Application of a Radar Echo Extrapolation-Based Deep Learning Method in Strong Convection Nowcasting

Jian Yin1, Zhiqiu Gao1,2, and Wei Han1

1Climate and Weather Disasters Collaborative Innovation Center, Key Laboratory for Aerosol-Cloud-Precipitation of China Meteorological Administration, School of Atmospheric Physics, Nanjing University of Information Science and Technology, Nanjing, China, 2State Key Laboratory of Atmospheric Boundary Layer Physics and Atmospheric Chemistry, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China, National Meteorological Center of Chinese Meteorological Administration, Beijing, China

Abstract  Strong convection nowcasting has been gaining importance in operational weather forecasting. Recently, deep learning methods have been used to meet the increasing requirement for precise and timely nowcasting. One of the promising deep learning models is the convolutional gated recurrent unit (ConvGRU), which has been proven to perform better than traditional methods in strong convection nowcasting. Despite its encouraging performance, ConvGRU tends to produce blurry radar echo images and fails to model radar echo intensities that have multi-modal and skewed distributions. To overcome these disadvantages, we tested the structural similarity (SSIM) and multiscale structural similarity (MS-SSIM) loss functions. The SSIM and MS-SSIM loss functions are composed of luminance, contrast, and structure, and provide more information about the intensity, grade, and shape of the radar echo, which can reduce blurring. Due to multi-layer downscaling, MS-SSIM extracted more radar echo characteristics, and its extrapolation was the most realistic and accurate among all of the loss function schemes. Only the MS-SSIM scheme successfully predicted strong radar echoes after 2 h, especially those at the rainstorm level.

Plain Language Summary  The Convolutional Recurrent Neural Network model has been considered a pioneering work in the field of nowcasting. However, this model tends to predict blurred radar echo images and fails to model the radar echo intensity that has a multi-modal and skewed distribution. In this paper, we adopt two measures of structural similarity (SSIM and MS-SSIM) as loss functions to reduce the blurring problem caused by the mean square error (MSE) loss function for precipitation nowcasting. We have shown that the MS-SSIM scheme is more efficient in capturing the spatiotemporal correlations, and has the best prediction performance, especially for heavier precipitation. Even for rotating moving radar echoes, the MS-SSIM scheme has the best forecast performance

1. Introduction

Precise and timely nowcasting is an increasingly important area of operational weather forecasting (Mecikalski et al., 2015), especially for convective weather, which has become more frequent in the context of global warming and causes considerable adverse socioeconomic impacts (Bouwer, 2019). Generally, the nowcasting technique refers to the provision of high spatial and temporal resolution predictions of local weather that substantially changes over a short-term period (0–2 h), such as thunderstorms, strong convection, and precipitation events (Browning & Collier, 1989). This technique has long played an important role in the generation of weather predictions that inform agricultural production, airport safety, and citywide rainfall alerts (Dancer & Tibblitts, 1973; Li, 2003; Wilson et al., 2010). Increasing the refinement and accuracy of convective weather predictions has proven quite challenging for nowcasting, and further study in this field is necessary (Sun et al., 2014).

Existing methods for strong convection nowcasting mainly include numerical weather prediction (NWP)-based methods (Wilson et al., 1998) and radar echo extrapolation-based methods (Sun et al., 2014). For the NWP method, predictions at the nowcasting timescale are difficult due to the complexity of simulating physical equations. Furthermore, the “spin up” problem of the NWP model often results in a large deviation in the early stage of the precipitation forecast, which makes it difficult to ensure the availability of...
nowcasting results (Li et al., 2005). Radar observations provide significant information for the prediction of convective-scale weather (Das et al., 2006; Sun, 2005) and capture a range of scales, including small-scale features, which helps to meet the ever-growing need for precise and timely precipitation nowcasting. Thus, the better-performing radar extrapolation-based method has been adopted as an advanced approach in operational nowcasting systems (Yang et al., 2018).

The Storm Cell Identification and Tracking (SCIT) algorithm (Johnson et al., 1998) and Tracking of Radar Echo with Correlations (TREC) algorithm (Rinehart & Garvey, 1978) are two representative traditional extrapolation methods that are used for the extrapolation of thunderstorms. The SCIT algorithm yields the extrapolation of thunderstorms by tracking the centroid position of the three-dimensional storm cell. The principle of the TREC algorithm is to track radar echoes based on the cross-correlation method. However, the SCIT algorithm only works for the extrapolation of convective weather systems (Han et al., 2008), and the TREC algorithm fails to complete the extrapolation when the precipitation cloud clusters change rapidly (Cao et al., 2015).

To improve the forecasting performance of radar echo extrapolation, the optical flow method from the computer vision field was introduced to radar echo tracking. Optical flow is defined as the apparent motion of individual pixels on the image plane (Whitehead & Gibson, 1981). Case analysis and batch experiments carried out by the Hong Kong Observatory (HKO) with 2 yr of data demonstrated the advantages of the optical flow method over the TREC method (Woo & Chun, 2014). Han et al. (2008) estimated the movement of radar echoes by using the optical flow method, and the results showed that it was effective for the extrapolation of strong convective precipitation. However, questions have been raised about the limitations of optical flow-based methods. Radar echo extrapolation may violate the assumption of constant intensity and space consistency in the optical flow method (Lucas & Kanade, 1997), as the intensity and shape of the radar echo often vary rapidly in strong convective weather systems. Other problems with this approach are that only two frames of radar echo images are taken into account and that a large number of historical radar echo map records in the database are not fully utilized (Tian et al., 2020).

In modern nowcasting systems, supervised deep learning approaches have also been developed to extrapolate the movements of radar echoes. Recent advances in recurrent neural network (RNN) models (Giles et al., 1994) and the long short-term memory (LSTM) framework (Hochreiter & Schmidhuber, 1997) have contributed to a significant breakthrough in spatiotemporal sequence prediction. Shi et al. (2015) creatively exploited convolutions to extend the LSTM framework to perform precipitation nowcasting. This convolutional LSTM (ConvLSTM) model has been considered a pioneering work in this field. Shi et al. (2017) further developed the method by replacing LSTM with a gated recurrent unit (GRU), yielding the convolutional GRU (ConvGRU) model, which has fewer parameters than ConvLSTM. Guo et al. (2019) applied the ConvGRU model in extrapolation experiments, and the results showed that the ConvGRU model greatly improved the prediction accuracy. Wang et al. (2017) proposed a new predictive RNN (PredRNN) with a new core named the Spatiotemporal LSTM (ST-LSTM) unit. Subsequently, a Memory in Memory (MIM) network for precipitation nowcasting was proposed (Wang et al., 2018). Both ST-LSTM and MIM were more effective than ConvLSTM in the extrapolation of radar echoes. Despite their encouraging performance, these models have disadvantages, as the mean square error (MSE) loss function employed by these models has two issues (Tian et al., 2020). First, the extrapolated images tend to become blurry as the forecast time increases. Second, the MSE loss function often fails to model radar echo intensities that have multi-modal and skewed distributions. Several attempts have been made to mitigate the blurry image issue. Previous research conducted by Tran and Song (2019) showed that the blurriness of images could be significantly reduced by employing a visual image quality assessment technique as the loss function. A generative adversarial ConvGRU (GA-ConvGRU) model proposed by Tian et al. (2020) has proven effective in overcoming the limitations of blurry extrapolated images.

