



OPEN

A method of radar echo extrapolation based on dilated convolution and attention convolution

Xiajong Shen^{1,2}, Kunying Meng¹, Lei Zhang^{1,2}✉ & Xianyu Zuo¹

The neural network method can obtain a higher precision of radar echo extrapolation than the traditional method. However, its application in radar echo extrapolation is still in the initial stage of exploration, and there is still much room for improvement in the extrapolation accuracy. To improve the utilization of radar echo information and extrapolation accuracy, this paper proposes a radar echo extrapolation model (ADC_Net) based on dilated convolution and attention convolution. In this model, dilated convolution, instead of the pooling operation, is used to downsample the feature matrix obtained after the standard convolution operation. In doing so, the internal data structure of the feature matrix is retained, and the spatial features of radar echo data from different scales are extracted as well. Besides, the attention convolution module is integrated in the ADC_Net model to improve its sensitivity to the target features in the feature matrix and suppress the interference information. The proposed model is tested in the extrapolation of radar echo images in the next 90 min from five aspects—extrapolated image, POD index, CSI index, FAR index, and HSS index. The experimental results show that the model can effectively improve the accuracy of radar echo extrapolation.

Nowcasting was first proposed by Browning¹ in 1982, and it is mainly used in early warning of disasters like thunderstorms, severe convective weather, rainstorm, and snowstorm. Accurate short-term nowcasting has attracted much attention in the field of meteorological services. Its goal is to accurately and timely predict the local weather in the next 0 to 2 h, enabling weather stations to issue Urban Emergency Disaster Alerts in time. There are three main methods of nowcasting—radar echo extrapolation method, weather situation analysis and forecast method, and comprehensive analysis method, which is the most accurate and widely used method. However, the accuracy of comprehensive analysis method is based on the accuracy of the radar echo extrapolation method. Therefore, the research on the radar echo extrapolation plays a significant role in the forecast and prevention of severe weather.

The traditional radar echo extrapolation methods mainly include tracking radar echoes by correlation (TREC)². The Strom cell identification and tracking (SCIT)³ and optical flow⁴. TREC divides two consecutive radar echo images at adjacent times into multiple image subsets, calculates the maximum correlation coefficient of different subsets in the two images, and uses the maximum correlation coefficient to determine the optimal matching area, which is the moving position of the image subset. The TREC method is applicable to extrapolation of steadily changing radar echoes, but not for rapidly changing ones, such as radar echoes of convective precipitation clouds⁵. Based on TREC, Zhang Yaping et al.⁶ proposed difference image based tracking radar echo by correlations (DITREC), which tracks radar echo motion on the basis of the correlation method of differential images. DITREC can eliminate the disordered vector in the TREC vector field due to the rapid change of echoes, but when using the derived DITREC field to predict the precipitation field, the prediction accuracy depends on the Z-I relationship it adopts^{7,8}. However, the strong dependence of the Z-I relationship method on a specific climatic background results in the model's weak generalization ability. Another radar echo extrapolation method, SCIT, calculates the centroid position and volume of the echo monomer, matches the calculated information with the scanned data before and after the echo monomer and track it, and then extrapolates with multiple consecutive moments of the matching and tracking results^{9,10}. However, SCIT is only applicable to the

¹School of Computer and Information Engineering, Henan University, Kaifeng, Henan, China. ²Henan Provincial Key Laboratory of Big Data Analysis and Processing, Henan University, Kaifeng, Henan, China. ✉email: zhanglei@henu.edu.cn

identification, tracking and nowcasting of thunderstorm monomers¹¹, which limits its application in weather forecasting to a certain extent. The third traditional radar echo extrapolation method, optical flow, finds the corresponding relationship between previous images and the current image by using the change of the target point in the image sequence in the time domain and the correlation between adjacent images. Then the motion information of objects between adjacent images is calculated according to their correspondence. Optical flow obtains the motion vector field of the echo by calculating the optical flow field of the radar echo, and extrapolates the radar echo based on the motion vector field. Different from TREC, the extrapolation of optical flow is based on the changes of radar echoes, rather than the invariant characteristics of the echoes¹². As a result, optical flow can better track radar echo movement. The extrapolation ability of optical flow is superior to other extrapolation algorithms, especially in rapidly changing weather conditions. However, natural movement is dynamic, so radar echoes will be split, disappear, merge, etc., which is contradictory to the invariance assumption of the generation and disappearance of radar echoes by optical flow, greatly weakening the extrapolation of optical flow. In general, the traditional radar extrapolation method could only infer the echo position of the next moment according to the extrapolated images of several past radar echoes. It ignores the motion nonlinearity of the small and medium-scale atmospheric systems existing in the radar echoes under actual conditions. Therefore, historical radar echo data have not been used to full advantage.

With the development of neural network, its powerful modeling ability and unique advantages in big data processing have attracted extensive attention of scholars, and it has been widely used in the field of meteorology¹³. The application of neural network in meteorology sheds new light on the radar echo extrapolation method, and some scholars have conducted research in this area. Chen Jiahui et al.¹⁴ used BP neural network model to extrapolate radar echoes and demonstrated the effectiveness of neural network in radar echo extrapolation. Guo Hanyang et al.¹⁵ used convolutional gated recurrent unit (CGRU) to extrapolate radar echoes and compared them with TREC. The results show that the neural network model has more advantages than traditional methods. Klein¹⁶ added a dynamic convolutional layer to the convolutional neural networks (CNNs) structure and achieved accurate prediction of precipitation echoes by generating two prediction probability vectors. Based on Klein's model¹⁷, added recurrent neural network (RNN) to the dynamic convolutional layer and successfully built recurrent dynamic convolutional neural networks (RDCNN). The neural network has achieved good results in both forecast accuracy and forecast timeliness. As traditional long short-term memory (LSTM) cannot extract spatial features, Shi et al.¹⁸ replaced the change of LSTM input to state and the change of state to state with convolution in their model and proposed ConvLSTM. To adapt to the spatial-temporal variableness, Shi et al. further improved ConvLSTM in 2017 and proposed TrajGRU¹⁹. To adapt to the spatio-temporal dependence of radar echo images, Singh et al.²⁰ added a convolution structure to recurrent neural network and realized the radar-echo-image-sequence-based prediction. Zhang Lei et al.²¹ proposed an EDP model for sustainable e-agriculture. This weather radar echo prediction method is based on convolution neural network and long short-term memory networks. The results show that the neural network can effectively improve the prediction accuracy and effectiveness. He Liang et al.²² regarded precipitation forecasting as a picture-to-picture conversion problem, and used a U-net-structured convolutional neural network for forecasting. To extrapolate radar echoes, Zhang Dezheng et al.²³ improved the model's ability to represent the details of spatial structure information by deepening the convolutional layers in the neural network, which is composed of CGRUs.

