 2. Firstly, the radar echo data  $R_t$ , the cell state  $C_{t-1}^l$ , and the hidden state  $H_{t-1}^l$  of the previous period enter into Fusion-Module, which decides which of the input data are important through the input gate  $i_t$ , which information should be retained or forgotten through the forgetting gate  $f_t$ , and generates candidate information to help the network update the new cell state  $C_t^l$  through the input modulation gate  $g_t$ . The specific formulas are as follows:

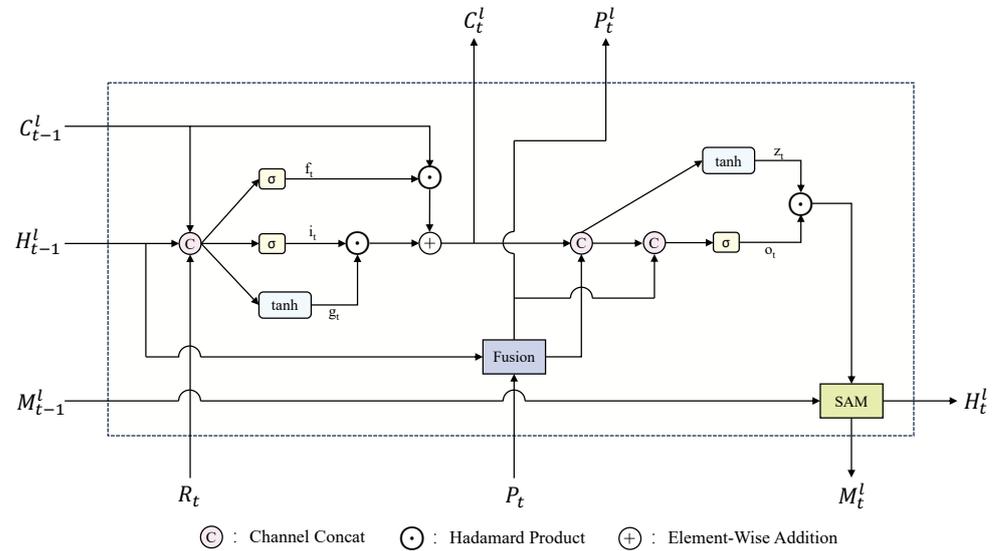
$$i_t = \sigma(Dy(W_{ri}, R_t) + Dy(W_{hi}, H_{t-1}^l) + W_{ci} \odot C_{t-1}^l + b_i) \quad (2)$$

$$f_t = \sigma(Dy(W_{rf}, R_t) + Dy(W_{hf}, H_{t-1}^l) + W_{cf} \odot C_{t-1}^l + b_f) \quad (3)$$

$$g_t = \tanh(Dy(W_{rg}, R_t) + Dy(W_{hg}, H_{t-1}^l) + b_g) \quad (4)$$

$$C_t^l = f_t \odot C_{t-1}^l + i_t \odot g_t \quad (5)$$

where  $\sigma$  is the Sigmoid activation function,  $Dy(\cdot)$  is the dynamic convolution,  $\odot$  is the Hadamard product,  $\tanh$  is the hyperbolic tangent activation function,  $W_{**}$  is the parameter to be learned, and  $b_*$  is the bias value.



**Figure 2.** The structure of Fusion-Module, where  $f_t$  is the forgetting gate,  $i_t$  is the input gate,  $g_t$  is the input modulation gate,  $z_t$  is the update gate, and  $o_t$  is the output gate. This module contains the Fusion-Module module and the SAM module.

Since radar maps can respond to climate changes such as cloud aggregation and dissipation, inspired by [27], we included the Fusion module in Fusion-Module, which fuses the hidden state  $H_{t-1}^l$ , which contains the previous radar information, with the current rainfall data  $P_t$  for extracting the radar and rainfall information, and then fuses the updated data with the cell state  $C_t^l$  containing the current radar information through the update gate  $z_t$  and the output gate  $o_t$ . This module not only extracts the evolution of the radar data in the two time periods, but also fuses the real rainfall data to improve the accuracy of the prediction. The specific formulas are as follows:

$$H_t^l, P_t^l = \text{Fusion}(H_{t-1}^l, P_t) \quad (6)$$

$$z_t = \tanh(Dy(W_{hz}, H_t^l) + W_{cz} \odot C_t^l + b_z) \quad (7)$$

$$o_t = \sigma(Dy(W_{po}, P_t^l) + Dy(W_{ho}, H_t^l) + W_{co} \odot C_t^l + b_o) \quad (8)$$

$$H_t^{l'} = z_t \odot o_t \quad (9)$$

where  $\sigma$  is the Sigmoid activation function,  $Dy(\cdot)$  is the dynamic convolution,  $\odot$  is the Hadamard product,  $\tanh$  is the hyperbolic tangent activation function,  $W_{**}$  is the parameter to be learned,  $b_*$  is the bias value, and  $\text{Fusion}(\cdot)$  denotes the fusion mechanism as shown in Figure 3. The specific formulas are as follows:

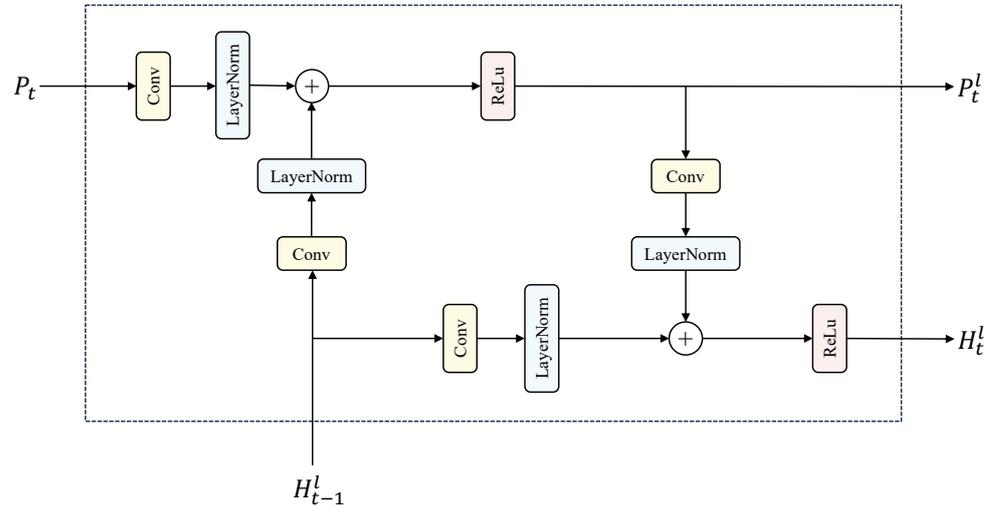
$$P_t^l = \text{Relu}(\text{LayerNorm}(\text{Conv}_{p2p}(P_t)) + \text{LayerNorm}(\text{Conv}_{h2p}(H_{t-1}^l))) \quad (10)$$

