MMST-LSTM: Leveraging Radar Echo Prediction for Emerging Consumer Applications in Edge Computing

Mengjia Wu, Bo Xiao*, Zhiyun Yang, Jiawei Sun, Qi Liu, *Senior Member, IEEE*, Yonghong Zhang, Xiaodong Liu, *Senior Member, IEEE*

***Abstract*—With the increasing frequency of extreme weather events, there is a growing demand from the public for rapid and accurate short-term heavy precipitation forecasts. This study proposes a lightweight deep learning model, MMST-LSTM, which integrates Multiscale Context Feature Fusion Mechanism (MCFFM) and Mixed-Domain Attention Fusion Mechanism (MAFUM). While maintaining high prediction accuracy, MMST-LSTM significantly improves forecast speed. The MMST-LSTM model is particularly suitable for deployment in Mobile Edge Computing (MEC) environments, enabling fast localized forecasting. Experimental results demonstrate MMST-LSTM's excellent predictive performance on two radar echo datasets, particularly in rapid response and handling localized data. Moreover, leveraging Smart Data-Driven Modeling (SDDM) technology with consumer-generated data enhances its application potential in smart consumer electronics products, providing an efficient tool for disaster weather alerts. This study introduces an innovative meteorological forecasting method and provides robust technical support for accurate weather warning systems, offering consumers timely and reliable weather information. This enables them to make more informed decisions, effectively reducing the potential risks and economic losses caused by extreme climate events.**

***Index Terms*—Radar Echo Extrapolation, Smart Data-driven Modeling, Spatiotemporal Sequence Prediction, Deep Learning.**

I. Introduction

In addressing the challenges posed by global climate change, short-term precipitation forecasting plays a crucial role. Its core objective is to predict the intensity and distribution of rainfall within specific regions in the next 0 to 6 hours, necessitating forecast systems with high spatial and temporal resolution [1]. In recent years, weather-related disasters have increased significantly around the world. A typical example is that short-range extreme precipitation can cause serious floods and other secondary disasters [2]. With the increasing frequency of extreme weather events, the public's demand for rapid and accurate near-term forecasts has grown significantly [3]. According to statistics from the World Meteorological Organization (WMO), weather-related natural disaster events have increased fivefold over the past 50 years, with cumulative economic losses exceeding $3.64 trillion. Furthermore, a survey in southern China revealed that more than 85% of the public seeks more accurate short-term precipitation forecasts to better cope with sudden weather changes. Accurate forecasts not only help in timely identification of destructive weather but also provide crucial early warning information for both the public and industries [4]. Moreover, rapid-response short-term forecasts exert a profound positive effect on safeguarding public safety and enhancing the overall quality of life. They provide real-time guidance for daily life arrangements, travel plans, and route planning for autonomous vehicles, effectively preventing life and property losses due to sudden weather changes [5]. Since data processing occurs closer to the data source, edge environments can significantly reduce the time required for data transmission, thereby decreasing both data transmission latency and response latency [6]. In environments requiring rapid and lightweight forecasting models, such capabilities are crucial for fast localized data processing and real-time response, this meets contemporary demands for immediate and personalized meteorological services. Consequently, developing a model for swift and precise short-term heavy precipitation forecasting holds significant scientific value in meteorology and enhances public disaster response capabilities. Due to the complexity of weather patterns and the rapid evolution of storm cloud systems, short-term precipitation forecasting is a task of high demands within the field of meteorology. As extreme weather events become more frequent, the demand for rapid and accurate short-term weather forecasts has become increasingly urgent. This not only imposes higher demands on meteorological research but also underscores the pivotal role of precise early warning systems in safeguarding public safety and protecting property

M. Wu is with the School of Software, Nanjing University of Information Science and Technology, Nanjing, 210044, China (e-mail: 202312210031@nuist.edu.cn).

B. Xiao is with the School of Chemistry and Materials Science, Nanjing University of Information Science and Technology, Nanjing, 210044, China (e-mail: 002479@nuist.edu.cn).

Z. Yang is with the School of Computer Science, Nanjing University of Information Science and Technology, Nanjing, 210044, China and also with Jiangsu Province Engineering Research Center of Advanced Computing and Intelligent Services, Nanjing 210044, China (e-mail: zhiyunyang@nuist.edu.cn).

J. Sun is with the School of Software, Nanjing University of Information Science and Technology, Nanjing, 210044, China (e-mail: 20211221032@nuist.edu.cn).

Q. Liu is with the School of Software, Nanjing University of Information Science and Technology, Nanjing, 210044, China. (e-mail: qi.liu@nuist.edu.cn).

Y. Zhang is with School of Automation, Nanjing University of Information Science and Technology, Nanjing, 210044, China. (e-mail: zyh@nuist.edu.cn).

X. Liu is with the School of Computing, Edinburgh Napier University, Edinburgh, EH10 5DT, U.K. (e-mail: x.liu@napier.ac.uk).

*Both are the correspondence authors of this paper.

from the adverse effects of disasters [7]. In precipitation event forecasting, two main approaches are commonly used: one relies on Numerical weather prediction-based (NWP-based) method, and the other approach utilizes predicting radar echo data into the future based on current observations. The essence of NWP-based methods is a set of nonlinear equations, and the forecast results are obtained by giving initial conditions and boundary values [8]. However, achieving accurate predictions may present significant challenges if the initial conditions are not appropriately set [9]. Furthermore, methods relying on Numerical Weather Prediction often perform poorly when high spatiotemporal resolution is required, and their high computational cost makes them unsuitable for immediate to near-future rainfall prediction [10]. techniques for forecasting radar echo progression predict the distribution of radar echoes at a future time by analyzing historical radar images captured by Doppler weather radar, then, the precipitation estimation results can be obtained according to the Z-R correlation [11]. Traditional radar echo extrapolation methods can be categorized into three types: centroid tracking, correlation analysis, and optical flow methods. The centroid tracking method predicts by tracing changes in the centroid of echo intensity distribution. While it is computationally simple and suitable for small-scale systems, it performs poorly under complex scenarios such as merging or dissipation [12], [13]. The correlation analysis method leverages the correlation of historical data to extract overall features, yet its capability to capture rapid local variations remains limited [14], [15]. The optical flow method calculates motion fields by analyzing the trajectories of feature points in image sequences. It is well-suited for handling complex motion patterns but exhibits limited adaptability to nonlinear changes and incurs high computational costs [16]–[18]. Each of these methods has its own advantages and limitations, however, they exhibit significant constraints in scenarios requiring high spatiotemporal resolution and rapid dynamic adaptability. Their core idea is to perform linear extrapolation by calculating the motion field between adjacent maps, but they have difficulty in obtaining satisfactory results when complex changes such as aggregation and dissipation of radar echoes happen or performing long-term extrapolation tasks [19]. Furthermore, these approaches fail to leverage the abundant historical observational data effectively, so the potential patterns of radar echo variations in specific regions are ignored [20].

Neural networks can automatically extract features, handle nonlinear relationships, and possess strong generalization capabilities, enabling efficient and accurate predictions of complex meteorological data [21] [22]. Recently, artificial neural network (ANN) represented by deep learning has developed rapidly as a new technique in the field of computer science [23], which can recognize potential patterns and model complex nonlinear relationships from the huge amount of existing training data [24]. This approach has attracted the attention of meteorological researchers due to its remarkable performance across a range of tasks, including semantic segmentation [25] and image quality control [26]. Researchers have explored its potential as a solution for radar echo extrapolation and precipitation estimation [27], [28]. At its core, the approach formalizes the radar echo extrapolation problem within the framework of deep learning as a spatiotemporal sequence prediction task, aiming to achieve precise predictions across spatial and temporal dimensions. Consequently, certain models originally designed for video prediction or other spatiotemporal sequence prediction tasks can be effectively applied to radar echo extrapolation tasks after structural optimization. For example, Shi et al. [29] introduced an innovative modification to LSTM by replacing its standard linear operations with convolutional operations. This advancement significantly enhanced the model's accuracy and efficiency in spatiotemporal sequence prediction, outperforming ROVER (a variational method-based real-time radar echo optical flow estimation technique) and fully connected networks. Klein et al. [30] introduced an innovative structure known as dynamic convolutional layers, which has been implemented in the realm of short-term weather prediction. Wang et al. [31] added a spatiotemporal memory cell to the ConvLSTM to constitute the Spatiotemporal-LSTM (LSTM) recurrent unit. This unit are used to build the PredRNN network, which obtain excellent results in several spatiotemporal prediction tasks. Guen et al. [32] integrated principles of partial differential equations (PDEs) with deep learning techniques to create a novel network architecture called PhyDNet, this architecture integrates physics-driven branches with residual learning branches. Compared to previous approaches, these novel methods have achieved significant advancements in enhancing the performance of radar echo extrapolation. However, they often suffer from underestimation of high reflectivity values when predicting radar echo regions. This problem becomes more noticeable as the prediction time frame extends further into the future. It's noteworthy that these high reflectivity echoes are closely associated with events that may lead to severe convective weather [33], [34]. The failure to predict them successfully may bring severe consequences, and therefore the successful prediction of high-intensity echo areas is very critical.

This problem is mainly caused by two factors. First, important contextual features are forgotten when most RNN-based or its variants-based methods perform memory update operations [35]. Although these key features are not entirely disregarded, models still face challenges in capturing and focusing on them, often resulting in underestimation of the intensity of high reflectivity regions in predicted radar echo maps. Moreover, during model training, loss functions such as Mean Squared Error (MSE) and Mean Absolute Error (MAE) are often employed to optimize prediction accuracy. However, this can sometimes result in predictions being biased toward the median or mean values. This bias is particularly evident in cases of highly uneven radar echo intensity distributions, where high reflectivity points constitute a very small proportion, thus causing predicted radar echo intensities to be generally underestimated.

Edge computing optimizes the MMST-LSTM model through localized data processing and storage, significantly

reducing latency and enhancing responsiveness. This makes it particularly suitable for resource-constrained real-time applications such as weather forecasting and intelligent transportation. Traditional centralized computing often fails to meet these demands due to high costs or inefficiencies. The MMST-LSTM model is specifically designed for resource-limited scenarios, leveraging model compression and efficient scheduling techniques to balance computational load and prediction accuracy within constrained resources. It dynamically adapts to real-time data inputs, providing precise and timely services.To address the issue of underestimating radar echo intensity, MMST-LSTM builds upon the ST-LSTM network and integrates the MCFFM and MAFUM modules. The MCFFM extracts multi-scale features and combines them with hidden states to retain critical information, while the MAFUM identifies and emphasizes key features, optimizing the gating mechanism. Additionally, the model employs the BMSE loss function [36], [37] , assigning higher weights to high-intensity echo regions. This resolves the underestimation of critical features inherent in traditional loss functions such as MSE and MAE, enhancing robustness to non-uniform data distributions and significantly improving the prediction accuracy of extreme weather events.

The development and application of edge intelligence are crucial for creating an efficient computing environment [38]. In edge computing (MEC) scenarios, the MMST-LSTM model efficiently processes radar echo data on edge devices, significantly reducing data transmission delays and network loads. The application of such models holds critical significance for achieving real-time environmental perception and prediction, providing robust technical support for fields such as intelligent traffic management, environmental monitoring, and disaster early warning, thereby driving the intelligentization and efficiency of related applications [39]. The proposed MMST-LSTM model integrates multi-scale feature extraction and hybrid-domain attention mechanisms, achieving high-precision environmental data processing and weather forecasting while meeting strict requirements for real-time performance and accuracy. The model excels in handling high-dimensional spatiotemporal data and rapidly changing environments, with significantly enhanced sensitivity to high-reflectivity regions, which greatly improves the prediction accuracy of high-intensity radar echoes, a critical factor for extreme weather forecasting.With its technical advantages, the MMST-LSTM model can provide real-time weather warnings and environmental monitoring support for smart devices in the consumer electronics sector, significantly enhancing the market competitiveness of related products and expanding their application prospects. Overall, the MMST-LSTM model combines advanced multi-scale feature extraction, hybrid-domain attention mechanisms, and the Balanced Mean Squared Error (BMSE) loss function to offer efficient and intelligent data processing solutions for edge computing and consumer electronics, driving innovation and development in the industry.

II. Related Work

Deep learning approaches exhibit superior performance over traditional methods in radar echo analysis, particularly in complex applications such as precipitation nowcasting [40]. This section of the paper will explore and provide a detailed overview of various deep learning-based radar echo progression forecasting methods, while also reviewing approaches used in video prediction and other spatiotemporal sequence forecasting tasks.

A. Deep Models Designed for Radar Echo Extrapolation

Deep learning strategies have fostered the creation of a multitude of model frameworks within the domain of radar echo forecasting, primarily utilizing Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs). Some studies have further integrated adversarial training and attention mechanisms with the aim of enhancing prediction accuracy [41].

Shi et al. [29] introduced the ConvLSTM recurrent unit as part of their proposal and build an encoder-prediction network to perform radar echo extrapolation tasks based on the data observed by the Hong Kong Observatory. Zhuang [42] introduced a groundbreaking model known as the spatio-temporal convolutional neural network (ST-CNN), offering an innovative approach to deeply explore spatiotemporal features in precipitation data for the prediction of extreme precipitation events. In addressing the challenge of position invariance within convolutional recurrent architectures, Shi et al. [36] developed the Trajectory-Gated Recurrent Unit (TrajGRU). This innovative mechanism, introduced to enhance ConvLSTM performance, represents a significant advancement in their research. This TrajGRU unit actively identifies positional differences in natural movements and variations, thereby effectively capturing spatio-temporal relationships in meteorological data. Wang et al. [31] enhanced ConvLSTM by integrating a spatio-temporal memory component, improving its ability to model the spatio-temporal dynamics in radar-captured images for echo detection. Building on this ST-LSTM unit, the PredRNN network further advances this by introducing skip connections across multiple time steps, increasing the network's depth and enhancing its ability to capture long-range dependencies. While ConvLSTM effectively integrates spatial and temporal information, it struggles with long-term dependencies. The addition of skip connections in PredRNN significantly enhances its ability to model complex, rapidly evolving systems, such as radar echo progression. Wang et al. [43] innovatively introduced the Memory in Memory (MIM) component for RNNs, aiming to overcome the limitations of traditional LSTM forget gates. These gates, being overly simplified, struggle to accurately capture complex higher-order non-stationary patterns in radar echo images. Agrawal et al. [44] treated radar echo prediction as a visual representation translation task and employed the U-Net architecture to forecast rainfall across the continental United States. Their methodology drew inspiration from U-Net and SegNet, originally intended for tasks that required dividing data into two distinct segments, Ayzel et al. [45] proposed the RainNet. Its extrapolation results surpass those of the operational nowcasting model provided by the German Weather Service (DWD). Jing et al. [46] developed the MLC-LSTM model, which can deeply

analyze the spatio-temporal relationships between radar echoes at different levels and model their dynamic characteristics. Han et al. [47] presented a 3D-CNN network that converts the radar echo extrapolation problem into the classification of reflectivity intensity. The experiment results demonstrate that its performance is superior to traditional machine learning methods [48]. Trebing et al. [49] introduced the depthwise separable convolution and attention mechanism into U-Net to constitute the SmaAt-UNet. Comparative experiment results show that its predictive performance is comparable to models which are of much larger scale [50]. Geng et al. [51] proposed an extrapolation model named MCCS-LSTM with Context Sensing Block and Multi-Scale Spatiotemporal block (MS block). This model effectively alleviates the blurring problem in long-term extrapolation tasks and enhances prediction accuracy. Yang et al. [52] utilized a self-attention mechanism by incorporating aglobal feature storage component into the ST-LSTM recurrent unit, overcoming the limitations of convolutional kernel size on the model's perceptual range. Huang et al. [53] designed a new Location-Refining (LR) neural network for heavy rainfall areas that are more threatened by human activities, including a localization network and a refinement network. The two networks are used to accurately predict the location of rainfall, and estimating the intensity, respectively. The application of deep learning in radar echo prediction has achieved substantial advancements, particularly within the context of edge computing. Modern approaches, ranging from convolutional neural networks (CNNs) to frameworks integrating attention mechanisms and physics-informed models, have significantly enhanced predictive accuracy and computational efficiency, meeting the stringent demands of real-time operations. The transition of radar echo prediction from theoretical research to practical deployment, particularly in edge computing environments, has enabled rapid, localized, and highly efficient forecasting. These advancements not only improve predictive precision but also enhance the real-time processing capabilities of models, laying a solid foundation for the development of intelligent meteorological early warning systems and driving innovation in extreme weather forecasting.

