



OPEN Advance drought prediction through rainfall forecasting with hybrid deep learning model

Brij B. Gupta^{1,2,3,4,15}✉, Akshat Gaurav^{5,6,15}, Razaz Waheeb Attar^{7,15}, Varsha Arya^{8,9,15}, Shavi Bansal^{10,11,12,15}, Ahmed Alhomoud^{13,15} & Kwok Tai Chui^{14,15}

Drought is a natural disaster that can affect a larger area over time. Damage caused by the drought can only be reduced through its accurate prediction. In this context, we proposed a hybrid stacked model for rainfall prediction, which is crucial for effective drought forecasting and management. In the first layer of stacked models, Bi-directional LSTM is used to extract the features, and then in the second layer, the LSTM model will make the predictions. The model captures complex temporal dependencies by processing multivariate time series data in both forward and backward directions using bi-directional LSTM layers. Trained with the Mean Squared Error loss and Adam optimizer, the model demonstrates improved forecasting accuracy, offering significant potential for proactive drought management.

Droughts are natural calamities that can lead to many socio-economic effects¹. Drought can cause significant losses. In 2012, almost 80% of the crops were damaged by the drought, which led to a loss of \$36 billion². Apart from this, the 2010-2011 East Africa drought, the 2011 Texas drought, and the 2012-2015 California drought caused considerable losses to society³. As represented in Figure 1, all the developing countries are affected by the drought⁴. Figure 1 shows that most African countries are affected by the drought.

Apart from the countries, Fig. 2 present the number of affected peoples by drought. There are several years where the impact of droughts spiked significantly compared to others. Notably, in 1998 and 2008, the numbers were exceptionally high, reaching around 74 million and 207 million people affected, respectively. The peak in 2008 is the highest on the graph, indicating a severe drought event during that year. In more recent years, the numbers remain high but somewhat less erratic than the highest peaks seen in the graph. However, drought still consistently affects a significant number of people.

Drought has many long-term effects, and we can only manage the effects of drought if a correct and efficient prediction is made. Hence, researchers are developing an efficient drought prediction model⁵. However, drought is a natural phenomenon, so it is challenging to predict it⁶. In this context, artificial intelligence (AI) has recently shown its significance in drought prediction. I models can easily find the relationship between different features, which improves accuracy.

Contribution

This paper makes several key contributions to the field of time series analysis and drought forecasting:

- Hybrid Model Design: We introduce a novel Hybrid Bi-directional LSTM and LSTM model that effectively captures complex temporal dependencies in multivariate time series data, enhancing the accuracy of rainfall predictions.

¹Department of Computer Science and Information Engineering, Asia University, Taichung 413, Taiwan. ²Kyung Hee University, 26 Kyungheedae-ro, Dongdaemun-gu, Seoul 02447, Korea. ³Symbiosis Centre for Information Technology (SCIT), Symbiosis International University, Pune, India. ⁴University of Economics and Human Science, Warsaw, Poland. ⁵Ronin Institute, Montclair, NJ, USA. ⁶CCRI, Asia University, Taichung, Taiwan. ⁷Management Department, College of Business Administration, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia. ⁸Hong Kong Metropolitan University (HKMU), Ho Man Tin, Hongkong. ⁹Department of Electrical and Computer Engineering, Lebanese American University, Beirut, Lebanon. ¹⁰Insights2Techinfo, Indore, India. ¹¹Center for Interdisciplinary Research, University of Petroleum and Energy Studies (UPES), Dehradun, India. ¹²UCRD, Chandigarh University, Chandigarh, India. ¹³Department of Computer Science, Faculty of Science, Northern Border University, Rafha 91431, Saudi Arabia. ¹⁴Hong Kong Metropolitan University (HKMU), Ho Man Tin, Hong Kong. ¹⁵These authors contributed equally: Brij B. Gupta, Akshat Gaurav, Razaz Waheeb Attar, Varsha Arya, Shavi Bansal, Ahmed Alhomoud and Kwok Tai Chui. ✉email: bbgupta@asia.edu.tw