To sum up, although the application of neural network in radar echo extrapolation is of great significance in utilizing historical radar echo data and improving the accuracy of radar echo extrapolation, the prediction results are still not satisfactory. They could be blurred or distorted, so the prediction accuracy needs to be improved. Therefore, this paper proposes a radar echo prediction model (ADC_Net) based on dilated convolution and attention convolution. To construct small object information and avoid the losing of internal data structure, the model substitutes dilated convolution for pooling in downsampling the feature matrix obtained after convolution. Also, the integration of an attention convolution module with an attention mechanism can improve the model's sensitivity to target features and suppress interference information, thus enhancing the extrapolation accuracy of radar echoes.

The model

The structure of the ADC_Net model based on dilated convolution and attention convolution proposed in this paper is shown in Fig. 1. It follows the encoder-decoder structure, and the network structure is completely symmetric. Firstly, the input image gets a hidden vector through convolution operation, activation function and downsample in the encoder. Then, the hidden vector is progressively upscaled to a higher spatial resolution by an upsampling operation in the decoder. The output is generated when the resolution of the hidden vector equals that of the input image.

As elementary building blocks, ADC_Net has nine recurrent convolution modules (RCNN), four dilated convolution (DCNN), four upsampling convolution modules (UPCNN), four attention convolution modules (ACNN), four skip connections, and a 1×1 standard convolution layer (cnn). RCNN is the basic module of the model, and its structure is shown in Fig. 2a. Because the key to radar echo extrapolation is the prediction of echo spatio-temporal sequence, RCNN has a recurrent memory unit that solves the temporal memory problem and a convolutional unit that extracts spatial features. Therefore, using the RCNN module in the model can enhance its ability to capture the changing laws of radar echoes. Dilated Convolution (DCNN) is used to downsample the feature matrix instead of pooling operation. In the training process, the model can not only extract the spatial features of radar echo data from different scales, but also retain the internal data structure of the feature matrix after downsample. It avoids the problems of the pooling operation, such as losing internal data structure and being unable to reconstruct information of small objects. The upsampling convolution module (UPCNN) is used to gradually restore the hidden vector obtained in the neural network encoder to the same size as the input

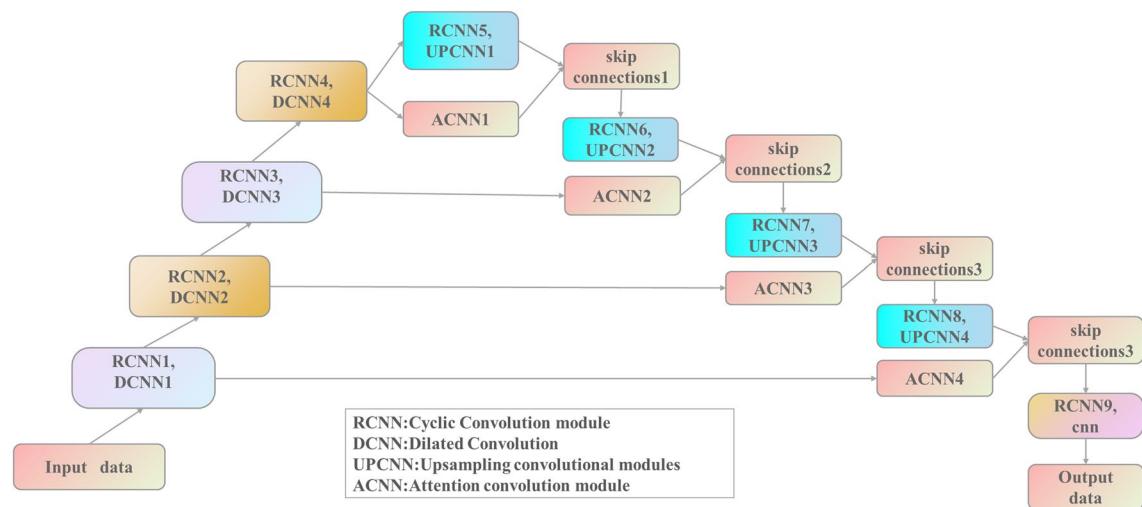


Figure 1. ADC_Net model structure.

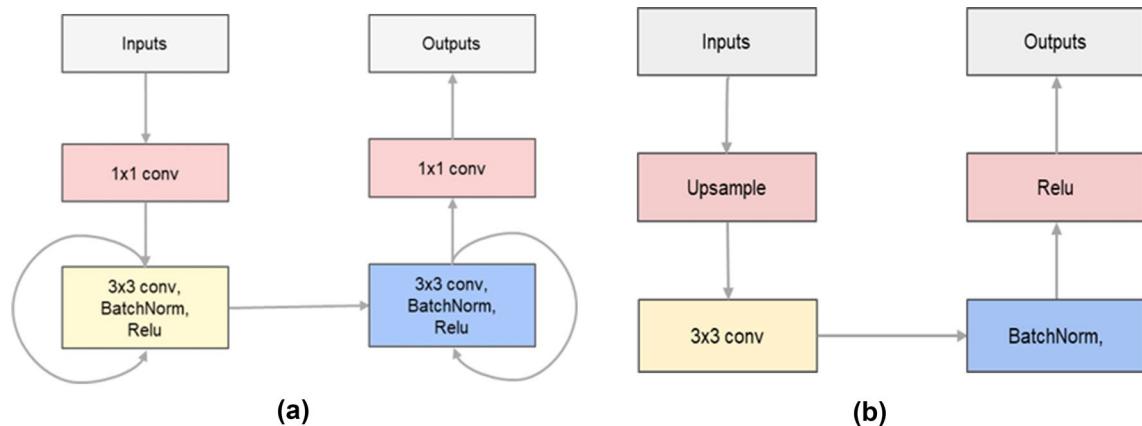


Figure 2. (a) RCNN module, (b) UPCNN module.

data. The specific module structure is shown in Fig. 2b. The attention convolution module (ACNN) assigns different weights to each pixel of the feature matrix during model training and prediction. The pixel information with stronger correlation with the current radar echo extrapolation task is given a higher weight value, and the pixel information with weaker correlation with the current radar echo extrapolation task is given a lower weight value. The skip connections combine the weight matrix obtained by ACNN with the high-level feature matrix in the decoder, and assign the attention weight to the high-level feature matrix, thereby improving the model's sensitivity to pixel information related to the current radar echo extrapolation task and suppressing irrelevant pixel information.

In the above module, dilated convolution instead of pooling operation is employed to downsample the feature matrix in the model and the attention convolution module is used to improve the model's sensitivity to the target features in the feature matrix and suppress the interference information. As the dilated convolution and the attention convolution module are the two main innovative structures of the ADC_Net model proposed in this paper, they will be illustrated in the following paragraphs.

