An IoT-Driven Federated Learning Method for Rainfall Prediction Employing Attention Convolutional Recurrent Networks and Golden Jackal Optimization Techniques

J. Jagadeesan
Department of Computer and Information Science, Annamalai University, Annamalai Nagar, Chidambaram, Tamil Nadu, India
jagatamil2016@gmail.com (corresponding author)

R. Nagarajan
Department of Computer and Information Science, Annamalai University, Annamalai Nagar, Chidambaram, Tamil Nadu, India
rathinanagarajan@gmail.com

Received: 6 May 2025 | Revised: 24 June 2025 and 4 July 2025 | Accepted: 8 July 2025

DOI: https://doi.org/10.48084/etasr.11957

ABSTRACT

Forecasting heavy rainfall poses a significant challenge for meteorological departments, as it is closely related to human life and the economy, being also a reason for the natural calamities, such as droughts and floods, that people face around the world each year. The precision of rainfall prediction is of enormous significance for countries such as India, whose economy is mainly dependent on agriculture. Rainfall prediction supports evading floods, saving properties, and human lives. There is a need to improve cooperative platforms for weather forecasting on large-scale meteorological data to address the global climate challenge. With the rapid improvement of Artificial Intelligence (AI), Machine Learning (ML) is progressively becoming popular for forecasting rainfall. Federated Learning (FL) is a subset of ML that has become a popular technology for data analysis for Internet of Things (IoT) applications. The innovative growth of IoT-based methods improves smart agriculture, leading to advanced agricultural systems that progressively increase crop yields, reduce irrigation waste, and make it more profitable. This study presents a Method for Rainfall Prediction by employing Attention Convolutional Recurrent Networks and the Golden Jackal Optimization (EMRF-ACRNGJO) approach. The proposed technique initially performs data preprocessing in various stages, such as missing value handling, duplicate removal, categorical to numerical encoding, and scaling. An Attention-based Convolutional Recurrent Neural Network (A-CRNN) is utilized for classification tasks, using the Golden Jackal Optimization (GJO) method for the hyperparameter tuning process. The EMRF-ACRNGJO method was tested on a benchmark dataset, demonstrating an accuracy of 73.30%.

Keywords-federated learning; rainfall prediction; Golden Jackal Optimization (GJO); attention convolutional recurrent network; Internet of Things (IoT)

I. INTRODUCTION

Rainfall plays a crucial role in human life, and its prediction is demanding and challenging [1]. Rainfall has a strong impact on drainage systems, traffic, and other human activities in city regions [2]. However, rainfall is one of the most complex and intricate elements of the hydrological cycle to understand and model, due to the complexities of atmospheric developments and the great degree of variability across a wider variety of scales in time or space [3]. Therefore, accurate rainfall forecasting is the ultimate challenge in operational hydrology, despite various developments in weather prediction in recent decades. Rainfall prediction is closely associated with agriculture—which contributes considerably to India's economy—and is important for protecting the properties and lives of humans. In addition, it assists in the management of water resources. Rainfall forecasting helps farmers manage their crops, which affects the development of a country's economy [4, 5]. Although rainfall forecasting is challenging, its impact can be mitigated with IoT technologies, which enable remote sensing and control in agriculture [6]. The IoT allows data gathering or control of objects remotely through various networks. IoT devices collect soil and environmental data, which are analyzed using Machine Learning (ML) and Deep Learning (DL) methods on large datasets [7].

In [8], a real-time rainfall prediction method combined data from different IoT sensors using a CNN with LSTM layers to forecast precipitation. In [9], the use of sensors was proposed to systematize weather prediction and area monitoring. Sensor data were sent wirelessly to a server, preprocessed for missing values and normalization, then used by GRU for weather prediction and ResNet50 for image classification and feature extraction. In [10], an IoT-based smart irrigating method was presented. In [11], a Crop Yield Prediction (CYP) technique was based on IoT. In [12], an IoT-based method utilized a Deep Reinforcement Learning (DRL) method with an LSTM to extract sequential dependencies. In [13], a weather monitoring method used the IoT to extract real-time data, send it to the cloud, and analyze it with a Support Vector Machine (SVM) model to predict rainfall. In [14], a Recurrent Attention Unit (RAU) was used to improve the GRU-Extreme Learning Machine (GRU-ELM) model, using the Gaussian Mutation-Alpine Skiing Optimization Algorithm (GM-ASOA) approach for improved feature selection. In [15], ConvLSTM was integrated with Self-Attention (SA) to improve extreme weather prediction by enhancing spatio-temporal feature extraction and capturing long-range patterns. In [16], a weather model utilized a stacked ConvLSTM network. In [17], a temporal convolutional network was proposed for seasonal short-term energy prediction. In [18], a DL ensemble model was proposed for precipitation forecasting, introducing a Stochastic Differential Equation U-Net (SDE U-Net). In [19], the DL Improved Whale Optimization Algorithm (DL-IWOA) integrated CNN, LSTM, and a self-attention mechanism optimized by the IWOA.

