

Article

Not peer-reviewed version

A Survey of Deep Learning-Based Lightning Prediction

[Xupeng Wang](#), [Keyong Hu](#)^{*}, Yongling Wu, Wei Zhou

Posted Date: 19 September 2023

doi: 10.20944/preprints202309.1232.v1

Keywords: lightning prediction; deep learning; spatio-temporal features; convolutional neural networks; long short-term memory networks




Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

A Survey of Deep Learning-Based Lightning Prediction

Xupeng Wang , Keyong Hu *, Yongling Wu and Wei Zhou

School of Information and Control Engineering, Qingdao University of Technology, 777 Jialingjiang Road, Huangdao District, Qingdao, People's Republic of China, 266520; wxp1130@yeah.net (X.W.); ylwu318@outlook.com (Y.W.); zhouwei@qut.edu.cn (W.Z.)

* Correspondence: hukeyong@qut.edu.cn

Abstract: The escalation of climate change and the increasing frequency of extreme weather events have amplified the importance of precise and timely lightning prediction. This predictive capability is pivotal for the preservation of life, protection of property, and maintenance of crucial infrastructure safety. Recently, the rapid advancement and successful application of data-driven deep learning across diverse sectors, particularly in computer vision and spatio-temporal data analysis, have opened up innovative avenues for enhancing both the accuracy and efficiency of lightning prediction. This article presents a comprehensive review of the broad spectrum of existing lightning prediction methodologies. Starting from traditional numerical forecasting techniques, we traverse the path to the most recent breakthroughs in deep learning research. We encapsulate these diverse methods, shedding light on their progression and summarizing their capabilities, while also predicting their future development trajectories. This exploration is designed to enhance our understanding of these methodologies, allowing us to better utilize their strengths, navigate their limitations, and potentially integrate these techniques to create novel and powerful lightning prediction tools. Through such endeavors, our aim is to bolster our preparedness against the growing unpredictability of our climate and ensure a proactive stance towards lightning prediction.

Keywords: lightning prediction; deep learning; spatio-temporal features; convolutional neural networks; long short-term memory networks

1. Introduction

Weather fluctuations have long been an unpredictable force, and humankind's desire to control such atmospheric caprice dates back to antiquity. Cinematic depictions such as in the "Thor" franchise often portray weather manipulation, particularly lightning, as a profound testament to power. The immense destructive force associated with lightning encapsulates its menacing presence in our collective consciousness. Presently, no effective mechanisms exist to prevent or mitigate the intensity of lightning, rendering lightning prediction an imperative field of study. Since the 1950s, technological advancements encompassing computers, radar, lasers, remote sensing, and artificial satellites have catalyzed meteorological developments, significantly enhancing our capabilities in lightning prediction.

Global satellite data indicates approximately 46 lightning strikes occurring every second across the globe, leading to over 10,000 fatalities and economic damages exceeding one billion USD annually. In response to these sobering statistics, a holistic meteorological monitoring system has been instituted, integrating atmospheric background stations, climate observatories, terrestrial automated meteorological stations, high-altitude meteorological stations, next-generation weather radars, and meteorological satellites. This amalgamation of technologies provides a comprehensive framework for lightning observations.

Traditional lightning prediction methodologies primarily pivot on the monitoring of atmospheric parameters and meteorological phenomena. By interpreting weather data, these approaches identify meteorological conditions conducive to lightning. Numerical model prediction, underpinning

most traditional lightning forecasting methods, exploits principles from atmospheric dynamics, thermodynamics, and related fields to formulate models that simulate atmospheric processes. However, the accuracy of these models requires enhancement due to the complex multi-scale physical processes implicated in lightning.

The advent of data-driven machine learning and the conceptualization of deep learning, proposed by Hinton G.E. et al. in 2006 [1], spurred a rapid evolution of machine learning methods. Deep learning, a subfield of machine learning, comprises multilayer perceptions (MLP) [2], with its primary objective being to mimic the workings of the human brain. It strives to learn data representation techniques, encompassing inherent data correlations, to attain the level of cognitive processing comparable to the human brain, thereby endowing machines with human-like learning and analytical capabilities [3].

Deep learning employs artificial neural networks (ANN) [4] for data learning and representation. Characterized by multiple layers, it represents data at varying levels of abstraction, hence capturing intricate data patterns and structures. Neurons, the basic computational units of deep learning, simulate certain functionalities of biological neurons. The input layer accepts raw data, the hidden layer extracts features and processes information, and the output layer delivers the final prediction. Forward propagation begins with the input layer feeding data into the neural network, progressing towards the output layer, computing layer by layer to yield prediction results. The accuracy of these results is evaluated using a loss function, such as mean squared error or cross-entropy. A smaller loss function indicates superior prediction performance. The loss function is minimized via backpropagation, an optimization algorithm based on gradient descent. To boost the model's generalization capabilities, regularization strategies and optimization algorithms can be integrated into the neural network training process.

With Professor Hinton's team securing the championship in the 2012 ImageNet image recognition competition, deep learning began to supersede traditional machine learning algorithms across many domains, emerging as a sought-after research direction within artificial intelligence. Deep learning is widely utilized in areas such as computer vision and time series data prediction, extending its reach to lightning prediction.

This article aims to provide a thorough review of lightning prediction methodologies. It begins with a fundamental understanding of lightning prediction principles, delves into traditional numerical prediction methods, explores lightning prediction through conventional machine learning, and then delves deeply into the crux of the article – deep learning-based lightning prediction, which currently dominates lightning prediction research. This includes the exploration of recurrent neural networks, convolutional neural networks, and hybrid neural network methodologies. The final section encapsulates the salient features of the discussed methodologies and envisages future trends in deep learning-based lightning prediction.

2. Numerical Prediction methods

Numerical prediction methodologies for forecasting lightning occurrences are grounded in the mechanisms governing lightning generation. Typically, these methodologies utilize parameterization schemes within the Weather Research and Forecasting (WRF) model. Serving as a versatile meteorological prediction tool, the WRF model encompasses numerous applications, ranging from real-time numerical weather prediction (NWP) and weather event research to the development of simulation models and parameterized physics. Additionally, it contributes to studies of regional climate simulations, air quality modeling, atmosphere-ocean coupling, and idealized atmospheric research [5]. The WRF model houses four critical lightning parameters, employed for estimating the Lightning Potential Index (LPI) and the frequency of lightning events. These parameters serve as key indicators in predicting the possibility and the frequency of lightning occurrence, thereby contributing significantly to the field of meteorological prediction and research.

