GAN-argcPredNet v2.0: A Radar Echo Extrapolation Model based on Spatiotemporal Process Intensification

Kun Zheng[1*], Qiya Tan[1*], Huihua Ruan[2], Jinbiao Zhang[2], Cong Luo[3], Siyu Tang[3], Yunlei Yi[4], Yugang Tian[1], Jianmei Cheng[5]

[1]School of Geography and Information Engineering, China University of Geosciences, Wuhan, 430074, China
[2]Guangdong Meteorological Observation Data Center, Guangzhou, 510080, China
[3]Guangdong Meteorological Observatory, Guangzhou, 510080, China
[4]Wuhan Zhaotu Technology Co. Ltd., Wuhan, 430074, China
[5]School of Environmental Studies, China University of Geosciences, Wuhan, 430074, China
*These authors have contributed equally to this work

Correspondence to: Kun Zheng (ZhengK@cug.edu.cn); Huihua Ruan (ruanhuihua@163.com)

Abstract. Precipitation nowcasting has important implications for urban operation and flood prevention. Radar echo extrapolation is the common method in precipitation nowcasting. Using deep learning models to extrapolate radar echo data has great potential. The increase of lead time leads to a weaker correlation between real rainfall evolution and generated images. The evolution information is easily lost during extrapolation, which is reflected as echo attenuation. Existing models, including Generative Adversarial Network (GAN)-based models, are difficult to reduce loss and curb attenuation, which results in insufficient rainfall prediction accuracy. Aim to the problem, a Spatiotemporal Process Intensification Network (GAN-argcPredNet v2.0) based on GAN-argcPredNet v1.0 is designed. GAN-argcPredNet v2.0 reduces the loss by intensifying the influence of the previously input evolution information. A Spatiotemporal Information Changes Prediction (STIC-Prediction) network is designed as generator. By intensifying the spatiotemporal evolution of the echo feature sequence with STIC Attention, the generator focuses on the spatiotemporal variation and generates more accurate images. Furthermore, discriminator is a Channel-Spatial Convolution (CS-Convolution) network. The discriminator enhances the discrimination of echo information by intensifying spatial information with CS Attention. Identification results are fed back to the generator, which reduces the loss of important evolutionary information. The experiments are based on the radar dataset of South China. The results show that GAN-argcPredNet v2.0 performs better than other models. In heavy rainfall prediction, compared with baseline, the Probability of Detection (POD), the Critical Success Index (CSI), the Heidke Skill Score (HSS) and Bias increase by 18.8 %, 17.0 %, 17.2% and 26.3 % respectively. The False Alarm Ratio (FAR) decreases by 3.0 %.

1 Introduction

Accurate precipitation nowcasting, especially heavy precipitation nowcasting, plays a key role in hydrometeorological applications such as urban-operation safety and flash-flood warnings (Liu et al., 2015). It can effectively prevent the hazards and losses caused by heavy precipitation to economy and people (Luo et al., 2020).

Radar echo extrapolation is the method most often used to nowcast precipitation (Reyniers, 2008). The essence is tracking areas of reflectivity to derive motion vectors, and then using the motion vectors to determine future location of the reflectivity (Austin and Bellon, 1974). Traditional radar echo extrapolation methods include cross-correlation, individual radar echo-tracking and the optical flow method (Bowler et al., 2004). As the storm evolution like merging, splitting, growth and decay, traditional methods are difficult to predict accurately. Using deep learning models to extrapolate has great potential (Foresti et al., 2019). Deep learning has powerful nonlinear mapping ability. By studying the motion process from a large number of historical radar echo images, deep learning has better results (Shi et al., 2015; Pan et al., 2021). Compared with other deep learning models, the Generative Adversarial Network (GAN)-based models have significant advantages in generating high-quality echo images (Tian et al., 2020; Xie et al., 2022).

The radar echo images are predicted for a future period based on the real echo sequence. In the deep learning models, the increase of lead time leads to a weaker correlation between the real images at the front of sequence and the generated images. The influence of real echo evolution is rapidly diminishing. In this process, the models lose rainfall evolution information. It is reflected as echo attenuation on generated images. Due to the smaller percentage of heavy rainfall areas, the attenuation is more severe. To the knowledge of authors, existing deep learning models, including GAN-based models, lack the method to curb attenuation, which leads to low accuracy in predicting heavy rainfall.

In this work, a Spatiotemporal Process Intensification Network (GAN-argcPredNet v2.0) is proposed based on Generative Adversarial Advanced Reduced-Gate Convolutional Deep Predictive Coding Network (GAN-argcPredNet v1.0) (Zheng et al., 2022), which aims at reducing loss and curbing attenuation. In GAN-argcPredNet v2.0, a Spatiotemporal Information Changes Prediction (STIC-Prediction) network is designed as the generator. The generator focuses on the spatiotemporal variation of radar echo feature sequence. Its purpose is to curb echo attenuation by intensifying the spatiotemporal evolution of previous inputs. Furthermore, a Channel-Spatial Convolution (CS-Convolution) network is designed as the discriminator. The discriminator aims to enhance the ability to identify echo information by focusing on radar echo features from spatial and channel dimensions. In this way, the generator can be guided to better retain evolution information. The generator and discriminator are trained against each other to have accurate rainfall prediction.

2 Related Work

2.1 Traditional extrapolation methods

Eulerian persistence is a naive extrapolation method, which is based on using the latest available observation as a prediction. It is quite a powerful model for very short lead times (Ayzel et al., 2019). Individual radar echo tracking and cross correlation are also traditional extrapolation techniques (Pierce et al., 2004; Liguori and Rico-Ramirez, 2014). Thunderstorm Identification, Tracking, Analysis, and Nowcasting (TITAN) is a classical centroid tracking algorithm (Dixon and Wiener, 1993). The algorithm achieves precipitation nowcasting through real-time tracking and automatic identification of individual storm. The tracking performance of TITAN is poor during multi-cell storms. Then, an enhanced TITAN (ETITAN) is proposed

(Han et al., 2009). By combining cross-correlation and individual radar echo-tracking, ETITAN achieves more accurate tracking and prediction. Cross-correlation method, however, has significantly lower prediction accuracy when echoes change rapidly. The optical flow method achieves local prediction by treating echoes motion as fluid (Sakaino, 2013). Pyramid Lucas-Kanade Optical Flow method (PPLK) achieves better extrapolation by introducing a typical local differential optical flow method (Liu et al., 2015). As radar echoes continuously evolve, the invariance assumption of the optical flow method cannot be satisfied. The extrapolation accuracy is affected. Besides, these traditional methods do not intend to utilize the large amounts of historical images.

2.2 Sequence prediction networks

Radar echo extrapolation can be regarded as an image sequence prediction problem. Therefore, the problem can be solved by implementing an end-to-end sequence learning method (Sutskever et al., 2014; Shi et al., 2015). ConvGRU learns video features through convolution operation, which realizes sparse connection of model units (Ballas et al., 2015). Convolution operation is also used in ConvLSTM. By replacing the steps of internal data state transformation in LSTM, ConvLSTM can better extract features (Shi et al., 2015). Convolutional recursive structure is position invariant, which is not consistent with the natural variation motion. Trajectory GRU (TrajGRU) is further proposed (Shi et al., 2017). Both LSTM and GRU models have long-term memory. However, this capability is limited to historical spatial information and the memory is also limited. RainNet builds a convolutional network architecture in precipitation nowcasting, which avoids the brittleness of LSTM structure (Ayzel et al., 2020). This new structure still fails to address the information loss. Meanwhile, the realistic details of images are also insufficient in these models.