Despite crucial improvements in using CNN, LSTM, GRU, and hybrid DL models for weather and crop forecasting, many methods lack robustness in handling heterogeneous sensor data and long-range temporal dependencies. Various models also do not utilize attention mechanisms for dynamic feature prioritization. There is a research gap in developing unified, efficient architectures that incorporate multi-modal data and improve interpretability for real-time prediction tasks. This study presents a Method for Rainfall Prediction by Employing Attention Convolutional Recurrent Networks and the Golden Jackal Optimization (EMRF-ACRNGJO) approach, trained and tested using a weather AUS benchmark dataset. The proposed method:

- Employs the multi-stage data preparation strategy, including null value handling, duplicate removal, categorical encoding, and feature scaling to improve model input quality.
- Utilizes an A-CRNN model to effectually handle regression and classification tasks by learning spatial and temporal dependencies in complex data structures, improving predictive accuracy and interpretability in sequential modeling.
- Implements the GJO technique to perform precise hyperparameter tuning, improving convergence speed and overall model performance by effectively navigating the search space and mitigating training loss in intrinsic learning scenarios.

The integration of A-CRNN with GJO presents a novel hybrid framework that efficiently captures spatial-temporal features while ensuring optimal parameter tuning. This integration significantly improves predictive accuracy.

II. MATERIALS AND METHODS

Figure 1 shows the architecture of the proposed EMRF-ACRNGJO model.

A. Data Preparation

The proposed method first prepares the data in various stages, such as handling null values, removing duplicates, converting categorical data to numerical encoding, and scaling. In preparing the dataset for rainfall prediction, various key preprocessing steps were applied to improve the accuracy and reliability of the model. Duplicate removal was performed to eliminate repeated records that could skew predictions by over-representing certain patterns, thereby enhancing model clarity and efficiency. One-hot encoding was utilized to make categorical variables, such as weather conditions, machine-readable, converting them into binary vectors that preserve the uniqueness of each category for accurate model interpretation. Finally, feature scaling was performed using the standard scaler method, which normalizes data to have a mean of zero and a standard deviation of one, ensuring that features with diverse ranges are treated equally, which is a crucial step for algorithms such as neural networks that are sensitive to input magnitude.

B. Classification Using A-CRNN

In the next stage, the A-CRNN technique is used for the classification task. The overall square convolution kernels (such as 3×3 kernels) can't correctly use feature information [20]. Dual convolution kernels are rectangular-shaped kernels (2×10, 10×2) that can take the frequency- and time-domain information from the input, respectively. Thus, the feature mapping dimension does not alter its dimensions after convolution. At first, convolution is applied in both frequency and time domains to produce feature maps of size 513×10×16, which are then concatenated to form a combined map of 513×10×32. This fused map passes through additional convolutional layers, expanding to 513×10×48 and 513×10×64. An attention mechanism is applied to emphasize important features and reduce redundancy, shrinking the feature map to $1\times1\times(64/r)$ and restoring it to 1×1×64 through a fully connected layer with leaky ReLU. A 2×1 pooling layer is then used to compress features in the frequency domain, and the output is fused with the pooled spectrogram to enrich feature representation and improve performance.

The RNN function utilizes feature data. Presently, dual variants are normally applied: GRU and LSTM. GRU generally associates the input and forget gates into a solitary update gate, which is easier than the typical LSTM approach. The reset and update gates (r_t, z_t) represent the dual gates of GRU. A reset gate, r_t, controls the amount of previous memory information to retain, with a small r_t denoting reduced information from the earlier condition. The z_t update gate is utilized for controlling the level to which the stated data from the former instant is directed into the present condition. A large z_t represents more stated information from the former instant acquired. Now, x_t denotes the input, h_t symbolizes the output featured information from the RNN Hidden Layer (HL) by the present time t, and h_{t-1} signifies the output featured information at the RNN HL by the earlier time $t-1$.