LPI: LPI [6] is defined as the volume integral of the total mass flux of liquid water and ice in the "charging zone" (0°C to -20°C) within a thundercloud. It typically represents the likelihood of a

thundercloud separating electric charges through the latent ice pellet mechanism. LPI varies over time, as it is estimated from the kinematic and microphysical model fields at each grid point and time step. In short, LPI measures the likelihood of charge generation and separation during lightning initiation. The basic formula is as follows:

$$LPI = \frac{1}{V} \iiint \varepsilon w^2 dx dy dz \quad (1)$$

where V represents the cloud volume in the "charging zone," w denotes the vertical wind component (ms⁻¹), and ε is a dimensionless number that depends on the mixing ratio of hydrometeor components, ranging from 0 to 1. The main data considered are the mixing ratios of cloud ice, snow, and graupel.

Lightning frequency: Price and Rind (1992, PR92) [7] developed a simple parameterization to simulate the global distribution of lightning. As lightning activity in convective clouds is positively correlated with the intensity of updrafts, and the intensity of updrafts is related to cloud top height, different lightning parameters were established for continental and oceanic storms due to the differences in their cloud characteristics. In both cases, convective cloud top height is considered as a variable. The parameterization for continental storms is defined as follows:

$$F_c = 3.44 * 10^{-5} H^{4.9} \quad (2)$$

$$w_c = 1.49 H^{1.09} \quad (3)$$

In this case, F_c represents the continental lightning frequency (in lightning strikes per minute), H is the height of the convective cloud top, and w_c is the maximum updraft strength. The lightning parameterization for oceanic convective clouds follows a similar form, which is:

$$F_m = 6.4 * 10^{-4} H^{1.73} \quad (4)$$

$$w_m = 2.86 H^{0.38} \quad (5)$$

Price and Rind (1993, PR93) [8] used radiosonde data collected from 17 observation stations in the western United States to perform lightning predictions. PR93 primarily focused on two parameters in lightning: the freezing level height (H_0) and the cold cloud thickness dH (the thickness from 0°C to the cloud top, which is: $dH = H_{top} - H_0$). The proportion Z is derived from the parameterization of the total flash (IC+CG) frequency and the observed cloud-to-ground flash frequency and is defined as follows:

$$Z = 0.021(dH)^4 - 0.648(dH)^3 + 7.49(dH)^2 - 36.54dH + 63.09 \quad (6)$$

Price and Rind (1994, PR94) [9] introduced the calibration factor parameter, specifically designed to address the dependency of PR92 on latitude, longitude, and season. They used the global convective cloud data from 1983 to 1990 at 3-hour intervals provided by the International Satellite Cloud Climatology Project. By studying different grid sizes, they reached the following conclusions (lightning at 5-km² resolution divided by lightning at lower resolution):

$$c = 0.97241 \exp(0.048203R) \quad (7)$$

where c represents calibration factor and R is the grid area in squared degrees.

Lynn B et al. modified the original LPI to predict the hourly lightning flash density [10]. They have made the LPI adaptable for grid scales from 1 to 4 km. Case studies confirm its applicability on both ultra-high resolution (1.33 km) and high-resolution (4 km) grids. Gharaylou M et al. examined the performance of potential difference (POT) and LPI in predicting lightning activity [11]. Derived from the WRF model, LPI showed superior predictive accuracy compared to POT, which uses the ELEC model package. Analyzing a decade of Tehran data, both indices matched actual lightning locations, with LPI being more precise. However, the simulated lightning flash count didn't correlate

significantly with WWLLN (The World Wide Lightning Location Network) data, possibly due to the absence of WWLLN observatories in the area.

3. Traditional Machine Learning Methods

Machine learning harnesses computational learning methodologies to augment system performance through experiential learning. This technology equips machines with the capability to discern patterns within complex datasets and to predict prospective behaviors, outcomes, and trends [12]. Traditional machine learning algorithms encompass Decision Trees (DT) [13], Support Vector Machines (SVM) [14], Random Forests (RF) [15], Naive Bayes [16], and simple ANN. These methods are heavily reliant on manual feature extraction, with the manual construction of features being an essential component when employing these traditional machine learning algorithms in lightning prediction.

Azad M et al. put forth a hybrid model aiming to predict the monthly frequency of lightning occurrences [17]. This model commences with a random forest to sieve out 11 impactful predictive features from a pool of 21 feature parameters. Subsequently, the Ensemble Empirical Mode Decomposition (EEMD) [18] is deployed to deconstruct the original time-series data into a finite set of Intrinsic Mode Functions (IMF) [19] and residuals. Irrelevant or redundant IMFs and residuals are discarded, with the selection of IMFs primarily being those that have a higher frequency band and display a significant correlation with the original sequence. These chosen IMFs, combined with other input parameters, serve as inputs to the ANN or SVM models. These models are then used to construct a predictive model for the frequency of lightning occurrences. This approach integrates the strengths of multiple techniques to increase the accuracy and robustness of lightning frequency prediction. The ANN adopts a backpropagation (BP) neural network, and its model mechanism is as follows:

$$y(x) = \sum_{j=1}^n w_j(p) * x_j(p) + c \quad (8)$$

The mechanism of the SVM model is as follows:

$$L(m, m_j) = \exp\left(-\frac{(|m - m_j|^2)}{2\phi^2}\right) \quad (9)$$

Schön C et al. proposed an innovative approach to lightning prediction that utilizes prediction errors derived from lightning forecasting models [20]. Traditional forecasting models for lightning generally rely on predictions of specific physical parameters, such as temperature and humidity. However, these models often face difficulties in accurately predicting the timing and location of lightning events. To address this issue, the authors have introduced a new methodology that employs prediction errors of lightning forecasting models as features. Machine learning algorithms are then utilized to predict lightning occurrences. The authors argue that these prediction errors encompass substantial information regarding atmospheric conditions and other influential factors, which could be pivotal in predicting the timing and location of lightning events. In their experiments, a machine learning model was developed based on the gradient boosting algorithm. This model used prediction errors and other relevant meteorological data to enhance the accuracy of lightning prediction.

