

Article

Short-Term and Imminent Rainfall Prediction Model Based on ConvLSTM and SmaAT-UNet

Yuanyuan Liao ^{1,*} , Shouqian Lu ¹ and Gang Yin ²¹ School of Computer Science and Technology, Xinjiang University, Urumqi 830049, China² College of Geography and Remote Sensing Sciences, Xinjiang University, Urumqi 830049, China; chinayg@xju.edu.cn

* Correspondence: liaoyuan@xju.edu.cn; Tel.: +86-133-1982-9162

Abstract: Short-term precipitation forecasting methods are mainly divided into statistical forecasting, numerical model-based forecasting, and radar image extrapolation techniques. The two methods based on statistical prediction and numerical model have the disadvantages of being unstable and generating large errors. Therefore, this study proposes the use of deep learning for radar image extrapolation for precipitation forecasting, in particular by developing algorithms for ConvLSTM and SmaAT-UNet. The ConvLSTM model is a fusion of a CNN (Convolutional Neural Network) and LSTM (Long Short-Term Memory network), which solves the challenge of processing spatial sequence data, which is a task that traditional LSTM models cannot accomplish. At the same time, SmaAT-UNet enhances the traditional UNet structure by incorporating the CBAM (Convolutional Block Attention Module) attention mechanism and replacing the standard convolutional layer with depthwise separable convolution. This innovative approach aims to improve the efficiency and accuracy of short-term precipitation forecasting by improving feature extraction and data processing techniques. Evaluation and analysis of experimental data show that both models exhibit good predictive ability, with the SmaAT-UNet model outperforming ConvLSTM in terms of accuracy. The results show that the performance indicators of precipitation prediction, especially detection probability (POD) and the Critical Success index (CSI), show a downward trend with the extension of the prediction time. This trend highlights the inherent challenges of maintaining predictive accuracy over longer periods of time and highlights the superior performance and resilience of the SmaAT-UNet model under these conditions. Compared with the statistical forecasting method and numerical model forecasting method, its accuracy in short-term rainfall forecasting is improved.

Keywords: short-term precipitation forecasting; deep learning; radar image extrapolation; ConvLSTM; SmaAT-UNet



Citation: Liao, Y.; Lu, S.; Yin, G. Short-Term and Imminent Rainfall Prediction Model Based on ConvLSTM and SmaAT-UNet. *Sensors* **2024**, *24*, 3576. <https://doi.org/10.3390/s24113576>

Academic Editor: Carmine Serio

Received: 22 April 2024

Revised: 13 May 2024

Accepted: 30 May 2024

Published: 1 June 2024



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

1. Introduction

Short-term precipitation is defined as rainfall events forecasted to occur within the ensuing few hours, typically constrained to a forecast horizon of up to 6 h. The precision of such forecasts can achieve remarkable granularity, extending to kilometer-level and minute-level accuracy [1]. The capacity to accurately predict short-term precipitation bears considerable significance across various sectors, including agriculture, transportation, and water management. By facilitating timely preemptive actions, it allows for effective mitigation of precipitation-related impacts.

With the advent of data mining and machine learning technologies, statistical forecasting methods have gained widespread applications [2]. Precipitation is a very complex nonlinear problem which involves the exchange of water, heat, and momentum between the ground and the air during the water cycle. However, statistical methods are criticized for their lack of a physical basis and inherent instability [3]. Given their linear nature, they struggle with the fundamentally nonlinear characteristics of meteorological elements, constraining forecast accuracy. This limitation has steered research towards nonlinear

forecasting methods, prominently featuring artificial neural networks (ANNs). ANNs excel in mapping nonlinear relationships between input and output variables, reliant solely on training samples and targets. This capability allows for the selection of an optimal network configuration through learning and training processes, enhancing the model's ability to generalize beyond the training data [4].

Numerical weather prediction is based on the input of radar, satellite, ground and upper air observations and topographic data, and the output of meteorological elements such as temperature and precipitation at a future time. Numerical weather precipitation forecasting offers several advantages but is not without drawbacks. Limitations such as the low spatial resolution of radar images, accumulative errors, and the influence of terrain and water bodies introduce significant uncertainty into forecast results. These factors adversely affect the practical application of this method's accuracy and reliability. An innovative approach involves using radar extrapolation algorithms to predict future precipitation amounts and distribution based on radar-derived reflectivity factor data.

Noteworthy contributions to the field include the work by Sohail et al., who employed a BP neural network and an ARMA linear model to successfully forecast precipitation in the Japanese region [5]. Similarly, Liu Yong and colleagues utilized a three-layer BP neural network, integrating principal component analysis with precipitation and runoff data, to predict the autumn flood season with high stability and satisfactory accuracy [6]. Despite the BP neural network's proven efficacy, its lengthy training time and the challenge of accurately approximating the error function due to discrepancies between input and target values remain significant obstacles [7–9].

A notable advancement is the application of the ConvLSTM model by Liang Zhenqing, Chen Sheng, and others for forecasting in the Guangzhou region [10]. ConvLSTM was proposed by Xingjian Shi, which synergizes the convolutional operations of CNNs with LSTM's temporal data processing capabilities, effectively addressing the spatiotemporal dimensions of precipitation forecasting [11]. The results show that the correlation coefficient between the forecast results and the actual observed values remains above 0.6 within 1 h, and the false positive rate is less than 40%. The ConvLSTM model demonstrates a sophisticated handling of time series and spatial data, facilitating more accurate predictions of short-term precipitation events. In addition, the MOML algorithm developed by Zhang Pingwen's team at Peking University demonstrates the potential of data mining and learning in intelligent and efficient forecasting, as demonstrated by its successful application in the Beijing 2022 Winter Olympics, where the algorithm improved the accuracy of model forecasting results for temperature, humidity, wind speed, and wind direction by more than 10% [12].