B. Deep Models for Other Spatiotemporal Sequence Prediction Tasks

Radar echo prediction fundamentally involves the forecasting of spatiotemporal sequences. Models in this domain share similarities with those used in tasks such as video prediction, traffic flow analysis, and typhoon path prediction, and are equally applicable to radar echo prediction tasks.

Srivastava et al. [54] used a network constructed by LSTM to learn the representation of video sequences for multi-frame video prediction. Wang et al. [55] introduced a Causal LSTM, which incorporates a Gradient Highway Unit (GHU) and a unique dual-cascade gating structure. GHU alleviates the vanishing gradient problem inherent in traditional recurrent neural networks by introducing new gradient propagation paths. This enables gradients to flow along shorter paths, enhancing the model's ability to capture both short- and long-term dependencies and improving the accuracy of short-term precipitation forecasting. These two structures are usually used to build the PredRNN++ network, which can further alleviate the vanishing gradient problem that often occurs in spatiotemporal sequence prediction tasks. Byeon et al. [56] crafted a comprehensively context-aware model framework that employs Parallel Multi-Dimensional LSTM units to seize historical context data for each pixel, and integrates these with hybrid units to enhance model clarity in video prediction tasks. In order to tackle the issue that earlier models had with grasping both immediate temporal linkages and extended, complex interdependencies within data sequences, Wang et al. [57] proposed the Eidetic 3D LSTM (E3D-LSTM), which employs 3D convolution to enhance the capture of short-term features by extracting information across temporal, spatial, and channel dimensions. Unlike traditional 2D convolution, 3D convolution simultaneously processes the dynamic variations of spatiotemporal data, thereby improving the predictive capability for the short-term evolution of meteorological events such as precipitation. Moreover, by employing an attention mechanism, the model facilitates the interaction between the historical state and the current memory state. This approach enables more effective integration and utilization of past information. Fan et al. [58] proposed a new video prediction unit (CubicLSTM) to process temporal and spatial states separately by two independent convolutional kernels to obtain better prediction performance. Proposing a novel approach, Lin et al. [59] developed the Self-Attention Memory (SAM) module. This component is meticulously designed to adeptly capture enduring interdependencies across spatial and temporal dimensions. The SAM component is embedded into a standard ConvLSTM unit to form the SA-ConvLSTM, which can be applied to human action prediction and traffic flow forecasting. Inspired by the idea of Partial Differential Equations (PDEs) describing physical knowledge, Guen et al. [32] designed a dual-branch network PhyDNet for unsupervised video prediction. The network has a physical branch composed of PhyCell and a residual branch composed of ConvLSTM units. This architecture enables PhyDNet to disentangle the dynamic prior knowledge inherent in video content from other variables required for video prediction, thereby achieving more accurate modeling capabilitie. Su et al. [60] proposed a higher-order convolutional LSTM model (Conv-TT-LSTM) to learn long-term spatiotemporal correlations from video frames in response to the poor performance of previous methods for tasks involving long-term forecasting in video sequences. Chai et al. [61] innovatively proposed the CE block focusing on contextual interactions and the SE block for multi-scale spatiotemporal representation. The two blocks are used to constitute the CMS-LSTM to tackle the challenge posed by the inherent uncertainty in predicting using continuous frames. Wu et al. [62] proposed a Motion RNN network, which is able to model complex changes of motion, while avoiding the problem of motion disappearance in stacked multi-layer models.

III. METHODOLOGY

To thoroughly investigate the architecture and operational mechanisms of the proposed components or models, this section covers the Multi-Source Context Feature Fusion Module

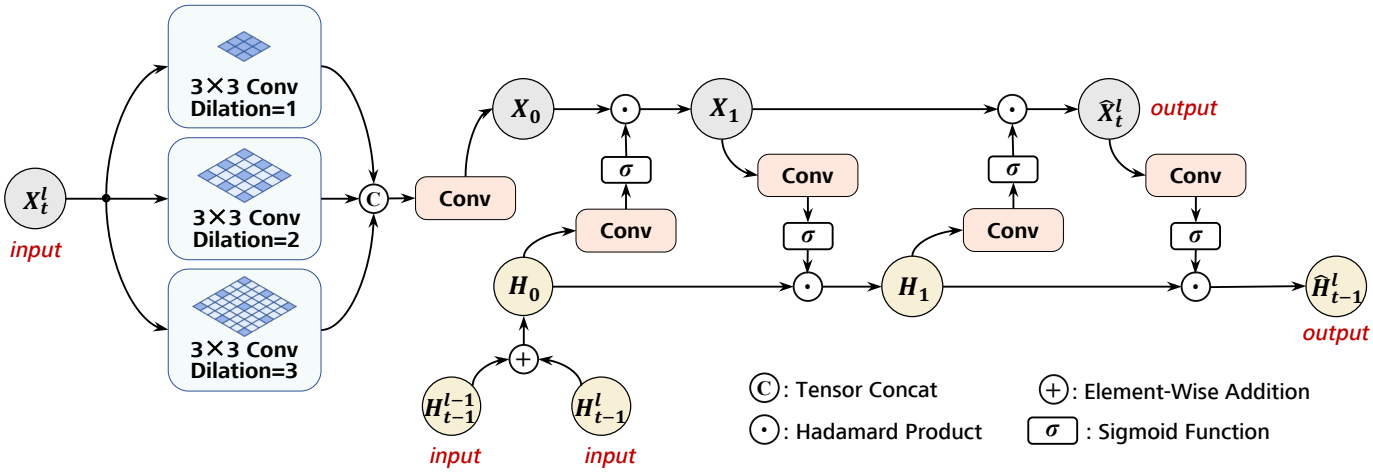

Fig. 1. The structure of Multi-scales Contextual Feature Fusion Module (MCFFM). The left part shows the extraction operation of multi-source contextual features, while the lower part is the obtained multi-order features. The right part shows the process of the fusion of these two types of features.

(MCFFM), the Feature Update Module incorporating Mixed-Domain Attention (MAFUM), the MMST-LSTM recurrent unit, and its corresponding data extrapolation network.

A. Module for fusing contextual features from multiple sources

To address the issue of potential neglect of high-intensity radar echo features due to insufficient contextual connections across adjacent frames in the gating process, this study introduces an innovative Multi-Source Context Feature Fusion Module (MCFFM). This module is positioned at the frontend of the recurrent unit's gating mechanism. Its structure is depicted in Fig. 1, where l denotes the l^{th} level, and t denotes the t^{th} time step. X is the input state, and H is the hidden state. The data fed into this module includes multi-scale features obtained from X, and multi-order features obtained from H. By introducing a multiscale feature extraction mechanism, this module integrates multilevel information from both input and hidden states to capture the spatiotemporal characteristics of high-intensity echoes at different scales. The module is designed to address the issue of significant feature loss caused by increased depth in deep neural networks. Unlike existing single-scale feature extraction methods, the Multiscale Context Feature Fusion Module (MCFFM) captures echo region details at varying resolutions, preventing the omission of critical information during the complex evolution of echoes. Experimental results demonstrate that MCFFM effectively enhances the prediction accuracy for high-intensity echo regions, particularly under complex weather patterns.

The operational process of the module begins with the use of three dilated convolution kernels, similar to the structure of Atrous Spatial Pyramid Pooling (ASPP), to extract and learn features of radar echo changes from the new input state X_t^l [63]–[65]. The size of these three dilated convolutional kernels is 3 × 3, and the dilation rates are specified as 1, 2, and 3, respectively. Therefore, the size of their receptive fields is 3 × 3, 5 × 5 and 7 × 7, respectively. This can prevent the holistic motion features of high-intensity radar echo areas from being missed due to the size limitation of a lone convolutional filter. The results from the dilated convolution operation, consisting of three separate feature maps, they are combined along the channel axis. Subsequently, the channel count is modified using a 1 × 1 convolutional kernel. This step can be represented by the equation given in Eq.(1). The convolutional kernel $w_{1\times1}$ is employed to modify the channel count, while w_{Dk} denotes the 3 × 3 convolutional kernel with a dilation rate k. The selection of dilation rates is crucial for multiscale feature capture. Smaller dilation rates (e.g., 1 or 2) focus on local features, suitable for high-resolution and short-time-scale radar echo data. Larger dilation rates (e.g., 4 or 8) expand the receptive field, capturing long-range dependencies and broader contextual information, thus improving global forecast accuracy. This approach enhances the model's ability to learn long-term dependencies without significantly increasing computational complexity, effectively balancing local details and global context to improve predictions for complex temporal data.

$$X_0 = w_{1\times1} * \left(\left[w_{D1} * X_t^l, w_{D2} * X_t^l, w_{D3} * X_t^l\right]\right) \quad (1)$$

Features of multiple orders are derived by summing the previous hidden state H_{t-1}^l with the state H_{t-1}^{l-1}. Previous models' recurrent units typically only captured features from H_{t-1}^l, while H_{t-1}^{l-1} was only accessible across different points in time and levels of the network. The proposed unit can directly access state H_{t-1}^{l-1}, thereby preventing the accidental omission of key feature information as with the expansion of the network's layers or temporal progression. The feature information from various levels is held in H_0. The information can be alternatively expressed as: the unit directly acquires the previous moment's state information, avoiding potential loss of critical feature information in deep networks or long time series, and accumulates these multi-level feature details in the initial state vector.

$$H_0 = H_{t-1}^{l-1} + H_{t-1}^l \quad (2)$$

Next, we integrate the input state X_0 that encompasses different scales or resolutions with the hidden state H_0 that

captures information across various orders or layers of complexity through a feature fusion operation. A convolutional kernel w_{hx} of size 5 × 5 is used for feature extraction from H_0, and the Sigmoid function is then utilized to scale the values within the feature map to a range between 0 and 1. The 5 × 5 convolution kernel w_{hx} in image processing strikes an optimal balance between computational efficiency and feature extraction capability. Compared to the 3 × 3 convolution kernel w_{hx}, the 5 × 5 kernel w_{hx} offers a broader receptive field, enabling it to capture more intricate local details and complex spatial patterns. The resulting map undergoes a Hadamard product operation with X_0. The outcome is denoted as X_1. Subsequently, we apply a second 5 × 5 convolutional kernel w_{xh} to extract features from X_1, and then normalizing using the function. The resulting map and H_0 undergo element-wise multiplication. The outcome is denoted as H_1. By employing this method, the early phases strengthen the contextual links between the input and hidden states. This process is illustrated as:

$$X_1 = \sigma\left(W_{hx} * H_0\right) \odot X_0 \tag{3}$$

$$H_1 = \sigma\left(W_{xh} * X_0\right) \odot H_0 \tag{4}$$

Lastly, the above procedures are iterated to compute the updated input state $\hat{X}_t^l$ and hidden state $\hat{H}_t^l$ using Equations Eq.(5) and Eq.(6). These two states are used as the input of the gating mechanism operation. The important features between them have been fused with each other. These two states together form the foundational input for the gating mechanism. Their key features have been mutually integrated, enhancing the contextual connection between states. This integration ensures that features linked to high-intensity echoes are comprehensively taken into account during subsequent gating processes, thus preventing any possible oversight.

$$\hat{X}_t^l = \sigma\left(W_{hx} * H_1\right) \odot X_1 \tag{5}$$

$$\hat{H}_t^l = \sigma\left(W_{xh} * \hat{X}_t^l\right) \odot H_1 \tag{6}$$

The integration of the 5 × 5 convolution kernel and multiscale feature extraction modules enhances the model's representation and feature-capturing capabilities, albeit with increased computational complexity. Optimized resource scheduling and model compression mitigate these demands in edge computing, ensuring real-time performance [66]. The model excels in prediction accuracy and responsiveness, particularly in disaster warning and intelligent transportation applications.