To overcome these two drawbacks, inspired by the previous research of Tran and Song (2019), the measure of structural similarity (SSIM) index and the multiscale structural similarity (MS-SSIM) index were tested as the loss function of the ConvGRU model in this study (Wang et al., 2003). SSIM and MS-SSIM are indexes that measure the similarity between two images by comparing local patterns of pixel intensities (Zhou et al., 2004). The SSIM and MS-SSIM indexes are composed of luminance, contrast, and structure and can express more details about the intensity, grade, and shape of the radar echo, which can result in more
realistic and accurate extrapolation results. Furthermore, we evaluated the extrapolation results of different loss functions with a prediction lead time of 2 h. Two representative nowcasting case studies of strong convective precipitation were performed using the ConvGRU model with different loss functions.

The remaining of this study is organized as follows: In Section 2, a specific explanation of the ConvGRU model, three types of loss functions, and the radar data set are presented. The extrapolation results using different loss functions are compared in Section 3. Finally, the discussion and conclusions are provided in Sections 4 and 5, respectively.

2. Methods and Data

Radar echo extrapolation-based methods use historical radar echo map sequences to forecast future fixed-length radar echo maps, which fall under the spatiotemporal sequence forecasting problem in the field of deep learning. From the spatial viewpoint, radar observations at a given time can be treated as a tensor $\chi$ of $M \times N$ grids consisting of $M$ rows and $N$ columns over a spatial region. From the temporal viewpoint, a sequence of tensors $X_1, X_2, \ldots, X_T$ can be obtained by collecting radar observations at fixed time intervals over time $T$. Thus, this spatiotemporal predictive learning problem can be illustrated as:

$$X_{t+1}, \ldots, X_{t+K} = \arg\max_{x_{t+1}, \ldots, x_{t+K}} p\left(X_{t+1}, \ldots, X_{t+K} \mid \hat{X}_{t-J+1}, \ldots, \hat{X}_t\right),$$

(1)

in which, $\{X_{t+1}, \ldots, X_{t+K}\}$ is the predicted sequence of length $K$, and $\{\hat{X}_{t-J+1}, \ldots, \hat{X}_t\}$ is the historical observation sequence data of length $J$.

According to the above modeling method of the spatiotemporal prediction, historical 0.5 h radar echo maps (one frame per 6 min and five consecutive frames) were used to predict radar echo maps for the next 2 h (20 consecutive frames). In the training set, each sample contained 25 consecutive frames of radar echo maps, and the first 5 frames were the inputs of the model. The task of this model is to output 20 radar echo maps that are as consistent as possible with the last 20 frames of the sample sequence.

2.1. Related Work

As one of the most common RNN models for general-purpose sequence modeling, LSTM can effectively overcome the problem of long-term dependence between sequences (Bengio et al., 1994). LSTM consists of a memory cell $c_t$, an input gate $i_t$, an output gate $o_t$, and a forget gate $f_t$. The memory cell $c_t$ is the core of LSTM and controls the information flow using the input gate $i_t$, the output gate $o_t$, and the forget gate $f_t$. Using memory units and gates to control information flow can effectively avoid the RNN problem in which the gradients vanish too quickly.

By replacing matrix multiplication in LSTM with the convolution operation, ConvLSTM can simultaneously extract the spatiotemporal structural attributes of sequence data. The ConvLSTM block is shown in Figure 1a, including the input $X_t$, the cell output $c_t$, the hidden state $H_t$, and the gates $(i_t, f_t, o_t)$. When new input $X_t$ arrives, the forget gate $f_t$ is activated to selectively discard some information of the last cell state $c_{t-1}$. Then, the input gate $i_t$ selectively retains the input information in the cell state at the current moment. Finally, the output gate $o_t$ controls how much information from the last cell state is output. The main equations of ConvLSTM are as follows:

$$i_t = \sigma\left(W_{ii}^{*} X_t + W_{io}^{*} H_{t-1} + W_o b_i\right),$$

$$f_t = \sigma\left(W_{fi}^{*} X_t + W_{fo}^{*} H_{t-1} + W_o b_f\right),$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tanh\left(W_{ic}^{*} X_t + W_{io}^{*} H_{t-1} + b_c\right),$$

$$o_t = \sigma\left(W_{oi}^{*} X_t + W_{io}^{*} H_{t-1} + W_o b_o\right),$$

$$H_t = o_t \odot \tanh(C_t),$$

(2)

where $*$ is the convolution operator, $\odot$ is the Hadamard product, and $\sigma$ is the sigmoid activation function.
2.2. ConvGRU

GRU is a variation of LSTM; it uses only the update and reset gates to control the information flow and removes the memory unit (Cho et al., 2014). Compared with LSTM, GRU is more concise and efficient, and its performance is similar or better (Chung et al., 2014). The training of GRU can achieve faster convergence. Hence, the convolution-based GRU (ConvGRU) is used to model the spatiotemporal sequence forecasting in this study (Figure 1b).

The key formulas of the ConvGRU are as follows:

\[
\begin{align*}
Z_t &= \sigma \left( W_{xz} * x_t + W_{hh} * h_{t-1} \right), \\
R_t &= \sigma \left( W_{xr} * x_t + W_{hr} * h_{t-1} \right), \\
H'_t &= f \left( W_{zh} * x_t + R_t \circ (W_{hh} * h_{t-1}) \right), \\
H_t &= (1 - Z_t) \circ H'_t + Z_t \circ h_{t-1},
\end{align*}
\]

(3)

where \(*\) and \(\circ\) are the convolution operation and Hadamard product, respectively. \(h_t, R_t, Z_t, \text{ and } H'_t \in R^{c \times H \times W}\) are the memory state, reset gate, update gate, and new information, respectively. \(C, H, \text{ and } W\) are the channel size, height, and width of the state tensor, respectively. \(x_t \in R^{c \times H \times W}\) is the input tensor. Similarly, \(C, H, \text{ and } W\) are the channel size, height, and width of the input tensors, respectively. \(f\) is the activation function of leaky ReLU (Maas et al., 2013) with a negative slope equal to 0.2. When the reset gate is activated, the previous state information is selectively forgotten, and new information is formed with the current input. Then, the update gate is activated, and the new information is updated to the current state.