Dilated convolution. Dilated convolution, also known as convolution kernel expansion operation, is a convolution method commonly used in pixel-wise output models²⁴. To increase the reception field, dilated convolution injects holes on the basis of the convolution map of standard convolution. Therefore, dilation rate, a hyper-parameter, is added to dilated convolution. The hyper-parameter refers to the number of intervals in the kernel, meaning filling the dilation rate-1 zeros in the convolution kernel. Figure 3 shows the difference between the convolution kernels of dilated convolution and standard convolution. The left side represents the 3×3 convolution kernel of standard convolution, and the right side represents the 3×3 convolution kernel of dilated convolution with the dilation rate of 2. The middle white part of the convolution kernel is filled with 0, and the unilateral size of the convolution kernel of dilated convolution becomes $2 * (3 - 1) + 1$ after filling, as is shown in formula (2).

Dilated convolution expands the convolution kernel to the scale of the dilated scale constraint, and fills the area not occupied by the original convolution kernel with zeros. This ensures the exponential expansion of the

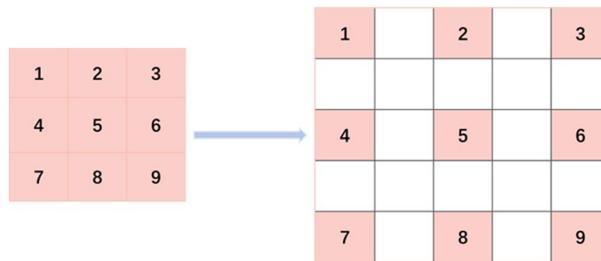


Figure 3. Difference between convolution kernels of dilated convolution and standard convolution.

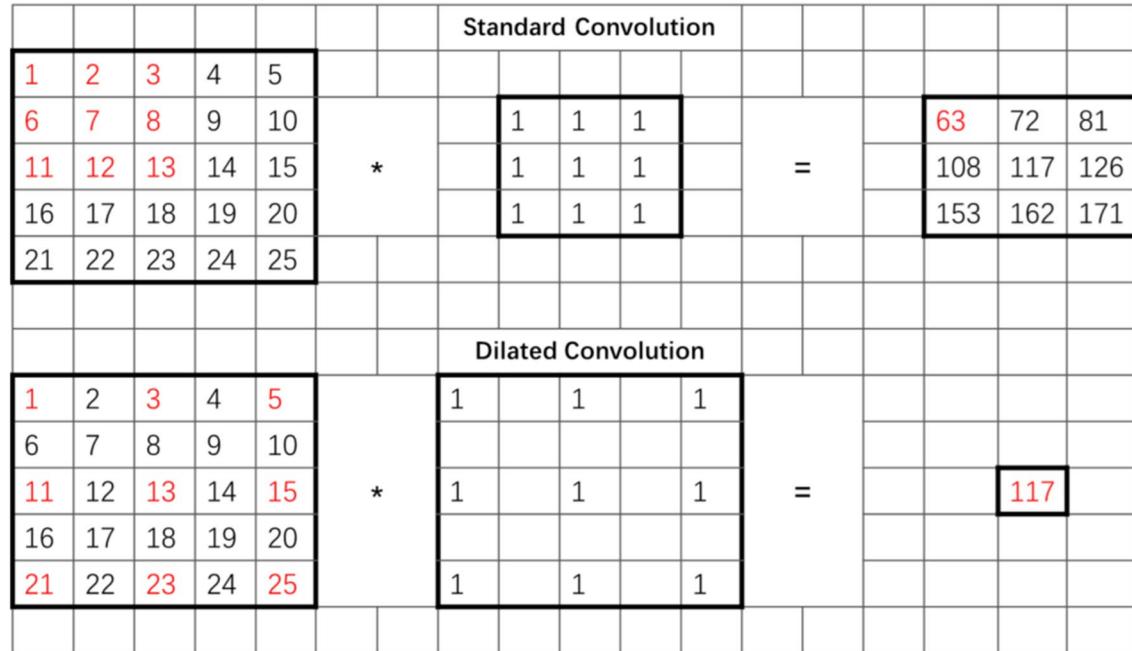


Figure 4. The difference between the reception field and the output of standard convolution and dilated convolution.

reception field by dilated convolution without losing resolution or coverage. Therefore, in dilated convolution, the size of the feature matrix before and after convolution is not only related to the convolution kernel, convolution step size, and padding, but also affected by the dilation rate. For standard convolution, the calculation method of the feature matrix size after the convolution layer is shown in Eq. (1). Here, h_1 represents the length (or width) of the input feature map; K is the size of the convolution kernel; p is the padding value; s represents the stride; and h_2 represents the length (or width) of the output feature map. When calculating the size of the feature map of dilated convolution after passing through the convolutional layer, the influence of the dilation rate on the size of the convolution kernel should be considered first. Equation (2) is the calculation formula of the size of the convolution kernel after expansion; ke is the size of the convolution kernel after the expansion; k is the size of the convolution kernel before the expansion; and r is the expansion coefficient. Equation (3), which combines (1) and (2), is the size calculation formula of the feature map after the convolution operation of dilated convolution. Figure 4²⁵ takes a 3×3 convolution kernel as an example to show the difference between standard convolution and dilated convolution in the reception field and output. The dilation rate of dilated convolution is 2; the red part of the two figures on the left is the eigenvalue of the convolution operation with the convolution kernel, and the red part on the right is the output result after convolution.