$$H_t^l = \text{Relu}(\text{LayerNorm}(\text{Conv}_{p2h}(P_t^l)) + \text{LayerNorm}(\text{Conv}_{h2h}(H_{t-1}^l))) \quad (11)$$

where  $Relu()$  is the Relu activation function,  $LayerNorm()$  is Layer Normalization, and  $Conv()$  is the convolution operation. Finally, to capture the long-range spatiotemporal dependencies, we added SAM to Fusion-Module, and the specific formula is as follows:

$$H_t^l, M_t^l = SAM(H_t^l, M_{t-1}^l) \quad (12)$$

where  $SAM(,)$  denotes self-attention memory module.



**Figure 3.** The structure of the Fusion. It fuses rainfall information with radar echo information.

### 3.2.3. Att-DyConvLSTM

To better decode the encoder-generated cell states  $C_t^l$ , hidden states  $H_t^l$ , and spatiotemporal memory cells  $M_t^l$ , we employed ConvLSTM as the baseline model and enhanced it by substituting its convolutional structures with dynamic convolution, aiming to better emphasize regions with intense precipitation. Similarly, we added the SAM module to the model, which was used to improve the ability to capture long-range spatiotemporal dependencies.

Att-DyConvLSTM, consisting of ConvLSTM, Dynamic Convolution, and the SAM module, constitutes the decoder module shown in Figure 4, which is formulated as follows:

$$i'_{t+1} = \sigma(Dy(W'_{pi}, H_{t+1}^{l+1}) + Dy(W'_{hi}, H_t^l) + W'_{ci} \odot C_t^l + b'_i) \quad (13)$$

$$f'_{t+1} = \sigma(Dy(W'_{pf}, H_{t+1}^{l+1}) + Dy(W'_{hf}, H_t^l) + W'_{cf} \odot C_t^l + b'_f) \quad (14)$$

$$g'_{t+1} = \tanh(Dy(W'_{pg}, H_{t+1}^{l+1}) + Dy(W'_{hg}, H_t^l) + b'_g) \quad (15)$$

$$o'_{t+1} = \sigma(Dy(W'_{po}, H_{t+1}^{l+1}) + Dy(W'_{ho}, H_t^l) + W'_{co} \odot C_t^l + b'_o) \quad (16)$$

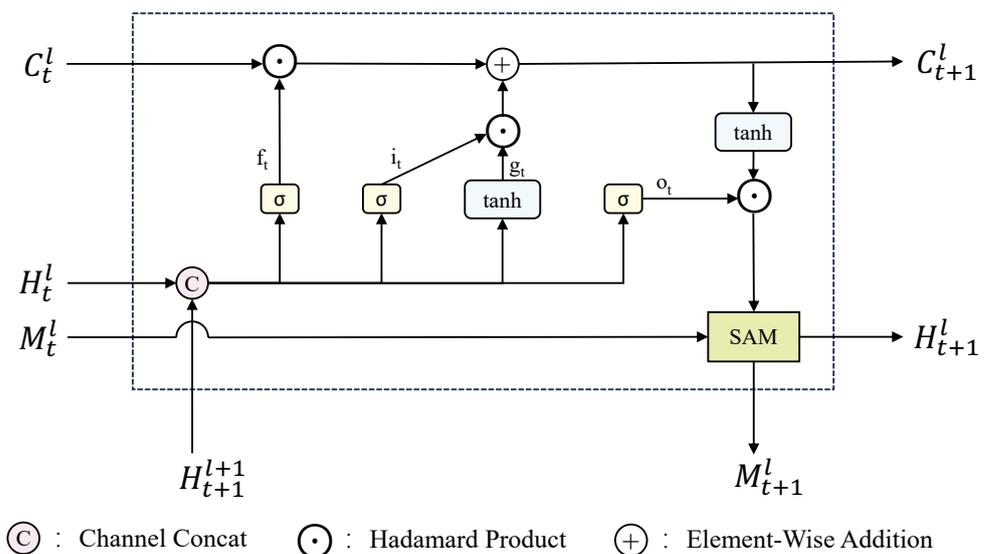
$$C'_{t+1} = f'_{t+1} \odot C_t^l + i'_{t+1} \odot g'_{t+1} \quad (17)$$

$$H'_{t+1} = o'_{t+1} \odot \tanh(C'_{t+1}) \quad (18)$$

$$H_{t+1}^l, M_{t+1}^l = SAM(H'_{t+1}, M_t^l) \quad (19)$$

where  $\sigma$  is the Sigmoid activation function,  $Dy(,)$  is the dynamic convolution,  $\odot$  is the Hadamard product,  $\tanh$  is the hyperbolic tangent activation function,  $W'_{**}$  is the parameter to be learned,  $b'_*$  is the bias value, and  $SAM(,)$  denotes the self-attention memory module. In Equation (19), if the Att-DyConvLSTM module is in the first layer of the decoder, the

output is the predicted rainfall  $\hat{P}_{t+1}$ , while in the other layers the output is the hidden state  $H_{t+1}^l$ .