2.3. Encoder-Decoder Structure

A complex RNN structure can be obtained by stacking and temporally concatenating multiple ConvGRU units (Figure 2a). However, this connection mode (all the input is connected to each output) may cause the problem of excessive training parameters. Another modeling problem that makes this network structure challenging is that the length of the input and output sequences may be variable. In the proposed model, an encoder-decoder structure is used to solve the multi-step prediction problem of nowcasting with equal intervals (Sutskever et al., 2014).

The encoder-decoder structure is shown in Figure 2b. A special design characteristic is that the order of the decoder structures is reversed. In the encoder module, the input image sequence \(\{X_{t-1}, \ldots, X_t\}\) is first convolved and then fed into a ConvGRU unit to produce the hidden feature map \(H'_t\) and output \(h'_t\); subse-
quently, $\mathcal{H}_t^1$ is further downsampled and fed into a convGRU unit to produce the next level hidden feature map $\mathcal{H}_t^2$; By a similar procedure, $\mathcal{H}_t^2$ and output $\mathcal{H}_t^3$ can be obtained. The decoder module first generates a prediction image sequence $\mathcal{H}_t^{K+1}$ by feeding $\mathcal{H}_t^K$ into the corresponding ConvGRU unit and performing an upsampling operator; then $\mathcal{H}_t^K$ and $\mathcal{H}_t^{K+1}$ are fed into the next level convGRU unit to produce $\mathcal{H}_t^{K+2}$; a similar procedure is carried out to produce the final output $\{X_{t+1}, \ldots, X_{t+K}\}$. The benefits of downsampling operator are reducing the redundancy information and forming a compact representation. Another advantage of this reverse structure is that, when additional RNN layers are added at the top, it is not necessary to aggregate low-level information through skip connections.

2.4. Loss Functions

Most RNN models employ the MSE as the loss function. The MSE loss function used in the ConvGRU model for radar echo extrapolation is defined as follows:

$$\text{MSE}(x, y) = \frac{1}{n} \sum_{i=1}^{n} (x_i - y_i)^2.$$  \hspace{1cm} (4)

where $y$ is the predicted radar echo image, $x$ is the true radar echo image, and $n$ is the number of samples. However, the MSE loss does not sufficiently reflect the properties of the human visual system in distinguishing between different pictures. This loss function has two disadvantages: (a) the extrapolated images tend to become blurry as forecast time increases and (b) the MSE loss function often fails to model radar echo intensities that have multi-modal and skewed distributions.

SSIM is a measure of the similarity between two pictures, and it is simple and correlates well with subjective evaluation. Thus, SSIM was used as a loss function to evaluate the similarity between real and predicted radar echo images in this study. The SSIM measurement consists of three contrast modules: brightness, contrast, and structure. The overall similarity measure $S(x, y)$ is as follows:

$$S(x, y) = f\left[l(x, y), c(x, y), s(x, y)\right].$$  \hspace{1cm} (5)

where $x$ and $y$ are two nonnegative radar reflectivity image signals. To thoroughly understand the similarity measure, it is necessary to first introduce the luminance function $l(x, y)$, contrast function $c(x, y)$, and structure function $s(x, y)$ as well as the combined function $f(\cdot)$.

The luminance comparison function is defined as follows:

![Diagram](https://example.com/diagram.png)

**Figure 2.** (a) The stacking convolution-based gated recurrent unit structure. (b) The encoder-decoder structure.
where \( \mu_x \) and \( \mu_y \) are the mean values of \( x \) and \( y \), respectively, and the constant \( C_1 \) is included to avoid instability when \( \mu_x^2 + \mu_y^2 \) is very close to zero. Specifically, \( C_1 \) is calculated as follows:

\[
C_1 = (K_1 L)^2, \tag{7}
\]

where \( L = 2^B - 1 \) is the dynamic range of pixel values (\( B \) is the bits per pixel. For example, \( L \) is \( 2^8 - 1 = 255 \) for normalized 8-bit images) and \( K_1 = 0.01 \) is a small default constant.

The contrast comparison function is defined as follows:

\[
c(x, y) = \frac{2\sigma_x \sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \tag{8}
\]

where \( \sigma_x \) and \( \sigma_y \) are the variances of \( x \) and \( y \), respectively, and \( C_2 = (K_2 L)^2, K_2 \ll 1 \).

Similarly, the structure comparison function is defined as follows:

\[
s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x \sigma_y + C_3}, \tag{9}
\]

where \( \sigma_{xy} \) is the covariance of \( x \) and \( y \), and \( C_3 = (K_3 L)^2, K_3 = 0.03 \).

Combining the three functions in (6), (8), and (9) and setting \( C_3 = C_2/2 \), the simplified SSIM index is as follows:

\[
SSIM(x, y) = \left[ l(x, y) \right]^{(\alpha \beta \gamma)} \cdot \left[ c(x, y) \right]^{\beta} \cdot \left[ s(x, y) \right]^{\gamma} \quad (\alpha = \beta = \gamma = 1)
= \frac{(2\mu_x \mu_y + C_1)(2\sigma_x \sigma_y + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{10}
\]

A larger SSIM value means that the images are more similar. \( SSIM = 1 \) when the two images are the same. Therefore, we set \( 1 - SSIM \) as the loss function.

MS-SSIM was developed to make the similarity index more consistent with the human visual system, and it supplies more flexibility than SSIM by incorporating variations in viewing conditions. MS-SSIM is a comprehensive evaluation metric that is, obtained by iteratively calculating the SSIM of different resolutions (scales; Wang et al., 2003):

\[
MS\cdot SSIM(x, y) = \left[ l(x, y) \right]^{(\alpha \beta \gamma)} \cdot \prod_{j=1}^{M} \left[ c_j(x, y) \right]^{\beta_j} \cdot \left[ s_j(x, y) \right]^{\gamma_j}, \tag{11}
\]

where \( M \) is the highest scale number selected for the reference image. After \( M - 1 \) iterations of low-pass filtering and downsampling, the final MS-SSIM value is obtained. Similarly, we set \( 1 - MS\cdot SSIM \) as the loss function.