$$h_2 = (h_1 - k + 2p)/s + 1 \quad (1)$$

$$ke = r(k - 1) + 1 \quad (2)$$

$$h_2 = (h_1 - ke + 2p)/s + 1 = [h_1 - r(k - 1) - 1 + 2p]/s + 1 \quad (3)$$

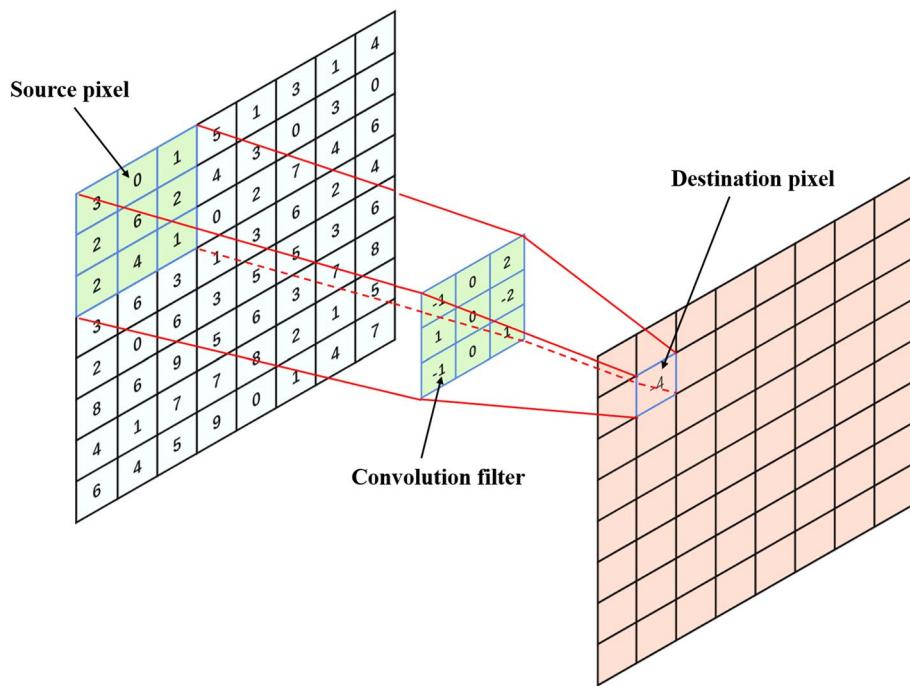


Figure 5. Example of information limitations of convolution operations.

When using a neural network to process images, features are generally extracted through convolution operations in the neural network, and then the convolutional feature matrix is downsampled by a pooling operation to reduce the image scale and increase the reception field. However, the loss of information caused by down-sampling in the pooling operation is irreversible, and this irreversible loss of information will cause the loss of the internal data structure of the feature matrix during the training process of the neural network. Usually only the probability of each class needs to be predicted in the classification model, so there is no need to consider the loss of image details caused by the pooling operation. When doing the pixel-level prediction task of radar echo extrapolation, it is necessary to restore the image to the same size as the input image through the upsampling operation, so the problem of irreversible information loss needs to be reconsidered. Therefore, in order to keep resolution and still expand the receptive field, in this paper, dilated convolution, instead of the pooling operation, is used to downsample the feature matrix obtained after convolution in the model, and the reception field of the convolution kernel is expanded through dilated convolution without losing resolution or coverage. In this way, dilated convolution, together with standard convolution, can not only perform convolution operation on data and extract spatial features from radar echo data from multiple scales, but also retain internal data structure, making data information reconstruction possible, which cannot be realized by pooling operation. This is useful in radar echo extrapolation tasks. On the one hand, a larger reception field makes it possible to better learn the overall change trend of radar echoes, and on the other hand, a higher resolution makes it possible to accurately locate the pixel information of smaller radar echoes.

Attention convolution module. The attention mechanism has gradually become a research hotspot and an important part of the neural network structure in recent years. Initially applied to machine translation, now it has been used in image processing and recommendation systems^{26,27}. The attention mechanism comes from the human visual attention mechanism. Humans usually observe and pay attention to a specific part of a scene according to the needs instead of seeing a scene throughout. Therefore, the essence of the attention mechanism is to filter out important information from a large amount and focus on them. In the neural network, a hidden vector is generated by the input image through convolution operation, activation function and pooling operation. The value of this hidden vector is calculated by referring to itself and several surrounding pixels. As a result, if standard convolution is employed, local information rather than global information will be used to calculate target pixels, which brings limitations to the training of neural networks, as is shown in Fig. 5. In the 3×3 convolution operation (Fig. 5), the convolution filter has 9 pixels, and the value of the target pixel is calculated only with reference to itself and the surrounding 8 pixels, while the global pixel information does not play a role, which causes deviation in prediction results. Consequently, with increasing neural network layers, the high-level feature map has strong predictive ability, but the target is blurred²⁸. In contrast, low-level feature maps have more accurate target location but poor prediction ability and more virtual target information²⁹.

In order to solve the above problems, this paper integrates the advantages of high-level feature maps and low-level feature maps. The attention convolution module is added to the ADC_Net neural network model. Considering each pixel in the feature map as a random variable, calculate the weight values of all pixels in the current prediction task, the pixels with high weights are used and the pixels with low weights are ignored in

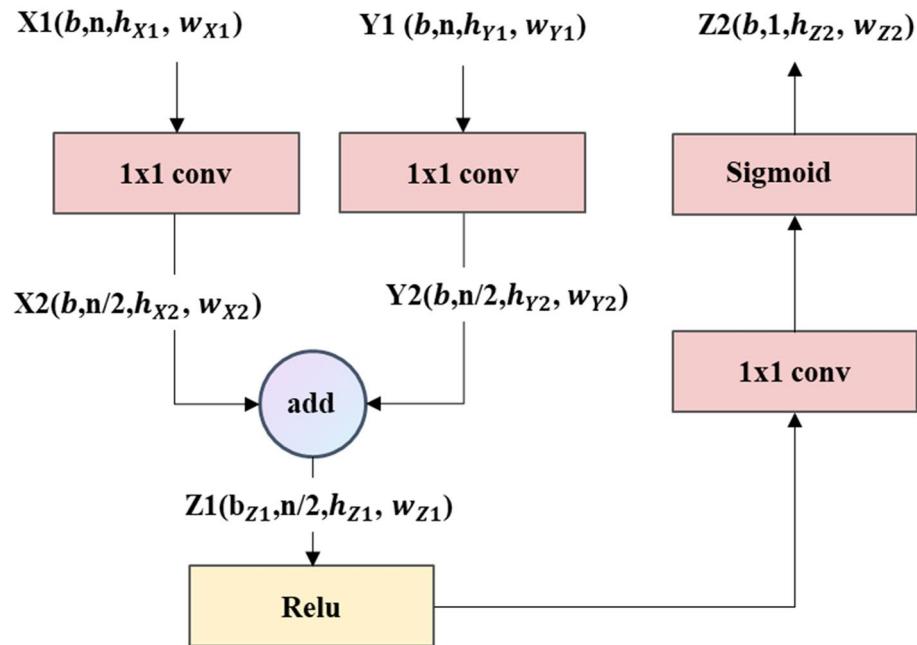


Figure 6. Attention Convolution module structure.