**Figure 4.** The structure of Att-DyConvLSTM. The SAM module is embedded in it.

## 4. Experiments

### 4.1. Dataset

In this paper, we used the dataset from the 2022 Jiangsu Meteorological AI Algorithm Challenge as the training data for the model. The dataset contains data from 2019 to 2021, from April to September. The radar echo dataset was sourced from the quality-controlled network of multiple S-band meteorological radars in Jiangsu, encompassing the entire regional area of Jiangsu Province. The radar echo data spanned a range of 0–70 dBZ with a horizontal resolution of  $0.01^\circ$  (approximately 1 km), a temporal resolution of 6 min, and a frame resolution of  $480 \times 560$  pixels. Higher atmospheric water droplet content correlated with increased radar fundamental reflectivity. Rainfall data were derived through the interpolation of data collected from automatic meteorological stations in Jiangsu and the surrounding regions onto a standardized grid, and the rainfall data were the 6-minute cumulative rainfall from the automatic stations, with a value range of 0–10 mm and a resolution of  $480 \times 560$  pixels per frame.

In terms of data preprocessing, considering the computational overhead of the experiment, we obtained radar echo maps and rainfall maps of  $480 \times 480$  pixels in the middle by cropping, and we then scaled the maps to  $128 \times 128$  according to bicubic interpolation. To enhance training effectiveness, we normalized the pixel values of both the radar echo and the rainfall maps. The experiment utilized radar echo and rainfall sequences from the previous hour to predict rainfall sequences for the subsequent hour. The dataset was divided into training and validation sets in a ratio of 8:2.

### 4.2. Implementation Details

Our model and comparison models were implemented using Pytorch and were run on NVIDIA TITAN RTX GPUs. The comparison models were ConvLSTM [10], ConvGRU [16], PredRNN [17], MIM [25], SA-ConvLSTM [11], IDALSTM [27], AF-SRNet [13], and SA-ConvGRU [26]. For consistency across the experiments, all our comparison models incorporated radar echo and rainfall data. We concatenated these data types in the channel dimension, aligning with the methodology employed in [13]. In this experiment, we utilized the Adam optimizer [31] with a learning rate set to 0.001, a batch size of 16, and trained for a total of 100 epochs. If the loss failed to decrease for 4 consecutive epochs, the learning rate was reduced to 0.1 times its original value. Moreover, if the loss did not decrease for 10 consecutive

epochs, the training process was halted. When training the model using the original MSE–MAE loss function, the model tends to overlook regions of intense precipitation, which makes the model more focused on the low-rainfall region [32], and the output image is more blurred. The formula for the MSE–MAE loss function is shown below:

$$MSE - MAE_{loss} = \frac{1}{N} \sum_j^H \sum_i^W ((y_{i,j} - \hat{y}_{i,j})^2 + |y_{i,j} - \hat{y}_{i,j}|) \quad (20)$$

where  $N$  is the number of predicted maps,  $H$  and  $W$  are the height and width of the maps, and  $y_{i,j}$  and  $\hat{y}_{i,j}$  are the actual and predicted maps values, respectively. Therefore, we used weight loss and the SSIM to build the loss function with the following formula:

$$loss = weight_{loss} + 1000 \times (1 - SSIM(y, \hat{y})) \quad (21)$$

$$weight_{loss} = \frac{1}{N} \sum_j^H \sum_i^W weight_{i,j} \times ((y_{i,j} - \hat{y}_{i,j})^2 + |y_{i,j} - \hat{y}_{i,j}|) \quad (22)$$

$$SSIM(y, \hat{y}) = \frac{(2\mu_y\mu_{\hat{y}} + C_1)(2\sigma_{y\hat{y}} + C_2)}{(\mu_y^2 + \mu_{\hat{y}}^2 + C_1)(\sigma_y^2 + \sigma_{\hat{y}}^2 + C_2)} \quad (23)$$

where  $N$  is the number of predicted maps,  $H$  and  $W$  are the height and width of the maps,  $y_{i,j}$  and  $\hat{y}_{i,j}$  are the actual and predicted maps values, respectively,  $\mu_y$ ,  $\mu_{\hat{y}}$  are the means of  $y$  and  $\hat{y}$ ,  $\sigma_y^2$ ,  $\sigma_{\hat{y}}^2$  are the variances of  $y$  and  $\hat{y}$ ,  $\sigma_{y\hat{y}}$  is the covariance of  $y$  and  $\hat{y}$ ,  $C_1$  and  $C_2$  are two non-zero constants, and  $weight_{i,j}$  is the weight value of the map at  $(i,j)$ . We analyzed the distribution of different rainfall amounts, and to ensure that areas of heavy precipitation were not overlooked, we assigned different weight classes to different rainfall amounts;  $weight_{i,j}$  was assigned according to the following equation:

$$weight_{i,j} = \begin{cases} 0.1 & 0.00 \leq y_{i,j} \leq 0.01 \\ 1.0 & 0.01 \leq y_{i,j} \leq 0.10 \\ 2.0 & 0.10 \leq y_{i,j} \leq 0.50 \\ 5.0 & 0.50 \leq y_{i,j} \leq 1.00 \\ 10.0 & 1.00 \leq y_{i,j} \leq 2.00 \\ 20.0 & 2.00 \leq y_{i,j} \leq 3.00 \\ 30.0 & 3.00 \leq y_{i,j} \leq 6.00 \\ 50.0 & 6.00 \leq y_{i,j} \leq 10.00 \end{cases} \quad (24)$$