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) = \sigma(x_t W_{xr} + h_{t-1} W_{hr} + b_r) \quad (1)$$

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) = \sigma(x_t W_{xz} + h_{t-1} W_{hz} + b_z) \quad (2)$$

$$\begin{aligned} \overline{h}_t &= tanh(W \cdot [r_t * h_{t-1}, x_t]) \\ &= tanh(x_t W_{xh} + r_t * h_{t-1} W_{hh} + b_h) \end{aligned} \quad (3)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * h_t \quad (4)$$

The x_t shape is *batch_size*, *time_step*, *input_dim*, which correspondingly denotes a sample input batch simultaneously, the maximal stage length of the input series, and each sequence size. In this work, these values were 10, 64, and 16,897, respectively, where 16,897=64×256+513. The shape of W_{xr} is *input_dim*, *num_hidden*), while *num_hidden* stands for the unit counts in the HL (1024). The shape of h_t is *batch_size*, *time_step, num_hidden*, the shape of b_r is *num_hidden*, and the shape of W_{hr} is *num_hidden*, *num_hidden.* A similar dimensional value over the dot product operation was removed. Similarly, W_r and W correspond to a similar operation.

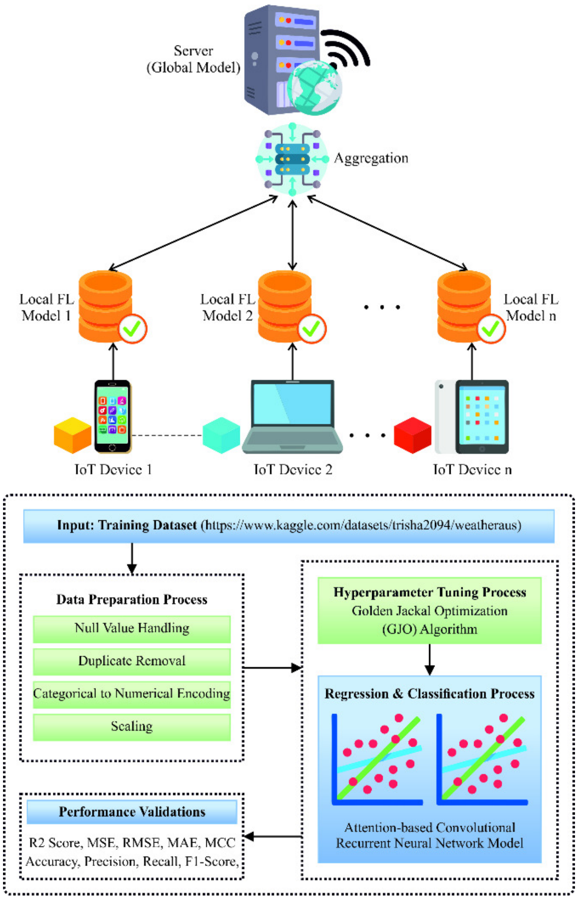

Fig. 1. Workflow of the EMRF-ACRNGJO model.

C. *Golden Jackal Optimization (GJO)-based Hyperparameter Tuning*

GJO is used to tune the hyperparameters of the model. Golden Jackals (GJs), medium-sized jackals, are mainly spread in Asia and Africa [21]. GJs primarily feed on small insects, birds, and animals, using predation within their usual territories to hunt for food. GJ pairs typically forage using the GJO method, with male jackals leading the search and female jackals following, often hunting in groups. Multiple GJ agents collaborate to enhance search success by adjusting tactics based on food availability and peer positions. The model involves three key phases: locating and approaching prey, surrounding it, and attacking the food source.

1) Initialization Stage

The formula for the procedure of GJ population initialization is:

$$Y_0 = Y_{\min} + rand(Y_{\max} - Y_{\min}) \quad (5)$$

where, $Y_{\min}$ and $Y_{\max}$ denote the lower and upper bounds of Y_0, and $rand$ signifies a random value distributed uniformly inside $(0, 1)$. The primary prey matrix is:

$$Prey = \begin{bmatrix} Y_{1,1} & Y_{1,2} & \cdots & Y_{1,d} \\ Y_{2,1} & Y_{2,2} & \cdots & Y_{2,d} \\ \vdots & \vdots & & \vdots \\ Y_{n,1} & Y_{n,2} & \cdots & Y_{n,d} \end{bmatrix} \qquad (6)$$

where n and d represent the population count and size, respectively.