B. Feature Update Module with Mixed-Domain Attention

Traditional models often exhibit limitations in effectively capturing and modeling the dynamic variations of features closely associated with high-intensity radar echoes [67]. Inspired by advancements in image segmentation and object detection fields [68], [69], to overcome this limitation, the recurrent unit incorporates a novel Feature Update Module guided by Mixed-domain Attention (MAFUM). Fig. 2 illustrates its structure. This module takes as input the hidden state $H_t^{\prime l}$, the spatiotemporal memory cell unit $M_t^{\prime l}$ obtained by the gating mechanism operation of the current unit. The element-wise addition of these is recorded as F_{raw}, which is then forwarded to the subsequent operation of the mixed-domain attention mechanism. Formulated as follows, this step involves:

$$F_{\text{raw}} = H_t^{\prime l} + M_t^l \tag{7}$$

Subsequently, a mechanism designed to enhance the model's focus across multiple data domains is applied to F_{raw}, incorporating a mixed-domain attention mechanism. This mechanism is specifically aimed at identifying key regions within the input feature maps that are highly correlated with high-intensity echoes. By combining spatial attention and channel attention, the module addresses the uneven distribution of echo intensities, thereby augmenting the model's capacity to emphasize critical regions. The hybrid attention system integrates channel attention, which selectively enhances certain feature channels, with spatial attention, which prioritizes specific spatial locations within the feature maps. These operations are performed in a sequential manner. In contrast to conventional single-domain attention mechanisms, MAFUM facilitates the refined processing of features across multiple dimensions, thereby significantly enhancing the model's performance, particularly in the prediction of complex meteorological phenomena. The input F_{raw} undergoes channel attention processing. Both average pooling and max pooling operations are concurrently conducted on it to gather spatial information from the feature maps. This enables comprehensive spatial coverage and improves the model's capacity to detect a variety of spatial patterns present in the input data. The results from these two pooling operations, once excited and squeezed, undergo an element-wise addition after being processed by a shared Multilayer Perceptron (MLP) structure. The obtained result by addition operation is applied Sigmoid function to get the heat map, and then is multiplied element-wise with the original input F_{raw}. The product result is the feature map F_C, which required for subsequent spatial attention operation. By incorporating the channel attention mechanism, the model is able to efficiently identify and prioritize key features within the input feature maps that are critical for the accurate prediction of high-intensity radar echoes. This process is depicted as follows:

$$M_c = \sigma\left(\text{MLP}\left(\text{AvgPool}\left(F_{\text{raw}}\right)\right) + \text{MLP}\left(\text{MaxPool}\left(F_{\text{raw}}\right)\right)\right) \tag{8}$$

$$F_c = M_c \odot F_{\text{raw}} \tag{9}$$

On this basis, a spatial attention mechanism is applied to F_C. This process prioritizes the input features by independently performing average pooling and maximum pooling across data points, followed by extracting the maximum values within the channel dimension of the visual data arrays obtained through convolution operations. Subsequently, these results are merged across the channel dimension. The concatenated result undergoes consecutive convolution operations followed by the Sigmoid function to produce a heatmap that focuses on key areas in the image. M_S is element-wise multiplied (Hadamard product) with the channel attention result F_C, yielding a grid

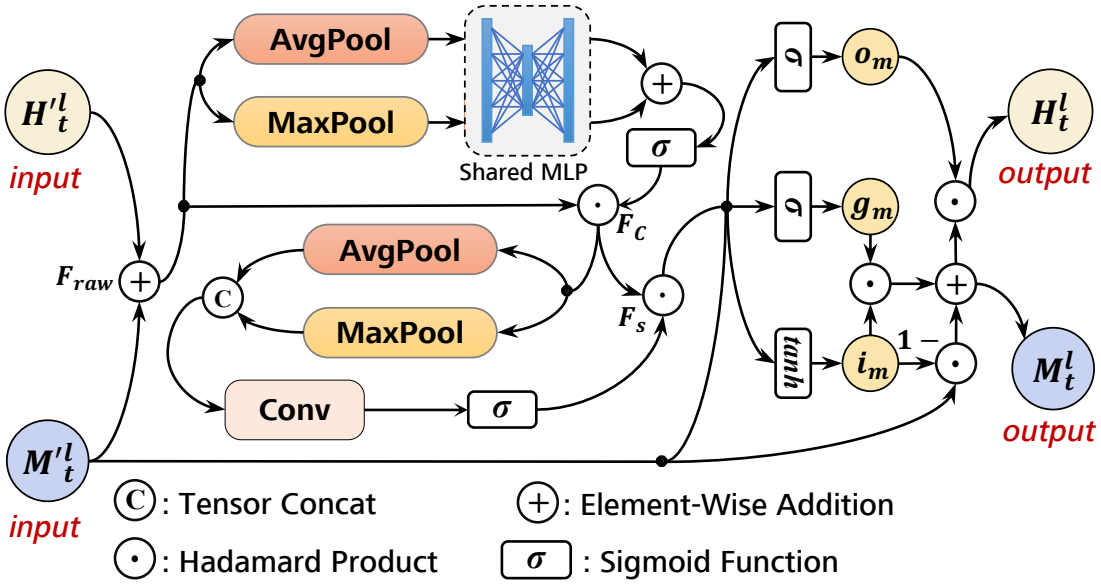

Fig. 2. The computation process within the Multi-Domain Attention-guided Feature Enhancement Unit (MAFUM). The left section illustrates the functioning of the mixed-domain attention mechanism, which includes both channel-wise and spatial focus, allowing the model to efficiently concentrate on and learn features crucial for high-intensity echoes, while the right part is the gating structure to update states and cells.

of detected features weighted by a mechanism that combines different types of attention, such as spatial and channel, to enhance feature processing. This series of operations can be represented by the following equations:

$$M_s = \sigma\left(W_s * \left[\text{AvgPool}\left(F_c\right), \text{MaxPool}\left(F_c\right)\right]\right) \tag{10}$$

$$F_s = M_s \odot F_c \tag{11}$$

Finally, the process of updating the hidden states and spatio-temporal memory units mimics a control system for feature flow found in GRU, as shown on the diagram's right-hand side. The calculations for the gating mechanism for input and input modulation follow Eq.(12) and Eq.(13).

$$i_m = \sigma\left(W_{m;fi} * F_s + W_{s;mi} * M_t^{\prime l} + b_{m;i}\right) \tag{12}$$

$$g_m = \sigma\left(W_{m;fg} * F_s + W_{s;mg} * M_t^{\prime l} + b_{m;ig}\right) \tag{13}$$

The obtained i_m and g_m are used to update the spatiotemporal memory cell:

$$M_t^l = i_m \odot g_m + (1 - i_m) \odot M_t^{\prime l} \tag{14}$$

Similar to the other two gating processes, the output gate o_m is calculated. The final hidden state is produced by taking the result of the previous steps and applying the output gate through an element-wise multiplication, which selectively updates the state based on the relevance of the information.

$$o_m = \sigma\left(W_{m;fo} * F_s + W_{s;mo} * M_t^l + b_{m;io}\right) \tag{15}$$

$$H_t^l = o_m \odot H_t^{\prime l} \tag{16}$$

After completing the above steps, the hidden states and spatio-temporal memory units will be updated. This allows the critical features that are closely associated with high-intensity radar echoes to be effectively propagated across different time steps and layers of the extrapolation network, thereby enhancing the accuracy of predicting high-intensity radar echoes. MMST-LSTM effectively addresses the limitations of short-term and long-term dependency modeling through multi-scale feature fusion and mixed-domain attention mechanisms. While traditional LSTM performs well in modeling short-term dependencies, it struggles to capture long-term dependencies [70]. By introducing an attention mechanism, MMST-LSTM balances the modeling of short- and long-term dependencies through weighted inputs across different time steps. Compared to traditional LSTM and GRU, the attention mechanism significantly enhances the model's performance in long-sequence tasks, demonstrating superior adaptability and accuracy, particularly in complex temporal problems such as extreme weather prediction.

C. Multi-Modal Spatio-Temporal LSTM Unit with its Projection Network

By leveraging the MCFFM and MA-FUM modules, the capabilities of the ST-LSTM unit in handling complex spatiotemporal data are significantly enhanced, particularly for radar echo intensity prediction, culminating in the proposed MMST-LSTM recurrent unit. Fig. 3 illustrates the architecture of a single MMST-LSTM recurrent unit. The unit's input comprises the input state X_t^l, the memory cell C_{t-1}^l from the previous time step, the unit's input encompasses the memory cell that amalgamates spatial and temporal data, represented as M_{t-1}^l, the encoded state H_{t-1}^{l-1} and H_{t-1}^l. It's important to recognize that for units positioned in the first layer of the network, their input is a tensor derived directly from the current radar echo measurements. In other network layers, the unit initiates its processing with the hidden state H_t^{l-1} inherited from the layer above. In other words, if the unit is positioned in the highest stratum of the network, its input is derived from the processed radar echo data. If the unit is in

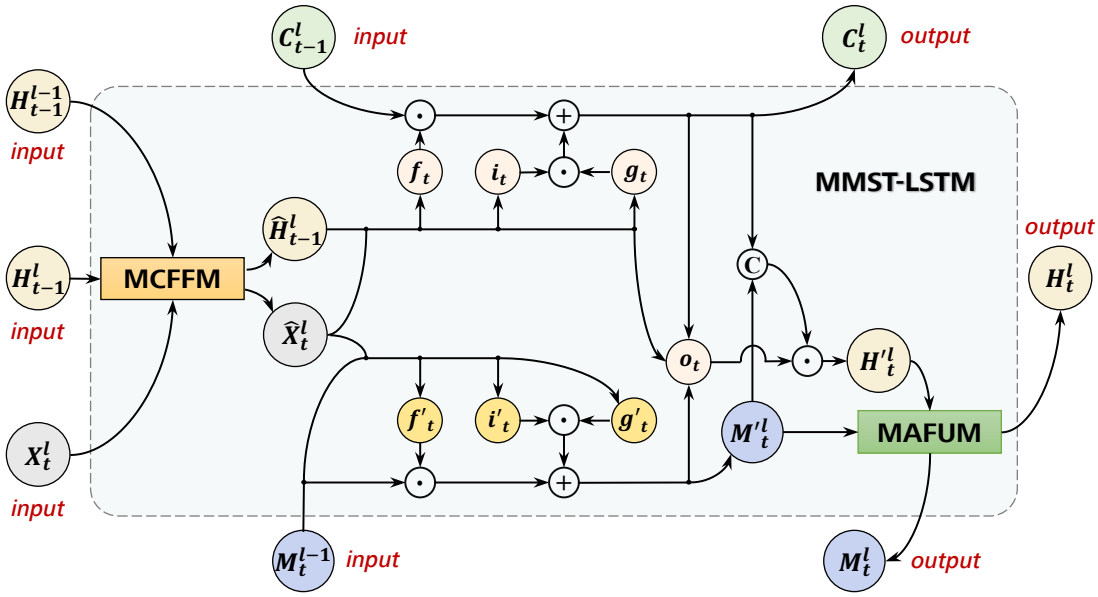

Fig. 3. The structure of a single MMST-LSTM recurrent unit. The MCFFM and MAFUM are the two components proposed in this paper. The upper part and lower part are the two parallel gating structure to update the memory cells, which are same as those in original ST-LSTM unit.

the middle or bottom layers, its input is based on the output state of the unit from the preceding layer.

The operational mechanism of the MMST-LSTM can be described in the following manne. First, the previous hidden statesH_{t-1}^{l-1}, the current hidden states H_{t-1}^{l} and the input at time t, X_t^l are all fed into the MCFFM for feature consolidation, this step aims to reinforce the connection between the current context and the most recent data input. The fusion result is $\hat{H}_{t-1}^{l}$ and $\hat{X}_t^l$. This step can be formulated as:

$$\hat{X}_t^l, \hat{H}_{t-1}^l = \text{MCFFM}\left(X_t^l, H_{t-1}^{l-1}, H_{t-1}^l\right) \tag{17}$$

Then, performing gated mechanism operation on $\hat{H}_{t-1}^{l}$, $\hat{X}_t^l$ and two memory cells. The calculation process is same as that in the original ST-LSTM cell. Two parallel gating mechanisms are performed to update the memory cells. These two gating structures each possess their own forget gate, data flow control mechanism, and input influence modulation gate, while the filter gate is shared. In the upper part of the recurrent unit, the gate that decides what to discard from memory, the gate for new information intake and the gate for adjusting input influence are obtained by four different convolution kernels from the input state and the gate for retaining and updating memory content, according to Eq.(18) to Eq.(21):

$$f_t = \sigma\left(W_{xf} * \hat{X}_t^l + W_{hf} * \hat{H}_{t-1}^l + b_f\right) \tag{18}$$

$$i_t = \sigma\left(W_{xi} * \hat{X}_t^l + W_{hi} * \hat{H}_{t-1}^l + b_i\right) \tag{19}$$

$$g_t = \tanh\left(W_{xg} * \hat{X}_t^l + W_{hg} * \hat{H}_{t-1}^l + b_g\right) \tag{20}$$

$$o_t = \sigma\left(W_{xo} * \hat{X}_t^l + W_{ho} * \hat{H}_{t-1}^l + b_o\right) \tag{21}$$

Upon successfully deriving the aforementioned four control gates, the unit utilizes mechanisms from the original Long Short-Term Memory (LSTM) network to systematically filter and discard obsolete features within the memory cell C that have become irrelevant over temporal progression. Meanwhile, through the gate that filters new inputs into the memory cell i_t and the gate that adjusts the influence of incoming data g_t, the unit can identify and absorb key features from the current state of data being processed. In other words, the unit uses these control gates to optimize its memory content, discarding outdated or irrelevant information and focusing on valuable features in the current input [71].

$$C_t^l = f_t \odot C_{t-1}^l + i_t \odot g_t \tag{22}$$

Consistent with the upper-layer structure of the unit, the spatio-temporal memory cell M and the input state X are also refreshed via an additional set of regulatory gates. This means that, similar to the operation of the upper-layer units, the memory cell and input state are updated by gating mechanisms that align them with current data trends. This group of gates computes by filtering and adjusting information flow is similar to the calculation method of f_t, i_t and g_t. Another group of convolutional kernels W_f', W_i' and W_g' are applied to $\hat{X}_t^l$ and $\hat{M}_t^{l-1}$ to get them, according to Eq.(23) to Eq.(25):

$$f_t' = \sigma\left(W_{xf}' * \hat{X}_t^l + W_{hf}' * \hat{H}_{t-1}^l + b_f\right) \tag{23}$$

$$i_t' = \sigma\left(W_{xi}' * \hat{X}_t^l + W_{hi}' * \hat{H}_{t-1}^l + b_i\right) \tag{24}$$

$$g_t' = \sigma\left(W_{xg}' * \hat{X}_t^l + W_{hg}' * \hat{H}_{t-1}^l + b_g\right) \tag{25}$$

The memory cell M is refreshed with key spatiotemporal features, following a similar approach to how memory cell C is updated:

$$M_{t-1}'^l = f_t' \odot C_{t-1}'^l + i_t' \odot g_t' \tag{26}$$

Subsequently, the feature representations of the two updated memory cells are concatenated along the channel axis. The output gate o_t is then employed for an element-wise multiplication, known as the Hadamard product on the concatenated

result to obtain the hidden state $H_t'^l$. This step can be encapsulated with the following equation:

$$H_t'^l = o_t \odot \left[C_t^l, M_{t-1}'^l\right] \tag{27}$$

Finally, the intermediate hidden state $H_t'^l$ and the modified memory $\hat{M}_t^{l-1}$ are directed to the MAFUM, which directs the model's focus to features closely associated with high-intensity radar echoes. Following this, the hidden states and spatiotemporal memory units are reconfigured and propagated to the subsequent recurrent unit, thereby enhancing the model's precision in predicting the trajectories of intense radar echoes. This can be formally represented with a mathematical formula:

$$H_t^l, M_{t-1}^l = \text{MAFUM}\left(\hat{X}_t^l, \hat{H}_{t-1}^l\right) \tag{28}$$

Following the settings in some previous works [10], [31], [52], [55], [59], [72], the proposed MMST-LSTM units are arranged in a stack to construct a network with four layers, and its architecture is shown in Fig. 4.