### 2.5. Data Set

The HKO-7 radar echo data set used in this study was downloaded from the HKO. This data set contains observed radar echo maps from 2009 to 2015. The radar echo maps are the radar reflectivity images (480 × 480 pixels) acquired in the Constant Altitude Plan Position Indicator (CAPPI) scanning mode, covering a 512 × 512 km\(^2\) area. The interval of radar reflectivity image data is 6 min, and there are 240 frames of reflectivity image data per day. First, the radar reflectivity data were preprocessed to convert radar reflectivity values to pixel values (0–255), which were then normalized to (0–1). Only radar data with precipitation events were required for training, so we traversed the entire data set and selected 46,814,640 echo maps containing precipitation events to form the final data set. Thus, we had 194,880 echo maps for training, 12,000 echo maps for validation, and 31,400 echo maps for testing.
2.6. Evaluation Methods

To quantitatively assess the forecast performance of all schemes, the skill scores at different rain rate thresholds \( r \) were calculated. The rain rate \( r \) is calculated by reflectivity factor \( Z \) by the formula
\[
10 \lg (10 \lg a \times 10 \lg b + 10 \lg c) = \times \times \times \times \times \times \times \times
\]
(where \( a = 58.53 \) and \( b = 1.56 \) are statistical constants). The selected rain rate thresholds are 2, 5, 10, and 30, which correspond to different rainfall levels, from light rain to rainstorms (Table 1). Four evaluation indicators, namely, the probability of detection (POD), Critical Success Index (CSI), false alarm rate (FAR), and Heidke Skill Score (HSS), were calculated at a specific rain rate threshold \( r \). Assuming that the rain rate threshold is \( K \), TP is the number that both the observation and prediction values are not less than \( K \), FN is the number that the observation values are not less than \( K \) and the prediction values are less than \( K \), FP is the number that the observation values are less than \( K \) and the prediction values are not less than \( K \), and TN is the number that both the observation and prediction values are less than \( K \). POD is a metric that measures the ability to identify “hits,” and it is calculated as follows:
\[
POD = \frac{TP}{TP + FN}
\]
The FAR measures the fraction of alarms that are triggered when the event does not occur, and it is calculated as follows:
\[
FAR = \frac{FP}{TP + FP}
\]
The CSI score is a metric that expresses the FAR based on POD. The CSI score is calculated as follows:
\[
CSI = \frac{TP}{TP + FP + FN}
\]
HSS is another compound metric that is used to quantitatively evaluate simulations with different schemes. The HSS score is calculated as follows:
\[
HSS = \frac{TP \times TN - FP \times FN}{(TP + FN)(FN + TN) + (TP + FP)(FP + TN)}
\]

3. Experiment

In this section, experiments of strong convective weather nowcasting using the ConvGRU model with three different loss functions are described.

3.1. Experimental Setting

The parameter values of the training process are as follows: the initial learning rate is \( 10^{-4} \), and the learning rate penalty factor is 0.5. That is, the learning rate is halved when the loss function value of the verification set stops decreasing. The batch size is 8 with 100,000 training iterations. The ConvGRU model is composed of a three-layer encoder-decoder structure, which is a self-coding network. The filter numbers of the three layers in the ConvGRU model are 64, 192, and 192, respectively. Accordingly, the kernel sizes are 5 \( \times \) 5, 5 \( \times \) 5, and 3 \( \times \) 3, respectively, for the ConvGRU models. As the model with the MSE loss function tends to produce blurred prediction results, SSIM and MS-SSIM were tested as the loss function of the model to mitigate this problem as much as possible. Two representative case studies of strong convective precipitation were carried out to verify the performance of the SSIM and MS-SSIM schemes.
3.2. Results

A squall line was selected as Case 1 to evaluate the prediction results of the ConvGRU model with three different loss functions for the extrapolation of linearly moving radar echoes. Figure 3 illustrates the 1 h nowcasts of three different loss function schemes and the ground truth in Case 1, in which Figure 3a is the ground truth of radar observations and Figures 3b–3d are the predicted results using the MSE, SSIM, and MS-SSIM loss functions, respectively. In the radar echo maps of Figure 3, different color ranges correspond to different detection targets. The blue colors (−5 to 15 dBZ) represent the echoes of water vapor and thin cloud, the green colors (15–30 dBZ) represent the echoes of laminar cloud, the yellow colors (30–45 dBZ) represent the echoes of convective cloud, and the red colors (45–70 dBZ) represent the echoes of the thunderstorm.

Figure 3. Comparison between radar observations (a) and 1 h nowcasting results by the convolution-based gated recurrent unit model with the mean square error loss function (b), structural similarity loss function (c), and multiscale-structural similarity loss function (d) in Case 1.
Compared with the ground truth, all three loss function schemes correctly predict the shape and orientation of the radar echoes after the 1 h extrapolation. However, the model with the MSE scheme fails to predict the intensity of the strong echo band with values greater than 45 dBZ. On the contrary, the echo fields predicted by the SSIM and MS-SSIM schemes contain more detail, and both schemes successfully predict the intensity and structure of the strong echo band. In the region outside the strong echo band, the SSIM scheme has the same problem: the predicted echo intensity is weaker than the ground truth. Although the radar echo intensity outside the strong echo band is over-predicted, the whole radar echo field predicted by the MS-SSIM scheme is closest to the ground truth. By comparing the performance of the 1 h forecast of the three schemes in Figure 3, it is concluded that for the extrapolation of linearly moving radar echoes, the schemes using the SSIM and MS-SSIM loss functions can effectively mitigate the blurry image problem that leads to the under-forecast of echo intensity. Compared with the SSIM loss function scheme, extrapolation by the MS-SSIM scheme is more consistent with the ground truth.

To further evaluate the nowcasting performance of the three different loss function schemes with longer prediction time steps, 2 h prediction results were compared between the three different loss function schemes. Figure 4 illustrates the 2 h nowcasts of the three different loss function schemes and the ground truth in Case 1, in which Figure 4a is the ground truth of radar observations and Figures 4b–4d are the predicted results using the MSE, SSIM, and MS-SSIM loss functions, respectively. As the forecast time increases, the performance of the model with the MSE scheme worsens. The 2 h prediction results of the MSE scheme in Figure 4b shows that echo intensity gradients in the whole predicted field are weaker, and the echo intensity is mostly concentrated in the range of 25–35 dBZ. Compared with the performance of the 1 h forecast, the accuracy of the 2 h forecast results using the SSIM scheme also decreases. Figure 4c shows that the forecast field outside the strong echo band generated by the SSIM scheme tends to be blurry, and there are some missed echo predictions in the area of the strong echo band. In terms of the strong echo band, the 2 h forecast performance of the MS-SSIM scheme is closest to the ground truth. Although there are false alarms outside the strong echo band in Figure 4d, the distribution of the echo intensity gradients predicted by the MS-SSIM scheme is the most accurate among the three schemes. For the extrapolation of the squall line, the SSIM and MS-SSIM schemes produce more realistic and more accurate extrapolation compared with the MSE scheme, and the position and intensity of the strong echoes predicted by the MS-SSIM scheme are closest to the ground truth.

All of the POD scores for Case 1 are shown in Figure 5. In terms of the POD metric, the performance of the MSE scheme is not very good for this task. When the threshold is >10, the POD scores of the model using the MSE scheme are less than 0.6, and when the threshold is 30, they decrease rapidly to 0 after 1 h of forecasting. Such poor POD scores indicate that the MSE scheme fails to predict precipitation in heavier rainfall after 1 h of forecasting. The schemes using the SSIM and MS-SSIM loss functions perform similarly at a precipitation threshold of 2. The POD scores of the SSIM scheme decrease rapidly when the precipitation threshold is >5. However, in contrast to the failed extrapolation of the MSE scheme for rainstorms, the SSIM scheme is able to partially simulate the rainstorm. The POD scores of the MS-SSIM scheme are the highest at all precipitation thresholds, and the forecast scores for thresholds under 10 are mostly above 0.8. When the threshold is 30, the MS-SSIM scheme still produces satisfactory forecasts. In terms of the POD scores, the MS-SSIM scheme has the best nowcast performance at different precipitation levels, especially at rainstorm levels.