Attention mechanism is also often used in sequential networks. By learning the importance of different image parts, attention mechanisms can improve prediction accuracy. For example, the self-attention mechanism combines spatial relationships of different locations and enhances important areas (Wang et al., 2018a). Eidetic 3D LSTM (E3D-LSTM) introduces self-attention to intensify long-term memory in LSTM (Wang et al., 2018b). However, it lacks attention in the channel dimension. Interaction Dual Attention LSTM (IDA-LSTM) expands spatial and channel attention based on self-attention to improve representation learning (Luo et al., 2021). As the high hardware load, self-attention is hard to train for high-resolution images. Convolutional Block Attention Module (CBAM) was developed simultaneously as a less computational attention mechanism. It can be flexibly applied in sequential networks (Woo et al., 2018). In sequence prediction, temporal information is also important, but these methods fail to intensify it. For radar echo extrapolation, it reflects as a lack of intensification to rainfall evolution information.

2.3 GAN-based radar echo extrapolation

At present, high-quality extrapolation is mostly achieved by GAN (Tian et al., 2020; Xie et al., 2022). GAN consists of a generator and a discriminator, which has powerful data generation capabilities (Goodfellow et al., 2020). This is because the model with anti-loss can better realize multi-modal modeling (Lotter et al., 2016). For instance, Deep Generative Models of

Rainfall (DGMR) generates more accurate reflectivity by adversarial training (Ravuri et al., 2021). GAN is also used to generate realistic details for broader extrapolation range (Chen et al., 2019). GA-ConvGRU uses ConvGRU as the generator. By implementing multi-modal data modeling, the image quality is far better than ConvGRU (Tian et al., 2020). A number of studies contribute to improving the stability of GAN training. Energy-Based Generative Adversarial Forecaster (EBGAN-Forecaster) combines convolution structure and codec framework to improve stability (Xie et al., 2022). Also, our proposed GAN-argcPredNet v1.0 has more advantages in improving the predicted echoes details and stabilizing GAN training (Zheng et al., 2022). As the lack of curbing echo attenuation, all these models have limited accuracy in rainfall prediction, especially heavy rainfall prediction.

3 Model

In GAN-argcPredNet v1.0, the generator generates predicted images according to input image sequences. Then, the predicted and real images are fed into discriminator with dual channel input. The discriminator makes judgments and the parameters are updated by adversarial loss optimization. Adam is used as the optimizer, which is an extension of stochastic gradient descent (Kingma Diederik and Adam, 2014). The generator parameters are updated once every five times. GAN-argcPredNet v2.0 model is constructed based on GAN-argcPredNet v1.0.

3.1 GAN-argcPredNet v2.0 overview

GAN-argcPredNet v2.0 consists of STIC-Prediction generator and CS-Convolution discriminator (Fig. 1). STIC-Prediction generator is designed to reduce information loss and echo attenuation by intensifying the spatiotemporal variations of the previous feature sequence. The generator is composed of argcPredNet and the STIC Attention module (Fig. 2). The argcPredNet is composed of a series of repeatedly stacked modules, with a total of four layers (Zheng et al., 2022). Each layer of the module including the input convolutional layer (A_l), the recurrent representation layer (R_l), the prediction convolutional layer ($\hat{A}_l$) and the error representation layer (E_l). R_l learns image features and generates the feature map $R_l^T \in R^{H\times W\times C}$, where l, T, H, W and C denote layer, current prediction time, map height, map width and feature channel respectively. The feature map guides the lower layers to generate images. STIC Attention is designed to assign weights to different rainfall regions in the second layer. The purpose of this weight assignment is to intensify the previous feature sequence in the spatiotemporal dimension, especially the heavy rainfall features. Then, the intensified R_l^T is fed to the next layer for more accurate image. The calculation method of STIC-Prediction is:

$$A_l^T = \begin{cases} x_T & if\ \ l = 0 \\ MAXPOOL(\gamma(f(E_{l-1}^T))) & 0 < l < L \end{cases}, \tag{1}$$

$$\hat{A}_l^T = \gamma(f(R_l^T)), \tag{2}$$

$$E_l^T = \left[\gamma\left(A_l^T - \hat{A}_l^T\right);\ \gamma(\hat{A}_l^T - A_l^T)\right], \tag{3}$$

$$R_l^T = \begin{cases} argcLSTM(E_l^{T-1},\ R_l^{T-1}) & if\ l = L \\ argcLSTM(E_l^{T-1},\ R_l^{T-1},\ UPSAMPLE(STIC(R_{l+1}^0 : R_{l+1}^T))) & if\ l = 2 \\ argcLSTM(E_l^{T-1},\ R_l^{T-1},\ UPSAMPLE(R_{l+1}^T)) & 0 \le l < 2\ and\ 2 < l < L \end{cases} \quad (4)$$

Here, x_T denotes the initial input, $MAXPOOL$ denotes the maximum pooling operation, γ denotes relu activation function, f denotes convolution operation, $argcLSTM$ denotes Advanced Reduced-Gate Convolutional LSTM (Zheng et al., 2022), $STIC$ denotes STIC Attention.

CS-Convolution discriminator is composed of a four-layer convolution structure and a CS Attention module. Convolution structure is responsible for extracting the echo features of input radar echo images. CS Attention is embedded after the first-layer convolution structure. The module is designed to intensify spatial information of echo features, especially heavy rainfall. The purpose is to enhance discriminative ability and provide better guidance for the generator. The hyperparameters of the generator and discriminator are provided in the supplement (Table S7 and S8).

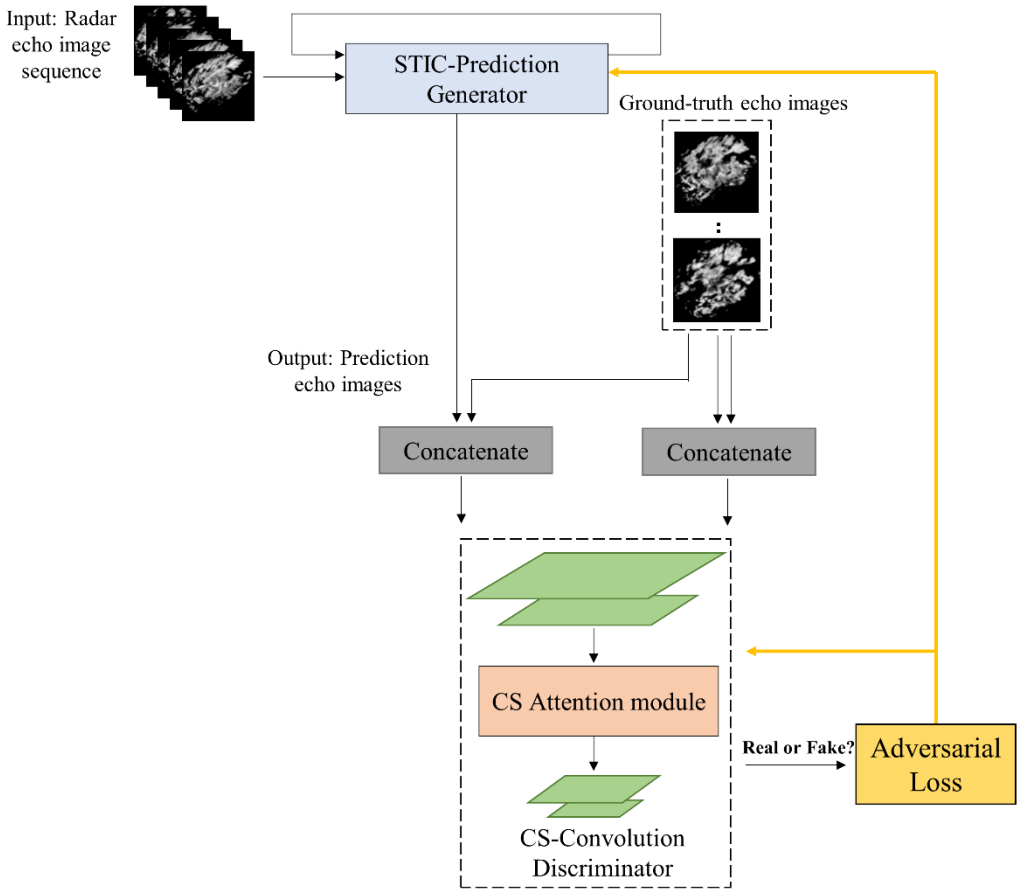

Figure 1: This is the structure of GAN-argcPredNet v2.0. Fifteen radar echo images are used in the testing set. Here, five images are used as the input sequence, and ten images are used as the ground-truth images.