In this study, we combined the previous models and used deep learning methods for radar image extrapolation and realized a radar image extrapolation method based on ConvLSTM and SmaAT-UNet. First, we will introduce the SmaAT-UNet Model.

2. Introduction to SmaAT-UNet

SmaAT-UNet, an advanced model derived from the foundational UNet architecture, incorporates CBAM (Convolutional Block Attention Module) to enhance its decoder [13,14]. This module enables adaptive adjustments of feature map weights, ensuring tailored processing for specific tasks. Additionally, SmaAT-UNet innovates by replacing traditional convolution methods with depthwise separable convolution, significantly enhancing model efficiency. The structural blueprint of SmaAT-UNet is depicted in Figure 1.

CBAM combines a channel attention mechanism and spatial attention mechanism, is designed for visual applications such as image classification and object detection [15]. It addresses the challenge of uneven feature significance across various channels and spatial regions in convolutional neural networks. Through CBAM, the model emphasizes crucial channels and spatial locations by dynamically weighting feature information, thereby refining the focus and effectiveness of the network.

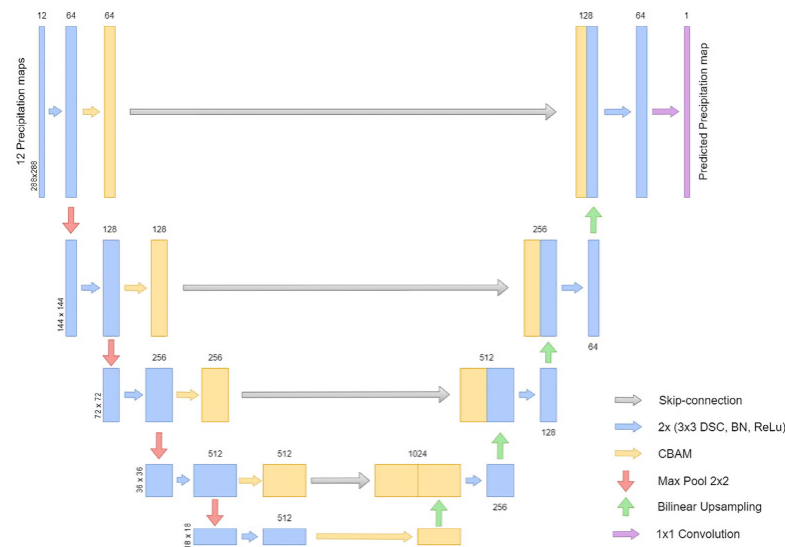


Figure 1. SmaAT-UNet network structure diagram.

In detail, CBAM’s channel attention mechanism evaluates and assigns weights to each channel’s features using global average pooling and fully connected layers, producing a set of optimized channel features. Concurrently, the spatial attention mechanism assigns weights to specific spatial locations through convolutional and fully connected layers, crafting a targeted spatial feature set.

For the channel attention module, input features F are derived from both MaxPool and AvgPool operations, symbolizing the highest and average aggregated features, respectively. These features are then processed through a shared, fully connected layer, resulting in combined and weighted feature vectors, as defined by Equation (1):

$$\begin{aligned} M_c(F) &= \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \\ &= \sigma(W_1(W_0(F_{avg}^c)) + W_1(W_0(F_{max}^c))) \end{aligned} \quad (1)$$

where F —Input feature, σ —Activation function and W_1 , W_0 —Weight value.

The spatial attention module utilizes the feature vector from the channel attention module as its input. Post-processing through max-pooling and average-pooling, the results are merged and further refined by standard convolution, culminating in the generation of spatial attention weights. The procedural formula for spatial attention is outlined in Equation (2).

$$\begin{aligned} M_s(F) &= \sigma(f^{7 \times 7}([AvgPool(F); MaxPool(F)])) \\ &= \sigma(f^{7 \times 7}([F_{avg}^s; F_{max}^s])) \end{aligned} \quad (2)$$

The CBAM model’s capability to adaptively discern and emphasize essential features across channels and spatial domains significantly bolsters SmaAT-UNet’s performance. Moreover, its low computational overhead allows for seamless integration into existing network frameworks.

SmaAT-UNet’s transition to depthwise separable convolution, comprising depthwise and pointwise convolutions, marks a departure from conventional convolution techniques. This modification enables individualized feature processing for each input channel and a subsequent amalgamation of these features, substantially reducing the model’s parameter count and computational demand. This strategic replacement underscores SmaAT-UNet’s streamlined and effective design, affirming its position as a lightweight yet powerful model for complex visual tasks.

3. Short-Term and Impending Precipitation Prediction Using Convolutional Long Short-Term Memory Network

3.1. Data Set Preparation

The efficacy of predictive models, particularly those extrapolating radar echo images for short-term precipitation forecasting, is fundamentally dependent on the quality of the dataset employed. Therefore, rigorous preprocessing of the dataset is indispensable to enhance the model's accuracy and generalization capabilities. The original dataset may include erroneous or incomplete data, such as default values and outliers, which can compromise model performance. Through careful preprocessing, these inaccuracies can be mitigated to improve data quality significantly. Using data standardization and normalization and image filtering can improve the quality of data sets.