2) *Search Stage*

GJ pairs hunt for prey inside their region and slowly move near the victim. In the hunting procedure for the target, the male GJ leads and the female GJ pursues. The computation technique of the relational location among the prey and the GJ is:

$$Y_1(t) = Y_M(t) - E * |Y_M(t) - rl * Prey(t)| \qquad (7)$$

$$Y_2(t) = Y_F(t) - E * |Y_F(t) - rl * Prey(t)| \qquad (8)$$

where t represents the present iteration counts, $Y(t)$ represents the location of the male GJ after t iterations, $Y(t)$ symbolizes the female GJ location after t iterations, $Prey(t)$ is the prey's location vector, rl stands for a random vector depending on the Levy distribution, $Y(t)$ indicates the male GJ location equivalent to the target once upgraded, and $Y(t)$ refers to the female GJ position equivalent to the prey afterwards an update. E denotes the energy throughout the escaping method of the prey, and its mathematical formulation is:

$$E = E_1 * E_0 \qquad (9)$$

where E_1 signifies the prey's energy reduction and E_0 indicates the prey's primary energy. The approximation equation of E_0 is:

$$E_0 = 2 * r - 1 \qquad (10)$$

where r refers to random values distributed uniformly inside [0,1]. The computation equation of E_1 is:

$$E_1 = c_1 * \left(1 - \frac{t}{T}\right) \qquad (11)$$

where c_1 denotes a constant $c_1 = 1.5$, and T signifies the maximal iteration counts. In the iterations, E_1 reduces linearly from 1.5 to 0. rl is represented in (12), and the function of Levy's flight is shown in (13).

$$rl = 0.05 * Levy(y) \qquad (12)$$

$$Levy\ (y) = 0.01 * \frac{\mu * \sigma}{v^{\frac{1}{\beta}}} \qquad (13)$$

where:

$$\sigma = \frac{T(1+\beta) sin\left(\frac{\pi * \beta}{2}\right)^{\frac{1}{\beta}}}{T\left(\frac{1+\beta}{2}\right) * 2^{\frac{\beta - 1}{2}}},$$

$\beta = 1.5$, and μ and v both symbolize random values distributed uniformly within [0,1]. The updated GJ equation is:

$$Y(t+1) = \frac{Y_1(t) + Y_2(t)}{2} \qquad (14)$$

3) *Siege Stage*

If the GJ injures the prey, its escaping energy will slowly reduce, and then the GJ will strongly besiege it. When the GJs besiege the target well, they will start an attack. During this phase, the location upgrades of the female and male GJ are considered as:

$$Y_1(t) = Y_M(t) - E * |rl * Y_M(t) - Prey(t)| \qquad (15)$$

$$Y_2(t) = Y_F(t) - E * |rl * Y_F(t) - Prey(t)| \qquad (16)$$

where t represents the present iteration count, $Y(t)$ denotes the male GJ position after t iterations, $Y(t)$ signifies the female GJ location after the t^{th} iteration, $Prey(t)$ indicates the prey's position after the t^{th} iteration, $Y(t)$ stands for the area parallel to the prey once the male GJ is upgraded during the t^{th} iteration, and $Y(t)$ refers to the location equivalent to the prey after the female GJ is upgraded throughout the t^{th} iteration.

4) *Transformation of Local and Global Search*

Parameter E represents the energy that escapes prey. During the GJ foraging procedure, the GJ state changes from the prey-searching to the prey-surrounding phase, defined by E. If $|E| \geq 1$, the GJ manages a global search to discover the prey's position. If $|E| < 1$, the GJ demeanours a local hunt to surround the prey.

The GJO is used to identify the hyperparameter intricate in the A-CRNN technique. MSE is used as the objective function:

$$MSE = \frac{1}{T} \sum_{j=1}^{L} \sum_{i=1}^{M} \left(y_j^i - d_j^i\right)^2 \qquad (17)$$

where M and L indicate the resulting layer value and data, respectively, y_j^i and d_j^i indicate the attained and suitable magnitudes for the j^{th} unit from the resulting network layer in time t, respectively.