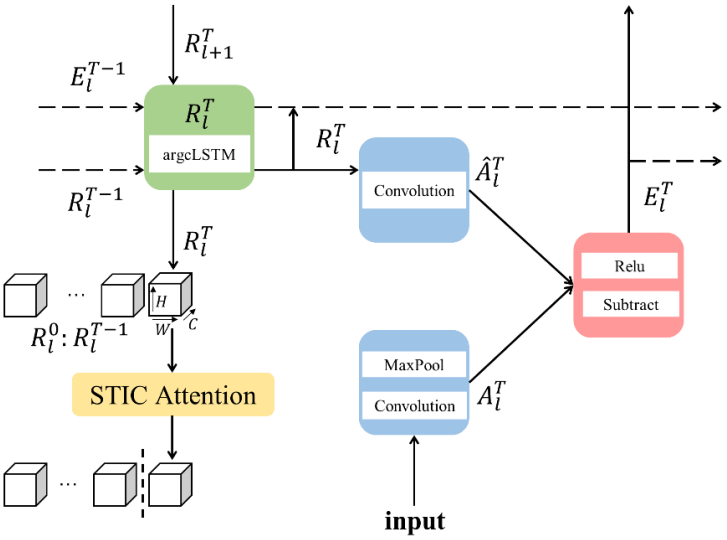

Figure 2: This is the structure of STIC-Prediction. Modules are in $l = 2$ at time T.

3.2 STIC Attention

The STIC Attention (Fig. 3) combines MaxPool3D (3D = map height, map width and time) and AvgPool3D to focus on the spatial information of feature sequences from both maximum and average perspectives. This step focuses on heavy rainfall echoes while considering non-heavy rainfall. The introduction of 3D convolution enables extraction of spatiotemporal changes in feature sequences. The module calculates weights for evolutionary information, which are then used to intensify the feature sequences. Following are the detailed steps.

Given the feature sequence $F \in R^{t\times H\times W\times C}$ as input, where t denotes time. Two feature sequences $F_{max}^{ts} \in R^{t\times H\times W\times 1}$ and $F_{avg}^{ts} \in R^{t\times H\times W\times 1}$, are obtained by pooling operation, which denote the maximum and average feature along the channel axis respectively. The feature sequences are then connected and 3D convoluted. The activation function is hard_sigmoid. By introducing linear behavior, hard_sigmoid allows gradients to flow easily in the unsaturated state, and provides a crisp decision in the saturated state, resulting in far less computational expense (Courbariaux et al., 2015; Gulcehre et al., 2016; Nwankpa et al., 2018). Then, the STIC Attention map sequence $M_{STIC} \in R^{t\times H\times W\times 1}$ is obtained. Finally, the output feature sequence $F_1 \in R^{t\times H\times W\times C}$, is calculated by element-wise multiplication of M_{STIC} and F. In short, the calculation method of STIC Attention is:

$$M_{STIC} = \sigma(f_{7\times7\times5}((MaxPool3D(F))concat(AvgPool3D(F)))) = \sigma(f_{7\times7\times5}(F_{max}^{ts}\text{concat}F_{avg}^{ts}))\,, \quad (5)$$

$$F_1 = M_{STIC}\otimes F\,, \quad (6)$$

Here, $concat$ denotes connection operation, $MaxPool3D$ denotes 3D maximum pooling operation, $AvgPool3D$ denotes 3D average pooling operation, $f_{7\times7\times5}$ denotes 3D convolution operation with convolution kernel of 7×7×5, and σ denotes hard_sigmoid activation function, ⊗ denotes element-wise multiplication.

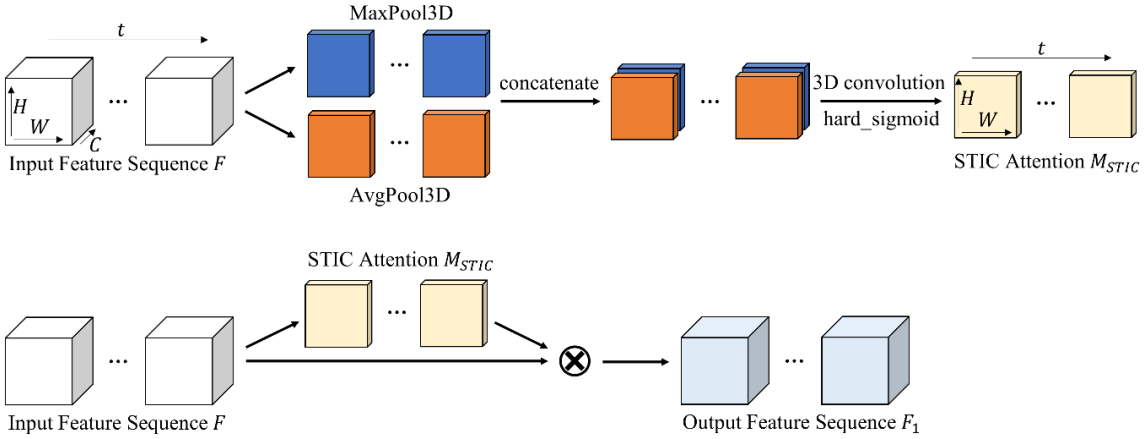

Figure 3: This is the structure of STIC Attention.

3.3 CS Attention

CS Attention consists of Channel Attention and Spatial Attention (Fig. 4). For input feature $F' \in R^{H\times W\times C}$, the channel attention map $M_c \in R^{1\times1\times C}$ is first generated. After element-wise multiplication with initial feature image, the spatial attention map $M_s \in R^{H\times W\times1}$, is generated by Spatial Attention module. Finally, the output feature $F_2' \in R^{H\times W\times C}$ is obtained in the same way. In short, the calculation process is as follows:

$$F_1' = M_c \otimes F' , \tag{7}$$

$$F_2' = M_s \otimes F_1' , \tag{8}$$

Here, ⊗ denotes element-wise multiplication, F_1' is the middle feature.

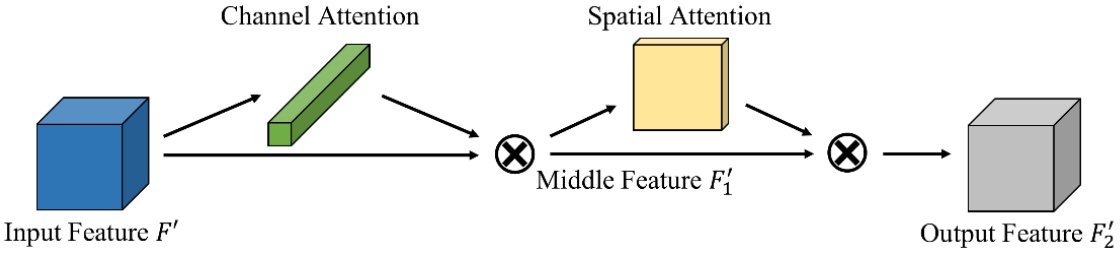

Figure 4: This is the structure of CS Attention.