Each vertical stripe in the figure represents the sequence of recurrent units present within each layer at a given time step, while each row depicts the temporal flow of memory cells and states within the current layer. At the first time step, all memory cells and hidden states of the initial unit are initialized with tensors comprising elements set to zero. For units in the l^{th} layer at time step t, the updated temporal memory cell C is passed laterally to the unit within the same layer at the subsequent time step, the revitalized spatiotemporal memory unit M is conveyed vertically to the unit in the layer immediately following, at the present time step. The memory unit propagates to the recurrent unit of the subsequent layer. For top-layer units, the memory unit M will propagate to the first-layer unit at the subsequent time step. M is conveyed simultaneously across spatial and temporal domains, thereby incorporating spatiotemporal characteristics. For the output hidden states, the proposed extrapolation network follows the original structure of ConvLSTM and other LSTM-based units, they are disseminated in both the horizontal and vertical dimensions. It is important to note that for unit which are not at the top layer, H is also transmitted to the unit at layer $l+1$ at moment $t+1$ as its input along the route of the red arrow in the figure. In this manner, the recurrent unit acquires multi-order features, enhancing contextual correlation and preventing the omission of features closely associated with high-intensity echoes.

IV. Experiments

This section will cover the dataset used for experiments, parameter configurations during the training phase, and evaluation criteria. Additionally, it will present and deeply analyze the quantitative results and visual outputs of the experiments.

A. Datasets

This study uses two real-world radar echo datasets to evaluate the proposed model's effectiveness and validate its new features. The experimental findings from these datasets underscore the model's substantial applicability in advancing intelligent applications, enabling secure and efficient data exchange processes [73].

The first dataset, referred to as the GZ dataset, comprises radar echo images captured by a Doppler weather radar in Guangzhou, China, over four consecutive rainy seasons. The data is partitioned into training, validation, and testing sets to

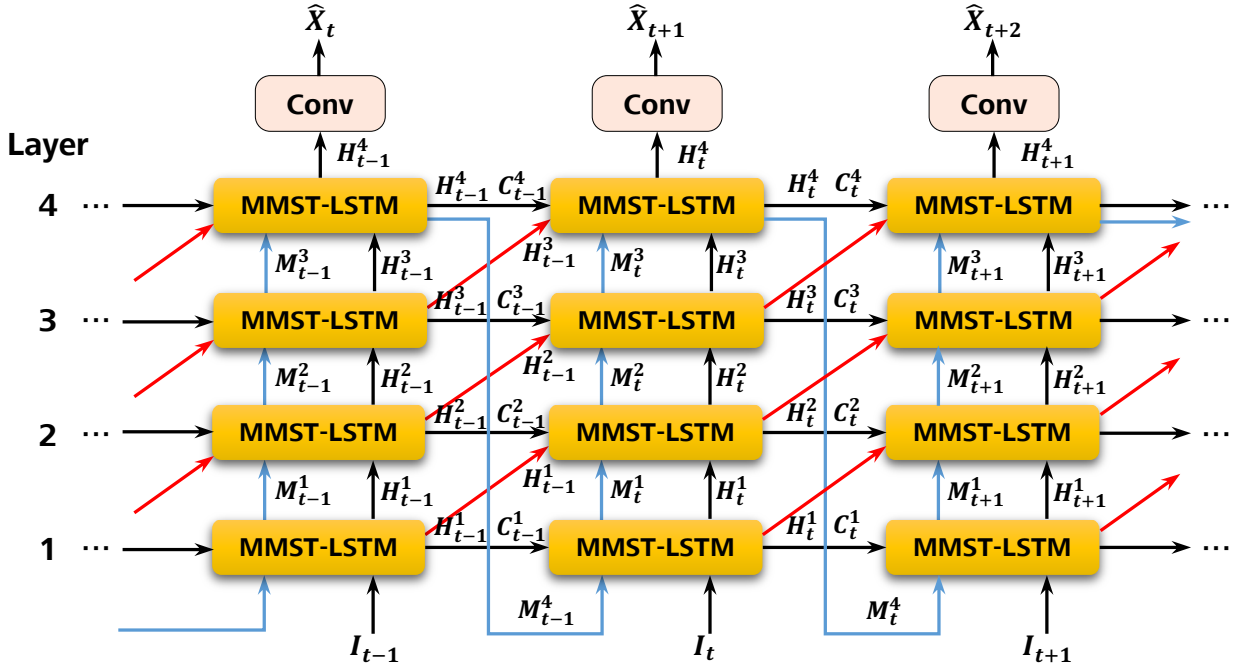

Fig. 4. The architecture of the four-layer network for radar echo extrapolation built by MMST-LSTM units. The blue arrows show the transmission route of spatiotemporal memory M, while the red arrows represent the transmission route of the hidden state H from the recurrent unit at the previous unit at the previous time step.

ensure the independence of each segment. The sequences in these three datasets consist of 5773, 670, the dataset is split into distinct groups for training and validation, each containing 800 data points, while the test group comprises a separate collection of instances, and the test set comprising 1000 instances, respectively. Each sample consists of a sequence of 20 composite reflectivity (CR) images captured by the meteorological radar within a 120-minute interval. In the experiments, each model predicts the subsequent 10 images based on the first 10 radar reflectivity maps in each sequence. Each map in its original form measures 500 $\times$ 500 pixels, with a spatial resolution of 1 kilometer per pixel. To alleviate computational demands, the maps in the GZ dataset were scaled down to 100 $\times$ 100 pixels.

The second dataset chosen is sourced from the 2017 CIKM AnalytiCup contest, which provides publicly available radar echo data. The dataset consists of 8,000 and 2,000 samples, with the training set comprising 4,000 sequences, along with validation and testing sets. Each group of data includes 15 radar echo images, captured at 6-minute intervals, with an observation period of 90 minutes per sequence. In these experiments, aevery model is tasked with forecasting the final 10 images of each sequence using the initial 5 radar echo images as a reference. Each radar echo map has dimensions of 101 $\times$ 101 pixels, corresponding to a coverage area of 101 by 101 kilometers in Shenzhen, China. To ease the model training process, the dataset's maps are partitioned into smaller sections, and the peripheral areas of each section are supplemented with zeros to maintain structure. Consequently, the dimensions of the newly acquired radar echo image are 104 $\times$ 104.

To avoid the influence of non-precipitation data on the extrapolated maps, points with intensity values below 10 dBZ or above 70 dBZ were replaced with 0. This indicates no data points in the two datasets [72], [74]. The conversion between intensity value and pixel value are implemented by Eq.(29):

$$\text{Pixel} = \left\lfloor 255 \times \left(\frac{\text{dBZ}}{70} \right) + 0.5 \right\rfloor \tag{29}$$

B. Experimental Setup

To evaluate the performance of the proposed model in radar echo prediction, a series of comparative experiments were conducted. Several representative and innovative deep learning models, published in top-tier computer science conferences, were selected as benchmarks. The selection of these models is attributed to their profound influence within both academic and industrial domains [75]. By comparing our model against these benchmarks, we aim to highlight the improvements and optimizations in our model's predictive performance, including ConvLSTM [29], PredRNN [31], PredRNN++ [55], Memory in Memory (MIM) [43], PhyDNet [32] and SA-ConvLSTM [59]. All the aforementioned models are arranged in a four-layer stack to form the radar echo prediction networks. The number of feature maps is 64, and each layer of them has a convolution kernel size of 5 $\times$ 5 and each layer of these networks contains 64 feature maps. All models are trained on Nvidia RTX 3080Ti with the optimizer Adam. Scheduled sampling [76], early stopping training [77] and layer normalization [78] strategies were used during the training process.

To address the issue of radar echo intensity being biased towards the mean or median when using Mean Squared Error (MSE) or Mean Absolute Error (MAE) as the loss function, this study adopts the Balanced Mean Squared Error (BMSE), as formulated in Eq.(30).

The dimensions of the radar echo map are represented by H for height and W for width. $X_{t+1:t+k}$ is the k observed maps, and $\hat{X}_{t+1:t+k}$ is the corresponding extrapolated maps. The coordinates of the points in the maps are represented as (i, j). The variable $w(i, j)$ represents the weight corresponding to the radar echo intensity at the coordinate (i, j). By judging the intensity range of each point, BMSE multiplies the corresponding weights with the original intensity value, the features of high-intensity echo areas that account for a small proportion can be amplified, and thus be paid attention to by the models. The weighting function $w(i, j)$ is formulated in the following manner:

$$\omega(i,j) = \begin{cases} 1, & X(i,j) \leq 20 \\ 2, & 20 \leq X(i,j) < 30 \\ 5, & 30 \leq X(i,j) < 40 \\ 10, & 40 \leq X(i,j) \end{cases} \tag{31}$$

C. Evaluation Metrics

When assessing the model's efficacy in forecasting radar echo images, this study introduced two widely recognized metrics from the field of computer vision: Mean Square Error (MSE) and Structural Similarity Index (SSIM). MSE, a numerical measure assesses the variance between the predicted radar imagery and the genuine observations, with smaller values indicating higher prediction accuracy. SSIM assesses the structural similarity of the images, enhancing MSE with a broader assessment of the model's predictive accuracy. SSIM evaluates the structural similarity between predicted and observed images, with higher SSIM values indicating greater similarity between the two images in terms of structure. These two metrics are calculated according to Eq.(32) and Eq.(33), where μ_x and μ_y are the mean values of the extrapolated map and observed map, respectively. The σ_x and σ_y are their

$$\text{BMSE} = \frac{\sum_{i=1}^{H} \sum_{j=1}^{W} w(i,j) \times \left(\hat{X}_{t+1:t+k}(i,j) - X_{t+1:t+k}(i,j) \right)^2}{H \times W} \tag{30}$$

variances, respectively. C_1 and C_2 are positive constants.

$$\mathrm{MSE} = \frac{\sum_{i=1}^{H}\sum_{j=1}^{W}\left(\hat{X}_{t+1:t+k}(i,j) - X_{t+1:t+k}(i,j)\right)^2}{H * W} \tag{32}$$

$$\mathrm{SSIM} = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{33}$$

Furthermore, in alignment with prior studies [9], [36], [79], [80], the Critical Success Index (CSI) and Heidke Skill Score (HSS), both widely recognized as standard evaluation metrics in meteorology, have been incorporated as supplementary assessment criteria. CSI measures the proportion of correctly predicted precipitation events among all predicted and observed precipitation events. It reflects the model's ability to capture precipitation events, which is especially important in forecasting extreme weather. HSS evaluates the skill of the model's predictions by comparing them to random predictions, removing the influence of accidental correct predictions. These two metrics effectively assess the model's performance in high-intensity precipitation and extreme weather events. Higher CSI and HSS scores indicate better predictive accuracy in high-reflectivity regions. The calculations for these two metrics are conducted through the subsequent procedures. Given a threshold τ of intensity, the Observation matrix O correspond to the observed map, and Prediction matrix P correspond to the extrapolated map. These two matrices are binary matrices. For the point pair with coordinates (i, j) in the real observed map X and the extrapolated map $\hat{X}$, comparing the value of the two points with the threshold τ. If the value of $X(i, j)$ is greater than τ, then the value of $O(i, j)$ is set to 1. If it is less than τ, then the value of $O(i, j)$ is set to 0. Similarly, the value of each point in the Prediction matrix can be obtained. The classification of the point pair is then determined based on the categorization methodology outlined in Table I, these metrics are based on four key prediction outcomes: correct identification of events (TP), correct identification of no events (TN), incorrect prediction of events (FP), and missed events (FN).

TABLE I
THE RULES FOR DETERMINING THE CATEGORY OF THE POINT PAIR $O(i,j)$AND $P(i,j)$.

The value of O(i,j)	The value of P(i,j)	Category
1	1	True Positive (TP)
0	0	True Negative (TN)
0	1	False Positive (FP)
1	0	False Negative (FN)

In conclusion, tally the count for each type of outcome pair. The CSI and HSS are subsequently calculated using the equations provided in Eq.(34) and Eq.(35):

$$\mathrm{CSI} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP} + \mathrm{FN}} \tag{34}$$

$$\mathrm{HSS} = \frac{2 \times (\mathrm{TP} \times \mathrm{TN} - \mathrm{FN} \times \mathrm{FP})}{(\mathrm{TP} + \mathrm{FN}) \times (\mathrm{FN} + \mathrm{TN}) + (\mathrm{TP} + \mathrm{FP}) \times (\mathrm{FP} + \mathrm{TN})} \tag{35}$$

D. Comparative Experiment

Table II presents the numerical results of all models on the GZ test set across the four selected evaluation metrics. The data demonstrate that the extrapolation network constructed with the proposed MMST-LSTM unit achieves the highest scores across all evaluation criteria. The second-ranked is the extrapolation network built by SAConvLSTM units, which precisely captures the progression of radar echoes, aided by the Spatial Attention Module (SAM), yielding superior outcomes. However, due to the lack of structures that can help it pay attention to the relevant features of high-intensity echoes, its CSI and HSS scores are lower than the proposed MMST-LSTM when the threshold τ is set to 40 dBZ. Compared with SAConvLSTM, the Critical Success Index (CSI) scores for the MMST-LSTM model at three thresholds are improved by 2.48%, 1.84% and 15.00%, respectively. The average CSI score improved by 4.68%. The Heidke Skill Score (HSS) increased by, 1.49% and 11.47% at the three thresholds, respectively. The average HSS was improved by 4.26%. Its MSE score decreased by 6.72% and the Structural Similarity Index Mean (SSIM) increased by 8.08%. Compared to the baseline model ST-LSTM, MMST-LSTM reduces the MSE score by 17.07% and improves the SSIM score by 14.02%. When the threshold is set to 20 dBZ, CSI and HSS improve by 5.05% and 4.82%, respectively; when the threshold is set to 30 dBZ, they improve by 13.91% and 11.57%, respectively; and when the threshold is set to 40 dBZ, CSI and HSS increase by 40.11% and 32.75%, respectively. Higher threshold settings prioritize high reflectivity regions, enhancing the model's ability to capture extreme weather phenomena. This is especially evident in regions where the threshold reaches 40 dBZ, where the model's performance shows significant improvement. The increased sensitivity to high reflectivity regions enables MMST-LSTM to achieve higher accuracy in predicting high-intensity radar echoes. By introducing physical constraints, particularly partial differential equation constraints in high-intensity echo regions, the model's ability to capture the evolution of echoes is enhanced, significantly improving prediction accuracy. Integrating the MCFFM and MAFUM components results in improved accuracy in high-intensity radar echo prediction. Nevertheless, this improvement may be less pronounced in low-reflectivity regions or rapidly changing conditions, and the focus on high-intensity areas could introduce trade-offs when handling lower-intensity patterns. However, in extreme weather scenarios, such as heavy rainfall events with reflectivity exceeding 40 dBZ, the errors are primarily concentrated near the boundaries of high-intensity echoes. This is hypothesized to be due to the relatively low proportion of such scenes in the training samples. Furthermore, the analysis suggests that the prediction errors in these regions may stem from insufficient feature learning, which results in the model's insufficient attention to the dynamic changes occurring in these areas.