The following fitness function is used for classification, with precision as the primary criterion to guide the GJO model. The encoded solution helps assess candidate efficiency, making fitness selection crucial for optimal hyperparameter tuning.

$$Fitness = \max(P) \qquad (18)$$

$$P = \frac{TP}{TP + FP} \qquad (19)$$

where TP and FP denote the true and false positives.

III. RESULT ANALYSIS AND DISCUSSION

The experimental analysis of the EMRF-ACRNGJO approach used the weather AUS benchmark dataset [22], which includes data from locations such as Sydney, Albury, Canberra, and Melbourne Airport. The dataset consists of 142151 samples in two classes, as shown in Table I.

TABLE I. DATASET DESCRIPTION

Class	Samples
RainToday-No	110696
RainToday-Yes	31455
Total	**142151**

IV. CONCLUSION

This research work presented a novel EMRF-ACRNGJO method to predict rainfall in agricultural land. This approach encompasses three processes: data preparation, classification, and hyperparameter tuning. At first, the proposed EMRF-ACRNGJO technique performed data preparation with various stages, such as null value handling, duplicate removal, categorical to numerical encoding, and scaling. The A-CRNN technique was utilized for the classification process, along with the GJO method for hyperparameter tuning. The EMRF-ACRNGJO method was tested using the weather AUS benchmark dataset. The experimental validation of the EMRF-ACRNGJO method demonstrated an accuracy of 73.30%. The limitations of the EMRF-ACRNGJO method comprise a reliance on a specific dataset that may not generalize well to diverse locations and potential challenges in handling highly imbalanced data. Future work may focus on incorporating more varied and massive datasets, exploring advanced data augmentation methods, and developing more interpretable models to enhance robustness and practical applicability.

REFERENCES

[1] B. Keswani, A. G. Mohapatra, P. Keswani, A. Khanna, D. Gupta, and J. Rodrigues, "Improving weather dependent zone specific irrigation control scheme in IoT and big data enabled self driven precision agriculture mechanism," *Enterprise Information Systems*, vol. 14, no. 9–10, pp. 1494–1515, Nov. 2020, https://doi.org/10.1080/17517575.2020.1713406.

[2] P. Schweizer, "Neutrosophy for physiological data compression: in particular by neural nets using deep learning," *International Journal of Neutrosophic Science*, pp. 74–80, 2020, https://doi.org/10.54216/IJNS.010203.

[3] E. Khosla, R. Dharavath, and R. Priya, "Crop yield prediction using aggregated rainfall-based modular artificial neural networks and support vector regression," *Environment, Development and Sustainability*, vol. 22, no. 6, pp. 5687–5708, Aug. 2020, https://doi.org/10.1007/s10668-019-00445-x.

[4] P. Shao, J. Feng, P. Zhang, and J. Lu, "Interpretable spatial-temporal attention convolutional network for rainfall forecasting," *Computers & Geosciences*, vol. 185, Mar. 2024, Art. no. 105535, https://doi.org/10.1016/j.cageo.2024.105535.

[5] S. Jayasree and K. R. Ananthapadmanaban, "Discrete Migratory Bird Optimizer with Deep Learning Driven Cyclone Intensity Prediction on Remote Sensing Images," *Engineering, Technology & Applied Science Research*, vol. 15, no. 2, pp. 21605–21610, Apr. 2025, https://doi.org/10.48084/etasr.8842.

[6] I. Salehin, I. M. Talha, Md. Mehedi Hasan, S. T. Dip, Mohd. Saifuzzaman, and N. N. Moon, "An Artificial Intelligence Based Rainfall Prediction Using LSTM and Neural Network," in *2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE)*, Bhubaneswar, India, Dec. 2020, pp. 5–8, https://doi.org/10.1109/WIECON-ECE52138.2020.9398022.

[7] R. Peeriga *et al.*, "Real-Time Rain Prediction in Agriculture using AI and IoT: A Bi-Directional LSTM Approach," *Engineering, Technology & Applied Science Research*, vol. 14, no. 4, pp. 15805–15812, Aug. 2024, https://doi.org/10.48084/etasr.8011.

[8] D. Kaplun *et al.*, "An intelligent agriculture management system for rainfall prediction and fruit health monitoring," *Scientific Reports*, vol. 14, no. 1, Jan. 2024, Art. no. 512, https://doi.org/10.1038/s41598-023-49186-y.