Figure 6 presents all CSI scores of the different schemes in Case 1. Similar to the POD scores of the MSE scheme, when the threshold is >30, the CSI scores of the model with the MSE scheme decrease rapidly to 0 after 1 h of forecasting. This performance indicates that the MSE scheme fails to predict precipitation in heavy rainfall after 1 h of forecasting. The schemes using the SSIM and MS-SSIM loss functions perform similarly in the first half an hour of the forecast at all precipitation thresholds. The CSI scores of the SSIM scheme markedly decrease after half an hour of forecasting when the precipitation threshold is >5. The CSI scores of the MS-SSIM scheme are the highest at all precipitation thresholds, and when the threshold is 30, the MS-SSIM scheme still produces adequate extrapolation after 2 h of forecasting. In terms of the CSI scores, the MS-SSIM scheme has the best nowcast performance at different precipitation levels, especially rainstorms.
Figure 7 shows all of the HSS scores of the different schemes in Case 1. The schemes using the SSIM and MS-SSIM loss functions perform similarly in the first half an hour of the forecast at all precipitation thresholds. When the threshold is greater than 30, the HSS scores of the model using the MSE scheme also decrease rapidly to 0 after 1 h of forecasting. The HSS scores of the SSIM scheme markedly decrease after half an hour of forecasting when the precipitation threshold is >5. The HSS scores of the MS-SSIM scheme are the highest when the precipitation threshold is >5, and when the threshold is 30, the MS-SSIM scheme still produces adequate extrapolation after 2 h of forecasting. In terms of the HSS scores, the MS-SSIM scheme has the best nowcast performance at different precipitation levels, especially at rainstorm levels.

Figure 8 shows all of the FAR scores of the different schemes in Case 1 (the missing data in the FAR curve of the MSE scheme at a threshold of 30 is because the TP is 0 after step 17 of extrapolation). The FAR results of all loss function schemes significantly differ from the outcomes for POD, CSI, and HSS. The FAR scores of the MSE scheme are the lowest at all thresholds, indicating that the MSE scheme outperforms the other
Figure 5. Comparison of probability of detection results between different schemes over lead time in Case 1.

Figure 6. Comparison of Critical Success Index results of different schemes over lead time in Case 1.
Figure 7. Comparison of Heidke Skill Score results of different schemes over lead time in Case 1.

Figure 8. Comparison of FAR results of different schemes over lead time in Case 1.
The FAR scores of the MS-SSIM scheme are the highest at almost all thresholds, which suggests that, among all loss function schemes, the MS-SSIM scheme results in the greatest over-forecast for the squall line extrapolation, especially at the heavier precipitation threshold of \( r > 10 \).

The performance of the model using the three different loss function schemes is discussed above for Case 1, involving linearly moving radar echo extrapolation. Case 2 was chosen to test the extrapolation of more complex rotating echo movement. Figure 9 illustrates the 1 h nowcasts of the three different loss function schemes and the ground truth in Case 2, in which Figure 9a is the ground truth of radar observations and Figures 9b–9d are the predicted results using the MSE, SSIM, and MS-SSIM loss functions, respectively. Compared with the ground truth, all three model schemes correctly predict the shape and orientation of the whole radar echo field after the 1 h extrapolation. However, the model with the MSE scheme fails to predict the intensity of the strong echo area in which values are greater than 40 dBZ and the whole predicted echo field tends to be blurry compared with the ground truth. In the strong echo area, the 1 h nowcast of

Figure 9. Comparison between radar observations (a) and 1 h nowcasting results by the convolution-based gated recurrent unit model with the mean square error loss function (b), structural similarity loss function (c), and multiscale-structural similarity loss function (d) in Case 2.
the SSIM is worse than its performance for linearly moving radar echoes. Moreover, the SSIM scheme does not effectively overcome the blurry image problem for the nowcasting of rotating echoes and the predicted intensity gradients of radar echoes are sparse. The MS-SSIM scheme outperforms the SSIM scheme. It successfully predicts the intensity gradients and structure of the strong radar echoes, and the whole radar echo field predicted by the MS-SSIM scheme is closest to the ground truth. By comparing the performance of the 1 h forecast of the three schemes in Figure 9, it is concluded that only the scheme with the MS-SSIM loss function can effectively mitigate the blurry image problem for the extrapolation of radar echoes with rotating movement.

To further evaluate the predictive performance of the three different loss function schemes in the extrapolation of rotating moving radar echoes with longer prediction time steps, their 2 h nowcasting results were compared. Figure 10 illustrates the 2 h nowcasts of the three different loss function schemes and the ground truth in Case 2, in which Figure 10a is the ground truth of radar observations and Figures 10b–10d are the
predicted results using the MSE, SSIM, and MS-SSIM loss functions, respectively. As the forecast time increases, the forecast errors of the MSE and SSIM schemes become larger. Compared with the ground truth, and Figures 9c and 10b show that the whole forecast field provided by the MSE scheme and SSIM scheme tends to blur, which causes weaker radar echo intensity gradients. In the nowcasting results of the MSE and SSIM schemes, there is a missed prediction in the area of strong echoes of more than 35 dBZ, and the echo intensity is mostly concentrated in the range of 25–35 dBZ. The 2 h forecast performance of the MS-SSIM scheme is closest to the ground truth. Notably, there are false alarms outside the strong echo circle (red box) in Figures 10b–10d, which indicates that all three loss functions poorly predict the orientation of small individual radar echoes. By comparing the performance of the 2 h forecasts of the three schemes in Figure 10, we can conclude that only the MS-SSIM performs well in the extrapolation of rotating motion radar echoes, especially strong radar echoes, and that the MSE and SSIM schemes fail in the extrapolation of strong radar echoes for longer forecast times (1–2 h forecast lead time).

All POD scores for the extrapolation of rotating moving radar echoes in Case 2 are shown in Figure 11. The performance of the MSE scheme is the worst for this task in terms of the POD metric. After 1 h of forecasting, the POD scores of the MSE scheme decrease rapidly at all thresholds. In particular, when the threshold is 30, the POD scores of the model using the MSE scheme decrease rapidly to 0 after 1 h of forecasting. The POD scores of the SSIM scheme markedly decrease when the precipitation threshold is >5, and when the threshold is 30, they decrease rapidly to 0 after 1 h of forecasting. The poor performance indicates that these two schemes fail to predict precipitation in heavy rainfall through the extrapolation of rotating moving radar echoes. When the threshold is 30, the prediction quality of the MS-SSIM scheme is still high after 2 h of forecasting. The comparison of these different schemes based on the POD metric over the forecast lead time shows that the MS-SSIM scheme has the best forecast performance in the extrapolation of rotating moving radar echoes, especially at the rainstorm level after a 1 h forecast.