training and prediction. The use of the attention convolution module can effectively suppress the pixel information irrelevant to the current prediction task and enhance the related pixel information. Figure 6 shows the structure of the attention module used in this paper, where b is the batch size of the model, i.e., the number of samples selected for one training, h and w the height and width of the feature matrix, and cat means matrix addition. First, two feature matrices with n channels, high-level feature matrix $X1(b,n,h_{X1},w_{X1})$ and low-level feature matrix $Y1(b,n,h_{Y1},w_{Y1})$ are input into the attention convolution module. In the attention convolution module, the feature information of different channels is integrated through 1×1 convolution, and the number of channels of the feature matrix is halved to reduce the amount of model computation. At this point, feature matrices $X2(b,n/2,h_{X2},w_{X2})$ and $Y2(b,n/2,h_{Y2},w_{Y2})$ are obtained. Then calculate the sum of the two characteristic matrices of $X2(b,n/2,h_{X2},w_{X2})$ and $Y2(b,n/2,h_{Y2},w_{Y2})$ to get the characteristic matrix $Z1(b,n/2,h_{Z1},w_{Z1})$. The nonlinearity of the neural network is then increased through the Relu activation function. Finally, $Z1$ undergoes a 1×1 convolution and a sigmoid activation function to obtain a weight matrix $Z2(b,1,h_{Z2},w_{Z2})$ with channel number of 1 and value range of (0, 1) and output. The specific calculation process of the weight of each pixel is shown in Eqs. (4) and (5). W_{X1}, W_{Y2}, W_{Z1} are the convolution kernels of the feature matrix $X1, X2, Z1$. $x \in X2(b,n/2,h_{X2},w_{X2}), y \in Y2(b,n/2,h_{Y2},w_{Y2}), Z1 \in Z1(b,n/2,h_{Z1},w_{Z1}), Z2 \in Z1(b,n/2,h_{Z2},w_{Z2})$. u, v, i, j are the indices of pixel x and pixel y in $X2$ and $Y2$, respectively, where $u \in (0, b)$, $v \in (0, n)$, $i \in (0, h)$, $j \in (0, w)$. The final weight of each pixel, $Z2_{i,j}^{u,v}$ is calculated by Eqs. (4) and (5). $Z2_{i,j}^{u,v}$ is used for each pixel to determine focus regions. The smaller the value of $Z2_{i,j}^{u,v}$ is, the smaller the pixel weight value of the position is. When the weight value of a certain pixel is close to 0, it is considered that the pixel has nothing to do with the current prediction task and is interference information. On the contrary, the larger the obtained value of $Z2_{i,j}^{u,v}$, the larger the weight value of the pixel at the position, and the position of the pixel is considered to be the focus region.

$$Z1_{i,j}^{u,v} = \max(0, (W_{X1}x_{i,j}^{u,v} + W_{Y1}Y_{i,j}^{u,v})) \quad (4)$$

$$Z2_{i,j}^{u,v} = \frac{1}{1 + \exp(-Z1_{i,j}^{u,v})} \quad (5)$$

To sum up, the application of the attention convolution module can simultaneously learn the advantages of high-level feature maps and low-level feature maps in the process of model training and prediction to achieve the global reference of feature values. According to the weight matrix obtained by referring to the global feature, each pixel of the feature matrix is given different weights. This allows the neural network to focus on the pixel information most relevant to the current task each time, making the neural network use pixel information more efficiently in the process of training and prediction²¹, so the prediction results are more accurate.

Experiments

Data. The data in the experiment are New generation of Doppler weather radar detection data, which are stored in polar coordinates. For the convenience of research, the radar echo data in polar coordinates are first stored in the three-dimensional Cartesian coordinate system by interpolation method. Considering the height of

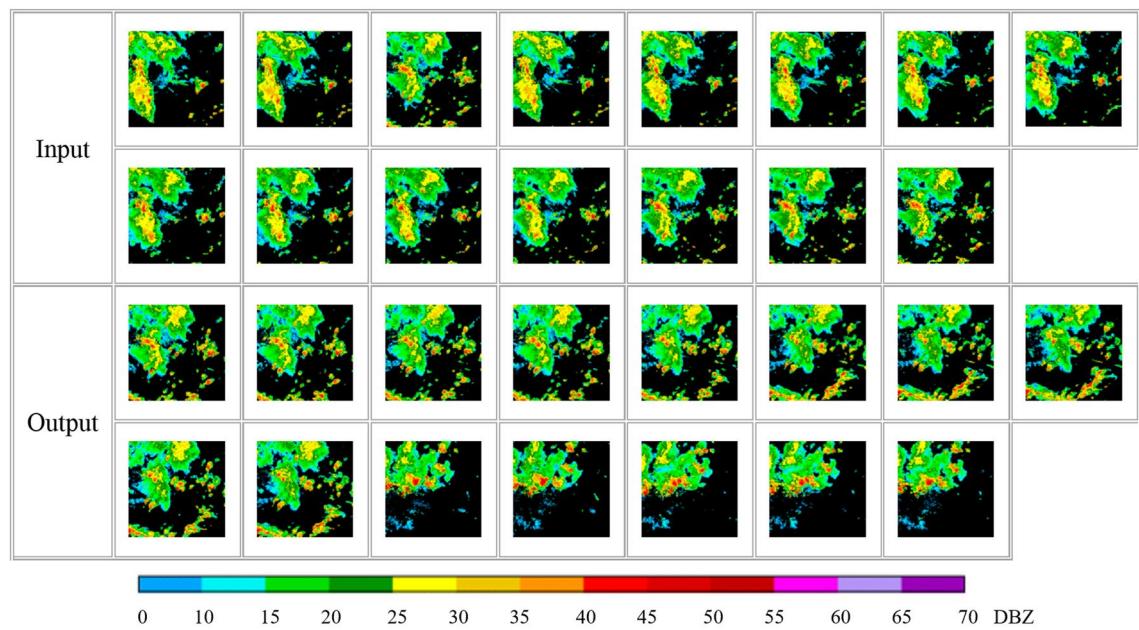


Figure 7. Image after visualization of a set of samples in the dataset.

the representative average guided airflow in the troposphere, we select the CAPPI radar echo image at a vertical height of 3KM as the experimental data. In order to reduce the parameters in the training process and improve the training efficiency, the CAPPI images are grayscaled before the experiment. The processed grayscale image is then cropped, keeping a 256×256 size area in the center of the image. Because the echoes with areas too small are generally weak and evolve rapidly, they have little guiding significance for forecasting work. In order to avoid the influence of the sample quality on the training results, each radar echo image is screened during the construction of the sample dataset. The cases with the echo area coverage greater than 1/10 in the effective detection range of the radar are selected into the sample data set, and the radar echo range is controlled within the range of 0–70 DBZ. Figure 7 shows the visualized images of a group of samples in the data set. The first two rows in Fig. 7 are fifteen input images, and the last two rows are fifteen label images. The time interval between each two adjacent radar echo images in the data used in this paper is 6 min. In this experiment, the first fifteen radar echo images are used to extrapolate the last fifteen radar echo images. That is, the radar echo image of the first 90 min is used to predict the radar echo image of the next 90 min.