To assess the robustness of the extrapolation performance across all models, the score curves for the selected metrics with respect to forecast lead time on the GZ test set are illustrated in Fig. 5. It is evident that the extrapolation network built

TABLE II
THE QUANTITATIVE RESULTS ON THE **GZ TEST SET** IN TERMS OF THE SELECTED FOUR METRICS IN THE COMPARATIVE EXPERIMENT.

Model	CSI ↑				HSS ↑				MSE ↓	SSIM ↑
	CSI-20	CSI-30	CSI-40	Average	HSS-20	HSS-30	HSS-40	Average		
ConvLSTM	0.578	0.479	0.222	0.427	0.645	0.598	0.347	0.529	184.256	0.426
PredRNN	0.605	0.470	0.227	0.434	0.679	0.585	0.349	0.538	176.266	0.439
PredRNN++	0.624	0.510	0.247	0.460	0.699	0.626	0.378	0.567	160.264	0.475
MIM	0.629	0.507	0.252	0.463	0.705	0.624	0.383	0.571	157.900	0.485
PhyDNet	0.608	0.507	0.242	0.452	0.683	0.627	0.373	0.561	159.537	0.497
SAConvLSTM	0.621	0.526	0.277	0.474	0.694	0.643	0.416	0.584	156.702	0.503
MMST-LSTM	0.636	0.536	0.318	0.497	0.711	0.653	0.464	0.609	146.169	0.5431

by the proposed MMST-LSTM outperforms other compared models for all metrics at all time steps, showing its improved extrapolation performance. Its curve exhibits a slower rate of decline compared to other models, indicating greater stability and sustained superior performance over time. In addition, in terms of 40 dBZ, the highest threshold, the CSI and HSS curves of MMST-LSTM have large distances from others', this suggests that the MCFFM and MAFUM empower the model to efficiently hone in on the characteristics linked to intense radar returns, thereby facilitating precise forecasts of their subsequent developmen.

Fig. 6 shows the visualization of a typical sequence selected from the results of comparative experiment conducted on the GZ Dataset. This sequence depicts the movement of echoes from the left part of the map toward the center. It should be noted that, as marked by the red box, there is a red band-area that represents high-intensity echoes, which is closely related to possible destructive precipitation events. It is crucial for models to predict it accurately. The ConvLSTM failed to correctly predict the yellow and red areas in the map because its structure is relatively too simple to capture features related to the variations of high-intensity echoes. The PredRNN built by ST-LSTM units, the baseline model, which are added the spatiotemporal memory cells, had improved extrapolation performance for high-intensity echoes. But notice that the red area marked by the red box shows a scattered shape rather than the band shape in the observed map. Although PredRNN++, MIM and SA-ConvLSTM predicted the band-like echo region successfully, the intensity of the echoes is still underestimated due to the lack of components that can help the model focus on features related to high-intensity echoes. The last row is the extrapolated maps of the network built by MMST-LSTM units. They are closest to the real observed echo maps compared to those of other models. This is attributed to the newly proposed components MCFFM and MAFUM. The MCFFM enhance the contextual feature association between the new input state at the current moment and the hidden state from the previous moment, thus preventing some important features from being forgotten during the gating operation and not being recalled again. The MAFUM tells the model which features in the newly obtained feature map are closely related to high-intensity echoes and need special attention, and where the model should find these features in the feature map for capture and learning, so as to achieve accurate extrapolation of variations of high-intensity echoes.

Despite the successes of PredRNN++, MIM, and SA-ConvLSTM in predicting banded echo regions, their predictions still fall short in echo intensity due to the lack of components specifically designed for high-intensity echo features. At the bottom row of the table are extrapolated images produced by networks using MMST-LSTM units, which closely resemble actual observed radar echo images. This advantage is primarily attributed to the newly introduced MCFFM and MAFUM components.The MCFFM fortifies the relationship between the current input's contextual features and those of the preceding hidden state, effectively preventing the loss of important features during gating processes. MAFUM guides the model in identifying features closely related to high-intensity echoes in new feature maps, indicating where the model should look and capture these features precisely to learn and predict changes in high-intensity echoes accurately.

Comparative experiments were also conducted on another publicly available radar echo dataset, CIKM 2017, to evaluate the model's generalization capabilities and substantiate its broad applicability. When contrasted with the GZ Dataset, the

TABLE III
THE QUANTITATIVE RESULTS ON THE **CIKM 2017 TEST SET** IN TERMS OF THE SELECTED FOUR METRICS IN THE COMPARATIVE EXPERIMENT.

Model	CSI ↑				HSS ↑				MSE ↓	SSIM ↑
	CSI-20	CSI-30	CSI-40	Average	HSS-20	HSS-30	HSS-40	Average		
ConvLSTM	0.394	0.214	0.087	0.232	0.466	0.332	0.159	0.319	140.614	0.486
PredRNN	0.389	0.227	0.097	0.238	0.445	0.345	0.174	0.321	157.030	0.479
PredRNN++	0.402	0.241	0.104	0.249	0.469	0.364	0.185	0.339	150.002	0.485
MIM	0.415	0.249	0.114	0.259	0.485	0.374	0.203	0.354	138.125	0.505
PhyDNet	0.367	0.185	0.093	0.215	0.432	0.289	0.168	0.296	171.097	0.417
SAConvLSTM	0.404	0.234	0.118	0.252	0.471	0.352	0.208	0.344	167.122	0.478
MMST-LSTM	0.424	0.271	0.149	0.281	0.499	0.405	0.257	0.387	134.015	0.507

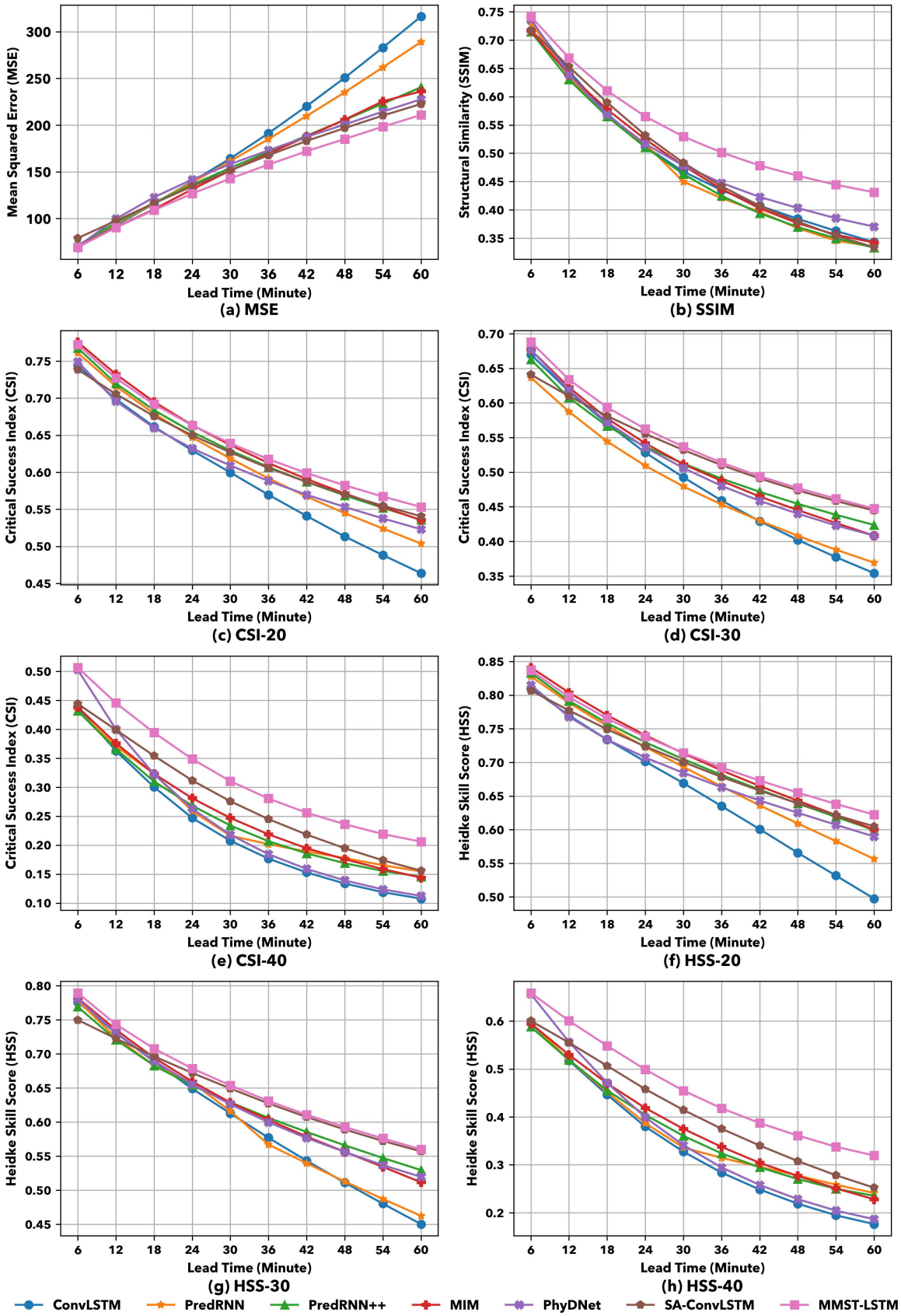

Fig. 5. The score curves of selected metrics on the **GZ test set** with respect to the forecast lead time.

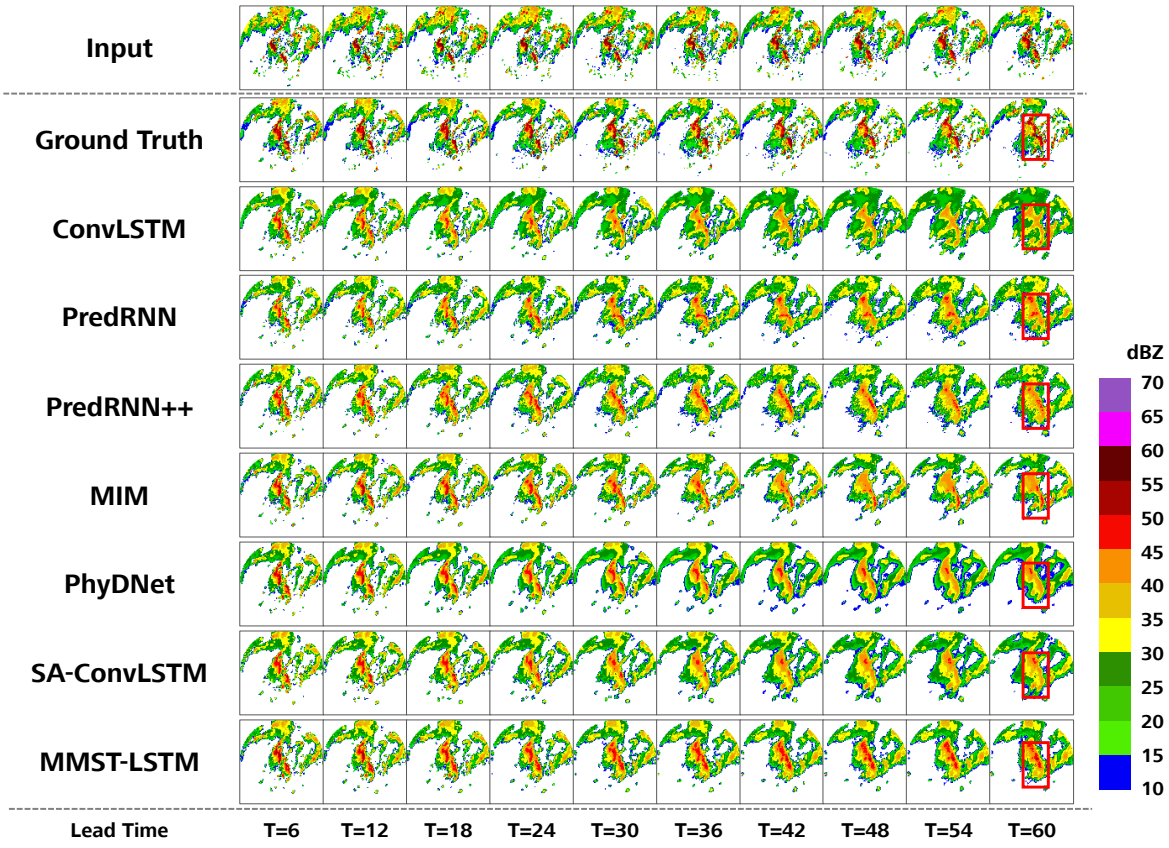

Fig. 6. The score curves of selected metrics on the **CIKM 2017 test set** with respect to the forecast lead time.

echo map in this dataset covers a smaller area, so models receive less features of radar echo evolution, which increases the challenge of the extrapolation task. Table III shows the quantitative result of all models on the CIKM 2017 test set in terms of the four selected metrics. The proposed MMST-LSTM still obtained the best results, and the second-ranked is MIM. Compared with it, the CSI scores of the proposed MMST-LSTM saw an enhancement of 2.15%, 9.09% and 30.41% for three thresholds, respectively. The proposed MMST-LSTM model achieved a 2.82% boost in its Heidke Skill Score (HSS), 8.28% and 26.70% for three thresholds, respectively. The model realized an 8.52% increase in the mean CSI score and a 9.30% increase in the mean HSS score, respectively. Its MSE score decreased by 2.98% and SSIM score improved by 0.48%. The SA-ConvLSTM followed the MIM, which performs slightly worse than MIM because the model gets less information about the holistic radar echo variation. Compared with PredRNN, the MMST-LSTM improves significantly. Its CSI scores improved by 8.86%, 19.48% and 53.56% under three different thresholds, respectively. There was an 18.32% rise in the average CSI score. the HSS score improved by 12.08%, 17.55% and 47.27% for the three different thresholds, respectively. The average HSS score improved by 20.40%. Its MSE score decreased by 14.66% and the SSIM score improved by 5.85%. The above results illustrate the general applicability and improved extrapolation performance of the proposed MMST-LSTM built network at different scales of radar echo datasets.

Fig. 7 shows all models' score curves of four metrics on the CIKM 2017 test set with respect to the forecast lead time. Compared to the GZ Dataset, the increased extrapolation difficulty causes a decrease in all models' scores for each metric. Although the distance between the curves of the proposed MMST-LSTM and other models smaller, its curves still decrease the slowest. It's worth noting that at the 60-minute lead time for the top threshold of 40 dBZ, the CSI and HSS scores for many models are almost negligible, while the curve of MMST-LSTM still lies above all the curves, indicating its ability to maintain a better extrapolation performance with increasing time.