Pakdaman M et al. employed a simple ANN and a decision tree as two binary classification algorithms to predict lightning occurrences [21]. Due to the imbalance in the acquired source data, down-sampling techniques were initially applied to rectify this issue, resulting in several balanced data subsets. These subsets were then input into a shallow neural network (with one hidden layer) and a decision tree, respectively, for classification predictions. The prediction results were evaluated based on metrics such as Precision, Recall, Accuracy, and F-measure. Johari D et al. proposed a simple ANN in predicting the occurrence of lightning [22]. A two layer back-propagation neural network has been developed to predict the occurrence of lightning at least four hours prior to its

arrival. Through continuous training and adjusting the activation functions and network structure, a network with high accuracy and convergence capability was achieved. Moon S. H. and colleagues employed machine learning techniques, specifically Support Vector Machines and Random Forests, to predict the likelihood of lightning occurrences within specific locations and time intervals [23]. Their data was sourced from the European Centre for Medium-range Weather Forecasts, focusing on the Korean Peninsula region. To enhance prediction accuracy, they undersampled the data points that were not linked with lightning events.

Overall, traditional machine learning methods have been shown to offer significantly greater accuracy and efficiency than conventional lightning prediction methodologies. These methods primarily revolve around constructing features from lightning observation data, which often necessitates specialized knowledge. In this respect, deep learning holds a significant advantage, due to its capacity to learn and extract meaningful features from complex datasets autonomously.

4. Deep Learning Methods

4.1. Convolutional Neural Network Methods

Convolutional Neural Networks (CNNs) represent a specialized category of feedforward neural networks characterized by their convolutional computations and deep structures. CNNs possess the ability to extract features directly from raw data, obviating the need for complex preprocessing, thereby finding extensive application in the field of computer vision. The distinctiveness of CNNs lies in two facets: first, the lack of full connectivity between neurons; second, the sharing of weights across neurons within the same layer. This configuration of partial connectivity and weight-sharing brings CNNs closer to the functioning of biological neural networks, effectively reducing both the complexity of the network model and the number of weights [24]. In modern meteorological systems, geostationary satellites hold a pivotal role, with satellite images often serving as a common medium for lightning prediction. A noteworthy example of this can be seen in the work of Sebastian Brodehl et al., who proposed an end-to-end CNN-based lightning prediction method that employs geostationary satellite images [25]. This approach allows for direct prediction of lightning events from satellite images. The network architecture is underpinned by U-Net [26] and complemented by ResNet-v2 [27] residual blocks, adapted to handle three-dimensional input - encompassing the height and width of an image, along with time-frames. In this model, stacked convolution layers and the pooling layer at each down/up sampling step are replaced with residual blocks. While convolutions are used for down-sampling, deterministic trilinear up-sampling operations are implemented. Batch normalization layers are supplanted with instance normalization, and instead of cropping the feature map of the contracting path, the feature map post the down-sampling operation is used. The model, demonstrated in Figure 1, maintains a fixed height and width of 256px*256px and is capable of handling imbalanced data without any requirement for pre-processing. Most information is drawn from structures in the visible spectrum, with infrared imaging providing some degree of classification performance during night-time. Furthermore, an attention mechanism-based enhanced model is proposed, which more effectively captures the spatio-temporal features of lightning events, thereby improving prediction accuracy and interpretability.

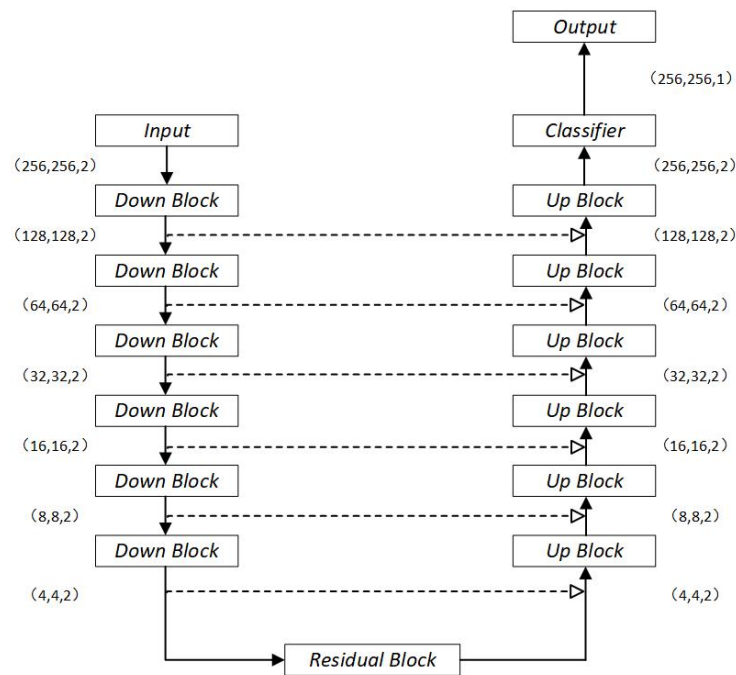


Figure 1. End-to-end CNN-based lightning prediction network architecture.

To accommodate time-series data as input for CNNs, Bao R et al. conducted feature reconstruction on time-series atmospheric electric field observation data [28]. The time-series observation data were treated as the horizontal coordinate and the observation sites as the vertical coordinate, forming a two-dimensional image as depicted in Figure 2. A straightforward three-layer network was utilized to convert the time-series data into image data, with the sigmoid activation function transforming the original data into a range of 0-1. The KL divergence [29] was integrated as a penalty mechanism in this process. The visualized data thus produced was used as input for the CNN, which was based on an enhanced ResNet50 architecture. The main branch employed 3×3 convolutional kernels for down-sampling, while the shortcut branch incorporated a 2×2 max-pooling layer. This network could discern between lightning and non-lightning weather. In instances of a lightning event, the location of the lightning could be determined using a MLP, which comprises an input layer, multiple hidden layers, and an output layer. The MLP, primarily used for classification and prediction of nonlinear relationships, takes the spatial position of the lightning as its input and produces the impending lightning location as its output.