The Channel Attention module (Fig. 5) studies the relationship between different feature channels. The global maximum and average pooling are used to gather spatial maximum and average information for each channel. The combination of these methods allows for a more comprehensive judgment of the importance of different feature channels. Then the correlation

between feature channels is obtained by learning the respective parameters in a dense layer. The Channel Attention assigns more weight to meaningful feature channels. The detailed steps are as follows:

The feature map F' is input into Channel Attention module. Two 1D feature maps $F_{max}^{c} \in R^{1\times1\times C}$ and $F_{avg}^{c} \in R^{1\times1\times C}$, are obtained by global pooling, which denote the global maximum and global average pooling features respectively. Then, the correlation between features is extracted through dense layers. In order to reduce parameter overhead, the number of neurons in the first dense layer is set to C/r, where r is the compression ratio. Finally, hard_sigmoid is used as the activation function and obtain final channel attention map M_c. In short, the channel attention map is calculated as follows:

$$M_c = \sigma(\varphi_{11}(\gamma(\varphi_{10}(GMP(F')))) + \varphi_{21}(\gamma(\varphi_{20}(GAP(F'))))) = \sigma(\varphi_{11}(\gamma(\varphi_{10}(F_{max}^{c}))) + \varphi_{21}(\gamma(\varphi_{20}(F_{avg}^{c})))) , \quad (9)$$

Here, GMP denotes global maximum pooling, GAP denotes global average pooling, φ_{10}, φ_{11} denotes the first and second dense layer of F_{max}^{c}, φ_{20}, φ_{21} denotes the first and second dense layer of F_{avg}^{c} and γ denotes relu activation function.

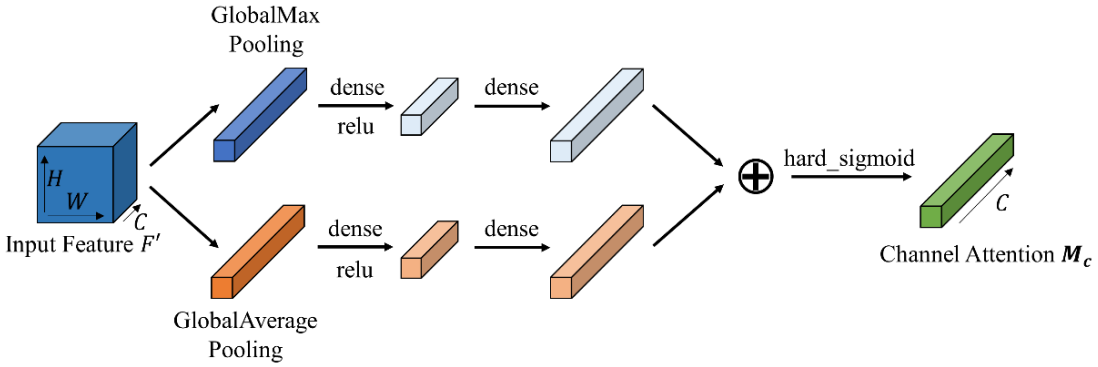

Figure 5: This is the structure of Channel Attention.

The Spatial Attention module (Fig. 6) studies the importance of each part in the same channel. The maximum and average pooling are used along the channel axis, which obtains echoes information of the feature image. The 2D convolution operation extracts feature and generates a spatial attention map with the same size as input image. The detailed steps are as follows:

After the Channel Attention module, the feature map $F_1' \in R^{H\times W\times C}$ is input into the Spatial Attention module. Two 2D feature maps $F_{max}^{s} \in R^{H\times W\times1}$ and $F_{avg}^{s} \in R^{H\times W\times1}$, are obtained by pooling operation, which denote the maximum pool feature and average pool feature on the channel respectively. The feature maps are then connected and 2D convoluted, using hard_sigmoid as the activation function to obtain final spatial attention map M_s. In short, the calculation method of spatial attention map is:

$$M_s = \sigma(f_{7\times7}((MaxPool2D(F_1'))concat(AvgPool2D(F_1')))) = \sigma(f_{7\times7}(F_{max}^{s}\text{concat}F_{avg}^{s})) , \quad (10)$$

Here, $MaxPool2D$ denotes 2D maximum pooling operation, $AvgPool2D$ denotes 2D average pooling operation, and $f_{7\times7}$ denotes 2D convolution operation with convolution kernel of 7×7.

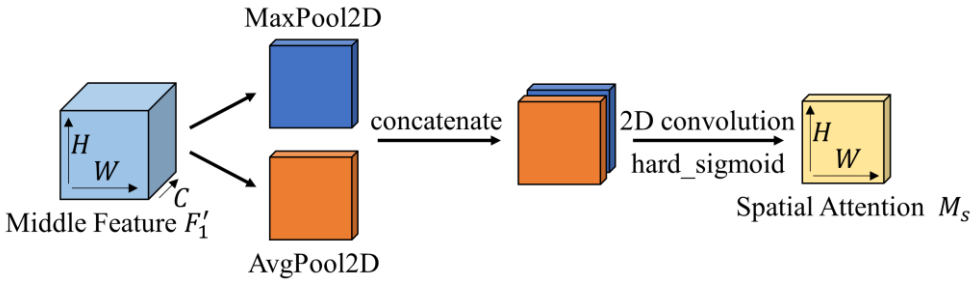

Figure 6: This is the structure of Spatial Attention.

4 Data and experimental setup

4.1 Dataset description

The paper uses the South China radar echo data provided by Guangzhou Meteorological Administration. The radar mosaic comes from 11 weather radars. The median filtering algorithm is used to control radar data quality, which eliminates errors caused by isolated clutter. In addition, the mirror filling and continuity checks are applied to remove traditional radar error sources. After quality control, there is only an extremely small amount of strong interference, which has negligible impact on the training of the model.

From 2015 to 2016, 32,010 consecutive echo images with rainfall are randomly selected as the training set. 7,995 consecutive images are randomly selected for testing from March to May 2017. The original resolution of the image is 1050 × 880, and each image covers 1050 km × 880 km. The pixel denotes the resolution of 1 km × 1 km. The reflectivity range is 0-80 dBZ, and the amplitude limit is between 0 and 255. The data is collected every 6 minutes, with the height of 1 km. To speed up the training and reduce the hardware load, the central 128 × 128 images are segmented.

Due to the relationship between radar reflectivity and rainfall type, the value on the radar echo image is converted to the corresponding rainfall. The calculation formula is as follows:

$$Z = 10\,log_{10}\,a + 10b\,log_{10}\,R\,, \tag{11}$$

Here, a is set to 58.53 and b is set to 1.56, Z denotes radar reflectivity intensity, R denotes rainfall intensity. The correspondence between rainfall, rainfall intensity and rainfall level are referred to Table 1 (Shi et al., 2017).

Table 1: This is the rainfall level table.

Rain Rate (mm h^{-1})	Radar Reflectivity Intensity (dBZ)	Rainfall Level
$0 \leq R < 0.5$	$Z < 12.98$	No / Hardly noticeable
$0.5 \leq R < 2$	$12.98 \leq Z < 22.37$	Light
$2 \leq R < 5$	$22.37 \leq Z < 28.58$	Light to moderate
$5 \leq R < 10$	$28.58 \leq Z < 33.27$	Moderate
$10 \leq R < 30$	$33.27 \leq Z < 40.72$	Moderate to heavy
$30 \leq R$	$40.72 \leq Z$	Rainstorm warning

4.2 Evaluation metrics

As for evaluation, the paper uses four metrics to evaluate the prediction accuracy of all 128 × 128 pixels, which are Probability of Detection (POD), False Alarm Ratio (FAR), Critical Success Index (CSI) and Heidke Skill Score (HSS). POD evaluates hit ability and FAR is the metric of false alarms. The combination of them can evaluate the model more objectively. CSI and HSS are two composite metrics that provide a direct judgment of model effectiveness. CSI measures the fraction of observed and/or forecast events that are correctly predicted. HSS measures the fraction of correct forecasts after eliminating those forecasts which would be correct due purely to random chance. To measure information loss, the paper also uses the Bias metric, which evaluates the ratio of the frequency of forecast events to the frequency of observed events. The formulas for calculating these five metrics are as follows:

$$POD = \frac{TP}{TP+FN}, \tag{12}$$

$$FAR = \frac{FP}{TP+FP}, \tag{13}$$

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

$$HSS = \frac{2(TP \times TN - FN \times FP)}{(TP+FN)(FN+TN)+(TP+FP)(FP+TN)}, \tag{15}$$

$$Bias = \frac{TP+FP}{TP+FN}, \tag{16}$$

Here, TP denotes that the real and predicted value reach specified threshold, FN denotes that the real value reaches the specified threshold and the predicted value does not reach, FP denotes that the real value does not reach specified threshold and the predicted value reaches, and TN denotes that the real value and predicted value do not reach specified threshold. In the study, we have applied threshold rain rates of 0.5, 2, 5, 10 and 30mm h^{-1} for calculating these metrics.