Model evaluation indicators. In the model training process, the prediction of strong convective weather is used as an example to evaluate the experimental results. The evaluation indicators are probability of detection (POD), false alarm rate (FAR), critical success index (CSI) and Heidke skill score (HSS). POD represents the probability that convective weather is predicted to occur and the prediction is successful. FAR represents the probability that convective weather is predicted to occur but the prediction fails. CSI is the prediction accuracy after weighing successful predictions and wrong predictions. HSS is a quantitative evaluation of the model after comprehensively considering various situations. Therefore, the higher the value of POD, CSI and HSS is, and the lower the value of FAR is, the more accurate the predicted result is, and the better the model effect is. Since strong convective weather usually occurs when the radar echo intensity is greater than 35DBZ, the threshold of the radar echo is set to 35DBZ in this experiment. The real label radar echo image and the predicted radar echo image are binarized respectively, with the pixels larger than 35DBZ in the radar echo image binarized to 1, and the pixels smaller than 35DBZ binarized to 0. When evaluating the experimental results, the binarized pixel value of the real image is named label, and the binarized pixel value of the predicted image is named fake, *hit* shows the times that both the predicted value fake and the real value label are 1, *far* is the times that the predicted value fake is 1 while the true value label is 0, *miss* the times the predicted value fake is 0 but the true value label is 1, and *cn* is the times that both the predicted value fake and the true value label are 0. The calculation formulas of the experimental evaluation indicators POD, FAR, CSI and HSS are shown in Eqs. (6), (7), (8), and (9), respectively.

$$POD = \frac{hit}{hit + miss} \quad (6)$$

$$FAR = \frac{far}{hit + far} \quad (7)$$

$$CSI = \frac{hit}{hit + miss + far} \quad (8)$$

$$HSS = \frac{2 * (hit * cn - miss * far)}{(hit + miss) * (miss + cn) + (hit + far) * (far + cn)} \quad (9)$$

Loss function. The loss function is the most basic and critical structure in deep learning, and is used to evaluate the degree of inconsistency between the predicted value of the model and the true value. It is a non-negative real-valued function, usually represented by $L(s, y)$, where s and y represent the predicted value and the true value, respectively. In the process of neural network training, the smaller the loss function value is, the more robust the model is. The prediction effect of the model can be measured through the definition and optimization of the loss function. In model training, considering that the use of the MSE loss function alone may cause a small loss value and result in a poor model training effect, the use of MAE alone may lead to the loss of data feature details. Therefore, in this paper, MSE + MAE is used as the loss function, which is conducive to the convergence of the function. Even if the learning factor is fixed, the function can quickly obtain the minimum value. The loss function of the proposed model is shown in Eq. (10).

$$L(S, Y) = MSELoss(s, y) + MAELoss(s, y) = \frac{1}{n} \sum_{t=1}^n (s_t - y_t)^2 + \frac{1}{n} \sum_{t=1}^n |s_t - y_t| \quad (10)$$

Optimizer. During the training and testing of the neural network, the choice of different optimizers will impact the prediction results differently. In this paper, Adam is used as the optimizer. Integrating the advantages of the two optimization algorithms AdaGrad and RMSProp, Adam calculates the update step size by considering the First Moment Estimation of the gradient (the mean of the gradient) and the Second Moment Estimation (the uncentered variance of the gradient). Its advantages include easy operation, high computational efficiency, and parameter updates immune to the scaling transformation of gradients. Also, because of the good interpretability of Adam hyperparameters and little or no need for parameter adjustment, it is applicable to large-scale data and parameter scenarios. In this experiment, the initial learning rate of Adam is set to 0.0001.

Analysis of experimental results. The language used in this experiment is Python, and the Pytorch framework is used for training on NVIDIA 2488 GPU. To evaluate the performance of the ADC_Net model based on dilated convolution and attentional convolution (ADC_Net) proposed in the paper, it is compared with the ADC_Net model based on dilated convolution (ADC_Net2), the ADC_Net model based on attention convolution module (ADC_Net3), the ADC_Net model based neither on dilated convolution nor on attentional convolution module (ADC_Net4), U_Net, EDP model, and ConvLSTM.

Evaluation of model prediction results. Figure 8 shows the comparison between the real radar echo image and the radar echo image predicted by each model. It can be seen that the predicted image of the ADC_Net model based on attention convolution and dilated convolution proposed in this paper is most similar to the real image. In comparison, although U_Net, EDP model, and ConvLSTM can also predict the shape of the radar echo with accuracy, they could only show a general outline with blurred details in terms of radar echo intensity. Compared with ADC_Net4, the prediction results of ADC_Net3 are more accurate in details, such as small gaps between radar echoes. This shows that adding an attention convolution module allows the model to learn more detailed features of the data. Compared with ADC_Net2, the prediction results of ADC_Net3 are not so impressive for the high radar echo intensity areas (shown in red in Fig. 8). It is because dilated convolution instead of pooling is used in ADC_Net2 to downsample and retain the features' internal data structure, which allows information reconstruction to keep detailed information of the image, while this cannot be realized by pooling.

Tables 1, 2, and 3 show the POD evaluation scores, CSI evaluation scores, FAR evaluation scores, and HSS evaluation scores of radar echo extrapolation at 30 min, 60 min, and 90 min for all models. With higher value of POD, CSI and HSS, and lower value of FAR, the predicted result is more accurate and the model is more effective. It can be clearly seen that the scores of each evaluation index of the ADC_Net model based neither on dilated convolution nor on attentional convolution module (ADC_Net4) are similar to U_Net, EDP model, and ConvLSTM. However, when dilated convolution (ADC_Net2) or attention convolution module (ADC_Net3) is added to the model, the evaluation scores of each indicator are significantly better than that of U_Net, EDP model and ConvLSTM. It can be seen from the experimental evaluation scores of ADC_Net2 and ADC_Net3 that both the dilated convolution and the attention convolution module can improve the accuracy of radar echo extrapolation in the radar echo extrapolation task. When the dilated convolution and attention convolution modules are added to the model (ADC_Net) at the same time, the POD, CSI, HSS evaluation scores of the model in the three time periods and the FAR evaluation score in 30 min are the best. Although the FAR values at 60 min and 90 min are not as good as ADC_Net2 and ADC_Net3, they are also much better than ADC_Net4 and U_Net, EDP, and ConvLSTM. This shows that for the radar echo extrapolation task, the dilated convolution and attention convolution modules can be put in the same model, and the integration of the two can achieve better results than using dilated convolution or attention convolution alone.