Fig. 8 is the visualization of a representative sequence selected from the test results of comparative experiment conducted on CIKM 2017 dataset. This sequence dipicts the aggregation process of the echo from the right part of the map towards the center. The red and orange areas representing possible heavy precipitation events are also marked by red box. Except for PhyDNet, all other models predicted the trend of radar echo movement relatively accurately. However, in the extrapolated maps of ConvLSTM and SA-ConvLSTM, the area of the echoes is smaller than that in the corresponding observed map. PredRNN, PredRNN++, and MIM failed to correctly predict the red area marked by the red box representing high-intensity echoes. Only the extrapolation network built by the proposed MMST-LSTM makes the most accurate predictions of both the trend of radar evolution and its intensity. The result again illustrates the improved ability

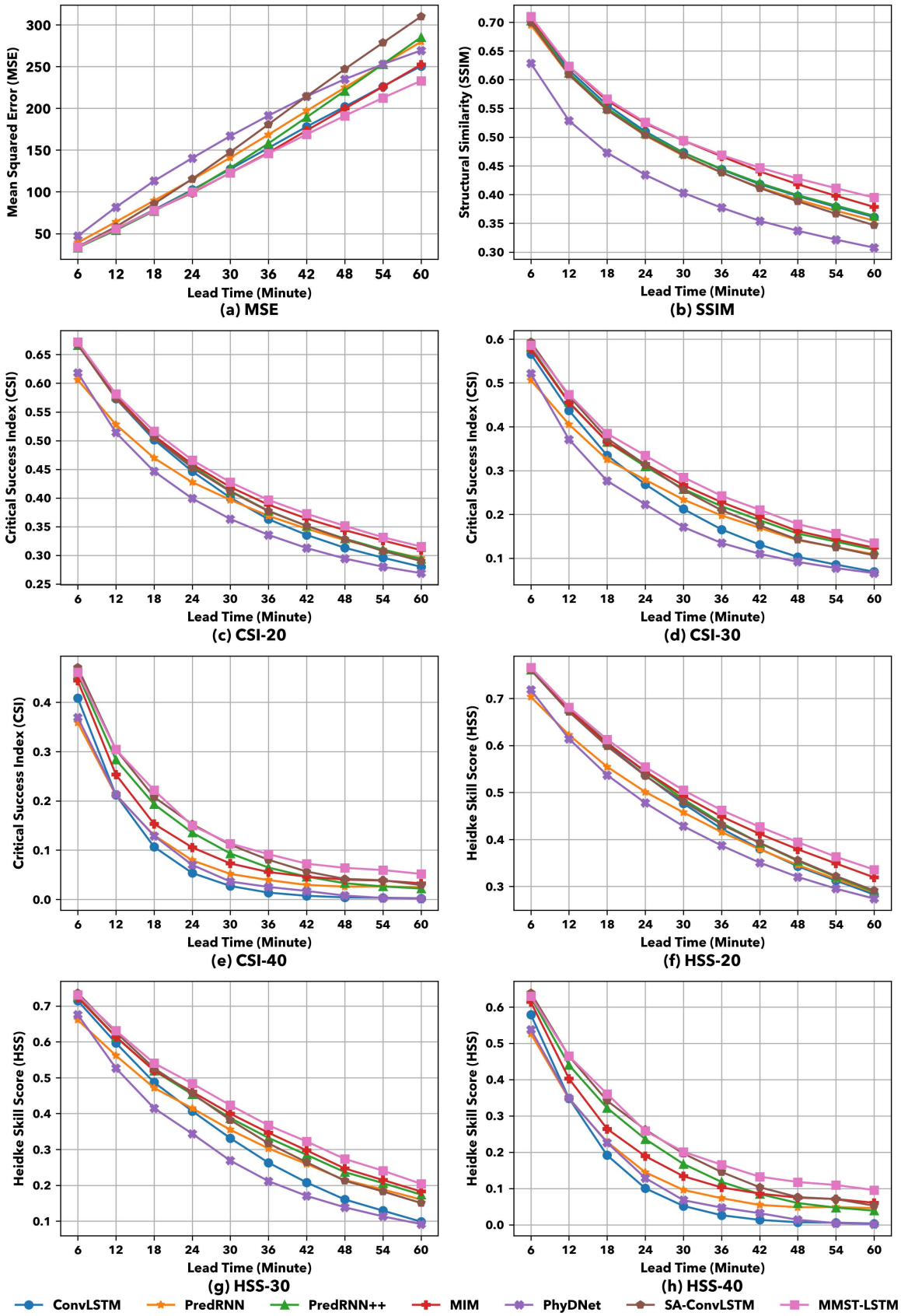

Fig. 7. The score curves of selected metrics on the **CIKM 2017 test set** with respect to the forecast lead time.

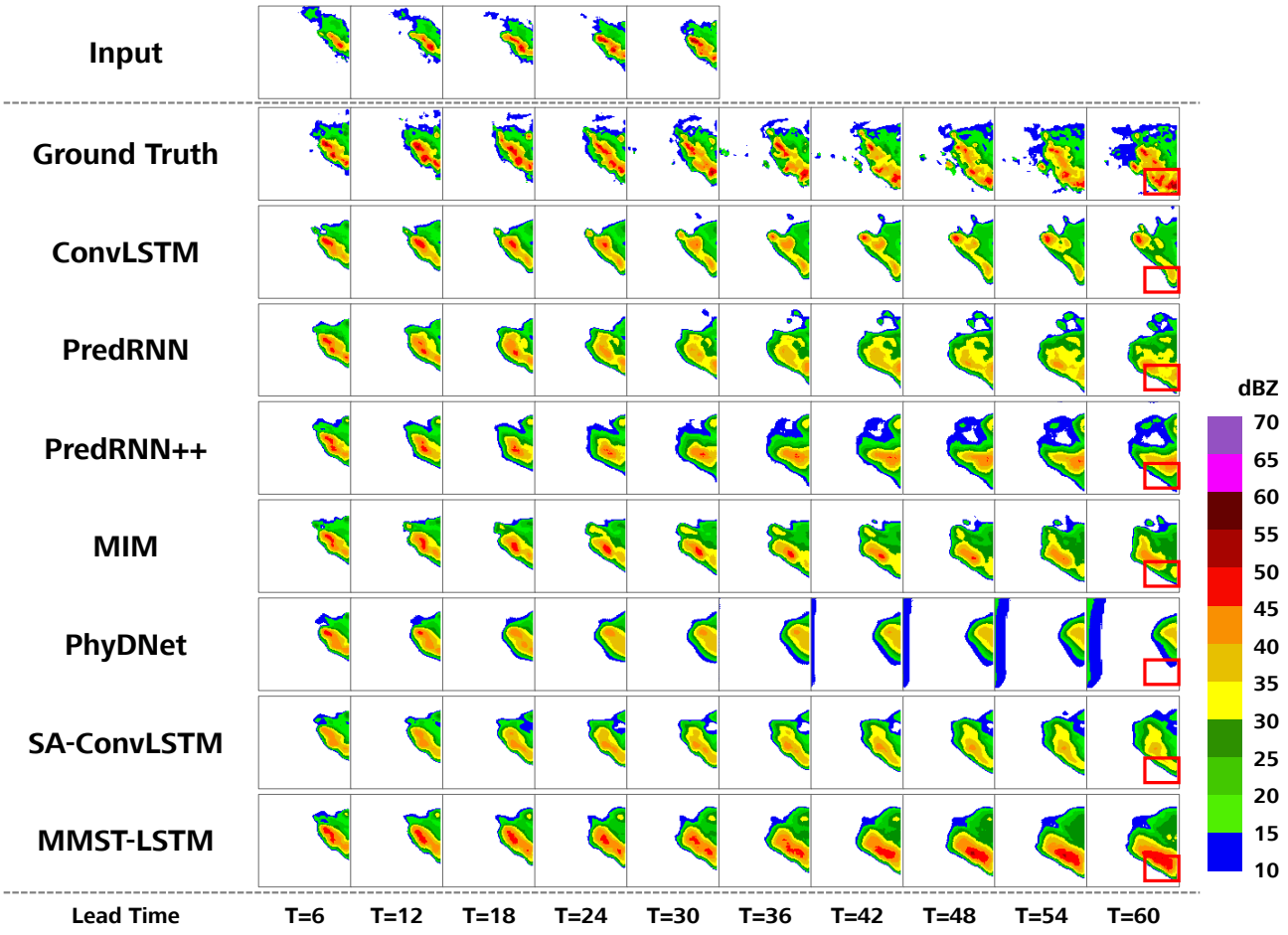

Fig. 8. The visualization of a typical sequence selected from the results of comparative experiment conducted on the **CIKM 2017 Dataset**. The extrapolation network built by the proposed MMST-LSTM units obtained the best results. It accurately predicted the evolution trends and intensity of echoes. In extrapolated maps of other models, the problem that echo dissipation or underestimation appeared.

TABLE IV
THE QUANTITATIVE RESULTS ON THE **GZ TEST SET** IN TERMS OF THE FOUR SELECTED METRICS IN THE ABLATION STUDY.

Model	CSI ↑				HSS ↑				MSE ↓	SSIM ↑
	CSI-20	CSI-30	CSI-40	Average	HSS-20	HSS-30	HSS-40	Average		
Baseline Model	0.605	0.470	0.227	0.434	0.679	0.585	0.349	0.538	176.266	0.439
wo. MAFUM	0.614	0.521	0.252	0.4622	0.686	0.638	0.385	0.570	160.987	0.501
wo. MCFFM	0.619	0.513	0.276	0.470	0.693	0.630	0.415	0.580	155.222	0.530
MMST-LSTM	0.636	0.536	0.318	0.497	0.711	0.653	0.464	0.609	146.169	0.543

of MMST-LSTM for high-intensity echoes extrapolation.

E. Ablation Study

To demonstrate the effectiveness of the two proposed components, the ablation study is conducted on the GZ Dataset. MCFFM and MAFUM were added to the original baseline model ST-LSTM unit, respectively. Therefore, the objects of the ablation study including the baseline model ST-LSTM, MMST-LSTM wo. MAFUM (MMST-LSTM without MAFUM), MMST-LSTM wo. MCFFM (MMST-LSTM without MCFFM), and the standard MMST-LSTM. They are respectively constructed the extrapolation network for performing extrapolation tasks. Table IV shows the quantitative result of the ablation study. The baseline model is added the MCFFM and MAFUM, respectively, both scores of the two obtained models in terms of all metrics are improved. With the MCFFM, the original model is able to enhance the contextual correlation between states and avoid some features from being forgotten, so its CSI and HSS scores are higher at some thresholds. With the help of MAFUM, the original model can focus on those features that are closely related to high-intensity echoes in the updated feature map, and then capture them, so the scores of MMST-LSTM without MCFFM under the highest threshold is significantly improved. The network built by standard MMST-LSTM units achieved the best performance, which is attributed to the cooperative work of these two components.Compared to the baseline models, MMST-LSTM demonstrates exceptional performance in predicting high-intensity echoes, proving the

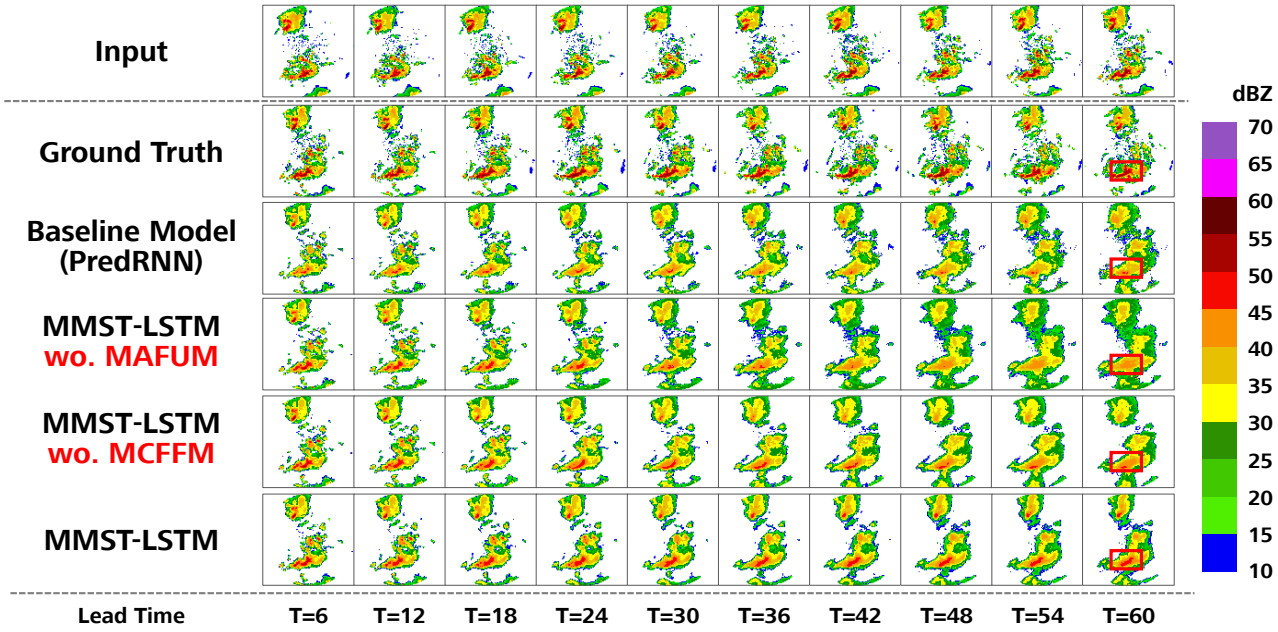

Fig. 9. The visualization of a typical sequence selected from the results of ablation study conducted on the **GZ Dataset**. In the echo map extrapolated by the proposed MMST-LSTM, the red area marked by the red box representing the potential precipitation events is closest to ground truth. It shows the effectiveness of MCFFM and MAFUM in improving extrapolation performance for high-intensity echoes.

effectiveness of the multi-scale feature extraction module in capturing complex features. Moreover, while MMST-LSTM slightly increases computational complexity, the significant improvement in prediction accuracy more than compensates for this cost.

Fig. 9 is the visualization of a representative sequence selected from the results of ablation study conducted on GZ dataset. This sequence presents the gradual aggregation process of echoes in the center of map. The red box marks a red echo area, and new echoes are continuously generated and aggregate around it. As can be seen from the extrapolated results, the baseline model failed to predict the red echo area correctly due to the lack of MCFFM and MAFUM, and it cannot focus on the important features in the feature map updated by the gating structure. The intensity of some other echoes in its extrapolated maps are also underestimated. The MMST-LSTM wo.MAFUM and MMST-LSTM wo. MCFFM only successfully predicted the orange echo area in the red box but failed to predict the red echo area. This is mainly because they have forgotten or failed to pay attention to important features related to high-intensity echoes. The last row shows the maps extrapolated by the standard MMST-LSTM. The red area marked by the red box are the closest to the real observed maps. With the help of MCFFM and MAFUM, the high-intensity echo-related features are prevented from being forgotten, and the features that evolve with the high-intensity echoes in the updated feature map are effectively learned and accurately predicted.