In order to evaluate the quality of generated images objectively, Mean Square Error (MSE) and Mean Structural Similarity (MSSIM) are also chosen for the experiment (Wang et al., 2004; Inoue and Misumi, 2022).

4.3 Experimental setting

Radar echo extrapolation is the prediction of future radar echo images based on real images. This paper sets the input sequence M and output sequence N to 5 and 10, respectively.

GAN-argcPredNet v2.0 is first compared with ConvLSTM, ConvGRU, GA-ConvGRU, and GAN-argcPredNet v1.0 in comparison experiments. The first three are common models in radar echo extrapolation, and the last one is the model we designed before. The hyperparameters for all models are provided in the supplement (Table S1-6). Then the ablation study is designed to verify the effectiveness of STIC Attention and CS Attention.

Before training, each pixel of the radar echo image is normalized to [0, 1]. All experiments are implemented by Python. Model training and testing based on the Keras deep learning library with Tensorflow as backend. The operating environment is a Linux workstation equipped with two NVDIA RTX 2080 Ti 11G GPU.

5 Results

5.1 Comparison study

In order to observe the performance of the models and evaluate information loss more easily, the average scores of POD, FAR, CSI, HSS and Bias were calculated for all lead time in the experiment. Fig. 7 shows that the FAR scores of GAN-argcPredNet v2.0 are slightly higher than those of ConvGRU and ConvLSTM at thresholds of 0.5 and 2 mm h^{-1}, respectively. However, the other scores are always the best. The comprehensive score of GAN-argcPredNet v1.0 is second only to GAN-argcPredNet v2.0. GA-ConvGRU performs better than the two non-GAN models in most cases. Among non-GAN models, all scores of ConvLSTM is better than ConvGRU when the threshold is 5 and 10 mm h^{-1}.

GAN-argcPredNet v2.0 shows excellent performance in heavy rainfall prediction. Using GAN-argcPredNet v1.0 as baseline, the POD, CSI, HSS and Bias scores of GAN-argcPredNet v2.0 increase by 20.1 %, 16.0 %, 15.4 % and 25.0 % when the threshold is 10 mm h^{-1}. The FAR score also decreases by 2.3 %. When the threshold is 30 mm h^{-1}, the POD, CSI, HSS and Bias scores increase by 18.8 %, 17.0 %, 17.2% and 26.3 % respectively. The FAR score decreases by 3.0 %.

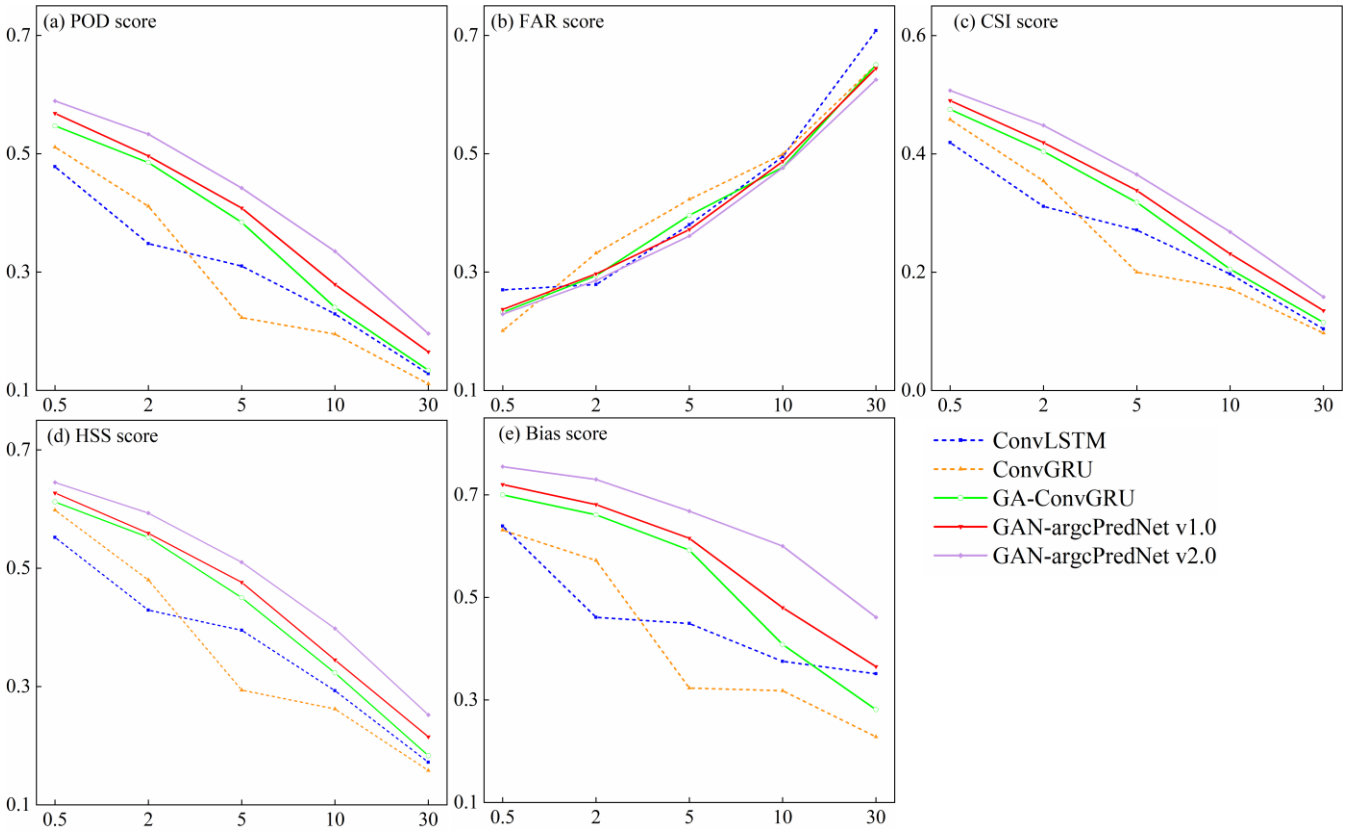

Figure 7: This is the average scores of (a) POD, (b) FAR, (c) CSI, (d) HSS and (e) Bias under different threshold. The horizontal axis represents the threshold, in units of mm h^{-1}. The perfect score of FAR is 0, and the others are 1.

The experiment selected a prediction example and drew the extrapolation comparison image (Fig. 8). As the lead time increases, all models exhibit the phenomenon of echo attenuation in the predicted images. The attenuation of GAN-argcPredNet v2.0 is obviously slower. In the circular and rectangular regions, the intensity and shape of high echoes are well retained until 60 minutes. There are some false predictions on the bottom right corner and the heavy rain area near the top right goes out of the domain, but the other predictions of GAN-argcPredNet v2.0 are more consistent with Ground-truth. The three GAN-based models also present better echo structure and prediction performance compared to the two non-GAN models. For the echo above 40 dBZ, ConvLSTM has the most severe attenuation.

