Prediction of Weather Radar Images via a Deep LSTM for Nowcasting

Guang Yao*†
yaoguang@em-data.com.cn

Zongxuan Liu*†
liuzongxuan@em-data.com.cn

Xufeng Guo†
guoxf3z@gmail.com

Chaoshi Wei‡
wcs81303@hotmail.com

Xinfeng Li‡
lixinfeng_2006@163.com

Zhihao Chen‡
chen_zhihao@126.com

* These two authors contribute equally to the work.
† Em-Data AI Research, Shanghai, China
‡ China Eastern Region Air Traffic Management Bureau(ECATMB), Shanghai, China

Abstract—Weather radar images provide critical information for mesoscale weather nowcasting which plays significant roles in a range of fields including civil aviation and navigation. Differed from traditional radar exploration methods, this paper presents a novel prediction model based on a deep recurrent neural network (DeepRNN). The approach converts the task of nowcasting to a task of image series prediction. We first design a new loss function that pays more attention to the changes of images in the input sequence. In the mean while, an image discriminator is incorporated into the model to improve the visual quality of predicted images. Furthermore, optical flow is explored to preserve the the motion information. The prediction results are evaluated based on widely used statistic scores. The experimental results show that the proposed model leads to significant improvement in tasks of 2 hours forecasting of radar echo.

Index Terms—Recurrent Neural Network (RNN), Nowcasting, Image Prediction

I. Introduction

Weather nowcasting usually refers to the next six hours weather forecast [1] and it has become a useful tool for releasing sever-impact weather warnings ahead. Its' accuracy and lead time are important aspects in practical applications. Flood prevention,airport traffic control, agricultural activities, and personal outdoor activities are closely linked to nowcasts. One hour or less ahead-forecasts of severe local-scale precipitation have great importance and practical interest.

Unlike long-term precipitation forecast which is based on atmospheric numerical weather prediction (NWP) models [2], precipitation nowcasting mainly relies on the extrapolation of radar images or neural network models with NWP models' outputs. Radar image extrapolation for rainfall forecast has been a research hotspot since 1960s [3]. At the early stage, simple linear methods were used to extrapolate radar images and failed to predict the motion and intensity change of rainfall band [4].

Then optical flow algorithms [5]–[9] were proposed after great efforts done in improving the precipitation nowcasting over several decades. The basic idea is estimating the rainfall band evolution according to the wind directions and magnitude calculated from a continuous radar images. These algorithms usually cost high running time and can't capture the rainfall intensity change well. It is hard to have a significant improvement of nowcasting accuracy via optical flow because of the atmospheric chaotic movement. Another technique for radar-based precipitation nowcasting is ensemble nowcasting. This idea was proposed by Andersson and Ivarsson [10] and they tried to provide the probability of precipitation forecasts. Statistical outcome can be obtained from the ensemble simulations by adopting small perturbation in the initial conditions. Many work has been studied [11]–[14]. Although NWP models have developed vigorously, its ability to nowcast strong convective weather events is still very limited [15]. Because for convective weather, its coverage is small and its growth rate is rapid and the current understanding of physics processes is not enough [16]. According to Zhang et al. [17], the accurate rate of short-range precipitation forecasting in Shenzhen (one of city in China) is only around 40% for the first hour and 15% for the second hour.

Besides this traditional meteorological schemes, with the development of computer science, machine learning is beginning to be used in the extrapolation of radars. At the early stage, artificial neural networks were used to do the predictions. French [18] used simulated precipitation field to train a single layer neural network and then predicted the 1-hour-later precipitation field. Random forests and logistic regression were also used to nowcast rainfall [19], [20]. Bhattacharjee et al. [?] overcome multiple close-by objects at different scales move in random directions with variable speeds by modeling the scene as a space-time graph with intermediate features from the pixels (or a local region) as vertices and the relationships among them as edges. Keren et al. [21] proposed a model to enhance the feature extraction process for the case of sequential data by inputting patches of data into a recurrent neural network and using the outputs or hidden states of the recurrent unit to calculate the extracted features. Today deep