V. Conclusion and Future Work

This study addresses a prevalent challenge in traditional radar echo prediction techniques: the systematic underestimation of high-density echo signals. This issue primarily arises from two critical factors: the inherent limitations of the model and the uneven distribution of intensity levels within real-world radar echo datasets, both of which significantly contribute to the complexity of the task. To address these model shortcomings, this article presents two novel components, the MCFFM and MAFUM are key enhancements to the prediction model. The MCFFM integrates features from various sources, while the MAFUM refines feature updates through cross-domain attention, both contributing to improved radar echo forecasting. These components form the core recurrent mechanism of the MMST-LSTM network. The MCFFM is designed to operate before the gating mechanism of the LSTM unit, enhancing the interconnection between input data and hidden states by integrating feature information from different scales and levels. This integration helps to effectively prevent the loss of high-intensity echo features.The MCFFM component is positioned before the gating mechanism of the LSTM recurrent unit. This design aims to strengthen the connection between input data and internal hidden states by integrating feature information from different levels and dimensions, ensuring that features related to high-density echoes are effectively preserved. The MAFUM is positioned subsequent to the LSTM unit's gating mechanism. Its purpose is to assist the model in identifying features closely related to high-intensity radar echoes and to guide the model in detecting and focusing on these features in critical areas. For the second factor, this paper uses BMSE to train all models, so that the features related to high-intensity echoes that account for a small proportion are amplified and thus attended to by the models. The MMST-LSTM model proposed in this study demonstrates excellent performance on specific radar echo datasets. In extreme weather conditions, particularly

in high-intensity precipitation regions, the model effectively predicts the dynamics of high-intensity echoes. The subsequent incorporation of low-intensity echo data can enhance the model's ability to capture the early stages of storm development, such as the formation of weak echoes before high-intensity signals. By learning these initial patterns, the model can better predict storm intensification and movement, aiding in the prediction of the spatial and temporal progression of extreme weather, leading to more accurate regional warnings. In edge computing environments, although inference efficiency has been optimized, computational complexity remains high when processing large-scale data. Future research will focus on enhancing the model's generalization ability, especially its adaptability across different climates and geographical regions. Transfer learning will be a key technique to improve cross-regional prediction capabilities, enabling rapid adaptation to climatic differences and data scarcity in target domains by transferring knowledge from the source domain. Furthermore, integrating partial differential equations (PDEs) for modeling the spatiotemporal evolution of echoes will further enhance the model's ability to handle extreme weather patterns, achieving a deep integration of physical constraints and data-driven methods to improve prediction accuracy and stability. In mobile edge computing and consumer electronics, this integrated model can efficiently operate in low-latency and low-power environments, providing real-time and accurate extreme weather predictions for smart devices. By incorporating transfer learning and physical constraints into edge computing, the model will adapt to complex climate conditions across regions. It will also provide efficient meteorological services on resource-constrained edge devices, driving the development of smart hardware towards higher precision and intelligence. Future research could combine physical constraints, such as partial differential equations, with data-driven methods to improve the accuracy and efficiency of extreme weather predictions. When applied to consumer electronics, this could enable smart devices to deliver more accurate real-time weather forecasts under low-latency, low-power conditions.

ACKNOWLEDGEMENTS

This work has received funding from National Natural Science Foundation of China (No. 42275157).

REFERENCES

[1] Z. Gao, X. Shi, H. Wang, D.-Y. Yeung, W.-c. Woo, W.-K. Wong, Deep learning and the weather forecasting problem: Precipitation nowcasting, Deep Learning for the Earth Sciences: A Comprehensive Approach to Remote Sensing, Climate Science, and Geosciences (2021) 218–239doi:https://doi.org/10.1002/9781119646181.ch15.
URL https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119646181.ch15

[2] Z. Yang, H. Wu, Q. Liu, X. Liu, Y. Zhang, X. Cao, A self-attention integrated spatiotemporal lstm approach to edge-radar echo extrapolation in the internet of radars, ISA Transactions 132 (2023) 155–166. doi:10.1016/j.isatra.2022.06.046.
URL https://www.sciencedirect.com/science/article/pii/S001905782200355X

[3] Q. Liu, J. Sun, Y. Zhang, X. Liu, Denmerd: A feature enhanced approach to radar beam blockage correction with edge-cloud computing, Journal of Cloud Computing: Advances, Systems and Applications 13 (2024) 32. doi:10.1186/s13677-024-00607-x.
URL https://doi.org/10.1186/s13677-024-00607-x

[4] X. Zhou, Q. Yang, X. Zheng, W. Liang, K. I.-K. Wang, J. Ma, Y. Pan, Q. Jin, Personalized federated learning with model-contrastive learning for multi-modal user modeling in human-centric metaverse, IEEE Journal on Selected Areas in Communications 42 (4) (2024) 817–831. doi:10.1109/JSAC.2023.3345431.

[5] X. Ren, X. Li, K. Ren, J. Song, Z. Xu, K. Deng, X. Wang, Deep learning-based weather prediction: a survey, Big Data Research 23 (2021) 100178. doi:https://doi.org/10.1016/j.bdr.2020.100178.
URL https://www.sciencedirect.com/science/article/pii/S2214579620300460

[6] X. Zhou, X. Zheng, X. Cui, J. Shi, W. Liang, Z. Yan, L. T. Yang, S. Shimizu, K. I.-K. Wang, Digital twin enhanced federated reinforcement learning with lightweight knowledge distillation in mobile networks, IEEE Journal on Selected Areas in Communications 41 (10) (2023) 3191–3211. doi:10.1109/JSAC.2023.3310046.

[7] J. Sun, M. Xue, J. W. Wilson, I. Zawadzki, S. P. Ballard, J. Onvlee-Hooimeyer, P. Joe, D. M. Barker, P.-W. Li, B. Golding, et al., Use of nwp for nowcasting convective precipitation: Recent progress and challenges, Bulletin of the American Meteorological Society 95 (3) (2014) 409–426. doi:https://doi.org/10.1175/BAMS-D-11-00263.1.
URL https://journals.ametsoc.org/view/journals/bams/95/3/bams-d-11-00263.1.xml

[8] P. Bauer, A. Thorpe, G. Brunet, The quiet revolution of numerical weather prediction, Nature 525 (7567) (2015) 47–55. doi:https://doi.org/10.1038/nature14956.
URL https://doi.org/10.1038/nature14956

[9] D. Niu, J. Huang, Z. Zang, L. Xu, H. Che, Y. Tang, Two-stage spatiotemporal context refinement network for precipitation nowcasting, Remote Sensing 13 (21) (2021) 4285. doi:https://doi.org/10.3390/rs13214285.
URL https://www.mdpi.com/2072-4292/13/21/4285

[10] C. Luo, X. Li, Y. Ye, Pfst-lstm: A spatiotemporal lstm model with pseudoflow prediction for precipitation nowcasting, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 14 (2020) 843–857. doi:https://doi.org/10.1109/JSTARS.2020.3040648.

[11] L. Alfieri, P. Claps, F. Laio, Time-dependent zr relationships for estimating rainfall fields from radar measurements, Natural Hazards and Earth System Sciences 10 (1) (2010) 149–158. doi:https://doi.org/10.5194/nhess-10-149-2010.
URL https://nhess.copernicus.org/articles/10/149/2010/

[12] A. del Moral, T. Rigo, M. C. Llasat, A radar-based centroid tracking algorithm for severe weather surveillance: Identifying split/merge processes in convective systems, Atmospheric Research 213 (2018) 110–120. doi:https://doi.org/10.1016/j.atmosres.2018.05.030.
URL https://www.sciencedirect.com/science/article/pii/S0169809518301091

[13] J. Johnson, P. L. MacKeen, A. Witt, E. D. W. Mitchell, G. J. Stumpf, M. D. Eilts, K. W. Thomas, The storm cell identification and tracking algorithm: An enhanced wsr-88d algorithm, Weather and forecasting 13 (2) (1998) 263–276. doi:https://doi.org/10.1175/1520-0434(1998)013¡0263:TSCIAT¿2.0.CO;2.
URL https://journals.ametsoc.org/view/journals/wefo/13/2/1520-0434_1998_013_0263_tsciat_2_0_co_2.xml

[14] Q. Liang, Y. Feng, W. Deng, S. Hu, Y. Huang, Q. Zeng, Z. Chen, A composite approach of radar echo extrapolation based on trec vectors in combination with model-predicted winds, Advances in Atmospheric Sciences 27 (5) (2010) 1119–1130. doi:https://doi.org/10.1007/s00376-009-9093-4.

[15] S. Otsuka, G. Tuerhong, R. Kikuchi, Y. Kitano, Y. Taniguchi, J. J. Ruiz, S. Satoh, T. Ushio, T. Miyoshi, Precipitation nowcasting with three-dimensional space–time extrapolation of dense and frequent phased-array weather radar observations, Weather and Forecasting 31 (1) (2016) 329–340. doi:https://doi.org/10.1175/WAF-D-15-0063.1.
URL https://journals.ametsoc.org/view/journals/wefo/31/1/waf-d-15-0063_1.xml

[16] W.-c. Woo, W.-k. Wong, Operational application of optical flow techniques to radar-based rainfall nowcasting, Atmosphere 8 (3) (2017) 48. doi:https://doi.org/10.3390/atmos8030048.
URL https://www.mdpi.com/2073-4433/8/3/48

[17] L. Li, Z. He, S. Chen, X. Mai, A. Zhang, B. Hu, Z. Li, X. Tong, Subpixel-based precipitation nowcasting with the pyramid lucas–kanade optical flow technique, Atmosphere 9 (7) (2018) 260. doi:https://doi.org/10.3390/atmos9070260.
URL https://www.mdpi.com/2073-4433/9/7/260

[18] G. Ayzel, M. Heistermann, T. Winterrath, Optical flow models as an open benchmark for radar-based precipitation nowcasting (rainymotion v0. 1), Geoscientific Model Development 12 (4) (2019) 1387–1402.

doi:https://doi.org/10.5194/gmd-12-1387-2019.
URL https://gmd.copernicus.org/articles/12/1387/2019/

[19] J. Jing, Q. Li, L. Ma, L. Chen, L. Ding, Remnet: Recurrent evolution memory-aware network for accurate long-term weather radar echo extrapolation, IEEE Transactions on Geoscience and Remote Sensing 60 (2022) 1–13. doi:https://doi.org/10.1109/TGRS.2022.3198851.

[20] Z. Yang, Q. Liu, H. Wu, X. Liu, Y. Zhang, Cema-lstm: Enhancing contextual feature correlation for radar extrapolation using fine-grained echo datasets, Computer Modeling in Engineering & Sciences 135 (1) (2023) 45–64. doi:https://doi.org/10.32604/cmes.2022.022045.
URL http://www.techscience.com/CMES/v135n1/50092

[21] X. Zhou, J. Wu, W. Liang, K. I.-K. Wang, Z. Yan, L. T. Yang, Q. Jin, Reconstructed graph neural network with knowledge distillation for lightweight anomaly detection, IEEE Transactions on Neural Networks and Learning Systems (2024) 1–12doi:10.1109/TNNLS.2024.3389714.

[22] X. Zhou, X. Zheng, T. Shu, W. Liang, K. I.-K. Wang, L. Qi, S. Shimizu, Q. Jin, Information theoretic learning-enhanced dual-generative adversarial networks with causal representation for robust ood generalization, IEEE Transactions on Neural Networks and Learning Systems (2023) 1–14doi:10.1109/TNNLS.2023.3330864.

[23] C. Meng, Z. Fang, Research on prediction and analysis of consumer behavior management based on deep learning, in: 2024 IEEE 2nd International Conference on Control, Electronics and Computer Technology, 2024, pp. 179–183. doi:10.1109/ICCECT60629.2024.10545671.

[24] X. Chen, M. Wang, S. Wang, Y. Chen, R. Wang, C. Zhao, X. Hu, Weather radar nowcasting for extreme precipitation prediction based on the temporal and spatial generative adversarial network, Atmosphere 13 (8) (2022) 1291. doi:https://doi.org/10.3390/atmos13081291.
URL https://www.mdpi.com/2073-4433/13/8/1291

[25] Q. Liu, J. Zhang, X. Liu, Y. Zhang, X. Xu, M. Khosravi, M. Bilal, Improving wireless indoor non-intrusive load disaggregation using attention-based deep learning networks, Physical Communication 51 (2022) 101584. doi:https://doi.org/10.1016/j.phycom.2021.101584.
URL https://www.sciencedirect.com/science/article/pii/S1874490721002834

[26] H. Wu, Q. Liu, X. Liu, Y. Zhang, Z. Yang, An edge-assisted cloud framework using a residual concatenate fcn approach to beam correction in the internet of weather radars, World Wide Web (2022) 1–27doi:https://doi.org/10.1007/s11280-021-00988-y.

[27] Y.-M. Chiang, F.-J. Chang, B. J.-D. Jou, P.-F. Lin, Dynamic ann for precipitation estimation and forecasting from radar observations, Journal of Hydrology 334 (1-2) (2007) 250–261. doi:https://doi.org/10.1016/j.jhydrol.2006.10.021.
URL https://www.sciencedirect.com/science/article/pii/S0022169406005439

[28] C.-C. Wei, Soft computing techniques in ensemble precipitation nowcast, Applied Soft Computing 13 (2) (2013) 793–805. doi:https://doi.org/10.1016/j.asoc.2012.10.006.
URL https://www.sciencedirect.com/science/article/pii/S1568494612004553

[29] X. SHI, Z. Chen, H. Wang, D.-Y. Yeung, W.-k. Wong, W.-c. WOO, Convolutional lstm network: A machine learning approach for precipitation nowcasting, in: Advances in Neural Information Processing Systems, Vol. 28, Curran Associates, Inc., 2015.
URL https://proceedings.neurips.cc/paper_files/paper/2015/file/07563a3fe3bbe7e3ba84431ad9d055af-Paper.pdf

[30] B. Klein, L. Wolf, Y. Afek, A dynamic convolutional layer for short range weather prediction, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 4840–4848. doi:10.1109/CVPR.2015.7299117.