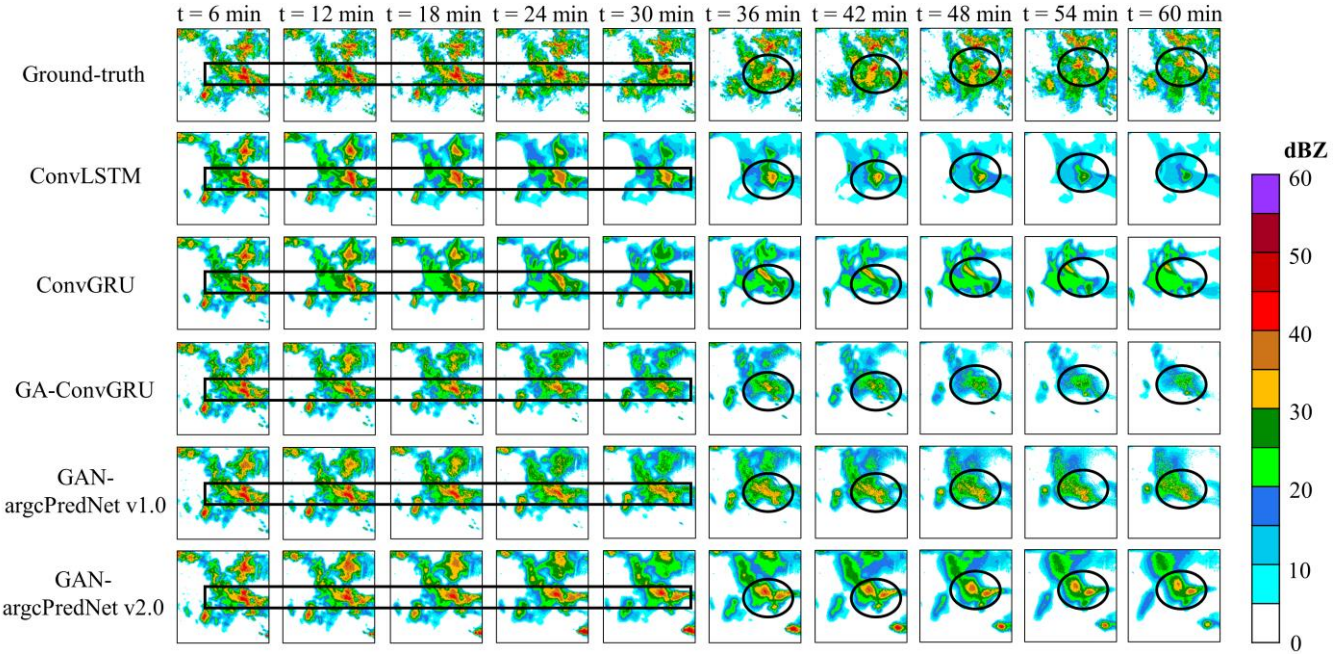

Figure 8: This is the example of radar echo exploration. The circular and rectangular regions represent heavy rainfall prediction.

To evaluate the quality of generated images objectively, the experiment also calculated the MSE and MSSIM metrics. Table 2 shows that GAN-based models have better scores in both MAE and MSSIM, with GAN-argcPredNet v2.0 achieving the best score. Compared to GAN-argcPredNet v1.0, the MAE metric of GAN-argcPredNet v2.0 decreased by 2.3 %, while the MSSIM metric increased by 1.25 %. In the non-GAN models, ConvLSTM scores better.

Table 2: This is the MSE and MSSIM scores of each model. The perfect score of MSE is 0 while the perfect score for MSSIM is 1. Bold represents the best score.

Model	MSE $\times 10^2$	MSSIM
ConvLSTM	0.230	0.775
ConvGRU	0.243	0.763
GA-ConvGRU	0.223	0.782
GAN-argcPredNet v1.0	0.218	0.797
GAN-argcPredNet v2.0	**0.213**	**0.807**

5.2 Ablation study

In the ablation study, we investigated the effects of STIC Attention and CS Attention. STIC-GAN and CS-GAN are constructed by adding STIC Attention module only in generator and CS Attention module only in discriminator. To more precisely observe the score differences, the experiment recorded the average score of each metric in a table format. Table 3 shows that both CS-GAN and STIC-GAN have better scores than GAN-argcPredNet v1.0. At the threshold of 0.5 mm h^{-1}, CS-GAN achieves the

best score in the FAR metric, but other metrics still perform worse than GAN-argcPredNet v2.0. At thresholds of 5, 10, and 30 mm h^{-1}, the metric scores of STIC-GAN exceed those of CS-GAN, which is closer to the GAN-argcPredNet v2.0.

Table 2: This is the average scores of POD, FAR, CSI, HSS and Bias under different threshold. Bold represents the best score.

Model	Threshold = 0.5 mm h^{-1}					Threshold = 2 mm h^{-1}				
	POD	FAR	CSI	HSS	Bias	POD	FAR	CSI	HSS	Bias
GAN-argcPredNet v1.0	0.568	0.237	0.490	0.627	0.720	0.496	0.297	0.419	0.559	0.681
CS-GAN	0.569	**0.221**	0.496	0.633	0.735	0.529	0.289	0.443	0.588	0.726
STIC-GAN	0.570	0.241	0.490	0.627	0.739	0.512	0.289	0.433	0.576	0.700
GAN-argcPredNet v2.0	**0.589**	0.229	**0.507**	**0.645**	**0.755**	**0.533**	**0.286**	**0.448**	**0.593**	**0.730**
Model	Threshold = 5 mm h^{-1}					Threshold = 10 mm h^{-1}				
	POD	FAR	CSI	HSS	Bias	POD	FAR	CSI	HSS	Bias
GAN-argcPredNet v1.0	0.408	0.372	0.338	0.476	0.615	0.279	0.487	0.231	0.345	0.480
CS-GAN	0.411	0.381	0.340	0.478	0.629	0.283	0.482	0.235	0.351	0.490
STIC-GAN	0.433	0.379	0.354	0.496	0.667	0.317	0.477	0.258	0.383	0.560
GAN-argcPredNet v2.0	**0.442**	**0.361**	**0.365**	**0.510**	**0.668**	**0.335**	**0.476**	**0.268**	**0.398**	**0.600**

Model	Threshold =30 mm h^{-1}				
	POD	FAR	CSI	HSS	Bias
GAN-argcPredNet v1.0	0.165	0.644	0.135	0.215	0.365
CS-GAN	0.171	0.637	0.141	0.224	0.383
STIC-GAN	0.194	0.632	0.156	0.248	0.457
GAN-argcPredNet v2.0	**0.196**	**0.625**	**0.158**	**0.252**	**0.461**

Fig. 9 shows that CS-GAN and STIC-GAN retain better echo intensity and shape compared to GAN-argcPredNet v1.0. For heavy rainfall, STIC-GAN shows better prediction with echoes closer to GAN-argcPredNet v2.0. In rectangular regions, the rainfall events are well predicted by STIC-GAN until 30 minutes. However, in circular regions, STIC-GAN overestimates the echo intensity compared to GAN-argcPredNet v2.0.