#### 4.3. Performance Metric

To assess the model's performance, we utilized four commonly employed evaluation metrics: the critical success index (CSI), the false alarm rate (FAR), the Heidke skill score (HSS), and the mean squared error (MSE). The values of the CSI, the FAR, and the HSS were in the range of 0-1. The closer the CSI and the HSS were to 1, and the closer the FAR was to 0, the better the prediction of the model was. The MSE was a value greater than or equal to 0. The closer it was to 0, the closer the predicted data were to the real data and the better the prediction was. To calculate the CSI, the FAR, and the HSS, we changed the rainfall map to a bivariate map based on thresholds, where the value was set to 1 if the rainfall value was greater than the threshold, and 0 otherwise, and the thresholds were set to 0.05 mm, 0.20 mm, and 0.50 mm, respectively. Using the bivariate map, we calculated the True Positive (TP), the False Negative (FN), the True Negative (TN), and the False Positive (FP) values. The CSI, FAR, HSS, and MSE metrics were defined as follows:

$$CSI = \frac{TP}{TP + FP + FN} \quad (25)$$

$$FAR = \frac{FP}{TP + FP} \quad (26)$$

$$HSS = \frac{2 \times (TP \times TN - FN \times FP)}{(TP + FN) \times (FN + TN) + (TP + FP) \times (FP + TN)} \quad (27)$$

$$MSE = \frac{1}{N} \sum_j^H \sum_i^W (y_{i,j} - \hat{y}_{i,j})^2 \quad (28)$$

where  $TP$  denotes the number of predicted values of 1 and true values of 1,  $FN$  denotes the number of predicted values of 0 and true values of 1,  $TN$  denotes the number of predicted values of 0 and true values of 0,  $FP$  denotes the number of predicted values of 1 and true values of 0,  $N$  is the number of predicted maps,  $H$  and  $W$  are the height and width of the maps, and  $y_{i,j}$  and  $\hat{y}_{i,j}$  are the true and predicted map values.

#### 4.4. Results and Analysis

##### 4.4.1. Comparative Experiments

We utilized the dataset provided by the 2022 Jiangsu Meteorological AI Algorithm Challenge to evaluate the performance of DSADNet proposed in this paper alongside other comparative models. The dataset used for evaluating the model was not used for the training of the model, and the evaluation results are more reflective of the model's generalization ability. Tables 1 and 2 show the experimental results of the comparison at 30 min and at 60 min, where the optimal metrics are in bold. In addition, for a more intuitive comparison between our model and other models, an example of a randomly selected visualization is given in Figure 5.

From the results of this experiment, DSADNet performed the best except for the FAR metric. In the field of precipitation nowcasting, the CSI usually stands for the hit rate of precipitation, i.e., the proportion of rainfall events that are successfully predicted. It is a statistical metric used to assess the accuracy of forecasting precipitation events, and the HSS is a statistical metric used to evaluate the performance of a classification model, especially for assessing the consistency between prediction results and observations. The higher of the two metrics means that our proposed DSADNet model can effectively extract information from radar and rainfall maps, which improves the accuracy of prediction. Not only that, our model had a lower MSE, which denotes the mean of the squares of the differences between the predicted and real data, indicating that the model's predictions fit the true values better than other models. In addition, the model improves the performance of maintaining long-range spatiotemporal dependencies by using the SAM module, which enables the model to maintain stable performance as the prediction time grows. As depicted in Figure 5, our model's prediction results exhibit a minimal shift in the overall rainfall distribution compared to the actual data, and the prediction stability is better than other comparative models. And our model replaces the convolution structure with dynamic convolution. As can be seen from the visualization in Figure 5, our model better predicts the heavy precipitation area (the red part in the figure) compared to other models, which expand or reduce the heavy precipitation area inconsistently with the reality.

Although our model has outstanding metrics on the CSI and the HSS, it is slightly below the optimal metrics on the FAR metrics, and, due to the model lacking in extracting the time evolution information and capturing the process of rainfall intensity change, the model focuses more on the area of heavy precipitation and ignores the information that the heavy precipitation decays into weak precipitation.

**Table 1.** Average evaluation results of comparison experiments within 30 min. The best metrics are highlighted in bold; “↑” signifies that higher scores are preferable, while “↓” indicates that lower scores are preferable.