[31] Y. Wang, M. Long, J. Wang, Z. Gao, P. S. Yu, Predrnn: Recurrent neural networks for predictive learning using spatiotemporal lstms, in: I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, R. Garnett (Eds.), Advances in Neural Information Processing Systems, Vol. 30, Curran Associates, Inc., 2017.
URL https://proceedings.neurips.cc/paper_files/paper/2017/file/e5f6ad6ce374177eef023bf5d0c018b6-Paper.pdf

[32] V. L. Guen, N. Thome, Disentangling physical dynamics from unknown factors for unsupervised video prediction, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 11474–11484.

[33] J. Yin, Z. Gao, W. Han, Application of a radar echo extrapolation-based deep learning method in strong convection nowcasting, Earth and Space Science 8 (8) (2021) e2020EA001621. doi:https://doi.org/10.1029/2020EA001621.
URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020EA001621

[34] F. Zhang, C. Lai, W. Chen, Weather radar echo extrapolation method based on deep learning, Atmosphere 13 (5) (2022) 815. doi:https://doi.org/10.3390/atmos13050815.
URL https://www.mdpi.com/2073-4433/13/5/815

[35] X. Zhou, W. Liang, A. Kawai, K. Fueda, J. She, K. I.-K. Wang, Adaptive segmentation enhanced asynchronous federated learning for sustainable intelligent transportation systems, IEEE Transactions on Intelligent Transportation Systems 25 (7) (2024) 6658–6666. doi:10.1109/TITS.2024.3362058.

[36] X. Shi, Z. Gao, L. Lausen, H. Wang, D.-Y. Yeung, W.-k. Wong, W.-c. Woo, Deep learning for precipitation nowcasting: A benchmark and a new model 30 (2017).
URL https://proceedings.neurips.cc/paper_files/paper/2017/file/a6db4ed04f1621a119799fd3d7545d3d-Paper.pdf

[37] L. Han, H. Liang, H. Chen, W. Zhang, Y. Ge, Convective precipitation nowcasting using u-net model, IEEE Transactions on Geoscience and Remote Sensing 60 (2022) 1–8. doi:https://doi.org/10.1109/TGRS.2021.3100847.

[38] J.-H. Syu, J. C.-W. Lin, G. Srivastava, K. Yu, A comprehensive survey on artificial intelligence empowered edge computing on consumer electronics, IEEE Transactions on Consumer Electronics 69 (4) (2023) 1023–1034. doi:10.1109/TCE.2023.3318150.

[39] Q. Lin, S. Jiang, Z. Zhen, T. Chen, C. Wei, H. Lin, Fed-pemc: A privacy-enhanced federated deep learning algorithm for consumer electronics in mobile edge computing, IEEE Transactions on Consumer Electronics 70 (1) (2024) 4073–4086. doi:10.1109/TCE.2024.3351648.

[40] Q. Liu, Z. Yang, R. Ji, Y. Zhang, M. Bilal, X. Liu, S. Vimal, X. Xu, Deep vision in analysis and recognition of radar data: Achievements, advancements, and challenges, IEEE Systems, Man, and Cybernetics Magazine 9 (4) (2023) 4–12. doi:10.1109/MSMC.2022.3216943.

[41] X. Zhou, W. Liang, K. I.-K. Wang, Z. Yan, L. T. Yang, W. Wei, J. Ma, Q. Jin, Decentralized p2p federated learning for privacy-preserving and resilient mobile robotic systems, IEEE Wireless Communications 30 (2) (2023) 82–89. doi:10.1109/MWC.004.2200381.

[42] W. Zhuang, W. Ding, Long-lead prediction of extreme precipitation cluster via a spatiotemporal convolutional neural network, in: Proceedings of the 6th International Workshop on Climate Informatics: CI, 2016, pp. 128–131.

[43] Y. Wang, J. Zhang, H. Zhu, M. Long, J. Wang, P. S. Yu, Memory in memory: A predictive neural network for learning higher-order non-stationarity from spatiotemporal dynamics, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 9154–9162.

[44] S. Agrawal, L. Barrington, C. Bromberg, J. Burge, C. Gazen, J. Hickey, Machine learning for precipitation nowcasting from radar images, arXiv preprint arXiv:1912.12132 (2019).

[45] G. Ayzel, T. Scheffer, M. Heistermann, Rainnet v1. 0: a convolutional neural network for radar-based precipitation nowcasting, Geoscientific Model Development 13 (6) (2020) 2631–2644. doi:https://doi.org/10.5194/gmd-13-2631-2020.
URL https://gmd.copernicus.org/articles/13/2631/2020/

[46] J. Jing, Q. Li, X. Peng, Mlc-lstm: Exploiting the spatiotemporal correlation between multi-level weather radar echoes for echo sequence extrapolation, Sensors 19 (18) (2019) 3988. doi:https://doi.org/10.3390/s19183988.
URL https://www.mdpi.com/1424-8220/19/18/3988

[47] L. Han, J. Sun, W. Zhang, Convolutional neural network for convective storm nowcasting using 3-d doppler weather radar data, IEEE Transactions on Geoscience and Remote Sensing 58 (2) (2019) 1487–1495. doi:https://doi.org/10.1109/TGRS.2019.2948070.

[48] C. Gianoglio, A. Rizik, E. Tavanti, D. D. Caviglia, A. Randazzo, On the edge recurrent neural network approach for ground moving fmcw radar target classification, IEEE Transactions on Consumer Electronics 70 (1) (2024) 522–534. doi:10.1109/TCE.2023.3343460.

[49] K. Trebing, T. Stanczyk, S. Mehrkanoon, Smaat-unet: Precipitation nowcasting using a small attention-unet architecture, Pattern Recognition Letters 145 (2021) 178–186. doi:https://doi.org/10.1016/j.patrec.2021.01.036.
URL https://www.sciencedirect.com/science/article/pii/S0167865521000556

[50] X. Zhou, X. Ye, K. I.-K. Wang, W. Liang, N. K. C. Nair, S. Shimizu, Z. Yan, Q. Jin, Hierarchical federated learning with social context clustering-based participant selection for internet of medical things applications, IEEE Transactions on Computational Social Systems 10 (4) (2023) 1742–1751. doi:10.1109/TCSS.2023.3259431.

[51] H. Geng, L. Geng, Mccs-lstm: Extracting full-image contextual information and multi-scale spatiotemporal feature for

radar echo extrapolation, Atmosphere 13 (2) (2022) 192. doi:https://doi.org/10.3390/atmos13020192.
URL https://www.mdpi.com/2073-4433/13/2/192

[52] Z. Yang, H. Wu, Q. Liu, X. Liu, Y. Zhang, X. Cao, A self-attention integrated spatiotemporal lstm approach to edge-radar echo extrapolation in the internet of radars, ISA transactions (2022). doi:https://doi.org/10.1016/j.isatra.2022.06.046.
URL https://www.sciencedirect.com/science/article/pii/S001905782200355X

[53] X. Huang, C. Luo, Y. Ye, X. Li, B. Zhang, Location-refining neural network: A new deep learning-based framework for heavy rainfall forecast, Computers & Geosciences 166 (2022) 105152. doi:https://doi.org/10.1016/j.cageo.2022.105152.
URL https://www.sciencedirect.com/science/article/pii/S0098300422001078

[54] N. Srivastava, E. Mansimov, R. Salakhudinov, Unsupervised learning of video representations using lstms, in: F. Bach, D. Blei (Eds.), Proceedings of the 32nd International Conference on Machine Learning, Vol. 37 of Proceedings of Machine Learning Research, PMLR, Lille, France, 2015, pp. 843–852.
URL https://proceedings.mlr.press/v37/srivastava15.html

[55] Y. Wang, Z. Gao, M. Long, J. Wang, P. S. Yu, PredRNN++: Towards a resolution of the deep-in-time dilemma in spatiotemporal predictive learning, in: Proceedings of the 35th International Conference on Machine Learning, Vol. 80 of Proceedings of Machine Learning Research, PMLR, 2018, pp. 5123–5132.
URL https://proceedings.mlr.press/v80/wang18b.html

[56] W. Byeon, Q. Wang, R. K. Srivastava, P. Koumoutsakos, Contextvp: Fully context-aware video prediction, in: Proceedings of the European Conference on Computer Vision, 2018, p. 753–769.

[57] Y. Wang, L. Jiang, M.-H. Yang, L.-J. Li, M. Long, L. Fei-Fei, Eidetic 3d LSTM: A model for video prediction and beyond, in: International Conference on Learning Representations, 2019.
URL https://openreview.net/forum?id=B1lKS2AqtX

[58] H. Fan, L. Zhu, Y. Yang, Cubic lstms for video prediction, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 33, 2019, pp. 8263–8270. doi:https://doi.org/10.1609/aaai.v33i01.33018263.
URL https://ojs.aaai.org/index.php/AAAI/article/view/4838

[59] Z. Lin, M. Li, Z. Zheng, Y. Cheng, C. Yuan, Self-attention convlstm for spatiotemporal prediction, Proceedings of the AAAI Conference on Artificial Intelligence 34 (07) (2020) 11531–11538. doi:https://doi.org/10.1609/aaai.v34i07.6819.
URL https://ojs.aaai.org/index.php/AAAI/article/view/6819

[60] J. Su, W. Byeon, J. Kossaifi, F. Huang, J. Kautz, A. Anandkumar, Convolutional tensor-train lstm for spatio-temporal learning, in: Advances in Neural Information Processing Systems, Vol. 33, Curran Associates, Inc., 2020, pp. 13714–13726.
URL https://proceedings.neurips.cc/paper/2020/file/9e1a36515d6704d7eb7a30d783400e5d-Paper.pdf

[61] Z. Chai, C. Yuan, Z. Lin, Y. Bai, Cms-lstm: context-embedding and multi-scale spatiotemporal-expression lstm for video prediction, arXiv preprint arXiv:2102.03586 (2021).

[62] H. Wu, Z. Yao, J. Wang, M. Long, Motionrnn: A flexible model for video prediction with spacetime-varying motions, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 15435–15444.

[63] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, H. Adam, Encoder-decoder with atrous separable convolution for semantic image segmentation, in: Proceedings of the European conference on computer vision, 2018, pp. 801–818.

[64] X. Tang, W. Tu, K. Li, J. Cheng, Dffnet: An iot-perceptive dual feature fusion network for general real-time semantic segmentation, Information Sciences 565 (2021) 326–343.

[65] Y. Ji, H. Zhang, F. Gao, H. Sun, H. Wei, N. Wang, B. Yang, Lgcnet: A local-to-global context-aware feature augmentation network for salient object detection, Information Sciences 584 (2022) 399–416.

[66] L. Zeng, Q. Liu, S. Shen, X. Liu, Improved double deep q network-based task scheduling algorithm in edge computing for makespan optimization, Tsinghua Science and Technology 29 (3) (2024) 806–817. doi:10.26599/TST.2023.9010058.

[67] X. Zhou, W. Liang, K. I.-K. Wang, L. T. Yang, Deep correlation mining based on hierarchical hybrid networks for heterogeneous big data recommendations, IEEE Transactions on Computational Social Systems 8 (1) (2021) 171–178. doi:10.1109/TCSS.2020.2987846.

[68] S. Woo, J. Park, J.-Y. Lee, I. S. Kweon, Cbam: Convolutional block attention module, in: Proceedings of the European conference on computer vision, 2018, pp. 3–19.

[69] L. Mou, Y. Zhao, L. Chen, J. Cheng, Z. Gu, H. Hao, H. Qi, Y. Zheng, A. Frangi, J. Liu, Cs-net: channel and spatial attention network for curvilinear structure segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2019, pp. 721–730.

[70] Q. Liu, O. F. Darteh, M. Bilal, X. Huang, M. Attique, X. Liu, A. Acakpovi, A cloud-based bi-directional lstm approach to grid-connected solar pv energy forecasting for multi-energy systems, Sustainable Computing: Informatics and Systems 40 (2023) 100892. doi:https://doi.org/10.1016/j.suscom.2023.100892.
URL https://www.sciencedirect.com/science/article/pii/S2210537923000471

[71] X. Zhou, Q. Yang, Q. Liu, W. Liang, K. Wang, Z. Liu, J. Ma, Q. Jin, Spatial–temporal federated transfer learning with multi-sensor data fusion for cooperative positioning, Information Fusion 105 (2024) 102182. doi:https://doi.org/10.1016/j.inffus.2023.102182.
URL https://www.sciencedirect.com/science/article/pii/S1566253523004980

[72] S. Zhong, X. Zeng, Q. Ling, Q. Wen, W. Meng, Y. Feng, Spatiotemporal convolutional lstm for radar echo extrapolation, in: 2020 54th Asilomar Conference on Signals, Systems, and Computers, IEEE, 2020, pp. 58–62. doi:https://doi.org/10.1109/IEEECONF51394.2020.9443310.

[73] X. Zhou, W. Huang, W. Liang, Z. Yan, J. Ma, Y. Pan, K. I-Kai Wang, Federated distillation and blockchain empowered secure knowledge sharing for internet of medical things, Information Sciences 662 (2024) 120217. doi:https://doi.org/10.1016/j.ins.2024.120217.
URL https://www.sciencedirect.com/science/article/pii/S0020025524001300

[74] C. Chen, C. M. Unal, A. C. O. Nijhuis, Jensen–shannon distance-based filter and unsupervised evaluation metrics for polarimetric weather radar processing, IEEE Transactions on Geoscience and Remote Sensing 60 (2022) 1–18. doi:https://doi.org/10.1109/TGRS.2022.3200731.

[75] C. Dong, M. Shafiq, M. M. A. Dabel, Y. Sun, Z. Tian, Dnn inference acceleration for smart devices in industry 5.0 by decentralized deep reinforcement learning, IEEE Transactions on Consumer Electronics 70 (1) (2024) 1519–1530. doi:10.1109/TCE.2023.3339468.

[76] S. Bengio, O. Vinyals, N. Jaitly, N. Shazeer, Scheduled sampling for sequence prediction with recurrent neural networks, arXiv preprint arXiv:1506.03099 (2015).

[77] G. Raskutti, M. J. Wainwright, B. Yu, Early stopping and non-parametric regression: an optimal data-dependent stopping rule, The Journal of Machine Learning Research 15 (1) (2014) 335–366.

[78] J. L. Ba, J. R. Kiros, G. E. Hinton, Layer normalization, arXiv preprint arXiv:1607.06450 (2016).

[79] S. Chen, S. Zhang, H. Geng, Y. Chen, C. Zhang, J. Min, Strong spatiotemporal radar echo nowcasting combining 3dcnn and bi-directional convolutional lstm, Atmosphere 11 (6) (2020) 569. doi:https://doi.org/10.3390/atmos11060569.
URL https://www.mdpi.com/2073-4433/11/6/569

[80] C. Luo, X. Li, Y. Wen, Y. Ye, X. Zhang, A novel lstm model with interaction dual attention for radar echo extrapolation, Remote Sensing 13 (2) (2021) 164. doi:https://doi.org/10.3390/rs13020164.
URL https://www.mdpi.com/2072-4292/13/2/164