[9] T. Akilan and K. M. Baalamurugan, "Automated weather forecasting and field monitoring using GRU-CNN model along with IoT to support precision agriculture," *Expert Systems with Applications*, vol. 249, Sep. 2024, Art. no. 123468, https://doi.org/10.1016/j.eswa.2024.123468.

[10] S. S. Srikar and P. Yellamma, "An IoT-based Intelligent Irrigation and Weather Forecasting System," *Recent Patents on Engineering*, vol. 18, no. 9, pp. 100–108, Dec. 2024, https://doi.org/10.2174/0118722121252705231009080251.

[11] F. M. Talaat, "Crop yield prediction algorithm (CYPA) in precision agriculture based on IoT techniques and climate changes," *Neural Computing and Applications*, vol. 35, no. 23, pp. 17281–17292, Aug. 2023, https://doi.org/10.1007/s00521-023-08619-5.

[12] J. Nithyashri, R. K. Poluru, S. Balakrishnan, M. Ashok Kumar, P. Prabu, and S. Nandhini, "IOT based prediction of rainfall forecast in coastal regions using deep reinforcement model," *Measurement: Sensors*, vol. 29, Oct. 2023, Art. no. 100877, https://doi.org/10.1016/j.measen.2023.100877.

[13] S. Indhumathi, S. Aghalya, S. J. A, and P. Aarthi M, "IoT-Enabled Weather Monitoring and Rainfall Prediction using Machine Learning Algorithm," in *2023 Second International Conference on Augmented Intelligence and Sustainable Systems (ICAISS)*, Trichy, India, Aug. 2023, pp. 1491–1495, https://doi.org/10.1109/ICAISS58487.2023.10250652.

[14] M. Ehteram, F. Barzegari Banadkooki, and M. Afshari Nia, "Gaussian mutation-alpine skiing optimization algorithm-recurrent attention unit-gated recurrent unit-extreme learning machine model: an advanced predictive model for predicting evaporation," *Stochastic Environmental Research and Risk Assessment*, vol. 38, no. 5, pp. 1803–1830, May 2024, https://doi.org/10.1007/s00477-023-02657-0.

[15] W. Zou, J. Ji, Y. Wang, J. Wang, Y. Qian, and J. Liu, "Convolutional LSTM with Self-Attention Mechanism for Extreme Weather Prediction," in *2023 China Automation Congress (CAC)*, Chongqing, China, Nov. 2023, pp. 6782–6787, https://doi.org/10.1109/CAC59555.2023.10451261.

[16] S. F. Tekin, A. Fazla, and S. Serdar Kozat, "Numerical Weather Forecasting Using Convolutional-LSTM With Attention and Context Matcher Mechanisms," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1–13, 2024, https://doi.org/10.1109/TGRS.2024.3409084.

[17] A. K. Shaikh, A. Nazir, N. Khalique, A. S. Shah, and N. Adhikari, "A new approach to seasonal energy consumption forecasting using temporal convolutional networks," *Results in Engineering*, vol. 19, Sep. 2023, Art. no. 101296, https://doi.org/10.1016/j.rineng.2023.101296.

[18] S. Monaco, L. Monaco, D. Apiletti, R. Cremonini, and S. Barbero, "Uncertainty-aware methods for enhancing rainfall prediction with deep-learning based post-processing segmentation," *Computers & Geosciences*, vol. 205, Nov. 2025, Art. no. 105992, https://doi.org/10.1016/j.cageo.2025.105992.

[19] W. Zhao, Z. Zhang, N. Khodadadi, and L. Wang, "A deep learning model coupled with metaheuristic optimization for urban rainfall prediction," *Journal of Hydrology*, vol. 651, Apr. 2025, Art. no. 132596, https://doi.org/10.1016/j.jhydrol.2024.132596.

[20] C. Sun *et al.*, "A convolutional recurrent neural network with attention framework for speech separation in monaural recordings," *Scientific Reports*, vol. 11, no. 1, Jan. 2021, Art. no. 1434, https://doi.org/10.1038/s41598-020-80713-3.

[21] S. Jiang, Y. Yue, C. Chen, Y. Chen, and L. Cao, "A Multi-Objective Optimization Problem Solving Method Based on Improved Golden Jackal Optimization Algorithm and Its Application," *Biomimetics*, vol. 9, no. 5, May 2024, Art. no. 270, https://doi.org/10.3390/biomimetics9050270.

[22] "weatherAUS." Kaggle, [Online]. Available: https://www.kaggle.com/datasets/trisha2094/weatheraus.