3.1.1. Standardization and Normalization of Data

Normalization plays a crucial role in aligning data across different scales to a uniform standard, thus neutralizing scale-induced discrepancies in data analysis. This process is particularly critical in deep learning, where it ensures the equal importance of diverse features by mitigating issues arising from variance in feature value ranges, which could otherwise lead to inaccurate or slow model convergence. Moreover, normalization enhances model stability and interpretability.

The Z–R relationship, foundational to the empirical correlation between radar echo intensity and precipitation rate, facilitates the conversion of radar echo intensity to a standard precipitation rate formula. This pivotal conversion process is encapsulated in Equation (3).

$$Z = aR^b \quad (3)$$

where Z—Radar echo intensity (in dBZ), R—Rainfall rate (in mm/h), a and b—Radar constant, with $a \approx 200$ and b between 1.5 and 2.

3.1.2. Image Filtering

The presence of noise in digital images can deteriorate image quality, adversely affecting subsequent image processing and analysis tasks. Excessive or irrelevant image details might also obscure experimental outcomes. To counter these issues, median filtering is employed to suppress noise, thereby augmenting image quality, eliminating superfluous details, and enhancing contrast. This technique, which replaces the gray value of each pixel with the median of the gray values in its vicinity, effectively highlights crucial image details for easier interpretation [16].

3.1.3. Radar Dataset

This study utilizes the HKO-7 radar dataset, generously provided and licensed by the Hong Kong Observatory. This dataset has been subjected to thorough data cleaning and quality assurance measures, ensuring the reliability and accuracy of the data for forecasting purposes. To access, see "<https://github.com/sxjsience/HKO-7> (accessed on 15 May 2024)".

The original radar echo images in this dataset are uniformly sized at 480×480 pixels. Considering the substantial volume and dimensions of the image data, which could prolong training durations, images are resized to 64×64 pixels. The 480×480 pixel radar echo images are illustrated in Figure 2.

Spanning from 2009 to 2015, the HKO-7 dataset encompasses radar echo data for the Hong Kong Special Administrative Region captured at 6-min intervals—equating to 240 data frames daily. The training model inputs 5 frames (30 min) to forecast the subsequent 20 frames (120 min). Due to computational constraints and the fact that rainfall in Hong Kong is concentrated from April to September, we selected two of these months: April and May [17]. A subset of data from April and May 2014, encompassing 61 days and 14,640 image frames, was randomly selected for model training.

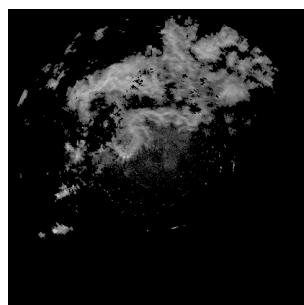


Figure 2. Radar gray echo data diagram.

3.2. Experimental Process

L1 regularization is the sum of the absolute values of each element in the weight vector w , and L2 regularization is the square root of the sum of squares in the weight vector. These two methods are to solve the overfitting problem, that is, to reduce the size or number of model parameters to alleviate the overfitting problem [18]. To mitigate the potential ambiguities associated with L1 and L2 regularization techniques, this study employs the Log-Cosh loss function. Known fully as “Logarithmic Hyperbolic Cosine Loss”, this function serves as a smooth alternative that effectively penalizes extreme values, thereby offering enhanced robustness compared to the Mean Squared Error (MSE) loss function. The mathematical representation of the Log-Cosh loss function is provided in Equation (4).

$$L(y, y^P) = \sum_{i=1}^n \log(\cosh(y_i^P - y_i)) \quad (4)$$

where y denotes the actual values, y^P signifies the predicted values, and n represents the total number of samples. The parameter training process involves the dynamic adjustment of model parameters through the use of iterative algorithms such as the backpropagation (BP) algorithm. To circumvent overfitting, an early stopping mechanism is employed; this involves monitoring the validation dataset’s accuracy at the conclusion of each training iteration. Training ceases when there is no observed improvement in accuracy for n consecutive iterations, with n typically set at 10. Accordingly, for this experiment, the training configuration was established with 10,000 epochs, a batch size of 2, and the early stopping parameter n also set at 10.

3.3. Meteorological Evaluation Standard

The threshold to distinguish positive and negative samples, set at default to 0.1. We use a meteorological default grid point ≥ 0.1 to determine the presence of precipitation.

In the field of meteorology, it is not possible to simply use accuracy to describe the quality of a model. Instead, we mainly use the idea of “two-class classification” to evaluate it, comparing the prediction results with the actual observed results. We divide both the observed values and the predicted values into two categories: precipitation and no precipitation. If the pixel value of a pixel point in the radar image extrapolation result is greater than the observed echo intensity threshold, then we binarize it to 1, otherwise we binarize it to 0 [19].

The purpose of this is to facilitate our analysis and processing of the prediction results in order to better evaluate the accuracy and reliability of the prediction. False alarms are when the predicted value has precipitation, but the true observed value has no precipitation. Misses are when the true observed value has precipitation, but the predicted value does not have precipitation. Hits are when both the observed and predicted values have precipitation, and are determined to have precipitation. The classification is shown in Table 1.

Table 1. Comparison of Rainfall Classification.