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

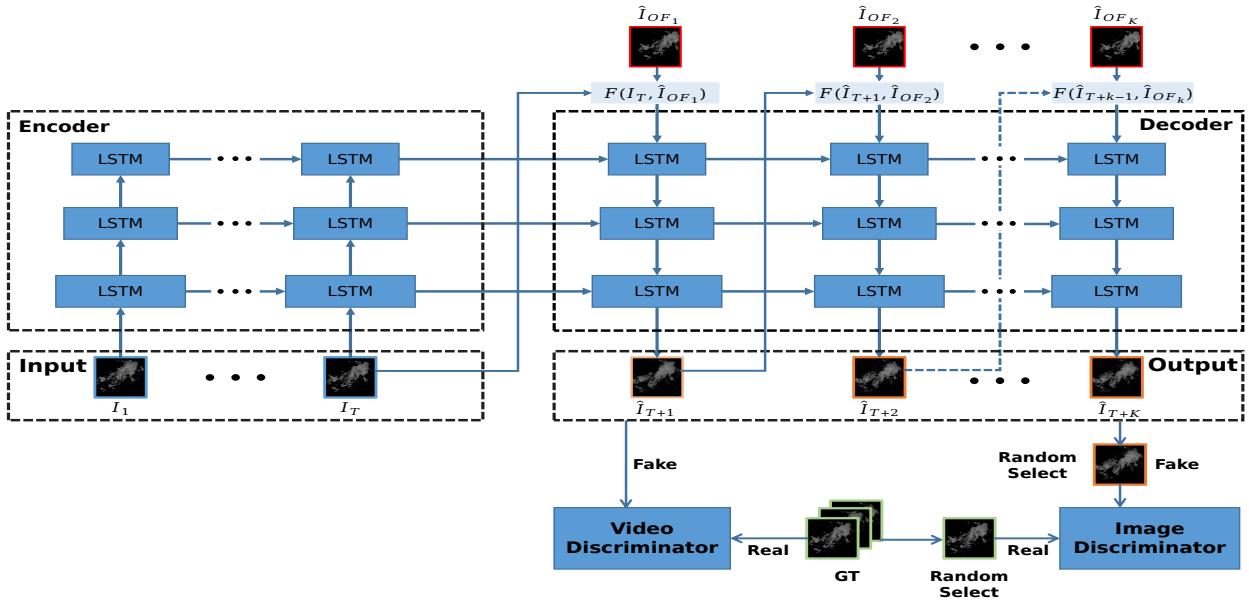


Fig. 1: The architecture of our model for the K frames prediction task with a T frames input sequence.

learning algorithms, the state-of-the-art machine learning, are applied in many areas. Some of the well known deep learning models are convolution neural networks(CNN), recurrent neural networks(RNN) and long short term memory(LSTM). From the view of deep learning point, precipitation prediction is similar to semantic segmentation situation, in which the input is a multichannel image and the output is assigned to every pixel. In recent years, convolutional neural networks have been the state-of the-art solution for semantic segmentation [22], so it is make sense to apply the similar neural networks for precipitation prediction. Zahraei [23] trained a CNN model by radar reflectivity images to predict rainfall filed. The result shows that deep learning algorithms perform better than traditional radar extrapolation methods. However, compared to traditional optical flow prediction methods, due to algorithmic reasons, machine learning methods are not as accurate as traditional optical flow prediction in very close steps. Moreover, when using the above method to predict, the realistic of the predicted images is not enough to produce convincing results.

The objective of this study is to implement a practical and refined neural network model got by deep learning for precipitation nowcasting based on radar imagery. Based on ConvLSTM, multi layer LSTM framework was adopted. we add two discriminator to ensure continuity of video and realistic of photo. In order to achieve better training effect and faster convergence speed, we improved the traditional MSE loss by using adaMSE loss.

This paper is organized as follows: Section 3 describes the tasks of our work. Section 4 give a detailed description the neural network model used in this study. Section 5 describes the training details of our experiment. Section 6 presents the results by using the new method on movingMNIST and nowcast. Conclution and discussion are showed in Section 7.

II. Prior Work

Shi et al. [24] proposed the Convolutional Long Short-Term Memory(ConvLSTM) model, which extends the LSTM by adding convolutional structures in transitions, changing LSTM from one-dimensional to two-dimensional. It makes LSTM more suitable for image sequence. Later, they proposed the trajectory GRU model [25] that can actively learn the location-variant structure for recurrent connections, which improved the results of the ConvLSTM. Wang et al. [26] proposed a new gradient highway structure and a new ST-LSTM, which made great improvement in weather prediction.

Generative adversarial networks is an excellent network for solving the image prediction quality problem. Based on GAN, conditional GAN [27] and others have proposed the classic structure of condition GAN. Nowadays, most varieties GAN are similar to it. CycleGAN [28] uses Two pairs of generators and discriminators to construct a ring-shaped network structure, which implements unsupervised ConditionGAN. MDGAN uses 3D convolution in the generator for video generation, two generators are used to ensure the clarity of the picture and the picture. MocoGAN [29] uses a dual discriminator method to simultaneously identify continuity and authenticity, which improves the quality of video generation. In order to solve the problem of video prediction, Wang added GAN to predRNN to solve the image generation of the LSTM network to a certain extent quality issue.

III. Proposed Method

A. Task Description

From the computer vision point of view, the nowcasting task can be converted to a task of K frames image series prediction given the historical T frames image series as input. We let $X^T = \{I_1, I_2, ..., I_T\}$ be a sequence of input series with T frames, $\hat{X}^K = \{\hat{I}_{T+1}, \hat{I}_{T+2}, ..., \hat{I}_{T+K}\}$ be the following K frames target future series prediction, and let $X^K = \{I_{T+1}, I_{T+2}, ..., I_{T+K}\}$ be the corresponding ground truth of the prediction.

Besides, our method uses the prediction of traditional optical flow algorithm [30] as prior knowledge, we denote $\hat{I}^K_{OF} = \{\hat{I}_{OF_1}, \hat{I}_{OF_2}, ..., \hat{I}_{OF_K}\}$ as the predicted sequence of the optical flow algorithms as part of our system input.

In the following subsections, we will present our new model for the nowcasting task described before. A new Multi-Layer LSTM structure tailored the this task will be introduced followed by some training details.