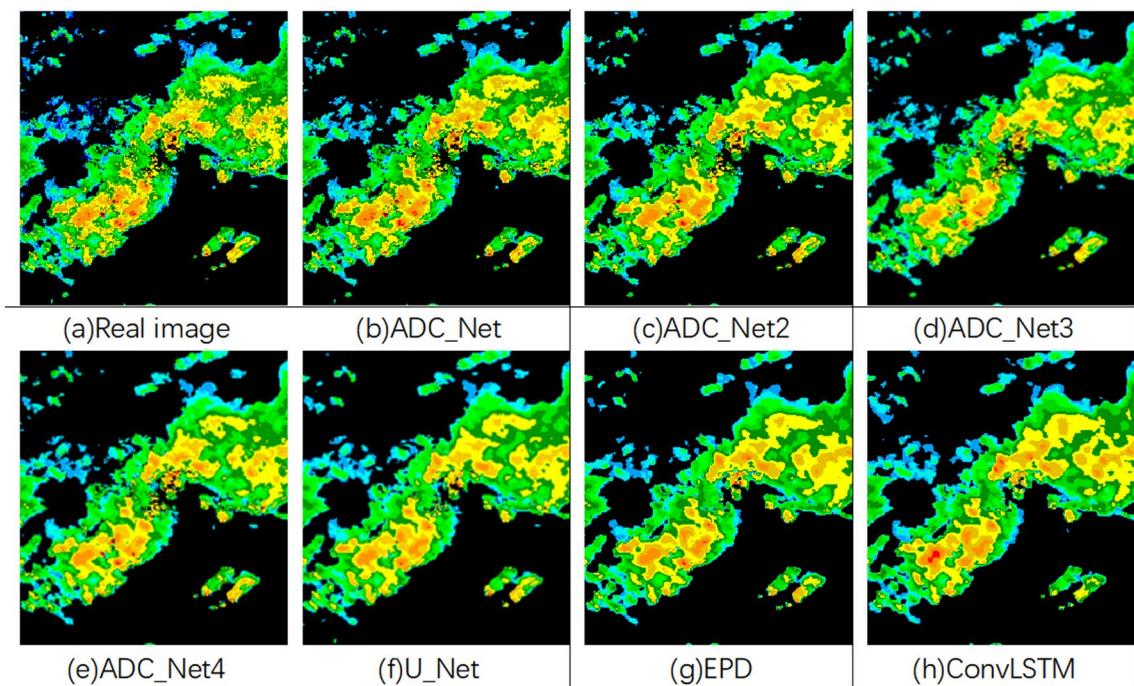


Figure 8. Comparison of model radar echo extrapolation results.

	30 min			
	POD	CSI	FAR	HSS
ADC_Net	0.76087586	0.58172320	0.22049363	0.60302877
ADC_Net2	0.72091937	0.52216042	0.24340544	0.59199967
ADC_Net3	0.70376852	0.53149514	0.23615935	0.55343548
ADC_Net4	0.62247112	0.47155796	0.27031753	0.53146832
U_Net	0.59286521	0.44365503	0.30021739	0.50110344
EDP	0.61163434	0.46923196	0.27989904	0.45632585
ConvLSTM	0.56044080	0.42366371	0.29077875	0.45303934

Table 1. 30-min radar echo extrapolation evaluation scores.

	60 min			
	POD	CSI	FAR	HSS
ADC_Net	0.73668836	0.55113560	0.25978267	0.54246716
ADC_Net2	0.68287509	0.49765180	0.25450861	0.54545296
ADC_Net3	0.65736475	0.48938651	0.25422885	0.53725577
ADC_Net4	0.59543897	0.42196212	0.32134102	0.50585725
U_Net	0.52786379	0.40746430	0.31627907	0.45462281
EDP	0.55468666	0.42866114	0.31237780	0.44512617
ConvLSTM	0.52867036	0.41107007	0.3188	0.41275907

Table 2. 60-min radar echo extrapolation evaluation scores.

Model parameter performance evaluation. This part compares the single training time and model parameters of ADC_Net model based on the ADC_Net model based on dilated convolution and attention convolution (ADC_Net), the ADC_Net model based on dilated convolution (ADC_Net2), the ADC_Net model based on attention convolution module (ADC_Net3), the ADC_Net model based neither on dilated convolution nor on attentional convolution module (ADC_Net4) and U_Net, EDP model, ConvLSTM. The details are shown in Table 4. As can be seen from Table 4, although the EDP model has the most parameters, the running time is the fastest and the

	90 min			
	POD	CSI	FAR	HSS
ADC_Net	0.69663783	0.50318873	0.297786948	0.51354401
ADC_Net2	0.60519550	0.45486333	0.276377305	0.50323715
ADC_Net3	0.61885704	0.42953144	0.307918722	0.50190045
ADC_Net4	0.54688346	0.40419994	0.349850822	0.47244314
U_Net	0.41271537	0.40040904	0.39193696	0.40240053
EDP	0.49411441	0.40056530	0.365196424	0.40518047
ConvLSTM	0.41460854	0.40004941	0.392364732	0.40312401

Table 3. 90-min radar echo extrapolation evaluation scores.

	ADC_Net	ADC_Net2	ADC_Net3	ADC_Net4	U_Net	EDP	ConvLSTM
Model parameters (unit: pcs)	39,444,603	40,485,711	38,483,036	39,093,071	34,534,863	79,713,680	29,396,039
Single training time (unit: s)	12.46	13.38	12.25	12.31	8.94	7.34	7.81

Table 4. Performance comparison.

training efficiency is the highest. Compared with U_Net and ConvLSTM, the model proposed in this paper can improve the prediction accuracy of radar echo extrapolation, but it has more model parameters and a longer single prediction time. The one with the most parameters and the slowest running is ADC_Net2, which may be a result of the preservation of the internal data structure when the data is downsampled with dilated convolution. However, by comparing the model parameters and running time of ADC_Net2 and ADC_Net, it can be found that when both dilated convolution and attention convolution are added to the model, the parameters in the model training process are reduced, and the training time is also significantly shortened. This shows that combining attention convolution with dilated convolution can not only improve the prediction performance of the model, but also reduce the parameters and running time added by dilated convolution to the model to a certain extent through the attention convolution module.

Conclusion

This paper proposes a radar echo extrapolation model (ADC_Net) based on dilated convolution and attention mechanism. Using dilated convolution instead of pooling operation can not only perform spatial feature extraction on radar echo data from multiple scales, but also preserve the internal data structure. It avoids the loss of the internal data structure in pooling and makes reconstruction of small object information possible. Also, by integrating an attention convolution module with attention mechanism into the model, the sensitivity to the target features is improved, and the interference information is suppressed, thereby improving the extrapolation accuracy of radar echoes. The radar echo extrapolation results of the ADC_Net model based on dilated convolution and attentional convolution (ADC_Net) proposed in the paper is compared with that of the ADC_Net model based on dilated convolution (ADC_Net2), the ADC_Net model based on attention convolution module (ADC_Net3), the ADC_Net model based neither on dilated convolution nor on attentional convolution module (ADC_Net4) and U_Net, EDP model, ConvLSTM. The results show that the proposed model has better prediction performance. Through comparative model experiments, it is found that combining dilated convolution with the attention convolution module can not only improve the prediction performance of the model, but also achieve better results than using dilated convolution or the attention convolution module alone. And through the use of the attention convolution module, the parameters and running time added by dilated convolution to the ADC_Net model can be reduced to a certain extent, and the running efficiency can be improved.