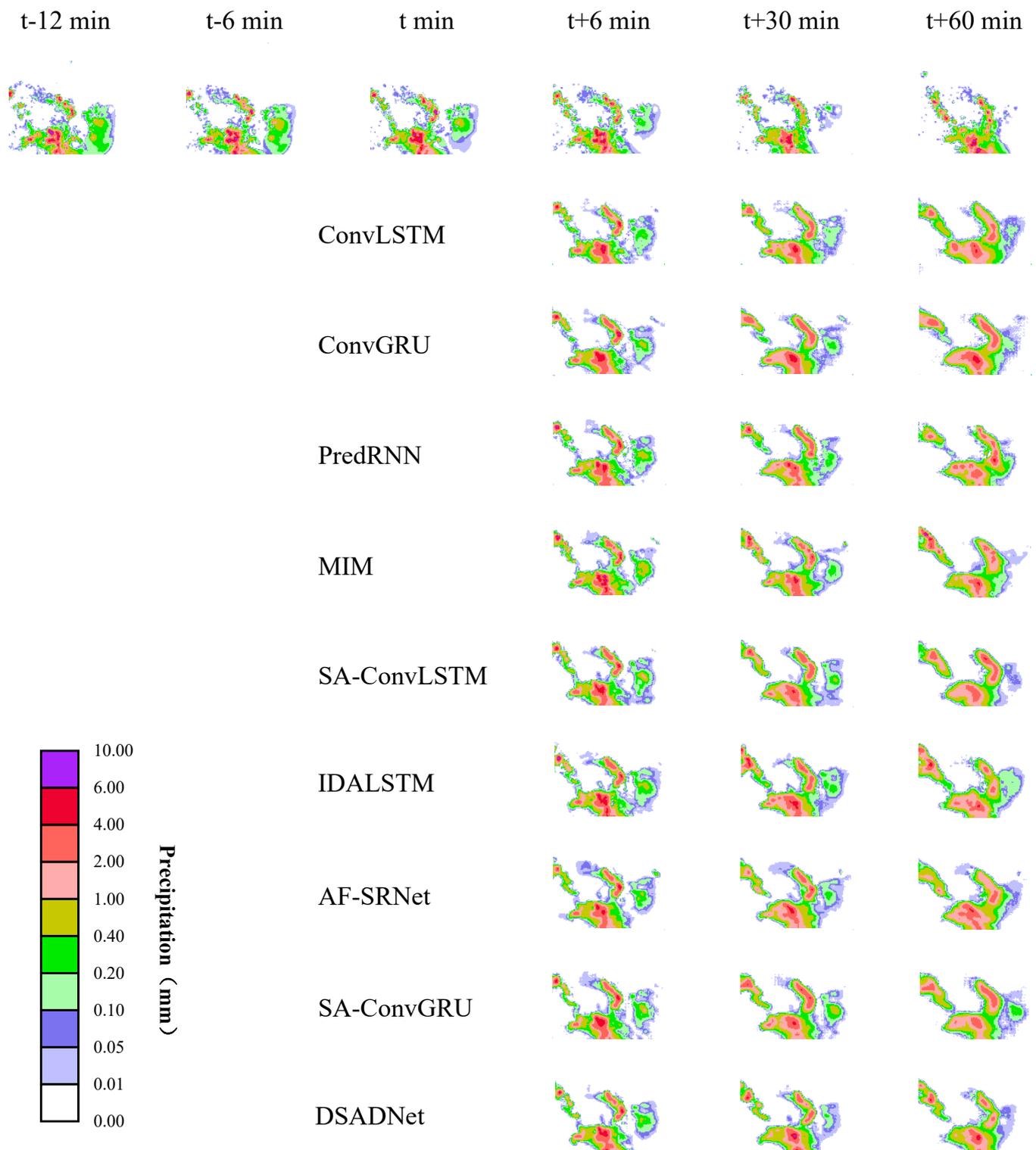
Model	0.05 mm			0.20 mm			0.50 mm			MSE↓
	CSI↑	FAR↓	HSS↑	CSI↑	FAR↓	HSS↑	CSI↑	FAR↓	HSS↑	
ConvLSTM	0.5329	<b>0.3124</b>	0.6461	0.4702	0.3951	0.5963	0.3953	0.4418	0.5252	7.5953
ConvGRU	0.5293	0.3668	0.6419	0.4646	0.4333	0.5897	0.3949	<b>0.4200</b>	0.5256	7.6868
PredRNN	0.5400	0.3396	0.6528	0.4721	0.3849	0.5986	0.3948	0.4307	0.5245	7.6427
MIM	0.5352	0.3315	0.6480	0.4762	0.3943	0.6035	0.4050	0.4452	0.5352	8.0327
SA-ConvLSTM	0.5404	0.3172	0.6532	0.4702	<b>0.3768</b>	0.5952	0.3965	0.4202	0.5269	7.2319
IDALSTM	0.5430	0.3584	0.6555	0.4747	0.4087	0.6017	0.3941	0.4371	0.5239	7.8811
AF-SRNet	0.5421	0.3529	0.6561	0.4721	0.4104	0.5996	0.3935	0.4319	0.5232	8.1311
SA-ConvGRU	0.5328	0.3399	0.6459	0.4653	0.3939	0.5910	0.3914	0.4305	0.5204	7.3832
DSADNet	<b>0.5516</b>	0.3189	<b>0.6641</b>	<b>0.4841</b>	0.3921	<b>0.6102</b>	<b>0.4066</b>	0.4217	<b>0.5364</b>	<b>7.1690</b>

**Table 2.** Average evaluation results of comparison experiments within 60 min. The best metrics are highlighted in bold; “↑” signifies that higher scores are preferable, while “↓” indicates that lower scores are preferable.

Model	0.05 mm			0.20 mm			0.50 mm			MSE↓
	CSI↑	FAR↓	HSS↑	CSI↑	FAR↓	HSS↑	CSI↑	FAR↓	HSS↑	
ConvLSTM	0.4848	0.3651	0.5946	0.4106	0.4459	0.5300	0.3264	0.4922	0.4445	10.3514
ConvGRU	0.4809	0.4117	0.5894	0.4076	0.4834	0.5268	0.3301	0.4736	0.4502	10.5840
PredRNN	0.4911	0.3785	0.6000	0.4126	0.4257	0.5328	0.3262	0.4870	0.4452	10.4432
MIM	0.4889	0.3730	0.5984	0.4160	0.4493	0.5359	0.3324	0.5083	0.4517	10.8839
SA-ConvLSTM	0.4884	<b>0.3547</b>	0.5981	0.4123	<b>0.4104</b>	0.5307	0.3322	<b>0.4715</b>	0.4524	9.9691
IDALSTM	0.4917	0.3971	0.6013	0.4120	0.4425	0.5316	0.3235	0.4989	0.4416	10.6774
AF-SRNet	0.4948	0.3960	0.6054	0.4160	0.4509	0.5369	0.3323	0.4882	0.4530	10.9695
SA-ConvGRU	0.4839	0.3821	0.5935	0.4095	0.4282	0.5286	0.3288	0.4834	0.4479	10.0829
DSADNet	<b>0.5010</b>	0.3548	<b>0.6111</b>	<b>0.4199</b>	0.4268	<b>0.5393</b>	<b>0.3347</b>	0.4758	<b>0.4538</b>	<b>9.8653</b>

As can be seen in the visualization example in Figure 5, our model performs better than the other models in heavy precipitation prediction. In addition, the prediction accuracy of ConvLSTM, ConvGRU, IDALSTM, and SA-ConvGRU keeps decreasing as the prediction time increases, and at 60 min, the regions of different precipitation intensities predicted by these models are quite different from the real situation. AF-SRNet also employs a fusion mechanism to predict rainfall, and the model first encodes the radar and rainfall maps through their respective encoders, and the radar and rainfall maps do not interact at that stage. After that, the radar and rainfall information are fused by the attention fusion module, which fuses the radar and rainfall information at the same stage, ignoring the evolution of the radar echo intensity at the two stages. Therefore this model incorrectly predicts the rainfall distribution, resulting in a large impact on the prediction.

Therefore, our model not only outperforms other models in terms of forecast accuracy, maintaining long-range spatiotemporal dependencies and focusing on regions of heavy precipitation, but is also at a superior level, although the FAR is not optimal.



**Figure 5.** Comparative experiments are visualized as follows. The rainfall maps in the first row depict the input and ground truth at t + 6 min, t + 30 min, and t + 60 min. Subsequent rows display the results of ConvLSTM, ConvGRU, PredRNN, MIM, SA-ConvLSTM, IDALSTM, AF-SRNet, SA-ConvGRU, and our proposed model, respectively.