B. System Overview

Generally the proposed system follows encoder-decoder architecture as demonstrated in Fig. 1. In the encoder part, T frames of historical images X_T is feed into the the model, and the model propagates hidden variable through two direction: vertically and horizontally. Vertically, the pyramid-like hidden variables mainly represent the abstraction of the input images, propagating from the bottom to the top with each channel of 128, 64 and 32 respectively. Horizontally, the model encodes the motion information through time.

The decoder of our model, on contrary with encoder, transmit the information in a top-down manner vertically. At each step, the input of the decoder is the output frame of the optical flow extrapolation method at this step, combined with the output frame of the decoder from the previous step. In the end, the results of the model are monitored by two discriminators to ensure the realistic of the prediction results.

C. Training Detail

The traditional optical stream extrapolation method [30] performs reasonably in the prediction effect of adjacent frames, we add its results as part of the input of decoder. The other part of the decoder input is the prediction of the previous frame. The decode input is the weighted sum of them as shown in Eq. 1.

$$F(\hat{I}_{t-1}, I_{OF_t}) = w_t * \hat{I}_{t-1} + (1 - w_t) * I_{OF_t} \tag{1}$$

$$w_t = min\{1.0, 0.5 + 0.5 * t/K\} \tag{2}$$

$\hat{I}_{t-1}$ is the previous predicted frame. Sequence result $\hat{I}_{OF_t}$ is the result of optical flow extrapolation method in this step. Weight w_t is a ratio that increases linearly with steps ranging within $[0, 1]$.

Standard mean square error (MSE) loss is not suitable for the task of nowcasting since it equally penalizes each pixel regardless its intensity value. For example, zero or low intensity pixels, such as those belongs to the image background of both ground truth and the prediction, will reduce the gradient, which makes training difficult. We propose an adaptive MSE loss (adaMSE) to reduce the influence of zero value pixels which can be formulated as:

$$adaMSE = \frac{\sum_{\forall i,j}(\hat{I}_t(i,j) - I_t(i,j))}{N - Count_{zero} + log(Count_{zero})}, \tag{3}$$

N represents the total number of pixels in a image. $Count_{zero}$ means the number of target image pixel with the values that close to 0 or the position where prediction is very similar to ground truth. When we calculate the adaMSE, We use the log of these counts as denominator. By reducing the background proportion, we try to let the model focus on prediction error region to speed up the model convergence.

To ensure the realistic of the output, the conditional image discriminator(DisI) is added to model. We also add the conditional video discriminator(DisV) to ensure the continuity of model prediction.

IV. Experiment

A. Evaluation Protocol

We chose HSS, average CSI, CORR, POD and FAR [25] as our evaluation indicators. Different from the evaluation matrics in [25], we used the mean of non-zero value of Ground truth as the threshold to avoid the extreme low score when the image intensity is small. At first we binarize the values of the predicted image and ground truth pixels with a threshold, and then calculated the TP, FN, FP and TN. The formulas for the evaluation metrics is shown below:

$$CSI = \frac{TP}{TP + FN + FP} \tag{4}$$

$$POD = \frac{TP}{TP + FN} \tag{5}$$

$$FAR = \frac{FP}{TP + FP} \tag{6}$$

$$HSS = \frac{TP \times TN - FN \times FP}{(TP + FN)(FN + TN) + (TP + FP)(FP + TN)}. \tag{7}$$

In addition, we used image correlation as an evaluation indicator:

$$CORR = \frac{\sum_{i,j}(Pixel^{prediction}_{i,j} * Pixel^{groundTruth}_{i,j})}{\sqrt{\sum_{i,j} Pixel^{prediction^2}_{i,j} \times \sum_{i,j} Pixel^{prediction^2}_{i,j}}}. \tag{8}$$

TABLE I: Quantitative Result of the MovingMNIST Ablation Experiments.

DisI	adaMSE	HSS	CSI	CORR	POD	FAR
	✓	0.60	0.51	0.58	0.57	0.20
✓		0.81	0.72	0.75	0.84	0.14
✓	✓	0.85	0.77	0.79	0.84	0.12

B. MovingMNIST

We firstly performed three sets of controlled experiments on the MovingMNIST ([24], [31]) dataset to evaluate the effect of discriminator, adaMSE as well as to estimate the best depth of the model. Three models are trained for the experiment: full model, model without adding discriminator (DisI) and model without adaMSE.

Dataset: The MovingMNIST dataset is divided into a series of continuous image sequences. The first $T = 10$ frames of images in each sequence is used as input, the following $K = 10$ frames are the ground truth of the expected prediction output. To facilitate the observation of experimental phenomena, each image of the MovingMNIST dataset that we used has only one moving hand-writing digit. The hand-writing digit in each group of data is given random initial position, movement direction, and moving speed. The edge of a frame is considered as a virtual 'wall' so that the digit will change the moving direction when it hits the 'wall'. Over the 2000 generated sequences, training, verification and testing split is $8 : 1 : 1$

Quantitative Results: Table. I shows the scores of three models on the test set. The model without image discriminator (DisI) has the worst results, with a CSI of only 0.51, and a HSS of 0.6, indicating the model generate the image with worst handwriting digit image structure. The model with DisI but without adaMSE has a CSI of 0.72, and a HSS of 0.81, which is a great improvement. When AdaMSE was added, the model's CSI and HSS reached the highest 0.77 and 0.85. Similar trends appear in the other three evaluation scores, which indicates a positive effect of the DisI module and adaMSE loss. The qualitative result are shown in Fig. 2. The following part of this section will discuss the qualitative result of the model without DisI module or adaMSE loss.