Figure 12 presents the CSI scores of the different schemes in Case 2. In terms of the CSI metric, the MSE scheme performs the worst at all rainfall levels in the extrapolation of rotating moving radar echoes. When
The threshold is 30, the CSI scores of the MSE and SSIM schemes decrease rapidly to 0 after 1 h of forecasting. The poor performance indicates that these two schemes fail to predict precipitation in heavy rainfall through the extrapolation of rotating moving radar echoes. When the threshold is 30, the prediction quality of the MS-SSIM scheme is still high after 2 h of forecasting. A comparison of the different schemes based on the CSI metric over the forecast lead time shows that the MS-SSIM scheme has the best forecast performance in the extrapolation of rotating moving radar echoes, especially at the rainstorm level after a 1 h forecast.

Figure 12 presents the HSS scores of the different schemes in Case 2. The HSS scores show that the MSE scheme performs the worst at all rainfall levels in the extrapolation of rotating moving radar echoes. The schemes using the SSIM and MS-SSIM loss functions perform similarly in the first hour of the forecast. When the threshold is 30, the HSS scores of the MSE and SSIM schemes decrease rapidly to 0 after a 1 h forecast. When the threshold is 30, the MS-SSIM scheme still yields better prediction quality after 2 h of forecasting. The comparison of the different schemes based on the HSS metric over the forecast lead time also shows that the MS-SSIM scheme has the best forecast performance in the extrapolation of rotating moving radar echoes, especially at the rainstorm level after a 1 h forecast.

Figure 13 shows the FAR scores of the different schemes in Case 2 (the missing data in the FAR curve of the MSE scheme at a threshold of 30 is because the TP is 0 after step 9 of extrapolation). The FAR results of all loss function schemes significantly differ from the outcomes for POD, CSI, and HSS. The FAR scores of the MSE scheme are the lowest at all thresholds, indicating that the MSE scheme outperforms the other two schemes in terms of false alarms. The FAR scores of the MS-SSIM scheme are the highest at almost all thresholds, which suggests that, among all of the loss function schemes, the MS-SSIM scheme results in the greatest over-forecast for Case 2, especially at the heavier precipitation threshold of \( r > 10 \). Most FAR values of the MS-SSIM scheme in Case 2 are higher than those in Case 1, which suggests that the MS-SSIM scheme results in greater over-forecasting in the extrapolation of rotating moving echoes compared with Case 1.
Figure 13. Comparison of Heidke Skill Score results of different schemes over lead time in Case 2.

Figure 14. Comparison of false alarm rate results of different schemes over lead time in Case 2.
Figure 15 compares the POD (with 80% confidence interval) results of different schemes over lead time for the entire test set. In terms of the POD metric, we can conclude that both the SSIM and MS-SSIM schemes perform better than the MSE scheme at all precipitation thresholds. For threshold = 2 and 5, the SSIM and MS-SSIM schemes perform similarly, but the MS-SSIM scheme is more accurate than the SSIM scheme. For threshold >10, the MS-SSIM scheme still performs well but the extrapolation quality of the SSIM scheme degrades rapidly. For threshold = 30, the extrapolation quality of the MSE scheme degrades rapidly, with the POD scores decreasing to nearly 0 after 60 min. For the POD results with the threshold of 30, the SSIM scheme performed the best in the first 30 min, and then the extrapolation quality of the SSIM scheme degrades rapidly. In general, the excellent extrapolation performance of the MS-SSIM scheme is reflected in its forecasting results with a longer lead time (after about 30 min) and heavier precipitation.

Figure 16 compares the CSI (with 80% confidence interval) results of different schemes over lead time for the entire test set. In terms of the CSI metric, we can conclude that both the SSIM and MS-SSIM schemes perform better than the MSE scheme at all precipitation thresholds. For the CSI results with threshold = 2, the SSIM scheme performed best, with longer lead times exhibit larger performance gains. For threshold = 5, both the SSIM and MS-SSIM schemes perform well, and the MS-SSIM scheme performed best. For threshold >10, the MS-SSIM scheme still performs well but the extrapolation quality of the SSIM scheme degrades rapidly. For threshold = 30, the extrapolation quality of the MSE scheme degrades rapidly, with CSI scores decreasing to nearly 0 after 60 min. In general, the excellent extrapolation performance of the MS-SSIM scheme is reflected in its forecasting results with a longer lead time (after about 30 min) and heavier precipitation.

Figure 17 compares the HSS (with 80% confidence interval) results of different schemes over lead time for the entire test set. In terms of the HSS metric, both the SSIM and MS-SSIM schemes perform better than the MSE scheme at all precipitation thresholds. For the HSS results with threshold = 2, the SSIM scheme performed best, with longer lead times exhibit larger performance gains. For threshold = 5, both the SSIM and MS-SSIM schemes perform well, and the MS-SSIM scheme performed best. For threshold >10, the
MS-SSIM scheme still performs well but the extrapolation quality of the SSIM scheme degrades rapidly. For threshold = 30, the extrapolation quality of the MSE scheme degrades rapidly, with the HSS scores decreasing to nearly 0 after 60 min. In general, the excellent extrapolation performance of the MS-SSIM scheme is reflected in its forecasting results with a longer lead time (after about 30 min) and heavier precipitation.

Figure 18 compares the FAR (with 80% confidence interval) results of different schemes over the lead time for the entire test set (the missing data in the FAR curve of the MSE scheme at a threshold of 30 is because the TP is 0 after step 17 of extrapolation). The FAR results of all of the loss function schemes significantly differ from the outcomes for POD, CSI, and HSS. The FAR scores of the MSE scheme are the lowest at all thresholds, indicating that the MSE scheme outperforms the other two schemes in terms of false alarms. The FAR scores of the MS-SSIM scheme are the highest at almost all thresholds after 30 min of extrapolation, which suggests that, among all loss function schemes, the MS-SSIM scheme results in greater over-forecasting, especially at the heavier precipitation threshold of $r > 10$.

The overall mean scores of the 20 predicted frames are summarized in Table 2. We also present the 80% confidence intervals of the mean scores with 500 bootstrap replications (Table 3). We find that the SSIM and MS-SSIM loss function schemes proposed in this study are essential for improving nowcasting performance. The ConvGRU model trained with the default MSE loss has a worse nowcasting score than that with the SSIM or MS-SSIM loss functions at the 10 and 30 mm/h thresholds. Among these loss function schemes, MS-SSIM performs the best, which verifies its effectiveness for rainstorm-level precipitation, at least for this task.

4. Discussion
As mentioned in the literature review, using the MSE loss function often yields blurry extrapolated images and causes the rapid degradation of the forecast image quality. In this study, the SSIM and MS-SSIM indexes were employed as loss functions to train the ConvGRU model, which was expected to yield more realistic
and more accurate extrapolation. We evaluated the extrapolation results of three different loss functions, namely, MSE, SSIM, and MS-SSIM.