#### 4.4.2. Ablation Study

To further demonstrate the influence of the fusion module, dynamic convolution, and SAM on the experiments, ablation experiments were conducted to assess the effectiveness

of the fusion module, dynamic convolution, and SAM individually. Tables 3 and 4 show the ablation results at 30 min and at 60 min, where the optimal metrics are in bold. In the table, ConvLSTM is the baseline model, DSNet is the fusion module added to ConvLSTM, DSDNet is the convolution structure on DSNet replaced with dynamic convolution, DSANet is the SAM module added to DSNet, and DSADNet is our proposed model. Similarly, Figure 6 gives an example of the visualization of the ablation experiment.

**Table 3.** Average evaluation results of ablation experiments within 30 min. The best metrics are highlighted in bold; “↑” signifies that higher scores are preferable, while “↓” indicates that lower scores are preferable.

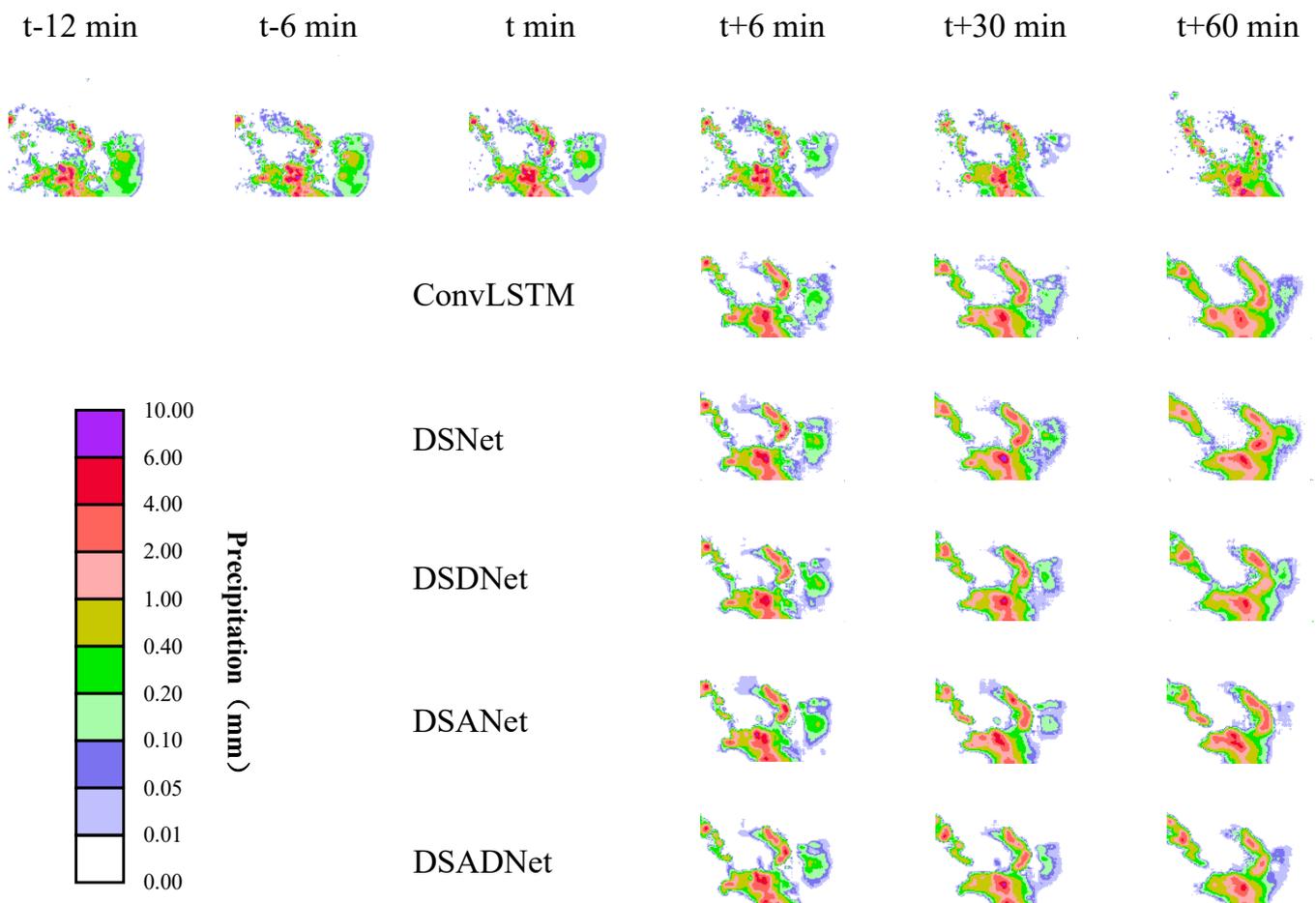
Model	0.05 mm			0.20 mm			0.50 mm			MSE↓
	CSI↑	FAR↓	HSS↑	CSI↑	FAR↓	HSS↑	CSI↑	FAR↓	HSS↑	
ConvLSTM	0.5329	<b>0.3124</b>	0.6461	0.4702	0.3951	0.5963	0.3953	0.4418	0.5252	7.5953
DSNet	0.5423	0.3504	0.6557	0.4728	0.4019	0.5989	0.3968	0.4505	0.5283	7.5271
DSDNet	0.5428	0.3651	0.6550	0.4784	0.4362	0.6058	0.4018	<b>0.4215</b>	0.5325	7.3648
DSANet	0.5438	0.3145	0.6564	0.4737	0.3976	0.5998	0.3934	0.4561	0.5244	7.3382
DSADNet	<b>0.5516</b>	0.3189	<b>0.6641</b>	<b>0.4841</b>	<b>0.3921</b>	<b>0.6102</b>	<b>0.4066</b>	0.4217	<b>0.5364</b>	<b>7.1690</b>

**Table 4.** Average evaluation results of ablation experiments within 60 min. The best metrics are highlighted in bold; “↑” signifies that higher scores are preferable, while “↓” indicates that lower scores are preferable.