DisI: In the predicted images from the model without DisI in Fig. 2, the shape of the digit deforms fast and significantly, while in the predicted image from the full model the digit looks more realistic. Moreover, the attenuation of the digit gray-scale intensity of the prediction is reduced on the model with DisI, meaning that the DisI will force the model to generate a more certain result.

To evaluate the influence of the discriminator to the model training, we hope to evaluate the quality of a series of intermediate models under different epochs. The quality of the intermediate prediction models can be reflected by the realistic level of the image sequences produced by the models. Therefore we used the trained discriminator from the DisI module as an image realistic regressor, and

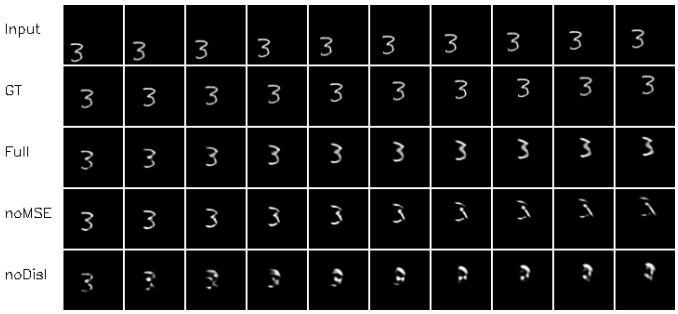


Fig. 2: Experimental results under different models configurations on the same MovingMNIST ([24], [31]) input sequence. We can see that, the model without adaMSE and the model without DisI generate blurry and deformed images.

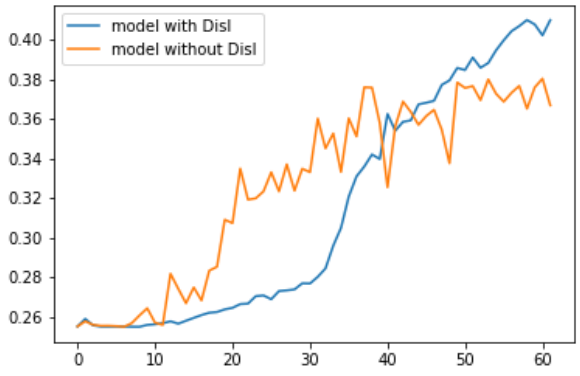


Fig. 3: The images realistic level (y-axis) from different models from different epochs(x-axis). The score of the model trained without DisI fluctuates through epochs and end up with low saturation point, whilst the full model can reach higher realistic level and stabler.

evaluate the checkpoints of the two groups of models: one trained with DisI, and the other one trained without DisI.

The resulting score is plotted in Fig. 3. From the figure, it can be seen that the score curve of model without DisI rises with strong fluctuation, and end up with a final saturation score of 0.36. The realistic score curve of model with DisI increases stably, and end up with the realistic level of 0.41 or above. The figure could lead to the conclusion that DisI module ensures the fidelity of prediction. It make the training more cautious, thereby prevent the model from dropping into a local minima or mode collapse.

Adaptive MSE (adaMSE) Loss: The main characteristic of our datasets is the presents of a large number of zero-value pixels in the images background. In Fig. 2, the prediction with traditional MSE loss become blur especially in the last several prediction steps, indicating that the original model cannot effectively converge under

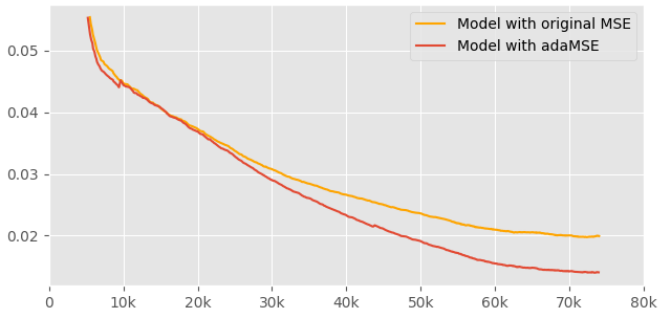


Fig. 4: The GDL loss during training for models using adaMSE and original MSE. The adaMSE could allow a final saturated loss to be about 0.07 compared to that uses original MSE of 0.2.

our scenario. Our full model with adaMSE, on the other hand, could produce clearer hand-writing digit shape. The adaMSE suppress the loss from background with zero intensity values therefore enhance the generation of foreground part. Fig. 4 shows the Gradient Difference Loss (GDL) [32] in models trained with original MSE and adaMSE. The adaMSE leads to a faster convergence as well as a lower saturation loss.