	No Precipitation Forecasted	Precipitation Forecasted
No precipitation observed	Correct negatives	False alarms
Precipitation observed	Misses	Hits

POD represents the ratio of the predicted precipitation area to the total observed precipitation area, as shown in Equation (5).

$$\text{POD} = \text{Hits} \div (\text{Hits} + \text{Misses}) \quad (5)$$

The False Alarm Rate (FAR) is the ratio of the predicted area of unobserved precipitation to the total area of observed precipitation, as shown in Equation (6).

$$\text{FAR} = \text{False alarms} \div (\text{Hits} + \text{False alarms}) \quad (6)$$

The Critical Success Index (CSI) represents the proportion of correctly predicted precipitation areas to the total predicted precipitation area, as shown in Equation (7).

$$\text{CSI} = \text{Hits} \div (\text{Hits} + \text{Misses} + \text{False alarms}) \quad (7)$$

3.4. Experimental Results

Evaluation of forecasting performance employs indicators such as Probability of Detection (POD), False Alarm Rate (FAR), and Critical Success Index (CSI).

Our findings are systematically compiled in Table 2. Additionally, the ConvLSTM model's predictive capabilities over a forthcoming two-hour interval, specifically regarding POD, FAR, and CSI changes, are graphically represented in Figure 3.

Table 2. The average rating of the future 20 frames predicted by the ConvLSTM model.

Model	Evaluation Indicators		
	POD	FAR	CSI
ConvLSTM	0.85	0.62	0.36

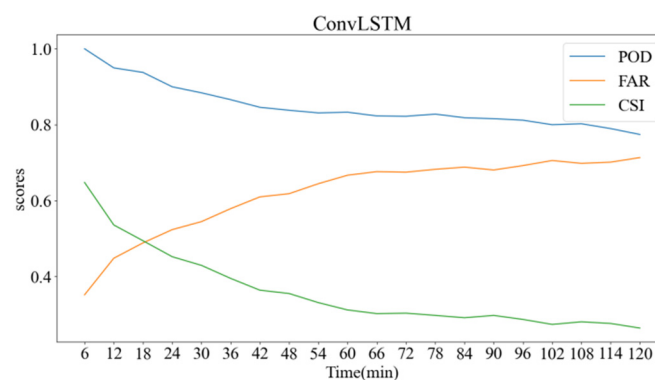
**Figure 3.** ConvLSTM predicts the change curve of three scores within two hours.

Figure 4 exhibits the observed radar imagery, providing a visual reference for the model's predictive accuracy.

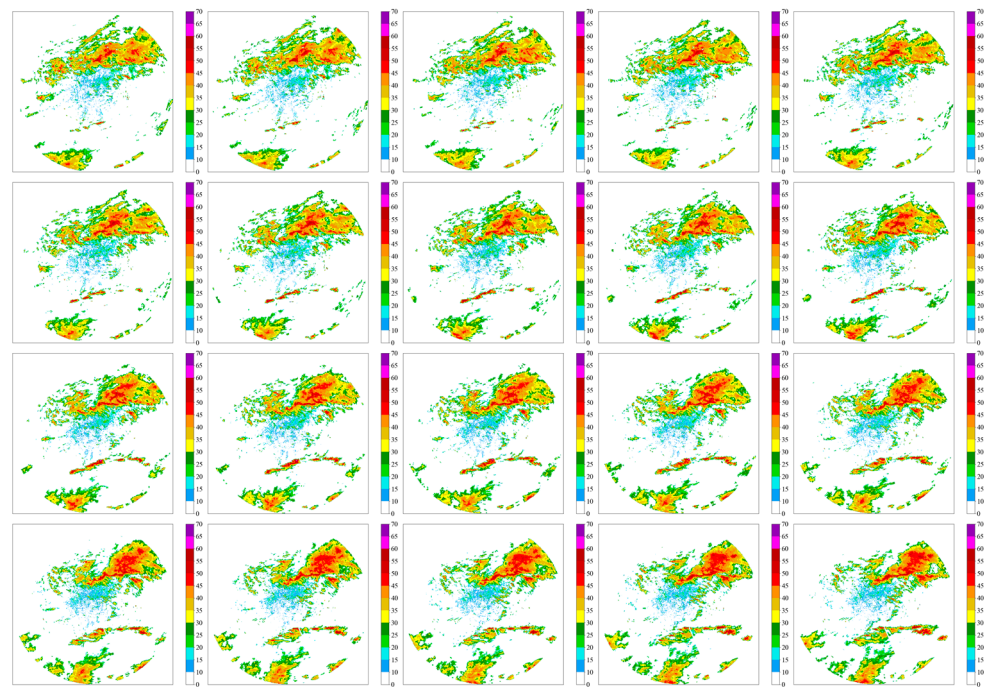


Figure 4. Observed radar image change diagram.

Figure 5 below shows the radar image changes predicted by the ConvLSTM model.