For a forecast lead time of 0–1 h, in terms of the POD, CSI, and HSS scores of the entire test set, the SSIM and MS-SSIM loss function schemes outperform the MSE scheme at all rainfall thresholds, and as the intensity of rainfall increases, the superiority of the SSIM and MS-SSIM schemes becomes more evident. In summary, the SSIM and MS-SSIM schemes can overcome the drawbacks of the MSE loss function in ConvGRU and thus yield more accurate extrapolation results in a 1 h forecast, which is consistent with previous findings by Tran and Song.

Extrapolation with a 1–2 h lead time was also performed in this study. For a 1–2 h forecast lead time, in terms of POD, CSI, and HSS scores of the entire test set, the MS-SSIM scheme outperforms the MSE and SSIM schemes at all rainfall thresholds, and as the intensity of rainfall increases, the superiority of the MS-SSIM scheme becomes more evident. For threshold >10, the POD, CSI, and HSS scores of the MSE and SSIM scheme degrade rapidly, especially at a threshold of 30, for which the MSE scheme scores are almost 0. In summary, the MS-SSIM scheme is best at overcoming the drawbacks of the MSE loss function in ConvGRU, thus yielding a more accurate prediction through extrapolation with a 1–2 h lead time.

However, the FAR scores of the MS-SSIM scheme indicate that it does not perform well with regard to false alarms. The FAR of the MS-SSIM scheme is the highest at almost all thresholds after 30 min of extrapolation, which suggests that it results in greater over-forecasting. The superiority of the POD, CSI, and HSS scores in the MS-SSIM scheme is partly a result of over-forecasting due to false alarms. Despite its poor false alarm performance, we conclude that the MS-SSIM scheme is a better option for extrapolation because of the importance of forecasting with longer lead times and heavy precipitation.

Furthermore, two typical case studies of strong convection nowcasting were carried out. For the extrapolation of a squall line, the SSIM and MS-SSIM schemes produce more realistic and more accurate extrapolation, and the position and intensity of the strong echoes predicted by the MS-SSIM scheme are closest to the...
ground truth. For the extrapolation of radar echoes with rotating motion, only the MS-SSIM performs well, especially for strong radar echoes, and the MSE and SSIM schemes fail in the extrapolation of strong radar echoes for a longer forecast time (1–2 h forecast lead time).

The reason behind the improved extrapolation may be that the MS-SSIM and SSIM schemes use three components (brightness, contrast, and structure, corresponding to the radar echo values, intensity gradients, and echo shapes) to train the model instead of just considering the echo values, as in the MSE scheme. In addition, multiscale downsampling may explain why the MS-SSIM scheme shows the best extrapolation performance when forecasting with longer lead times and heavier precipitation.

However, the MS-SSIM scheme also has some shortcomings. On the echo map of radar observations, there are several different thunderstorm cells in adjacent locations, which is difficult for the ConvGRU model to distinguish. In the predicted radar echo map, several different thunderstorm cells are merged into a larger echo, causing false alarms. The likely reason for this phenomenon is the inherent uncertainty in the weather forecasting problem (nonlinear rapid evolution of strong convective systems). The uncertainty increases

![Figure 18. Comparison of false alarm rate (with 80% confidence interval) results of different schemes over lead time for the entire test set.](image)

| Table 2 | The Overall Evaluation Results |
|---------|-------------------------------|
|         | POD          | CSI          | HSS          |
| Schemes | r = 2 | r = 5 | r = 10 | r = 30 | r = 2 | r = 5 | r = 10 | r = 30 | r = 2 | r = 5 | r = 10 | r = 30 |
| MSE     | 0.78 | 0.68 | 0.40 | 0.09 | 0.71 | 0.61 | 0.35 | 0.06 | 0.39 | 0.35 | 0.23 | 0.05 |
| SSIM    | 0.85 | 0.81 | 0.60 | 0.37 | **0.76** | 0.65 | 0.45 | 0.26 | 0.40 | 0.37 | 0.28 | 0.19 |
| MS-SSIM | **0.86** | **0.83** | **0.78** | **0.58** | 0.74 | **0.66** | **0.50** | **0.31** | 0.40 | **0.37** | **0.30** | **0.23** |

Note. Each cell contains the mean score of the 20 predicted frames, and the best result is marked in boldface. CSI, Critical Success Index; HSS, Heidke Skill Score; MSE, mean square error; MS-SSIM, multiscale-structural similarity; POD, probability of detection; SSIM, structural similarity.
The Confidence Intervals of the Mean Scores With Bootstrap Replications

| Confidence interval | POD       | CSI       | HSS       |
|---------------------|-----------|-----------|-----------|
|                     | r = 2     | r = 5     | r = 10    | r = 30    | r = 2     | r = 5     | r = 10    | r = 30    | r = 2     | r = 5     | r = 10    | r = 30    |
| MSE                 | [0.75, 0.80] | [0.66, 0.68] | [0.38, 0.40] | [0.05, 0.13] | [0.69, 0.75] | [0.58, 0.63] | [0.34, 0.35] | [0.05, 0.11] | [0.38, 0.41] | [0.34, 0.37] | [0.22, 0.23] | [0.04, 0.08] |
| SSIM                | [0.84, 0.87] | [0.77, 0.82] | [0.58, 0.64] | [0.28, 0.47] | [0.72, 0.79] | [0.64, 0.68] | [0.44, 0.50] | [0.19, 0.32] | [0.39, 0.42] | [0.36, 0.39] | [0.28, 0.31] | [0.14, 0.22] |
| MS-SSIM             | [0.84, 0.87] | [0.84, 0.86] | [0.78, 0.79] | [0.41, 0.61] | [0.71, 0.78] | [0.65, 0.71] | [0.49, 0.58] | [0.23, 0.38] | [0.38, 0.42] | [0.36, 0.40] | [0.30, 0.34] | [0.17, 0.26] |

CSI, Critical Success Index; HSS, Heidke Skill Score; MSE, mean square error; MS-SSIM, multiscale-structural similarity; POD, probability of detection; SSIM, structural similarity.

Data Availability Statement

The data set of radar echo images used in this study comes from Hong Kong Observatory and is included in the literature (Shi et al., 2017). This data set can be downloaded after submitting a data application form to the Hong Kong Observatory (swirls@hko.gov.hk). The data application form is uploaded as supporting information, and it also can be downloaded through https://github.com/sxjscience/HKO-7.

5. Summary

To minimize the value of the MSE loss function, the ConvGRU model tends to use the mean value of the sample probability distribution as the prediction result, which leads to smoother local strong echo values. In this study, we adopted two measures of structural similarity (SSIM and MS-SSIM) as loss functions for precipitation nowcasting to mitigate the blurry image issue caused by the MSE loss function. The results show that the MS-SSIM scheme is a more efficient means of spatiotemporal sequence prediction and that it has the best prediction performance among all loss function schemes, especially for heavier precipitation forecasting with longer lead times. The MS-SSIM scheme has the best forecast performance, even for rotating moving radar echoes.