Layers Number Tuning: In order to verify the effect of different layers on this multilayer LSTM structure, we perform a series of experiments and record the system performance. To simplify the experimental environment, we fix the channel shape to a size of 64×64. The layer size has seven possible values $nLayer \in \{1, 2, 3, 4, 6, 8, 10\}$. After 600 epoch, the model was tested on the test set, and the results are shown in Fig. 5.

From the figure, it can be seen that the minimum layer number for an acceptable result is $nLayer = 3$. When the LSTM layer is chosen to be 1 or 2, the prediction result decays quickly and become unrecognizable at the end, indicating that the shadow-layer LSTM structure is easy to lose information.

We consider $nLayer = 3$ to be a most suitable model depth since it is deep enough to describe the sequence and reduce the risk of over-fitting under a limited size of dataset. Moreover, the radar image dataset has a larger image size, it is computational expensive to build a prediction model with unnecessary additional layers.

C. Nowcasting

In the nowcasting task, the input sequence is radar echo gray scale image of $T = 10$ frames, and the expected output is a sequence of K frames future frames prediction.

The rest of this subsection will introduce our nowcasting dataset, and the experimental result compared to the baseline system in [24] and [25] as well as traditional optical flow algorithm in [30]. Note that since the baseline system [24], [25] could only predict 10 future frames, the comparison with them will base on a 10 steps prediction

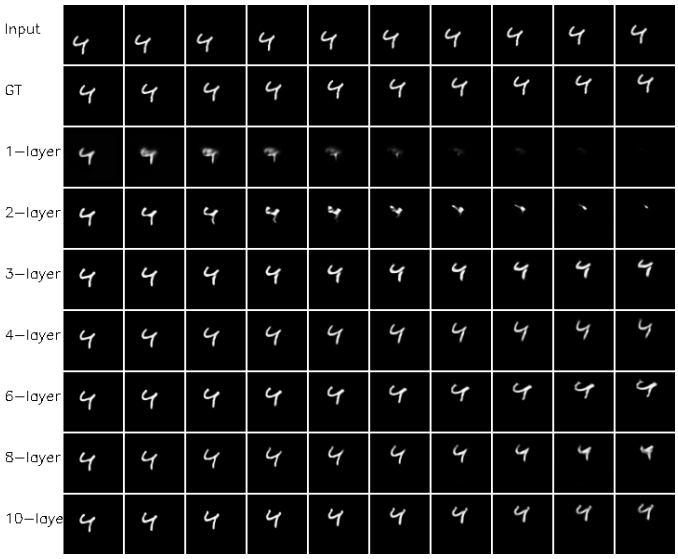


Fig. 5: MovingMNIST Prediction results of model with different layers. We consider $nLayer = 3$ to be the minimum acceptable model depth.

TABLE II: Quantitative Result of ConvLSTM, TrajGRU and Our Model on the nowcasting dataset.

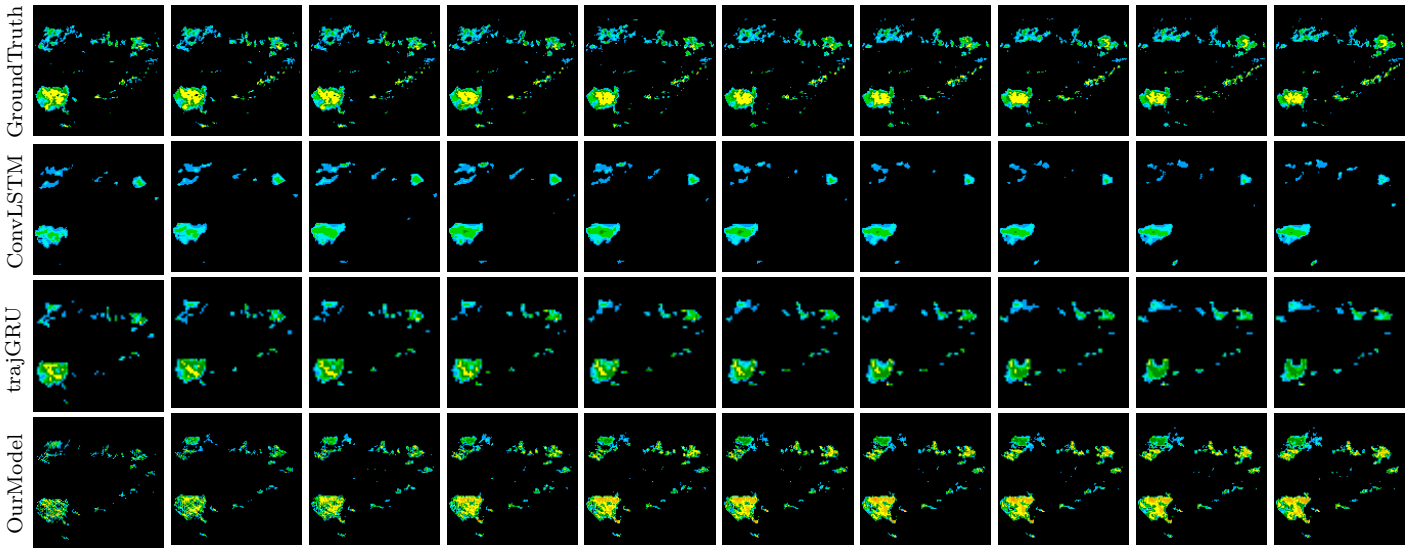
Model	HSS	CSI	FAR	POD	CORR
ConvLSTM	0.52	0.41	0.32	0.52	0.68
TrajGRU	0.55	0.46	0.34	0.61	0.72
OurModel	0.61	0.55	0.29	0.69	0.75

task, and compare to the optical flow algorithm on a 20 steps prediction task. Each step represents a 6 minutes time step.