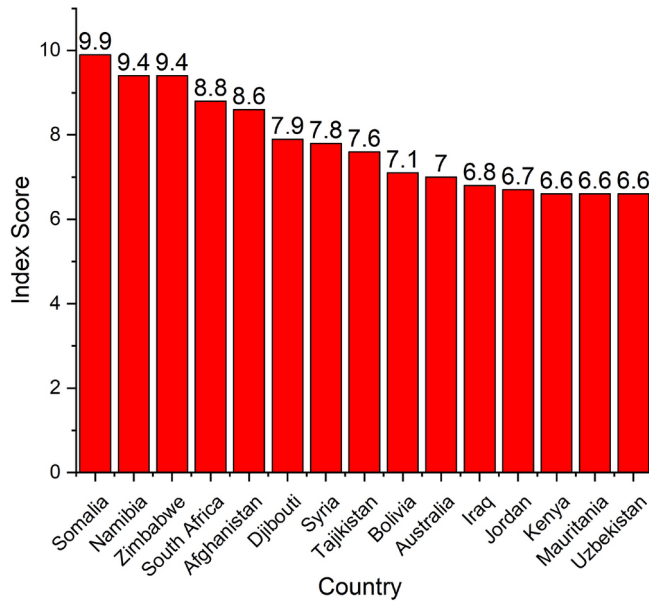


Fig. 1. Countries with highest drought exposure in 2024.

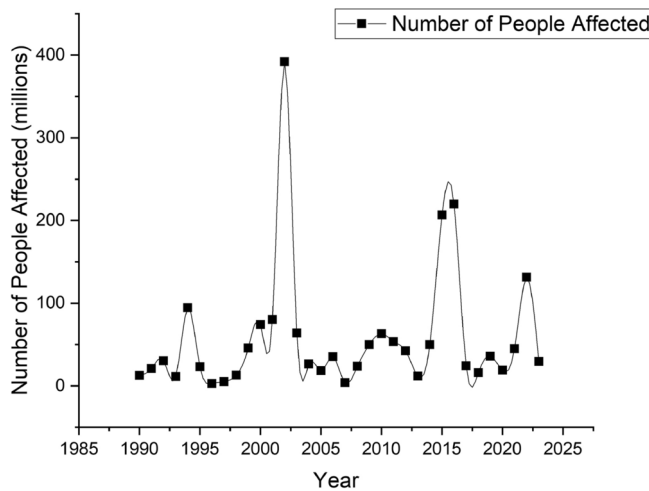


Fig. 2. People affected by drought over the year.

- **Improved Forecasting Accuracy:** By incorporating both forward and backward sequence processing through Bi-directional LSTM layers, our model improves the ability to learn from contextual information, leading to more accurate predictions.
- **Robustness Against Overfitting:** Using a Dropout layer within the architecture enhances the model's robustness, preventing overfitting and promoting the learning of generalized features.
- **Real-world Application:** Our model demonstrates significant potential for improving drought management strategies by providing reliable rainfall forecasts and aiding decision-makers in resource allocation and planning.
- **Future Research Directions:** We outline possibilities for extending the model to incorporate additional meteorological variables and apply transfer learning techniques, paving the way for further advancements in adaptable climate modeling across diverse regions.

Organization

The rest of the paper is organized as follows: “Related work” presents the details of related work. “Proposed work” details our proposed work, and the simulation results are presented in “Simulation results”. Finally, “Conclusion” concludes the paper.

Related work

Deep learning models for drought prediction

Wang, X. et al.⁷ proposed a deep learning model (K-parallel recurrent neural networks (KRNN), convolutional neural networks (CNN), graph convolutional networks (GCN), and multilayer perceptron (MLP)) for the prediction of sub-seasonal drought. The proposed approach used soil moisture to predict drought. However, the primary limitation of this study is its reliance on historical ERA5-Land soil moisture records, which may not fully capture future changes in climate variability. Additionally, the quality and resolution of input data might influence the models' performance, potentially affecting the accuracy of drought predictions in regions with less detailed soil moisture records.

Dikshit, A. et al.¹ used LSTM to predict the Standard Precipitation Evaporation Index (SPEI) through the data collected from the Climatic Research Unit (CRU) that range from 1901 to 2018. However, the study's reliance on historical data from the Climatic Research Unit (CRU) dataset may limit its ability to account for future climate variability and unforeseen changes in drought patterns.