Model	0.05 mm			0.20 mm			0.50 mm			MSE↓
	CSI↑	FAR↓	HSS↑	CSI↑	FAR↓	HSS↑	CSI↑	FAR↓	HSS↑	
ConvLSTM	0.4848	0.3651	0.5946	0.4106	0.4459	0.5300	0.3264	0.4922	0.4445	10.3514
DSNet	0.4931	0.3987	0.6033	0.4126	0.4443	0.5321	0.3300	0.5065	0.4502	10.2246
DSDNet	0.4930	0.4146	0.6018	0.4174	0.4924	0.5385	0.3338	0.4916	<b>0.4547</b>	10.0921
DSANet	0.4947	0.3597	0.6048	0.4166	0.4411	0.5378	0.3309	0.5212	0.4528	10.0208
DSADNet	<b>0.5010</b>	<b>0.3548</b>	<b>0.6111</b>	<b>0.4199</b>	0.4268	<b>0.5393</b>	<b>0.3347</b>	0.4758	0.4538	<b>9.8653</b>

From the results of the ablation experiments: Firstly, by comparing the evaluation indexes of the benchmark model ConvLSTM and DSNet, Fusion-Module was added to help the model better extract the radar information and rainfall information, and the CSI and HSS indexes of DSNet were better than those of the benchmark model, but, due to the weak ability of Fusion-Module in extracting information, strong precipitation decayed into weak precipitation, which led to the FAR being on the high side. So, Fusion-Module helps to improve prediction performance, but the FAR index needs to be reduced. Secondly, to improve the focus on strong precipitation, we replaced the convolution structure on DSNet with dynamic convolution. As shown in Figure 6, DSDNet paid more attention to the region of strong precipitation than DSNet (the red part in the figure). Next, to improve the model’s ability to maintain long-range spatiotemporal dependencies, we added the SAM module to DSNet, and, as observed in Tables 3 and 4, DSANet with the addition of SAM outperforms DSNet in all the metrics as the prediction time grows. Our model, DSADNet, which combines the fusion module, dynamic convolution, and the SAM module, is basically optimal in all metrics. With the addition of dynamic convolution, the SAM module, and Fusion-Module, basically, the MSE metric decreases and the model predictions improve. According to Figure 6, DSADNet not only outperforms other models in capturing heavy precipitation areas, but also maintains stable model performance as the prediction time increases. Although our model combines the fusion module, dynamic convolution, and the attention mechanism, Fusion-Module plays a dominant role. Fusion-Module improves model prediction more significantly than dynamic convolution and the attention mechanism, which can be seen in Tables 3 and 4. While our model generally outperforms

the baseline model across all metrics, the improvement is not substantial. Enhancing the model's prediction accuracy remains the primary focus of future work.



**Figure 6.** Ablation experiments are visualized as follows. The rainfall maps in the first row depict the input and ground truth at  $t + 6$  min,  $t + 30$  min, and  $t + 60$  min. Subsequent rows display the results of ConvLSTM, DSNet, DSDNet, DSANet, and our proposed model, respectively. In this figure, ConvLSTM serves as the baseline model. DSNet represents the fusion module added to ConvLSTM, DSDNet denotes the replacement of the convolution structure on DSNet with dynamic convolution, and DSANet signifies the SAM module added to DSNet.

## 5. Conclusions

This paper introduces a novel precipitation nowcasting model, DSADNet. The model endeavors to maximize the utilization of valuable information from multiple data sources through the development of a novel fusion module. This module effectively captures variations in radar echo intensity and extracts useful information from both radar echo and rainfall data. To improve prediction, we replaced the convolution in the model with dynamic convolution, which can pay more attention to areas of heavy precipitation. The added SAM module allows prediction to remain stable as prediction time increases and improves the model's ability to capture long-range spatiotemporal dependencies. Our proposed model has higher CSI and HSS metrics and lower MSE metrics, but our model focuses more on the effective combination of radar and rainfall information and is weaker in extracting time-evolution information and capturing the rainfall intensity change process. The model predicts weak precipitation regions as strong precipitation regions, which ultimately leads to high FAR metrics. In future endeavors, our emphasis will be on refining Fusion-Module to minimize the false alarm rate (FAR) and enhance the accuracy of precipitation nowcasting.

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

**Funding:** This research was funded by the Zhejiang Provincial Basic Public Welfare Research Project under Grant LGG20F020012

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data used in this study can be obtained by contacting the corresponding author upon request. The data are not publicly accessible due to the confidentiality policy of Jiangsu Meteorological Observatory.