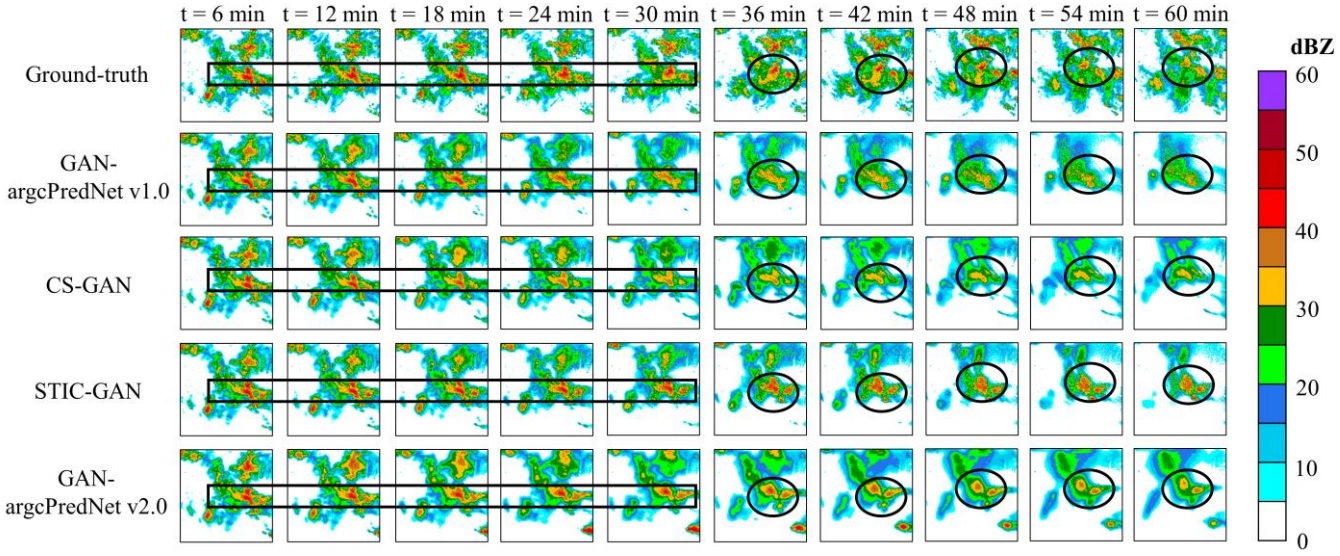

Figure 9: This is the example of radar echo exploration. The circular and rectangular regions represent heavy rainfall prediction.

6 Discussions and conclusions

6.1 Discussions

In the comparison study, the three GAN-based models have higher prediction accuracy and better image quality compared to the two non-GAN models (Fig. 7 and Table 2). This indicates that GAN structure has more advantages with its powerful image generation capability. Although ConvLSTM and ConvGRU sometimes have lower FAR scores, according to metrics such as POD, this is due to the fact that they fail to predict a large number of rainfall events. As the number of predicted rainfall events decreases, the false alarm rate also decreases. Combined with the Bias metric, ConvLSTM and ConvGRU generally predict lower rainfall frequency than other models. This indicates that they suffer from more severe information loss. GA-ConvGRU and GAN-argcPredNet v1.0 improve the prediction accuracy with GAN structure, but the phenomenon of information loss also cannot be ignored.

GAN-argcPredNet v2.0 has the highest prediction accuracy and quality while predicting more rainfall events, which demonstrates that it successfully curbs information loss and echo attenuation compared to other models. This is because important information is given more weight by intensifying the previous rainfall evolution process, making it easier to be retained. Meanwhile, the intensification of spatial information in the discriminator can also better guide the generator to retain information. According to the scores at the thresholds of 10 and 30 mm h^{-1}, it can be observed that GAN-argcPredNet v2.0 exhibits the best improvement in predicting heavy rainfall.

GAN-argcPredNet v2.0 maintains the evolution trend well, but there are still some special cases (Fig. 8). There are some false predictions on the bottom right corner, and the rain area near the top right goes out of the domain. This is because in the input sequence (Fig. 10), the rainfall on the bottom right corner shows a growing trend, and the rain area near the top right shows a

trend of moving towards the upper left. Our model retains the information and maintains the trend through the attention mechanism, but there are still some deviations from the Ground-truth. This may be overcome by increasing the dataset and strengthening training on special cases. Compared to other models, the overall result of GAN-argcPredNet v2.0 is more competitive.

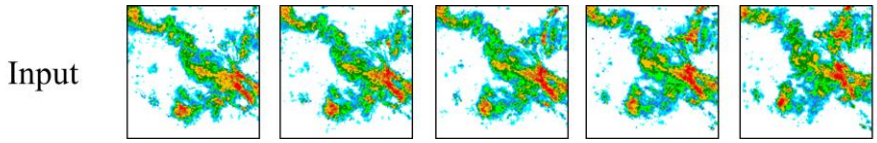

Figure 10: This is the input radar images sequence. The example in Fig. 8 is extrapolated based on these five images.

In the ablation study, we find that the scores of CS-GAN and STIC-GAN are mostly better than GAN-argcPredNet v1.0 (Table 3 and Fig. 9). This indicates that both STIC Attention and CS Attention can reduce information loss and improve prediction accuracy. As the threshold increases, the scores of STIC-GAN become closer to that of GAN-argcPredNet v2.0. This is because the STIC Attention focuses more on the rapidly evolving areas of the radar images. The CS Attention helps GAN-argcPredNet v2.0 achieve better overall performance.

6.2 Conclusions

The study improves precipitation nowcasting by reducing information loss and echo attenuation. With the intensification of the rainfall evolution information, GAN-argcPredNet v2.0 reduces information loss and improves the rainfall prediction accuracy, especially for heavy rainfall. STIC Attention intensifies the previously input feature sequence, allowing the generator to curb echo attenuation. CS Attention intensifies the spatial information of features, enabling the discriminator to better guide the generator to retain information. Meanwhile, the model is designed based on the generative adversarial structure, which achieves high-quality radar echo extrapolation.

In practice, a predictive software has been developed based on our model. After the software accesses the radar data and establishes a prediction task, rainfall prediction results are output as dataset. Then the dataset can be fed into the urban flood warning system. The improvement of rainfall prediction has a positive impact on flood prediction and urban-operation safety. Overall, GAN-argcPredNet v2.0 is a spatiotemporal process intensification model based on GAN, which achieves more accurate rainfall prediction.

Future work can be considered from two aspects. The prediction accuracy of the model proposed in the study still has room for improvement. False predictions may be reduced by increasing the dataset and strengthening training on special cases. High-resolution prediction is often limited by hardware, such as graphics card. Therefore, it is possible to reduce the need for hardware by optimizing the algorithm complexity and the number of parameters.

Code and data availability

The radar data used in the paper comes from Guangdong Meteorological Administration. Due to the confidentiality policy, we only provide a sequence of 12 images. If you need to access more data, please contact Kun Zheng (ZhengK@cug.edu.cn) and Qiya Tan (ses_tqy@cug.edu.cn). The GAN-argcPredNet v2.0 model is open source. You can find the source code from https://doi.org/10.5281/zenodo.7505030.

Supplement

Table S1-8 are provided in Supplement.

Author contributions

Kun Zheng and Qiya Tan were responsible for developing models and writing manuscripts; Huihua Ruan, Jinbiao Zhang, Cong Luo, Siyu Tang Yunlei Yi, Yugang Tian and Jianmei Cheng were responsible for data screening and preprocessing.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

This work was supported by Science and Technology Planning Project of Guangdong Province, China [grant No.2018B020207012].

References

Austin, G. and Bellon, A.: The use of digital weather radar records for short-term precipitation forecasting, Quarterly Journal of the Royal Meteorological Society, 100, 658-664, 1974.

Ayzel, G., Heistermann, M., and Winterrath, T.: Optical flow models as an open benchmark for radar-based precipitation nowcasting (rainymotion v0.1), Geoscientific Model Development, 12, 1387-1402, 10.5194/gmd-12-1387-2019, 2019.

Ayzel, G., Scheffer, T., and Heistermann, M.: RainNet v1.0: a convolutional neural network for radar-based precipitation nowcasting, Geoscientific Model Development, 13, 2631-2644, 10.5194/gmd-13-2631-2020, 2020.

Ballas, N., Yao, L., Pal, C., and Courville, A.: Delving deeper into convolutional networks for learning video representations, arXiv preprint arXiv:1511.06432, 2015.