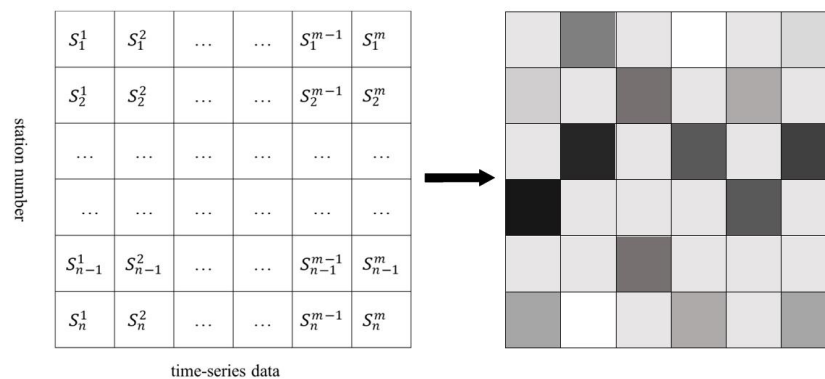


Figure 2. Time-series observation data visualization.

In their study, Sashiomarda J A et al. leveraged electromagnetic and acoustic signals from lightning in tandem with a CNN to predict the precise location of lightning events [30]. They converted a

medley of random noise, lightning sounds, and background noise into spectrograms and subsequently sectioned these into a multitude of millisecond-scale audio signals. Each of these slices underwent a discrete fourier transform, a process that was also applied to the electromagnetic signals emanating from the lightning. These transformed spectrograms served as inputs for the CNN, which utilized max-pooling and ReLU activation functions. The culmination of this process yielded three distinct categories: "quiet", "unknown", and "lightning". By pinpointing the origin of the detected sounds, the researchers were able to determine not only the occurrence of a lightning event but also its exact location.

Lu M et al. used 3D weather radar data to predict the location of lightning strikes [31]. They used a $M \times N$ sliding window to obtain radar data samples, one of which contains nine layers. If the lightning data point is located in the center grid of the sliding window, it is considered that the lightning occurred in that area; otherwise, it didn't. In this way, predicting the location of lightning becomes a binary classification problem. They used methods such as CNN, logistic regression [32], random forest, and k-nearest neighbor [33] for prediction. Among them, the CNN performed the best. The structure of the CNN is shown in Figure 3. The CNN consists of seven layers, with two convolutional layers and two pooling layers appearing alternately. There are three fully connected (FC) layers connected to the last feature map.

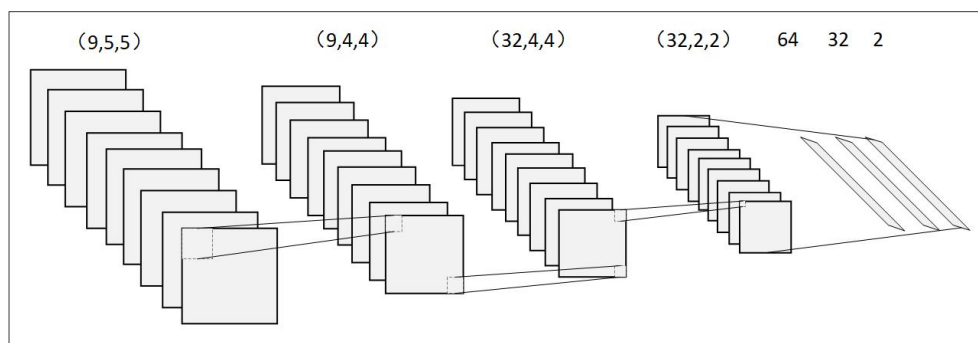


Figure 3. [32] CNN architecture.

4.2. Recurrent Neural Network Methods

Recurrent Neural Networks (RNNs) are a type of recursive neural networks specifically designed to process sequence data. They accomplish this by iteratively processing the sequence in a sequential direction, with all nodes (or recurrent units) connected in a chain-like pattern. The unique trait of RNNs is that their input comprises not only the current data but also information from previous steps. When it comes to lightning prediction, RNNs often employ Long Short-Term Memory (LSTM) networks. LSTMs were created to tackle the challenge of long-term dependencies often encountered in standard RNNs [34], shown as Figure 4. By feeding the spatio-temporal features of lightning data into the LSTM, reliable lightning predictions can be made. Presently, LSTM is the most frequently used deep learning network model for lightning prediction.

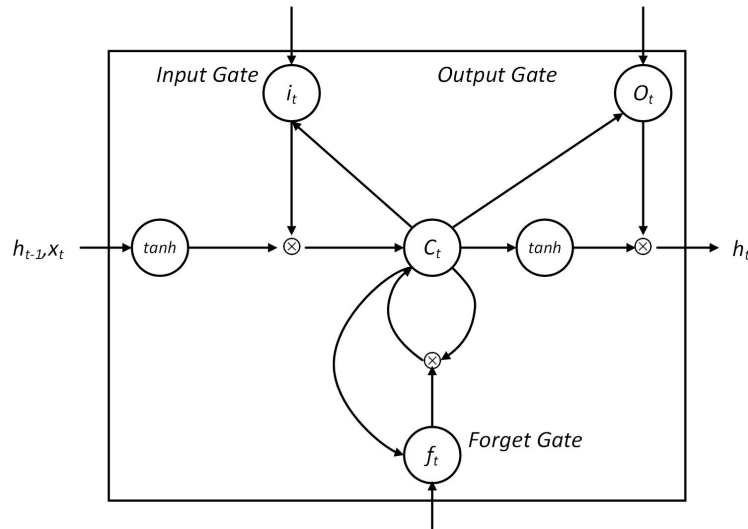


Figure 4. Typical LSTM block.

Bao R et al. employed a spatio-temporal localization method leveraging LSTM neural networks and interpolation techniques to predict and monitor lightning activity [35]. The team utilized time-series data captured from 30 atmospheric electric field instruments as LSTM input, with a softmax function categorizing the results into five classes to predict lightning timing. In order to predict the location of the lightning, they enhanced the prediction accuracy by incorporating data from multiple networked atmospheric electric field instruments. They also employed the Ordinary Kriging (OK) interpolation method [36] — one of the most commonly used Kriging methods — to derive the electric potential distribution and infer the likely area of impending lightning. By integrating these two methodologies, they achieved satisfactory prediction outcomes.

Fukawa M et al. proposed a novel method for lightning prediction by direct electric field data [37]. They collected data on both fair-weather and lightning-weather as the input of the many-to-one network. This network model is based on LSTM, and the specific network structure is shown in Figure 5. Mean Absolute Percentage Error (MAPE) is applied to measure accuracy and anomaly score α to control the output. If fair-weather α should be 0, it should be exposed to sudden changes in the case of lightning. Meanwhile Electric field sum β is intended to accurately predict the electric field value in a fair-weather.