Data availability

All data generated or analysed during this study are included in this published article [and its [supplementary information](#) files].

Received: 12 March 2022; Accepted: 31 May 2022

Published online: 22 June 2022

References

1. Browning, K. A. *Nowcasting* 20–23 (Academic Press, 1982).
2. Chen, M., Wang, Y. & Yu, X. Improvement of cross-correlation extrapolation algorithm and its application in convective nowcasting. *Chin. J. Appl. Meteorol.* **18**(5), 12 (2007).
3. Johnson, J. T. *et al.* The storm cell identification and tracking algorithm: An enhanced WSR-88D algorithm. *Weather Forecast.* **13**(2), 263–276 (1998).
4. Ayzel, G., Heistermann, M. & Winterrath, T. Optical flow models as an open benchmark for radar-based precipitation nowcasting (rainymotion v0.1). *Geoscientific Model Dev. Discuss.* **12**(4), 1387–1402 (2019).

5. Wu, J. *et al.* Variational echo tracking algorithm and its application in convective nowcasting. *Chin. J. Meteorol.* **77**(6), 16 (2019).
6. Zhang, Y. *et al.* Estimation of weather radar echo motion field and its application in precipitation nowcasting. *Acta Meteorol.* **64**(5), 631–646 (2006).
7. Hou, J. & Situ, G. Image encryption using spatial nonlinear optics. *eLight* **2**, 3. <https://doi.org/10.1186/s43593-021-00010-y> (2022).
8. Yuan, Y. *et al.* A fully phase-modulated metasurface as an energy-controllable circular polarization router. *Adv. Sci.* **7**(18), 2001437. <https://doi.org/10.1002/advs.202001437> (2020).
9. Zhang, K., Wang, Y., Burokur, S. N. & Wu, Q. Generating dual-polarized vortex beam by detour phase: From phase gradient metasurfaces to metagratings. *IEEE Trans. Microwave Theory Tech.* **70**(1), 200–209. <https://doi.org/10.1109/TMTT.2021.3075251> (2022).
10. Zhang, K. *et al.* Polarization-engineered noninterleaved metasurface for integer and fractional orbital angular momentum multiplexing. *Laser Photonics Rev.* **15**, 2000351 (2021).
11. Lin, J. Technical analysis of thunderstorms and severe convection near-term weather forecasting. *Sci. Public Sci. Technol. Innov.* **2**, 1 (2020).
12. Chen, X. *et al.* Research on nowcasting method based on convolutional gated recurrent unit neural network. *Plateau Meteorol.* **40**(02), 411–423 (2021).
13. Dong, S., Wang, P. & Abbas, K. A survey on deep learning and its applications. *Comput. Sci. Rev.* **40**(1), 100379 (2021).
14. Chen, J. Artificial neural network method for nowcasting using Doppler weather radar data. *Nanjing Meteorol. Inst.* **23**(2), 283–287 (2000).
15. Guo, H. *et al.* High-resolution nowcasting experiment of strong convection based on deep learning. *Chin. J. Meteorol.* **77**(4), 13 (2019).
16. Klein, B., Wolf, L. & Afek, Y. A dynamic convolutional layer for short rangeweather prediction. In *Computer Vision & Pattern Recognition* 4840–4848 (IEEE, 2015).
17. Shi, E. *et al.* Weather radar echo extrapolation method based on convolutional neural networks. *J. Comput. Appl.* **38**(3), 661–665 (2018).
18. Shi, X. *et al.* *Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting* 802–810 (MIT Press, 2015).
19. Shi, X. *et al.* Deep learning for precipitation nowcasting: a benchmark and a new model. In *Proceedings of the 31st International Conference on Neural Information Processing Systems* 5617–5627 (Curran Associates Inc., 2017).
20. Singh, S., Sarkar, S. & Mitra, P. Leveraging convolutions in recurrent neural networks for doppler weather radar echo prediction. In *Proceedings of the 14th International Symposium on Neural Networks* 310–317, (2017).
21. Zhang, L. *et al.* Weather radar echo prediction method based on convolution neural network and long short-term memory networks for sustainable e-agriculture. *J. Clean. Prod.* **298**, 126776 (2021).
22. Liang, H., Chen, H., Zhang, W., Ge, Y. & Han, L. Convective precipitation nowcasting using U-Net Model. In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 7134–7137, <https://doi.org/10.1109/IGARSS47720.2021.9553617> (2021).
23. Zhang, D. *et al.* Video frame prediction based on deep convolutional long short-term memory neural network. *J. Comput. Appl.* **39**(6), 1657–1662 (2019).
24. Zhang, S., Shao, M., Qi, G. & Xu, X. Crop pest detection based on multi-scale attention convolution network. *Jiangsu J. Agric. Sci.* **37**(03), 579–588 (2021).
25. hiudawn, Pytorch ordinary convolution and atrous convolution. <https://blog.csdn.net/hiudawn/article/details/84500648> (2018).
26. Lv, L. *et al.* An Edge-AI Based Forecasting Approach for Improving Smart Microgrid Efficiency. *IEEE Transactions on Industrial Informatics.* **1**, 1 <https://doi.org/10.1109/TII.2022.3163137> (2022).
27. Lv, L., Wu, Z., Zhang, J., Tan, Z., Zhang, L., & Tian, Z. A VMD and LSTM Based Hybrid Model of Load Forecasting for Power Grid Security. *IEEE Transactions on Industrial Informatics.* **1**, 1 <https://doi.org/10.1109/TII.2021.3130237> (2021).
28. Lv, L., Chen, J., Zhang, L., Zhang, F. Gradient-based neural networks for solving periodic Sylvester matrix equations. *J. the Franklin Institute* **1**, 1 <https://doi.org/10.1016/j.jfranklin.2022.05.023> (2022).
29. Chen, Y. & Tan, X. Application of convolutional neural network based on attention mechanism in image classification. *Sci. Technol. Innov.* **34**, 166–169 (2021).

Author contributions

Conceptualization, X.S. and K.M.; methodology, K.M.; software, K.M.; validation, X.S., K.M. and L.Z.; formal analysis, X.Z.; investigation, X.S.; resources, K.M.; data curation, L.Z.; writing—original draft preparation, X.S.; writing—review and editing, K.M.; visualization, L.Z.; supervision, X.Z.; project administration, K.M. All authors have read and agreed to the published version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-13969-6>.

Correspondence and requests for materials should be addressed to L.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022