Bowler, N. E., Pierce, C. E., and Seed, A.: Development of a precipitation nowcasting algorithm based upon optical flow techniques, Journal of Hydrology, 288, 74-91, 10.1016/j.jhydrol.2003.11.011, 2004.

Chen, H. G., Zhang, X., Liu, Y. T., and Zeng, Q. Y.: Generative Adversarial Networks Capabilities for Super-Resolution Reconstruction of Weather Radar Echo Images, Atmosphere, 10, 10.3390/atmos10090555, 2019.

Courbariaux, M., Bengio, Y., and David, J. P.: Binaryconnect: Training deep neural networks with binary weights during propagations, Advances in neural information processing systems, 28, 2015.

Dixon, M. and Wiener, G.: TITAN: Thunderstorm identification, tracking, analysis, and nowcasting-A radar-based methodology, Journal of Atmospheric and Oceanic Technology, 10, 785-797, 1993.

Foresti, L., Sideris, I. V., Nerini, D., Beusch, L., and Germann, U.: Using a 10-year radar archive for nowcasting precipitation growth and decay: A probabilistic machine learning approach, Weather and Forecasting, 34, 1547-1569, 10.1175/WAF-D-18-0206.1, 2019.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y.: Generative adversarial networks, Communications of the ACM, 63, 139-144, 2020.

Gulcehre, C., Moczulski, M., Denil, M., and Bengio, Y.: Noisy activation functions, International conference on machine learning, 3059-3068, 2016.

Han, L., Fu, S. X., Zhao, L. F., Zheng, Y. G., Wang, H. Q., and Lin, Y. J.: 3D Convective Storm Identification, Tracking, and Forecasting-An Enhanced TITAN Algorithm, Journal of Atmospheric and Oceanic Technology, 26, 719-732, 10.1175/2008jtecha1084.1, 2009.

Inoue, T. and Misumi, R.: Learning from Precipitation Events in the Wider Domain to Improve the Performance of a Deep Learning-based Precipitation Nowcasting Model, Weather and Forecasting, 37, 1013-1026, 10.1175/WAF-D-21-0078.1, 2022.

Kingma Diederik, P. and Adam, J. B.: A method for stochastic optimization, arXiv preprint arXiv:1412.6980, 2014.

Liguori, S. and Rico-Ramirez, M. A.: A review of current approaches to radar-based quantitative precipitation forecasts, International Journal of River Basin Management, 12, 391-402, 2014.

Liu, Y., Xi, D.-G., Li, Z.-L., and Hong, Y.: A new methodology for pixel-quantitative precipitation nowcasting using a pyramid Lucas Kanade optical flow approach, Journal of Hydrology, 529, 354-364, 2015.

Lotter, W., Kreiman, G., and Cox, D.: Deep predictive coding networks for video prediction and unsupervised learning, arXiv preprint arXiv:1605.08104, 2016.

Luo, C., Li, X., and Ye, Y.: PFST-LSTM: A spatiotemporal LSTM model with pseudoflow prediction for precipitation nowcasting, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 843-857, 10.1109/JSTARS.2020.3040648, 2020.

Luo, C. Y., Li, X. T., Wen, Y. L., Ye, Y. M., and Zhang, X. F.: A Novel LSTM Model with Interaction Dual Attention for Radar Echo Extrapolation, Remote Sensing, 13, 10.3390/rs13020164, 2021.

Nwankpa, C., Ijomah, W., Gachagan, A., and Marshall, S.: Activation functions: Comparison of trends in practice and research for deep learning, arXiv preprint arXiv:1811.03378, 2018.

Pan, X., Lu, Y. H., Zhao, K., Huang, H., Wang, M. J., and Chen, H. A.: Improving Nowcasting of Convective Development by Incorporating Polarimetric Radar Variables into a Deep-Learning Model, Geophysical Research Letters, 48, 10.1029/2021gl095302, 2021.

Pierce, C. E., Ebert, E., Seed, A. W., Sleigh, M., Collier, C. G., Fox, N. I., Donaldson, N., Wilson, J. W., Roberts, R., and Mueller, C. K.: The nowcasting of precipitation during Sydney 2000: An appraisal of the QPF algorithms, Weather and Forecasting, 19, 7-21, 2004.

Ravuri, S., Lenc, K., Willson, M., Kangin, D., Lam, R., Mirowski, P., Fitzsimons, M., Athanassiadou, M., Kashem, S., Madge, S., Prudden, R., Mandhane, A., Clark, A., Brock, A., Simonyan, K., Hadsell, R., Robinson, N., Clancy, E., Arribas, A., and Mohamed, S.: Skilful precipitation nowcasting using deep generative models of radar, Nature, 597, 672-+, 10.1038/s41586-021-03854-z, 2021.

Reyniers, M.: Quantitative precipitation forecasts based on radar observations: Principles, algorithms and operational systems, Institut Royal Météorologique de Belgique Brussel, Belgium, 2008.

Sakaino, H.: Spatio-Temporal Image Pattern Prediction Method Based on a Physical Model with Time-Varying Optical Flow, IEEE Transactions on Geoscience and Remote Sensing, 51, 3023-3036, 10.1109/tgrs.2012.2212201, 2013.

Shi, X. J., Chen, Z. R., Wang, H., Yeung, D. Y., Wong, W. K., and Woo, W. C.: Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting, 29th Annual Conference on Neural Information Processing Systems (NIPS), Montreal, CANADA, Dec 07-12, 2015.

Shi, X. J., Gao, Z. H., Lausen, L., Wang, H., Yeung, D. Y., Wong, W. K., and Woo, W. C.: Deep Learning for Precipitation Nowcasting: A Benchmark and A New Model, 31st Annual Conference on Neural Information Processing Systems (NIPS), Long Beach, CA, Dec 04-09, 2017.

Sutskever, I., Vinyals, O., and Le, Q. V.: Sequence to sequence learning with neural networks, Advances in neural information processing systems, 27, 2014.

Tian, L., Li, X. T., Ye, Y. M., Xie, P. F., and Li, Y.: A Generative Adversarial Gated Recurrent Unit Model for Precipitation Nowcasting, IEEE Geoscience and Remote Sensing Letters, 17, 601-605, 10.1109/lgrs.2019.2926776, 2020.

Wang, X., Girshick, R., Gupta, A., and He, K.: Non-local neural networks, Proceedings of the IEEE conference on computer vision and pattern recognition, 7794-7803, 2018a.

Wang, Y., Jiang, L., Yang, M.-H., Li, L.-J., Long, M., and Fei-Fei, L.: Eidetic 3D LSTM: A model for video prediction and beyond, International conference on learning representations, 2018b.

Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P.: Image quality assessment: from error visibility to structural similarity, IEEE transactions on image processing, 13, 600-612, 10.1109/TIP.2003.819861, 2004.

Woo, S., Park, J., Lee, J.-Y., and Kweon, I. S.: Cbam: Convolutional block attention module, Proceedings of the European conference on computer vision (ECCV), 3-19, 2018.

Xie, P. F., Li, X. T., Ji, X. Y., Chen, X. L., Chen, Y. Z., Liu, J., and Ye, Y. M.: An Energy-Based Generative Adversarial Forecaster for Radar Echo Map Extrapolation, IEEE Geoscience and Remote Sensing Letters, 19, 10.1109/lgrs.2020.3023950, 2022.

Zheng, K., Liu, Y., Zhang, J. B., Luo, C., Tang, S. Y., Ruan, H. H., Tan, Q. Y., Yi, Y. L., and Ran, X. T.: GAN-argcPredNet v1.0: a generative adversarial model for radar echo extrapolation based on convolutional recurrent units, Geoscientific Model Development, 15, 1467-1475, 10.5194/gmd-15-1467-2022, 2022.