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{\hat{v}_i - v_i}{v_i} \right| \quad (10)$$

$$\alpha = \log_2 \left(\sum_i^{i+300} |\hat{v} - v| \right) \quad (11)$$

$$\beta = \log_2 \left(\sum_i^{i+300} |v| \right) \quad (12)$$

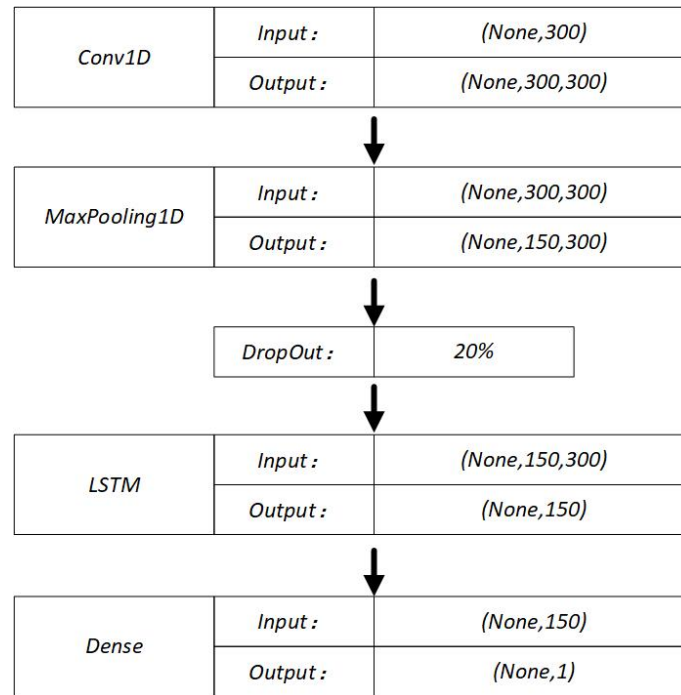


Figure 5. [37] Network architecture.

Extending upon the popular fully connected LSTM, Shi X et al. introduced a Convolutional LSTM (ConvLSTM) model, primarily to cater to spatio-temporal data features [38]. The ConvLSTM morphs the 2D input in LSTM into a 3D tensor, with the last two dimensions symbolizing spatial dimensions (rows and columns). For data at each time instance 't', ConvLSTM replaces certain connection operations in LSTM with convolution operations. This implies that ConvLSTM predictions are based on the current input and the past states of its local neighbors. In ConvLSTM, $X_1 \dots X_t$ represent inputs, $C_1 \dots C_t$ represent cell outputs, and $H_1 \dots H_t$ represent hidden outputs. g_t , i_t and f_t represent the three gates and are three-dimensional tensors, with the last two dimensions being spatial components. The calculation formulas for the outputs of each layer and the related gates are as follows:

$$g_t = \tanh(W_{xg} * X_t + W_{hg} * H_{t-1} + b_i) \quad (13)$$

$$i_t = \sigma(W_{xi} * X_t + W_{hi} * H_{t-1} + W_{ci} \circ C_{t-1} + b_i) \quad (14)$$

$$f_t = \sigma(W_{xf} * X_t + W_{hf} * H_{t-1} + W_{cf} \circ C_{t-1} + b_f) \quad (15)$$

$$C_t = f_t \circ C_{t-1} + i_t \circ g_t \quad (16)$$

$$o_t = \sigma(W_{xo} * X_t + W_{ho} * H_{t-1} + W_{co} \circ C_t + b_c) \quad (17)$$

$$H_t = o_t \circ \tanh(C_t) \quad (18)$$

where σ represents the sigmoid activation function, $*$ denotes the convolution operation, and \circ signifies the Hadamard product. The input gate i_t , forget gate f_t , output gate o_t , and input modulation gate g_t control the flow of information in the memory cell C_t .

Building upon the ConvLSTM, Geng Y et al. have proposed a composite architecture called LightNet, which integrates the Weather Research and Forecasting (WRF) model [39]. The LightNet structure is divided into four components: the WRF encoder, the observation encoder, the fusion module, and the prediction decoder, shown in Table 1. The WRF encoder is primarily responsible for processing simulation data produced by the WRF model. This data includes various parameters, such as the cloud ice mixing ratio, snow mixing ratio, graupel mixing ratio, radar reflectivity, and the maximum vertical wind component. On the other hand, the observation encoder is designed to

handle the processing of observational data. Both encoders adopt the same approach, going through a convolutional layer and inputting the convolution results into the ConvLSTM network to obtain two tensors C^{wrf}, H^{wrf} and C^{obs}, H^{obs} . The fusion module applies a convolution operation to the WRF data and observation data, respectively, resulting in the final initialized input tensors C^{fuse} and H^{fuse} . The prediction decoder acquires the fused features and makes predictions. All variables are initialized as C^{fuse} and H^{fuse} , as well as the input $Conv_5(L_{-1})$. The input and output at time $t > 0$ are:

$$X_t = Conv_5(\hat{L}_{t-1}) \quad (19)$$

$$\hat{L}_t = softmax(Conv_6(DeConv(Y_t))) \quad (20)$$

where \hat{L}_t is the final result.

Table 1. LightNet architecture.

Module	Notation	Size	Stride
WRF Encoder	$Conv_1$	[7x7,64]	2
	ConvLSTM	[5x5,128]	1
Obs. Encoder	$Conv_2$	[7x7,4]	2
	ConvLSTM	[5x5,8]	1
Fusion Module	$Conv_3$	[1x1,64]	1
	$Conv_4$	[1x1,64]	1
Pred. Decoder	$Conv_5$	[7x7,4]	2
	ConvLSTM	[5x5,64]	1
	DeConv	[7x7,64]	2
	$Conv_6$	[1x1,1]	1

Geng Y et al. introduced a heterogeneous spatio-temporal network (HSTN) for lightning prediction, aimed at extracting knowledge from multiple heterogeneous spatio-temporal data sources [40]. The HSTN consists of three modules: the Gaussian Diffusion module, the Spatio-temporal Encoder, and the Spatio-temporal Decoder. the Gaussian diffusion module converted the sparse tensor \tilde{S} into a dense form S (the weather station observation data, real four-dimensional data). The Gaussian diffusion module, founded upon the principles of Gaussian distribution, undergoes a series of mathematical transformations and derivations to yield the resultant dense tensor. Three spatio-temporal encoders extract information from W (the WRF simulation data, real four-dimensional data), L (the lightning observation data, binary three-dimensional data) and S . The spatio-temporal encoder is built based on ConvLSTM through initializing all states of ConvLSTM as zero. The spatio-temporal decoder merges all information and produces lightning prediction. Meanwhile the multi-scale pooling loss is employed to address the shortsightedness issue caused by grid-wise losses. The multi-scale pooling involves applying maximum pooling separately to both the predicted results and the actual outcomes, and employing weighted cross entropy to balance the disparities between lightning and non-lightning grids. The architecture of HSTN is shown as Figure 6.