Dataset: The radar data is provided by China Eastern Region Air Traffic Management Bureau (ECATMB), which covers a region of $250km$ centered in Shang Hai Hong Qiao Airport on a 2 years duration from 2017 to 2019. The data is converted to the CAPPI (Constant Altitude Plan Position Indicator) format which consists of 40 layers representing different altitudes. To simplify the demonstration, we only select the bottom layer as the representative in this section, if other layers results are of interested, please contact the corresponding author via email. The raw CAPPI data is converted to images format with a resolution of 720×720, the radar echo intensity, in the unit of dBZ within the range of $[-32, 95.5]$, is linearly mapping to the range of $[0, 255]$ gray scale by the mapping equation of $I_{gray} = (I_{dBZ} + 32) * 2$. The CAPPI data is processed by quality control methods including noise removal and interpolation.

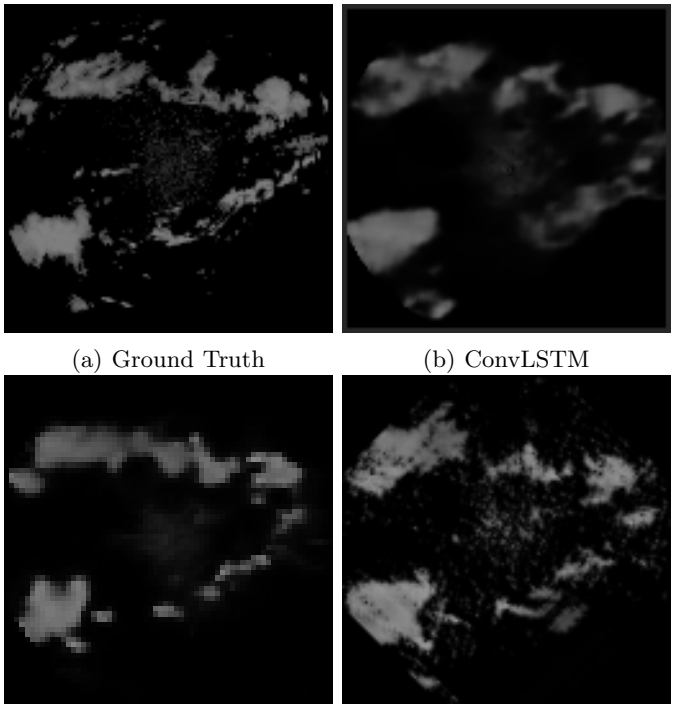
Evaluation Results: Table II shows the evaluation result of the baseline [24], [25] and that of our system. Generally, the CSI index of our model is the highest among all the candidate models.

The qualitative result of nowcasting compared to the baseline was showed in Fig. 6, in which gray scale images are converted to heat map for a clearer visualization



(a) step 0 (b) step 1 (c) step 2 (d) step 3 (e) step 4 (f) step 5 (g) step 6 (h) step 7 (i) step 8 (j) step 9

Fig. 6: Experimental results of different models in the 1 hours (10 steps) radar image prediction task. The time difference between each consecutive step is 6 minutes. It is clearly demonstrated that, the state-of-the-art systems produce blurry result even under the task of only 1-hour prediction, whilst our model produce certain and more realistic prediction.



(a) Ground Truth (b) ConvLSTM

(c) TrajGRU (d) Our Model

Fig. 7: Comparison of the step-9 frame generated by three models. It can be seen that the intensity decay and the texture loss of baseline systems in Fig. 6 is mainly due to the blurry of predictions. Our model produces images that are closer to the the real radar echo image.

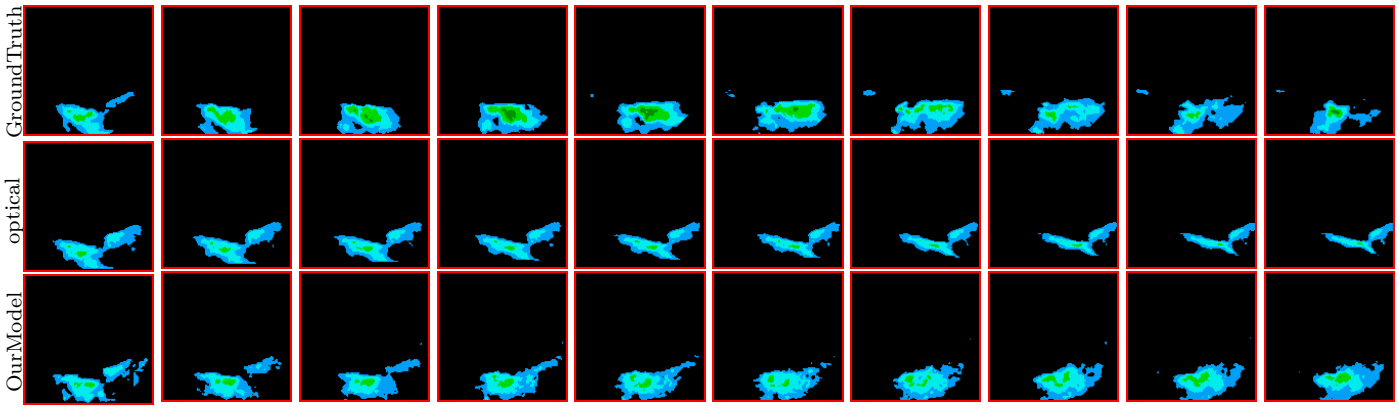
purpose. The warmer color represent stronger radar echo and cooler color represents the weaker radar echo which is less of interested.