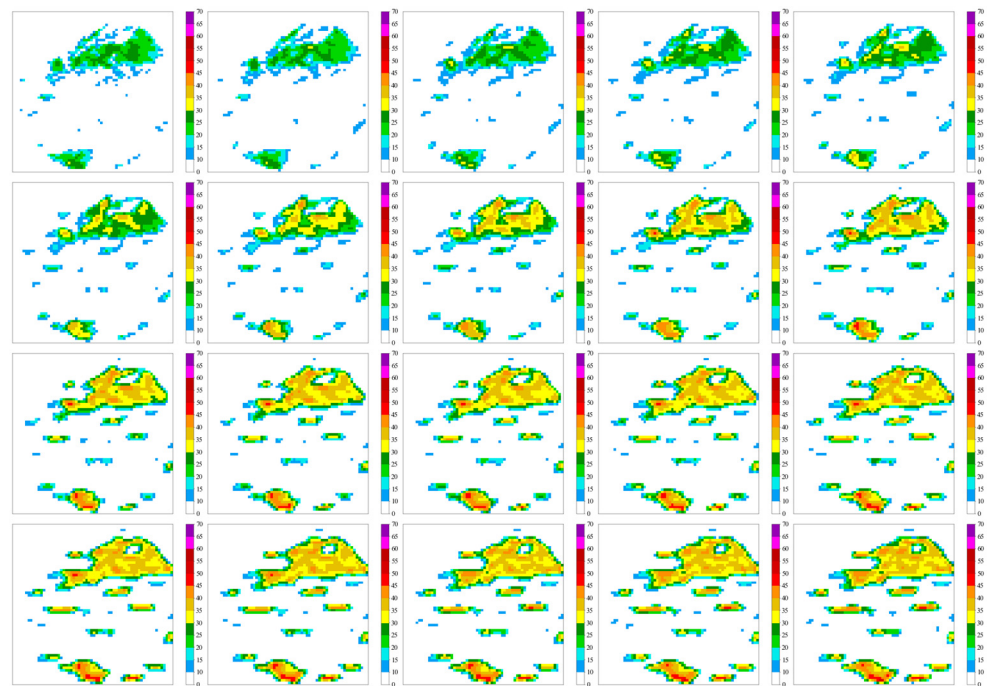


Figure 5. Predicted radar image changes using the ConvLSTM model.

4. Short-Term and Impending Precipitation Prediction Based on SmaAT-UNet

4.1. Backpropagation Optimization Algorithm—Adam Algorithm

The Adam algorithm stands as a pivotal optimization method for updating parameters within neural networks, distinguished by its adaptive learning rate mechanism. This algorithm, designed to dynamically adjust the learning rate based on the gradient variations of different parameters, significantly enhances both the efficiency and the generalizability

of model training. By amalgamating the strengths of momentum gradient descent and the RMSProp algorithm, Adam achieves superior convergence rates and effects, as has been evidenced by studies [20,21].

Adam employs both first-order and second-order momentum in the parameter update process, facilitating an accelerated gradient descent approach. This dual momentum strategy allows Adam to not only swiftly converge towards optimal solutions but also to fine-tune the update trajectory and speed of model parameters with high precision [22].

Momentum Update Formula: The momentum update, integral to achieving accelerated descent, is formalized in Equation (8).

$$v(t) = \mu v(t-1) + (1 - \mu)g(t) \quad (8)$$

where $v(t)$ denotes the momentum of t times, μ refers to hyperparameters, and $g(t)$ denotes the gradient value.

Adaptive Learning Rate Calculation: Adam calculates the adaptive learning rate by averaging the gradients and squared gradients, ensuring that the learning rate is optimally adjusted in response to parameter-specific gradient behaviors. The detailed calculation process is outlined in Equations (9)–(11).

$$m(t) = \beta_1 m(t-1) + (1 - \beta_1)g(t) \quad (9)$$

$$s(t) = \beta_2 s(t-1) + (1 - \beta_2)g^2(t) \quad (10)$$

$$\alpha(t) = \frac{\eta \mu^t}{\sqrt{s(t)} + \varepsilon} \quad (11)$$

where $m(t)$, $s(t)$, and $\alpha(t)$ represent the mean average gradient moving average, square gradient moving average, and adaptive learning rate values, respectively; β_1 and β_2 are control gradient hyperparameters; μ and η are the momentum hyperparameters and initial learning rate, respectively; $g(t)$ is the current gradient value of the batch; and ε is the offset term.

The algorithm's ability to adaptively modulate the learning rate across different parameter gradients contributes to enhanced training efficiency and model generalization. Moreover, the incorporation of both first-order and second-order momentum enables Adam to adeptly manage the directionality and velocity of parameter updates, fostering improved model convergence dynamics.

4.2. Experimental Process

In the regression prediction of radar images, the distribution of radar reflectivity values is uneven, so the BMSE loss function is commonly used. In the regression prediction task, our goal is to predict continuous data images based on the input data, so we need to define a loss function to evaluate the gap between the predicted value and the true value. The BMSE loss function can achieve this [23]. The calculation formula of the BMSE loss function is shown in Equation (12).

$$\text{BMSE} = \frac{1}{N} \sum_{n=1}^N \sum_{i=1}^{600} \sum_{j=1}^{600} \omega_{n,ij} \left(x_{n,ij} - y_{n,ij} \right)^2 \quad (12)$$

The Adam optimization algorithm, known for its adaptive learning rate, was utilized with specific parameters: learning rate (lr) of 0.001, betas of (0.9, 0.999), epsilon (eps) of 1×10^{-8} , weight decay of 0, and AMSGrad set to False. An early stopping mechanism was implemented to prevent overfitting, with training parameters set to 10,000 epochs, a batch size of 4, and a patience parameter (n) of 10.

4.3. Experimental Result Comparison, Evaluation, and Analysis

Utilizing a consistent test dataset, we embarked on predicting precipitation for the forthcoming two hours through distinct methodologies. The acquired performance metrics of the models are systematically organized in the ensuing Table 3.

Table 3. Average scores of ConvLSTM and SmaAT-UNet models in predicting the next 20 frames.