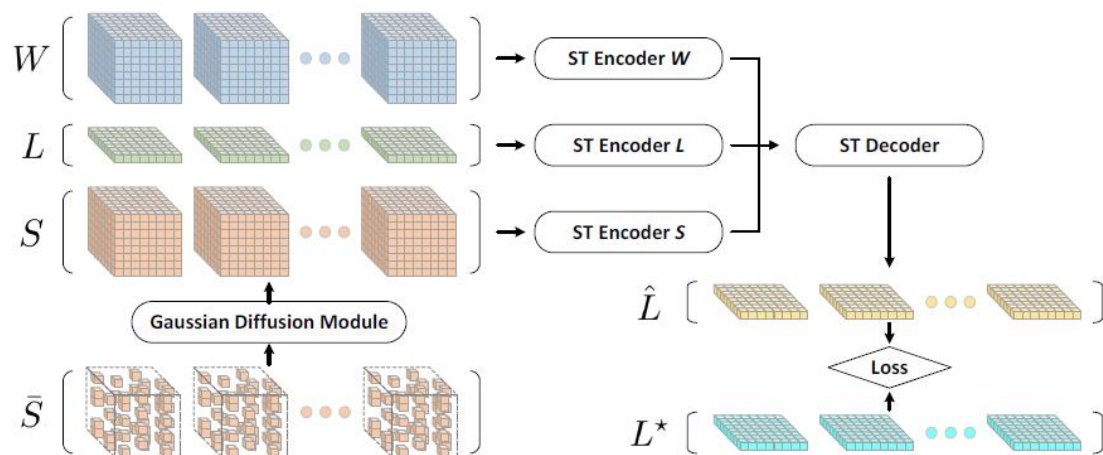


Figure 6. HSTN architecture.

Essa Y et al. conducted a study on the predictive capability of LSTM model in predicting short-term lightning flash density within South Africa [41]. The research focused on predicting lightning flash densities over intervals of one hour, three hours, and 24 hours. The architecture of their LSTM model comprised an initial layer with 50 units, followed by two layers each containing 25 units, and a final dense layer employing a Leaky ReLU activation function. The model optimization was performed using the Adam optimizer, targeting the minimization of the Mean Square Error (MSE). However, the model's performance was evaluated based on both Mean Absolute Error (MAE) and MSE.

5. Hybrid Neural Network Methods

Hybrid neural networks combine different types of neural networks, leveraging the strengths of each to better accommodate varying data types and task requirements, thereby enhancing the model's performance. Given the complexity and variety of lightning observation data, employing hybrid neural network models for lightning prediction is a prevalent approach. The most common amalgamation involves coupling CNNs with RNNs or LSTM networks.

Guastavino S et al. put forth a Long-Term Recurrent Convolutional Network (LRCN) which fuses CNN and LSTM networks to devise spatio-temporal deep learning models using radar data for predicting lightning occurrences [42]. The CNN is employed to extract spatial features from the dataset. These features are subsequently broken down into sequential components and supplied to the LSTM network for analysis. Finally, the output from the LSTM layer is passed into the fully connected layer, where the sigmoid activation function is used to generate the probability distribution of the positive class.

Zhou X et al. extended upon lightnet, proposing lightnet+ [43]. This network represents a sophisticated hybrid neural network. It is composed of three modules: the SpatioTemporal (ST) encoder module for lightning observations, the Bi-Directional SpatioTemporal (BDST) propagator module for WRF simulations, and the Non-Local SpatioTemporal (NLST) for lightning forecasts. The ST module initially employs a convolution layer to subsample the raw data, followed by a ConvLSTM that extracts trend information from the compact lightning observation features. The BDST module is bifurcated into two components: the subsampling module and the Dual ConvLSTM (DCLSTM). This module also uses a convolution layer for subsampling the WRF simulation data, followed by the DCLSTM to fully harness information both preceding and succeeding the t -th hour. The structure of the DCLSTM is illustrated in Figure 7. For the NLST module, a Non-Local Fusion Unit (NLFU) is introduced. The NLFU initially uses a convolution operation with a kernel size of 1×1 to extract features after amalgamating the historical lightning state trend, the past WRF trend information, and the future trend data into a comprehensive trend profile. This is followed by a two-stage attention

mechanism to uncover informative long-range correlations. Finally, the data passes through the ConvLSTM module and a Deformable CNN, and it is subsampled using deconvolution (deConv) to produce the final output.

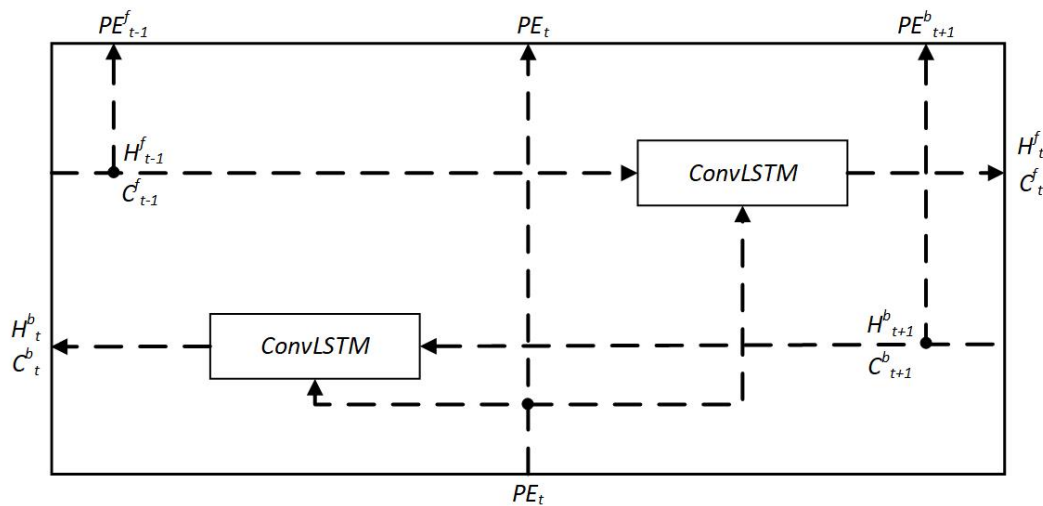


Figure 7. DCLSTM architecture: PE represents the result after subsampling.