Since the data set used in this study does not contain any information from radar stations or latitude-longitude coordinates, it is difficult to examine the generalization ability of the model using only this data set. In future work, we will attempt to build an operational nowcasting system using the proposed algorithm and will aim to use radar observations covering other surrounding areas to test the operational application capabilities of the model. Although the value of radar echo maps for severe convective systems nowcasting, quantitative estimation of precipitation is one of the most important tasks in nowcasting. How to use the predicted sequence of radar echo maps to accurately estimate the precipitation is also the next target of this study. Updating and setting the weight of the MS-SSIM loss function according to different precipitation intensity thresholds will be the focus of future research to reduce false alarms.

References

Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks, 5*(2), 157–166. https://doi.org/10.1109/72.279181

Bouwer, L. M. (2019). Observed and projected impacts from extreme weather events: Implications for loss and damage. In R. Mechler, L. Bouwer, T. Schinko, S. Surminski, & J. Linnerooth-Bayer (Eds.), *Loss and damage from climate change: Concepts, methods and policy options* (pp. 63–82). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-72026-5_3

Browning, K. A., & Collier, C. G. (1989). Nowcasting of precipitation systems. *Reviews of Geophysics, 27*, 345–370. https://doi.org/10.1029/ rg027i003p0034

Cao, C. Y., Chen, Y. Z., Liu, D. H., Li, C., Li, H., & He, J. (2015). The optical flow method and its application to nowcasting. *Acta Meteorologica Sinica, 73*(3), 471–480.
Cho, K., Merrienboer, B. V., Gulcehre, C., Bougares, F., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. Cornell University.

Chung, J., Gulcehre, C., Cho, K. H., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. Eprint Arxiv.

Dancer, W., & Tibbitts, T. (1973). Impact of nowcasting on the production and processing of agricultural crops (pp. 1–47). University of Wisconsin-Madison, Space Science and Engineering Center.

Das, S., Abhilash, S., & Gupta, M. D. (2006). Simulation of Nor’westers using Doppler weather radar wind observations in a mesoscale model. Proceedings of SPIE 6404, Remote Sensing and Modeling of the Atmosphere, Oceans, and Interactions (p. 64040L). International Society for Optics and Photonics. https://doi.org/10.1117/12.694006

Giles, C. L., Kuhn, G. M., & Williams, R. J. (1994). Dynamic recurrent neural networks: Theory and applications. IEEE Transactions on Neural Networks, 5(2), 153–156. https://doi.org/10.1109/70.353425

Guo, H., Chen, M., Han, L., Zhang, W., Qin, R., & Song, L. Y. (2019). High resolution nowcasting experiment of severe convections based on deep learning. Acta Meteorologica Sinica, 077(004), 715–727.

Han, L., Wang, H., & Lin, Y. (2008). Application of optical flow method to nowcasting convective weather. Beijing Daxue Xuebao (Ziran Ke Xue Ban)/Acta Scientiarum Naturalium Universitatis Pekinesis, 44, 751–755.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neo.1997.9.8.1735

Johnson, J. T., Mackeen, P. L., Witt, A., Mitchell, E. D. W., Stumpf, G. J., Eilts, M. D., & Thomas, K. W. (1998). The storm cell identification and tracking algorithm: An enhanced WSR 88D algorithm. Weather and Forecasting, 13(2), 263–276. https://doi.org/10.1175/1520-0434(1998)013<0263:tsiataa>2.0.co;2

Li, P. W. (2003). Development of a thunderstorm nowcasting system in support of air traffic management. Proceedings of the National Academy of Sciences of the United States of America, 100(22), 13030–13035.

Li, P. W., Wong, W. K., Lai, T. S., Li, P. W., & Lai, E. S. T. (2005). RAPIDS—A new rainstorm nowcasting system in Hong Kong. Paper presented at WMO/WWRP International Symposium on Nowcasting and Very-Short-Range Forecasting.

Lucas, B., & Kanade, T. (1997). An iterative image registration technique with an application to stereo vision. Paper presented at the Proceedings of the 7th International Joint Conference on Artificial Intelligence.

Maas, A. L., Hannun, A. Y, & Ng, A. Y. (2013). Rectifier nonlinearities improve neural network acoustic models. Paper presented at International Conference on Machine Learning.

Mecikaliski, J. R., Williams, J. K., Jewett, C. P., Ahijevych, D., LeRoy, A., & Walker, J. R. (2015). Probabilistic 0–1-h convective initiation nowcasting. Paper presented at the Proceedings of the 31st International Conference on Neural Information Processing Systems.

Meehl, G. A., & Bane, J. M. (1994). An enhanced WSR 88D algorithm. Weather and Forecasting, 9(2), 207–212. https://doi.org/10.1175/1520-0434(1994)009<0207:AEWSRA>2.0.CO;2

Shi, X., Chen, Z., Wang, H., Yeung, D. Y., Wong, W. K., & Woo, W. C. (2015). Rectifier nonlinearities improve neural network acoustic models. Paper presented at International Conference on Machine Learning.

Sun, J., Xue, M., Wilson, J. W., Zawadzki, I., Ballard, S. P., Onvlee-Hooimeyer, J., et al. (2014). Use of NWP for nowcasting convective precipitation: Recent progress and challenges. Bulletin of the American Meteorological Society, 95(3), 409–426. https://doi.org/10.1175/bams-d-11-00263.1

Sutskever, I., Vinyals, O., & Le, Q. (2014). Sequence to sequence learning with neural networks. Eprint Arxiv.

Tian, L., Li, X., Ye, Y., Xie, P., & Li, Y. (2020). A generative adversarial gated recurrent unit model for precipitation nowcasting. IEEE Geoscience and Remote Sensing Letters, 17(4), 601–605. https://doi.org/10.1109/LGRS.2019.2926776

Tran, Q. K., & Song, S. K. (2019). Computer vision in precipitation nowcasting: Applying image quality assessment metrics for training deep neural networks. Atmosphere, 10(5), 244. https://doi.org/10.3390/atmos10050244

Wang, Y., Long, M., Wang, J., Gao, Z., & Yu, P. S. (2017). PredRNN: Recurrent neural networks for predictive learning using spatiotemporal LSTMs. Proceedings of the 31st International Conference on Neural Information Processing Systems.

Yu, J., Chang, N., & Zadeh, H. (2015). Memory in memory: A predictive neural network for learning higher-order non-stationarity from spatiotemporal dynamics. Retrieved from https://ui.adsabs.harvard.edu/abs/2015arXiv15107490W

Zhou, W., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing, 13(4), 600–612. https://doi.org/10.1109/TIP.2003.819861