Evaluation Indicators		POD	FAR	CSI
Model				
	ConvLSTM	0.85	0.62	0.36
	SmaAT-UNet	0.95	0.61	0.38

Furthermore, Figure 6 delineates the prediction dynamics of Probability of Detection (POD), False Alarm Rate (FAR), and Critical Success Index (CSI) over the specified period.

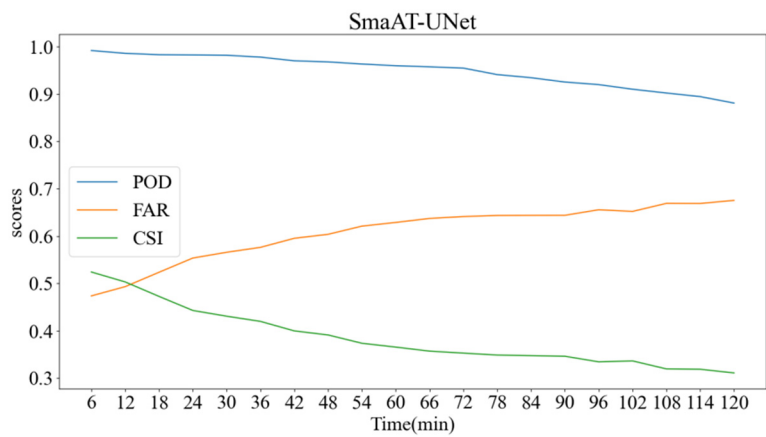


Figure 6. Three scoring curves of SmaAT-UNet predictions within two hours.

Figure 7 illustrates the variations in radar imagery as forecasted by the SmaAT-UNet model.

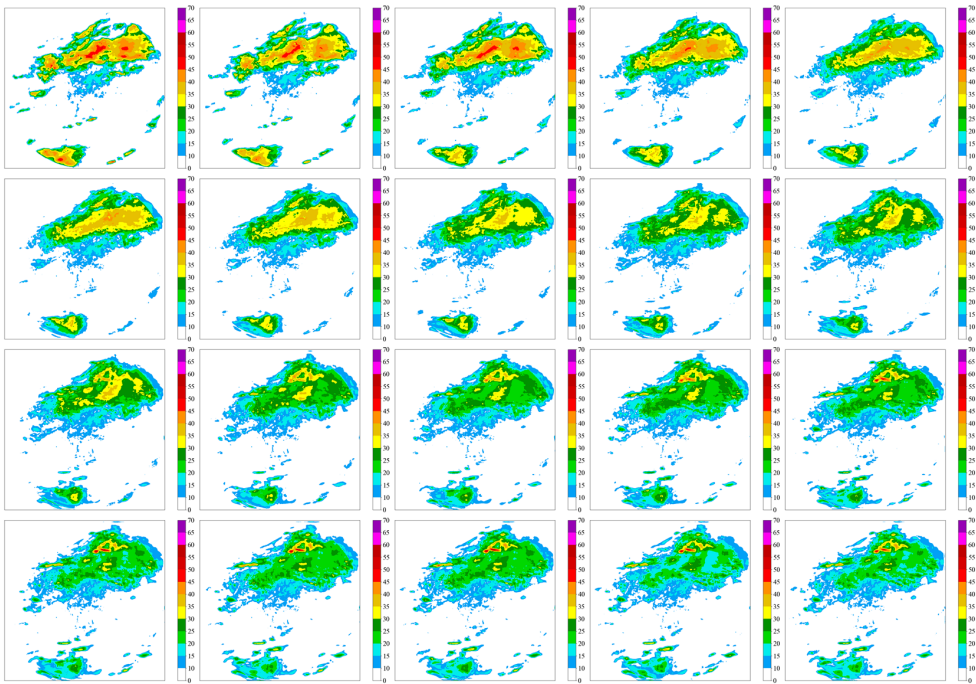


Figure 7. SmaAT-UNet predicted radar change map.

An examination of the radar echoes, as depicted in Figure 8, reveals that within each six-minute increment, a total of 20 images span the two-hour forecast interval. A notable trend is the gradual decline in CSI scores over time, signifying a diminishing prediction accuracy with the extension of the forecast period. This trend underscores an increasing discrepancy between observed values and those predicted by the models. In the comparative analysis, the ConvLSTM model exhibited superior performance to SmaAT-UNet in the initial 30 min based on CSI metrics. Conversely, from the 30-minute mark to the 120-minute endpoint, SmaAT-UNet outperformed ConvLSTM. The average CSI metrics for both models, predicting the subsequent 20 frames, were approximately 0.38 and 0.36, respectively, with the CSI change curve indicating greater stability in the SmaAT-UNet model.

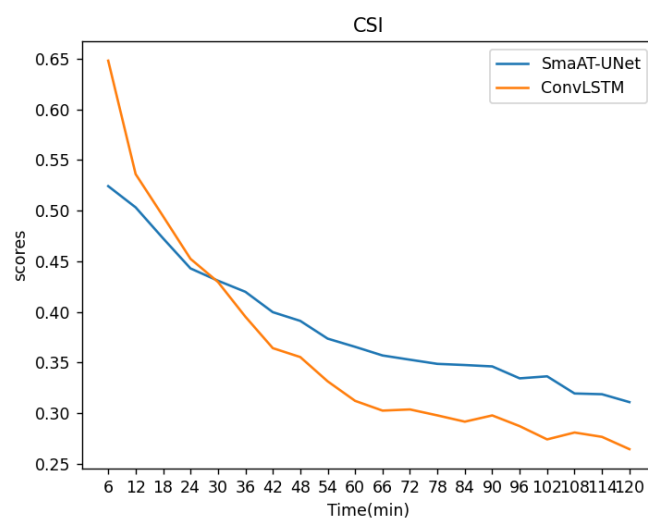


Figure 8. CSI change curves of the two methods.