6. Conclusions

From the above description, we can delve deeply into the role of data-driven deep learning in the field of lightning prediction. Evidently, deep learning methods, especially CNN and RNN, have demonstrated exceptional performance when dealing with complex multi-dimensional spatial and temporal data.

Traditional lightning prediction methods might rely on heuristic rules or simplified physical models, which might be inadequate to capture the full complexity of lightning formation. In contrast, CNNs and RNNs are capable of learning and extracting meaningful features from vast historical and real-time data, leading to more accurate predictions of lightning occurrences. This capability significantly surpasses the accuracy of conventional methods.

Furthermore, as more meteorological data becomes available and the resolution and quality of the data continually improve, we can anticipate that deep learning methods will continue to be optimized and adapt to these new data sources. This implies that in the future, we can rely on these advanced algorithms to provide us with earlier and more precise lightning alerts, thereby safeguarding lives and property.

In conclusion, data-driven deep learning, especially CNNs and RNNs, has not only made significant contributions to lightning prediction but also holds tremendous potential and prospects for future applications. Table 2 provides a summary of the lightning prediction methods outlined in this article.

Table 2. Summary of lightning prediction methods.

Methods	Generalize
Numerical Prediction methods	In accordance with the principles of lightning genesis, relevant observation parameters are used to calculate the Lightning Potential Index (LPI) and Positive Ratio (PR), determining the probability of lightning occurrence.
Traditional Machine Learning Methods	Employing manual calculations, key lightning data features are hand-extracted, and then traditional machine learning methods such as Support Vector Machines (SVM) and Simple Artificial Neural Networks (Simple ANN) are utilized based on the extracted features for classification.
Convolutional Neural Network Methods	Integrating image data such as satellite images, electromagnetic and acoustic signals, or converting lightning data into image form allows the utilization of convolution to extract features and conduct predictions.
Recurrent Neural Network Methods	The most commonly used deep learning method for lightning prediction primarily processes sequential data, often combining with existing methodologies in a variant form to achieve more accurate predictions.
Hybrid Neural Network Methods	Combining multiple neural network models often involves the use of one network model for feature extraction and another for prediction. With more information contained in the features, this approach aids in improving the accuracy of predictions.

7. Discussion

Deep learning-based lightning prediction primarily focuses on directly analyzing observed parameters, often without selection or filtering of these parameters. The accuracy of such models remains an area that warrants enhancement. This straight-forward methodology could overlook intricate interrelationships among the observational parameters or the underlying physics driving lightning occurrences.

1. Parameter Selection Based on Lightning Mechanisms: Rather than indiscriminately using all observed parameters, a more discerning approach could involve parameter selection based on the underlying mechanisms of lightning formation. This would involve a deeper understanding of meteorology and the physics of atmospheric electricity, thereby ensuring that the most relevant and influential parameters are input into the model. Such a focused approach might eliminate noise and irrelevant information, potentially enhancing prediction accuracy.
2. Multimodal Data Fusion: Another promising avenue is the fusion of multimodal data. Instead of analyzing disparate data forms independently, combining various forms of data organically can provide a holistic view. By coalescing data from different sensors or observational platforms, deep learning networks can be exposed to a richer set of inputs, potentially enhancing their capacity to discern patterns and make accurate predictions.
3. Leveraging Adjacent Observational Nodes: Incorporating data from neighboring observational nodes can enrich the contextual understanding of local atmospheric conditions. This spatial embedding ensures that the localized conditions leading to lightning formation are not viewed in isolation but are seen in the context of broader atmospheric dynamics. Integrating neighboring node data with the current node can offer a more comprehensive input for the network, potentially boosting predictive accuracy.
4. Expert-Guided Deep Learning Prediction: Lastly, instead of relying solely on algorithmic predictions, integrating expert knowledge from traditional lightning prediction methods can guide and refine deep learning-based forecasts. Such an interdisciplinary approach can combine the strengths

of empirical knowledge and computational prowess. Furthermore, predictions can be provided with expert interpretations, lending more credibility and context to the forecasts.

In conclusion, while deep learning offers a promising toolset for lightning prediction, its integration with domain expertise, data fusion techniques, and a keen understanding of the intricacies of observational parameters can push the boundaries of prediction accuracy. Leveraging these strategies can pave the way for more reliable, insightful, and actionable lightning forecasts in the future.

Author Contributions: Conceptualization, Xupeng Wang; investigation, Keyong Hu; resources, Yongling Wu; writing—original draft preparation, Xupeng Wang; writing—review and editing, Xupeng Wang and Wei Zhou; supervision, Keyong Hu; funding acquisition, Keyong Hu.

Funding: This work was supported by National Natural Science Foundation of China under Grants 61902205 and 42201506.