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

## References

1. Yang, Z.; Liu, Q.; Wu, H.; Liu, X.; Zhang, Y. CEMA-LSTM: Enhancing contextual feature correlation for radar extrapolation using fine-grained echo datasets. *Comput. Model. Eng. Sci.* **2022**, *135*, 45–64. [\[CrossRef\]](#)
2. Kimura, R. Numerical weather prediction. *J. Wind. Eng. Ind. Aerodyn.* **2002**, *90*, 1403–1414. [\[CrossRef\]](#)
3. Bauer, P.; Thorpe, A.; Brunet, G. The quiet revolution of numerical weather prediction. *Nature* **2015**, *525*, 47–55. [\[CrossRef\]](#)
4. Bowler, N.E.; Pierce, C.E.; Seed, A. Development of a precipitation nowcasting algorithm based upon optical flow techniques. *J. Hydrol.* **2004**, *288*, 74–91. [\[CrossRef\]](#)
5. Ayzel, G.; Heistermann, M.; Winterrath, T. Optical flow models as an open benchmark for radar-based precipitation nowcasting (rainymotion v0. 1). *Geosci. Model Dev.* **2019**, *12*, 1387–1402. [\[CrossRef\]](#)
6. Woo, W.C.; Wong, W.K. Operational application of optical flow techniques to radar-based rainfall nowcasting. *Atmosphere* **2017**, *8*, 48. [\[CrossRef\]](#)
7. Medsker, L.R.; Jain, L. *Recurrent Neural Networks. Design and Applications*; CRC Press: Boca Raton, FL, USA, 2001; Volume 5.
8. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Cho, K.; Van Merriënboer, B.; Bahdanau, D.; Bengio, Y. On the properties of neural machine translation: Encoder-decoder approaches. *arXiv* **2014**, arXiv:1409.1259.
10. Shi, X.; Chen, Z.; Wang, H.; Yeung, D.Y.; Wong, W.K.; Woo, W.C. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Adv. Neural Inf. Process. Syst.* **2015**, *28*, 802–810.
11. Lin, Z.; Li, M.; Zheng, Z.; Cheng, Y.; Yuan, C. Self-attention convlstm for spatiotemporal prediction. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 11531–11538.
12. Hu, Q.; Li, Z.; Wang, L.; Huang, Y.; Wang, Y.; Li, L. Rainfall spatial estimations: A review from spatial interpolation to multi-source data merging. *Water* **2019**, *11*, 579. [\[CrossRef\]](#)
13. Geng, L.; Geng, H.; Min, J.; Zhuang, X.; Zheng, Y. AF-SRNet: Quantitative Precipitation Forecasting Model Based on Attention Fusion Mechanism and Residual Spatiotemporal Feature Extraction. *Remote Sens.* **2022**, *14*, 5106. [\[CrossRef\]](#)
14. Chen, Y.; Dai, X.; Liu, M.; Chen, D.; Yuan, L.; Liu, Z. Dynamic convolution: Attention over convolution kernels. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 11030–11039.
15. Shi, X.; Gao, Z.; Lausen, L.; Wang, H.; Yeung, D.Y.; Wong, W.K.; Woo, W.C. Deep learning for precipitation nowcasting: A benchmark and a new model. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 5617–5627.
16. Siam, M.; Valipour, S.; Jagersand, M.; Ray, N. Convolutional gated recurrent networks for video segmentation. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 3090–3094.
17. Wang, Y.; Long, M.; Wang, J.; Gao, Z.; Yu, P.S. Predrnn: Recurrent neural networks for predictive learning using spatiotemporal lstms. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 879–888.
18. Bengio, Y.; Simard, P.; Frasconi, P. Learning long-term dependencies with gradient descent is difficult. *IEEE Trans. Neural Netw.* **1994**, *5*, 157–166. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Pascanu, R.; Mikolov, T.; Bengio, Y. On the difficulty of training recurrent neural networks. In Proceedings of the International Conference on Machine Learning. PMLR, Atlanta, GA, USA, 17–19 June 2013; pp. 1310–1318.
20. Wang, Y.; Gao, Z.; Long, M.; Wang, J.; Philip, S.Y. Predrnn++: Towards a resolution of the deep-in-time dilemma in spatiotemporal predictive learning. In Proceedings of the International Conference on Machine Learning. PMLR, Stockholm, Sweden, 10–15 July 2018; pp. 5123–5132.
21. Wang, Y.; Wu, H.; Zhang, J.; Gao, Z.; Wang, J.; Philip, S.Y.; Long, M. Predrnn: A recurrent neural network for spatiotemporal predictive learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 2208–2225. [\[CrossRef\]](#) [\[PubMed\]](#)

22. Wu, D.; Wu, L.; Huang, J.; Wang, X. ISA-PredRNN: An Improved Self-Attention PredRNN Network for Spatiotemporal Predictive Learning. In Proceedings of the 2022 International Conference on Image Processing, Computer Vision and Machine Learning (ICICML), Xi'an, China, 28–30 October 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 432–440.
23. Tuyen, D.N.; Tuan, T.M.; Le, X.H.; Tung, N.T.; Chau, T.K.; Van Hai, P.; Gerogiannis, V.C.; Son, L.H. RainPredRNN: A new approach for precipitation nowcasting with weather radar echo images based on deep learning. *Axioms* **2022**, *11*, 107. [[CrossRef](#)]
24. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th international Conference, Munich, Germany, 5–9 October 2015; proceedings, part III 18; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.
25. Wang, Y.; Zhang, J.; Zhu, H.; Long, M.; Wang, J.; Yu, P.S. Memory in memory: A predictive neural network for learning higher-order non-stationarity from spatiotemporal dynamics. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 9154–9162.
26. Zhou, M.; Wu, J.; Chen, M.; Han, L. SA-ConvGRU: A Method for Short-Duration Heavy Rainfall Warning with a Self-Attention Memory. In Proceedings of the 2023 IEEE 6th International Conference on Big Data and Artificial Intelligence (BD AI), Jiaxing, China, 7–9 July 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 202–207.
27. Luo, C.; Li, X.; Wen, Y.; Ye, Y.; Zhang, X. A novel LSTM model with interaction dual attention for radar echo extrapolation. *Remote Sens.* **2021**, *13*, 164. [[CrossRef](#)]
28. Yao, J.; Xu, F.; Qian, Z.; Cai, Z. A Forecast-Refinement Neural Network Based on DyConvGRU and U-Net for Radar Echo Extrapolation. *IEEE Access* **2023**, *11*, 53249–53261. [[CrossRef](#)]
29. Zhang, Y.; Geng, S.; Tian, W.; Ma, G.; Zhao, H.; Xie, D.; Lu, H.; Lim Kam Sian, K.T.C. Weather Radar Echo Extrapolation with Dynamic Weight Loss. *Remote Sens.* **2023**, *15*, 3138. [[CrossRef](#)]
30. Zhang, Z.; Luo, C.; Feng, S.; Ye, R.; Ye, Y.; Li, X. RAP-Net: Region attention predictive network for precipitation nowcasting. *Geosci. Model Dev. Discuss.* **2022**, *2022*, 1–19. [[CrossRef](#)]
31. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
32. Guo, S.; Sun, N.; Pei, Y.; Li, Q. 3D-UNet-LSTM: A Deep Learning-Based Radar Echo Extrapolation Model for Convective Nowcasting. *Remote Sens.* **2023**, *15*, 1529. [[CrossRef](#)]

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