The evaluation of POD, a metric representing the hit rate, is visualized in Figure 9.

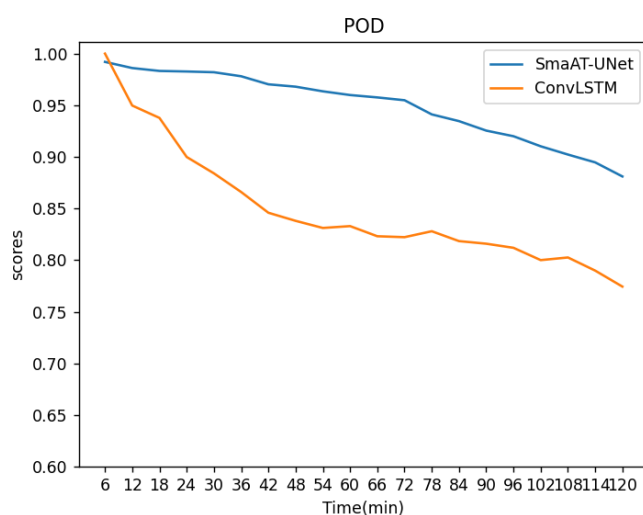


Figure 9. POD change curves of the two methods.

This analysis indicates a declining hit rate for both models as time progresses, with SmaAT-UNet consistently outpacing ConvLSTM. Specifically, the average POD index for the first 20 frames was 0.85 for ConvLSTM and 0.95 for SmaAT-UNet. The FAR metrics, assessing the False Alarm Rate of radar echo extrapolation images predicted by the models, are detailed in Figure 10.

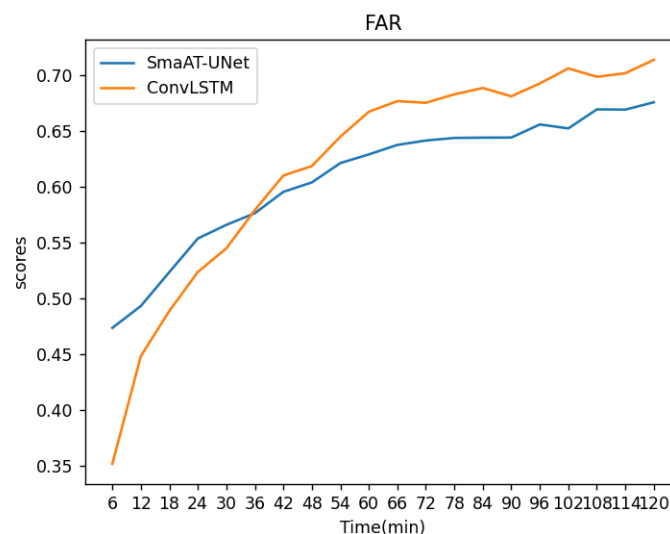


Figure 10. FAR change curves of the two methods.

Initial observations up to the 36-minute mark suggest a higher false alarm rate for SmaAT-UNet compared to ConvLSTM, implying inferior performance. However, subsequent predictions show an inversion in this trend, with ConvLSTM exhibiting a higher false alarm rate than SmaAT-UNet. The average FAR indices for the first 20 frames were 0.62 for ConvLSTM and 0.61 for SmaAT-UNet, respectively.

5. Conclusions

In contrast to traditional radar image extrapolation methods that rely on physical models demanding extensive historical data, deep learning-based algorithms offer a robust alternative. These advanced algorithms require significantly less data for training, yet they yield superior predictive outcomes, demonstrating enhanced applicability in complex radar scenarios.

The ConvLSTM model merges the capabilities of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks to overcome LSTM's limitations in processing long-term dependencies. While LSTM excels in handling time series data, CNNs adeptly manage spatial information. The SmaAT-UNet model further advances this approach by integrating the Convolutional Block Attention Module (CBAM) with UNet and adopting depthwise separable convolution, substantially reducing the model's parameter count while maintaining efficacy for regression-based prediction tasks.

In summary, both the SmaAT-UNet and ConvLSTM models present promising solutions for short-term precipitation forecasting, with SmaAT-UNet exhibiting slight advantages in performance and stability. These findings advocate for the continued exploration and integration of deep learning methodologies in meteorological forecasting, paving the way for more accurate and reliable prediction models.

Future Directions:

- **Exploration of Neural Network Architectures:** Given the study's scope constraints, only two deep learning models were examined. Future research should extend to comparing various neural network architectures known for efficacy in short-term precipitation forecasting. While the model parameters were optimized for the HKO-7 dataset, further refinement is recommended.
- **Integration of Multi-Source Data:** The dataset's quality is pivotal to a model's generalization capabilities. Deep learning models, typically data-hungry, face limitations due to the geographical and temporal constraints in meteorological data collection. Future endeavors could explore the incorporation of multi-source data, such as satellite and radar data, to enhance forecasting accuracy.