In the result of the ConvLSTM model, the intensity quickly decay from the very first frame compared to the ground truth. The yellow area which represents higher echo intensity disappears from the first frame. Moreover, the median intensity parts (green color) are losing their texture, and merging to a group of textureless area through time. Finally, the overall colored area is shrinking quickly representing a falsely prediction of intensity reduction. The intensity decay, detail missing and area shrinking reflect the weakness of the prediction model to generate a clear sequence of images with rich local details. We can see the detail of the step-9 frame in the enlarged gray scale images in Fig. 7, the baselines produce blurry result which leads to the false predictions mentioned before.

The prediction of TrajGRU [25] is better than ConvLSTM [24], it preserves more high intensity areas and more local details as shown in Fig. 6, however, the intensity decay as well as blurry prediction can still be found in Fig. 6 and Fig. 7.

In our model, the high intensity areas in yellow are well preserved in terms of their size and texture details in Fig. 6. The gray scale image in Fig. 7 illustrates that our model is capable to produce images with rich local detail which is closer to the ground truth.

We also compare our model with the optical flow extrapolation method [30] on the task of predicting 20 frames with 10 input frames. Fig. 8 shows the experimental results, we only show the odd number steps to save room.



(a) step 1 (b) step 3 (c) step 5 (d) step 7 (e) step 9 (f) step 11 (g) step 13 (h) step 15 (i) step 17 (j) step 19

Fig. 8: Experimental results of our model and the traditional well-accepted optical flow method for the task of 2 hours (20 steps radar image prediction task. We only show the odd number steps so the time difference between two frames is 12 minutes. Optical flow algorithm could not predict the intensity change through time, whilst our method provide the prediction that is close to the ground truth in terms of intensity, structure and location.

For traditional optical flow, because there is no additional information added, the existing image is extrapolated only by optical flow. Although the performance will be good during the first few frames, after around 5 frames, which is equivalent to 30 minutes time, distortion and tearing become obvious in the images. Our model generate reasonable image content even in the final steps.

V. Conclusion

In this paper, we discussed a deep recurrent neural network model for predicting image sequences of weather radars. We designed image discriminators to ensure continuity of sequences and visual quality of images. A new adaptive loss was proposed to train the new model, which accelerated the convergence rate during training and improved the performance of the model. The optical extrapolation method was used to capture the motion information of sequential images, and it was fed into the DeepRNN model. The experimental results substantiated that the new model outperformed the conventional optical flow method and the deep learning-based ConvLSTM and TrajGRU methods. Future research may be directed to the exploration of additional meteorological elements to make the prediction model more accurate.

References

[1] James W Wilson, N Andrew Crook, Cynthia K Mueller, Juanzhen Sun, and Michael Dixon, "Nowcasting thunderstorms: A status report," Bulletin of the American Meteorological Society, vol. 79, no. 10, pp. 2079–2100, 1998.

[2] Peter Lynch, "The origins of computer weather prediction and climate modeling," Journal of Computational Physics, vol. 227, no. 7, pp. 3431–3444, 2008.

[3] Neil I Fox and James W Wilson, "Very short period quantitative precipitation forecasting," Atmospheric Science Letters, vol. 6, no. 1, pp. 7–11, 2005.

[4] A Bellon and GL Austin, "The accuracy of short-term radar rainfall forecasts," Journal of hydrology, vol. 70, no. 1-4, pp. 35–49, 1984.

[5] Neill EH Bowler, Clive E Pierce, and Alan Seed, "Development of a precipitation nowcasting algorithm based upon optical flow techniques," Journal of Hydrology, vol. 288, no. 1-2, pp. 74–91, 2004.

[6] Frank Steinbrücker, Thomas Pock, and Daniel Cremers, "Large displacement optical flow computation withoutwarping," in 2009 IEEE 12th International Conference on Computer Vision. IEEE, 2009, pp. 1609–1614.

[7] Yu Liu, Du-Gang Xi, Zhao-Liang Li, and Yang Hong, "A new methodology for pixel-quantitative precipitation nowcasting using a pyramid lucas kanade optical flow approach," Journal of Hydrology, vol. 529, pp. 354–364, 2015.

[8] Till Kroeger, Radu Timofte, Dengxin Dai, and Luc Van Gool, "Fast optical flow using dense inverse search," in European Conference on Computer Vision. Springer, 2016, pp. 471–488.