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

References

1. Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets. *Neural computation*. **2006**, 18(7), 1527–1554.
2. Taud H, Mas J F. Multilayer perceptron (MLP). *Geomatic approaches for modeling land change scenarios*. **2018**, 451–455.
3. LeCun Y, Bengio Y, Hinton G. Deep learning. *nature*. **2015**, 521(7553), 436–444.
4. Krogh A. What are artificial neural networks? *Nature biotechnology*. **2008**, 26(2), 195–197.
5. Skamarock W C, Klemp J B, Dudhia J, et al. A description of the advanced research WRF model version 4. *NCAR tech. note ncar/tn-556+ str.* **2019**, 145.
6. Yair Y, Lynn B, Price C, et al. Predicting the potential for lightning activity in Mediterranean storms based on the Weather Research and Forecasting (WRF) model dynamic and microphysical fields. *Journal of Geophysical Research: Atmospheres*. **2010**, 115(D4).
7. Price C, Rind D. A simple lightning parameterization for calculating global lightning distributions. *Journal of Geophysical Research: Atmospheres*. **1992**, 97(D9), 9919–9933.
8. Price C, Rind D. What determines the cloud-to-ground lightning fraction in thunderstorms? *Geophysical Research Letters*. **1993**, 20(6), 463–466.
9. Price C, Rind D. Possible implications of global climate change on global lightning distributions and frequencies. *Journal of Geophysical Research: Atmospheres*. **1994**, 99(D5), 10823–10831.
10. Lynn B, Yair Y. Prediction of lightning flash density with the WRF model. *Advances in Geosciences*. **2010**, 23, 11–16.
11. Gharaylou M, Farahani M M, Hosseini M, et al. Numerical study of performance of two lightning prediction methods based on: Lightning Potential Index (LPI) and electric POTential difference (POT) over Tehran area. *Journal of Atmospheric and Solar-Terrestrial Physics*. **2019**, 193, 105067.
12. Jordan M I, Mitchell T M. Machine learning: Trends, perspectives, and prospects. *Science*. **2015**, 349(6245), 255–260.
13. Kingsford C, Salzberg S L. What are decision trees? *Nature biotechnology*. **2008**, 26(9), 1011–1013.
14. Keerthi S S, Shevade S K, Bhattacharyya C, et al. Improvements to Platt's SMO algorithm for SVM classifier design. *Neural computation*. **2001**, 13(3), 637–649.
15. Cutler A, Cutler D R, Stevens J R. Random forests. *Ensemble machine learning: Methods and applications*. **2012**, 157–175.
16. Webb G I, Keogh E, Miiikkulainen R. Naïve Bayes[J]. *Encyclopedia of machine learning*. **2010**, 15(1), 713–714.
17. Azad M A K, Islam A R M T, Rahman M S, et al. Development of novel hybrid machine learning models for monthly thunderstorm frequency prediction over Bangladesh. *Natural Hazards*. **2021**, 108, 1109–1135.
18. Wang W, Chau K, Xu D, et al. Improving forecasting accuracy of annual runoff time series using ARIMA based on EEMD decomposition. *Water Resources Management*. **2015**, 29, 2655–2675.
19. Sharpley R C, Vatchev V. Analysis of the intrinsic mode functions. *Constructive Approximation*. **2006**, 24, 17–47.

20. Schön C, Dittrich J, Müller R. The error is the feature: How to forecast lightning using a model prediction error. Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery and data mining, Anchorage AK, USA, August 4-8, 2019, 2979–2988.
21. Pakdaman M, Naghab S S, Khazanedari L, et al. Lightning prediction using an ensemble learning approach for northeast of Iran. *Journal of Atmospheric and Solar-Terrestrial Physics*. **2020**, 209, 105417.
22. Johari D, Rahman T K A, Musirin I. Artificial neural network based technique for lightning prediction. 2007 5th Student Conference on Research and Development. Selangor, Malaysia, December 11-12, 2007, 1-5.
23. Moon S H, Kim Y H. Forecasting lightning around the Korean Peninsula by postprocessing ECMWF data using SVMs and undersampling. *Atmospheric research*. **2020**, 243, 105026.
24. Gu J, Wang Z, Kuen J, et al. Recent advances in convolutional neural networks. *Pattern recognition*. **2018**, 77, 354–377.
25. Brodehl S, Müller R, Schömer E. End-to-End Prediction of Lightning Events from Geostationary Satellite Images. *Remote Sensing*. **2022**, 14(15), 3760.
26. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference. Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. Springer International Publishing, 2015, 234–241.
27. He K, Zhang X, Ren S, et al. Identity mappings in deep residual networks. Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14. Springer International Publishing, 2016, 630–645.
28. Bao R, Zhang Y, Ma B J, et al. An Artificial Neural Network for Lightning Prediction Based on Atmospheric Electric Field Observations. *Remote Sensing*. **2022**, 14(17), 4131.
29. Kullback S, Leibler R A. On information and sufficiency. *The annals of mathematical statistics*. **1951**, 22(1), 79–86.
30. Sashiomarda J A, Aji A S, Handoko D. Implementation of convolutional neural networks to determine lightning location. Journal of Physics: Conference Series. IOP Publishing, 2020, 1528(1), 012009.
31. Lu M, Zhang Y, Ma Z, et al. Lightning strike location identification based on 3D weather radar data. *Frontiers in Environmental Science*. **2021**, 9, 714067.
32. LaValley M P. Logistic regression. *Circulation*. **2008**, 117(18), 2395–2399.
33. Peterson L E. K-nearest neighbor. *Scholarpedia*. **2009**, 4(2), 1883.
34. Graves A, Graves A. Long short-term memory. *Supervised sequence labelling with recurrent neural networks*. **2012**, 37–45.
35. Bao R, He Z, Zhang Z. Application of lightning spatio-temporal localization method based on deep LSTM and interpolation. *Measurement*. **2022**, 189, 110549.
36. Wackernagel H, Wackernagel H. Ordinary kriging. *Multivariate geostatistics: an introduction with applications*. **2003**, 79–88.
37. Fukawa M, Deng X, Imai S, et al. A Novel Method for Lightning Prediction by Direct Electric Field Measurements at the Ground Using Recurrent Neural Network. *IEICE TRANSACTIONS on Information and Systems*. **2022**, 105(9), 1624–1628.
38. Shi X, Chen Z, Wang H, et al. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems*. **2015**, 28.
39. Geng Y, Li Q, Lin T, et al. Lightnet: A dual spatiotemporal encoder network model for lightning prediction. Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery and data mining. Anchorage AK, USA, August 4-8, 2019, 2439–2447.
40. Geng Y, Li Q, Lin T, et al. A heterogeneous spatiotemporal network for lightning prediction. 2020 IEEE international conference on data mining (ICDM). Sorrento, Italy, November 17-20, 2020, 1034-1039.
41. Essa Y, Ajoodha R, Hunt H G P. A LSTM recurrent neural network for lightning flash prediction within southern africa using historical time-series data. 2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE). Gold Coast, Australia, November 16-18, 2020, 1-6.

42. Guastavino S, Piana M, Tizzi M, et al. Prediction of severe thunderstorm events with ensemble deep learning and radar data. *Scientific Reports*. **2022**, 12(1), 20049.
43. Zhou X, Geng Y, Yu H, et al. LightNet+: A dual-source lightning forecasting network with bi-direction spatiotemporal transformation. *Applied Intelligence*. **2022**, 52(10), 11147–11159.

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