- **Addressing ConvLSTM's Limitations:** Despite its advanced memory capabilities, ConvLSTM's complex structure and substantial training demands limit its efficiency, particularly in managing long-term dependencies and processing radar image rotations and diffusions. The model's sensitivity to input data variations also warrants attention. Future work should focus on addressing these challenges to improve model robustness and performance.

Author Contributions: Methodology and analysis, Y.L.; software and writing—original draft preparation, S.L.; writing—review and editing, G.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science Foundation of the Xinjiang Uygur Autonomous Region, grant number 2021D01C080.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wang, T.; Liu, Y.; Dong, C.; Li, J. Review of Short-term and Imminent Precipitation Forecasting Methods and Their Applications. *Electron. World* **2019**, 11–13.
2. Chai, D.; Song, X.; Li, G.; Wang, J. Application of Doppler radar for observation and analysis of rainstorm and its approaching warning. *Meteorol. Sci.* **2007**, 27, 57–64. [CrossRef]
3. Hao, Y.; Yao, Y.; Zheng, Y.; LU, J. Multi-scale analysis and near-term warning of short-term heavy precipitation. *Meteorology* **2012**, 38, 903–912.
4. Zhang, J.; Wang, P.; Zhang, L.; Wang, Y. Application of artificial neural network in short-term precipitation prediction. *Sci. Technol. Wind.* **2016**, 123, 124. [CrossRef]
5. Sohail, A.; Watanabe, K.; Takeuchi, S. Runoff Analysis for a Small Watershed of Tono Area Japan by Back Propagation Artificial Neural Network with Seasonal Data. *Water Resour. Manag. Int. J. Publ. Eur. Water Resour. Assoc. (EWRA)* **2008**, 22, 1–22. [CrossRef]
6. Liu, Y.; Wang, Y.; Chen, Y.; Wang, Z.; Hu, J.; Feng, X. Long-term runoff forecasting for the autumn flood season in the Danjiangkou Reservoir. *Adv. Water Sci.* **2010**, 21, 771–778.
7. Chen, X.; Liu, J.; Zheng, Q.; Li, X.; Liu, J.; Ji, X.; Chen, Y.; Ye, Y. Nowforecasting method based on convolutional gated recurrent unit neural network. *Plateau Meteorol.* **2021**, 40, 411–423.
8. Han, F.; Long, M.; Li, Y.; Xue, F.; Wang, J. Application of recurrent neural network in radar nowcasting. *J. Appl. Meteorol.* **2019**, 30, 61–69.
9. Shi, E.; Li, Q.; Gu, D.; Zhao, Z. Radar echo extrapolation method based on convolutional neural network. *Comput. Appl.* **2018**, 38, 661–665, 676.
10. Liang, Z.; Chen, S. Evaluation of short-term and imminent prediction accuracy based on deep learning and radar observations in South China. *Meteorol. Res. Appl.* **2020**, 41, 41–47. [CrossRef]
11. Shi, X.; Chen, Z.; Wang, H.; Yeung, D.-Y. *Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting*; MIT Press: Cambridge, MA, USA, 2015.
12. Academician Zhang Pingwen: Artificial Intelligence Makes Weather Forecasts for the Winter Olympics More Accurate. 2022. Available online: <http://bda.pku.edu.cn/info/1003/2291.htm> (accessed on 15 May 2024).
13. Zhang, Y.; Kers, J.; Cassol, C.A.; Roelofs, J.J.; Idrees, N.; Farber, A.; Haroon, S.; Daly, K.P.; Ganguli, S.; Chitalia, V.C.; et al. U-Net-and-a-half: Convolutional network for biomedical image segmentation using multiple expert-driven annotations. *arXiv* **2021**, arXiv:2108.04658.
14. Pamuk, S. A Case Study: Graduate student's perspective on faculty technology mentoring within the CBAM framework. In Proceedings of the Society for Information Technology & Teacher Education International Conference, Phoenix, AZ, USA, 1–5 March 2005; Association for the Advancement of Computing in Education (AACE): Waynesville, NC USA, 2005.
15. Zhao, H.; Xue, L. Research on Stock Prediction based on LSTM-CNN-CBAM Model. *Comput. Eng. Appl.* **2019**, 57, 203–207.
16. Milanfar, P. A tour of modern image filtering: New insights and methods, both practical and theoretical. *IEEE Signal Process. Mag.* **2013**, 30, 106–128. [CrossRef]
17. Zhang, R. Analysis of precipitation trend and its evolution process in Hong Kong. *J. Hohai Univ. Nat. Sci. Ed.* **2010**, 5, 505–510.
18. Zhou, Z. *Machine Learning*; Tsinghua University Press: Beijing, China, 2019.
19. Dou, H. *Research on Short-Term and Imminent Precipitation Forecasting Based on ConvLSTM Deep Spatiotemporal Network*; Nanjing University of Posts and Telecommunications: Nanjing, China, 2022. [CrossRef]

20. Duchi, J.; Hazan, E.; Singer, Y. Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. *J. Mach. Learn. Res.* **2011**, *12*, 2121–2159.
21. Ruder, S. An overview of gradient descent optimization algorithms. *arXiv* **2016**, arXiv:1609.04747.
22. Yang, G.; Yang, J.; Li, S.; Hu, J. Based on Dropout and ADAM optimizer improve CNN algorithm. *J. Huazhong Univ. Sci. Technol. (Nat. Sci. Ed.)* **2018**, *46–48*, 122–127.
23. Molnar, C. *Interpretable Machine Learning*; Publishing House of Electronics Industry: Beijing, China, 2021.

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