[9] Wang-chun Woo and Wai-kin Wong, "Operational application of optical flow techniques to radar-based rainfall nowcasting," Atmosphere, vol. 8, no. 3, pp. 48, 2017.

[10] Tage Andersson and Karl-Ivar Ivarsson, "A model for probability nowcasts of accumulated precipitation using radar," Journal of Applied Meteorology, vol. 30, no. 1, pp. 135–141, 1991.

[11] W Schmid, S Mecklenburg, and J Joss, "Short-term risk forecasts of severe weather," Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere, vol. 25, no. 10-12, pp. 1335–1338, 2000.

[12] Alan W Seed, Clive E Pierce, and Katie Norman, "Formulation and evaluation of a scale decomposition-based stochastic precipitation nowcast scheme," Water Resources Research, vol. 49, no. 10, pp. 6624–6641, 2013.

[13] Aitor Atencia and Isztar Zawadzki, "A comparison of two techniques for generating nowcasting ensembles. part i: Lagrangian ensemble technique," Monthly Weather Review, vol. 142, no. 11, pp. 4036–4052, 2014.

[14] Loris Foresti, Luca Panziera, Pradeep V Mandapaka, Urs Germann, and Alan Seed, "Retrieval of analogue radar images for ensemble nowcasting of orographic rainfall," Meteorological Applications, vol. 22, no. 2, pp. 141–155, 2015.

[15] Agnieszka Barszcz, Jason A Milbrandt, and Julie M Thériault, "Improving the explicit prediction of freezing rain in a kilometer-scale numerical weather prediction model," Weather and Forecasting, vol. 33, no. 3, pp. 767–782, 2018.

[16] Xunlai Chen, Guangjun He, Yuanzhao Chen, Shuting Zhang, Jinsong Chen, Jing Qian, and Haicong Yu, "Short-term and local rainfall probability prediction based on a dislocation support vector machine model using satellite and in-situ observational data," IEEE Access, 2019.

[17] L Zhang, M Wang, and H Li, "Discussion of relative accuracy

of short-range heavy rain nowcasting (in chinese)," Guangdong Meteorol., vol. 37, pp. 1–6, 2015.
[18] Mark N French, Witold F Krajewski, and Robert R Cuykendall, "Rainfall forecasting in space and time using a neural network," Journal of hydrology, vol. 137, no. 1-4, pp. 1–31, 1992.
[19] John R Mecikalski, John K Williams, Christopher P Jewett, David Ahijevych, Anita LeRoy, and John R Walker, "Probabilistic 0–1-h convective initiation nowcasts that combine geostationary satellite observations and numerical weather prediction model data," Journal of Applied Meteorology and Climatology, vol. 54, no. 5, pp. 1039–1059, 2015.
[20] Aditya Nagarajan, "Explorations into machine learning techniques for precipitation nowcasting," 2017.
[21] Gil Keren and Björn Schuller, "Convolutional rnn: an enhanced model for extracting features from sequential data," in 2016 International Joint Conference on Neural Networks (IJCNN). IEEE, 2016, pp. 3412–3419.
[22] Jonathan Long, Evan Shelhamer, and Trevor Darrell, "Fully convolutional networks for semantic segmentation," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 3431–3440.
[23] Ali Zahraei, Kuo-lin Hsu, Soroosh Sorooshian, JJ Gourley, Valliappa Lakshmanan, Yang Hong, and Tim Bellerby, "Quantitative precipitation nowcasting: a lagrangian pixel-based approach," Atmospheric Research, vol. 118, pp. 418–434, 2012.
[24] Xingjian Shi, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai kin Wong, and Wang chun Woo, "Convolutional lstm network: A machine learning approach for precipitation nowcasting," .
[25] Xingjian Shi, Zhihan Gao, Leonard Lausen, Hao Wang, Dit-Yan Yeung, Wai kin Wong, and Wang chun Woo, "Deep learning for precipitation nowcasting: A benchmark and a new model," .
[26] Yunbo Wang, Zhifeng Gao, Mingsheng Long, Jianmin Wang, and Philip S. Yu, "Predrnn++: Towards a resolution of the deep-in-time dilemma in spatiotemporal predictive learning," .
[27] Mehdi Mirza and Simon Osindero, "Conditional generative adversarial nets," .
[28] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," .
[29] Sergey Tulyakov, Ming-Yu Liu, Xiaodong Yang, and Jan Kautz, "Mocogan: Decomposing motion and content for video generation," .
[30] L. Foresti, M. Reyniers, A. Seed, and L. Delobbe, "Development and verification of a real-time stochastic precipitation nowcasting system for urban hydrology in belgium," Hydrology and Earth System Sciences, vol. 20, no. 1, pp. 505–527, jan 2016.
[31] Nitish Srivastava, Elman Mansimov, and Ruslan Salakhudinov, "Unsupervised learning of video representations using lstms," in International conference on machine learning, 2015, pp. 843–852.
[32] Michael Mathieu, Camille Couprie, and Yann LeCun, "Deep multi-scale video prediction beyond mean square error," arXiv preprint arXiv:1511.05440, 2015.