Anshuka, A. et al.⁸ proposed an LSTM-based temporal hydrological extreme forecasting. The results are related to the South Pacific region and are based on rainfall data. The study's reliance on satellite rainfall estimates and sea surface temperature (SST) anomalies may limit its ability to capture localized hydrological extremes influenced by factors not accounted for in the datasets. Additionally, the model's predictive accuracy might vary with different spatial resolutions and data quality, impacting the generalizability of results across diverse geographical regions.

Maity, R. et al.⁹ used a CNN model to predict basin-scale drought assessment using air temperature, surface pressure, wind speed, relative humidity, evaporation, soil moisture, and geopotential height. The study's reliance on primary hydrometeorological precursors may not fully capture the complex interactions and feedback mechanisms influencing droughts, particularly in regions with unique climatological or topographical characteristics. Additionally, the model's performance may vary when applied to basins with different data availability or quality, potentially impacting its generalizability and effectiveness across diverse environments.

Hybrid models for enhanced drought forecasting

Abbes, A. et al.¹⁰ integrates LSTM and Multi-Resolution Analysis Wavelet Transform (MRA-WT) for forecasting drought. This research also focuses on the prediction of the SPEI index. However, the study's effectiveness may be limited by the availability and quality of input data, such as station data and NDVI from Landsat images, which could affect the model's accuracy in regions with sparse or unreliable data. Additionally, the model's performance in capturing non-stationary drought indices might be challenged by extreme weather events not represented in the historical dataset.

Vo Q. T. et al.¹¹ proposed drought prediction using LSTM and climate model. The model's performance may be limited by biases in the input data and the inherent uncertainties in the climate model GloSea5 (GS5), which could affect the accuracy of drought predictions, especially for long-lead-time cases. Additionally, the model's effectiveness might vary across different geographical regions due to variations in climate and environmental conditions not fully captured by the input variables.

Al-Ayyoub, M. et al.¹² presents a hybrid parallel implementation of the Fuzzy C-Means (FCM) algorithm on a GPU for segmenting 3D medical images from DICOM files. The approach significantly reduces processing time, achieving a 5x speedup compared to the traditional sequential version, making it more feasible for handling large medical datasets.

Moteri, M. A. A. et al.¹³ proposed a convolutional long short-term memory with self-attention (SA-CLSTM) to predict the SPEI index. However, the model's reliance on historical data from the Climatic Research Unit (CRU) dataset may not fully capture the unique environmental and climatic variations specific to coastal arid regions, potentially affecting prediction accuracy. Additionally, data quality and availability could challenge the model's effectiveness, particularly in areas with sparse meteorological monitoring systems.

Explainable AI and neural network approaches

Dikshit, A. et al.¹⁴ proposed a deep learning model for drought prediction, in addition to using explainable artificial intelligence (XAI) for explaining the working of deep learning. The study's reliance on historical data and climatic variables may limit its ability to predict droughts accurately under rapidly changing climate conditions and unforeseen environmental factors. Additionally, the model's applicability across diverse regions may be constrained by variations in data quality and the specific climatic characteristics of each area.

Poornima S. et al.¹⁵ proposed an LSTM-based drought prediction model. The study's reliance on historical meteorological data for forecasting drought indices may limit its predictive accuracy in rapidly changing climate conditions or regions with sparse data. Additionally, the performance of the LSTM model could be affected by the quality and completeness of the input variables, potentially impacting its generalizability across diverse geographical areas.

IoT and data-driven frameworks

Kaur, A. et al.¹⁶ proposed an IoT-based framework for collecting parameters for predicting drought. The author used an Artificial Neural Network (ANN), ANN optimized with Genetic Algorithm (ANN-GA), DNN (Deep Neural Network), and SVM for drought prediction. However, the model's reliance on IoT data for drought prediction may be limited by the availability and quality of IoT infrastructure in different regions, potentially affecting the accuracy and reliability of forecasts. Additionally, variations in regional climate conditions and data collection standards could impact the model's generalizability across diverse geographical areas.

Alsmirat M. A et al.¹⁷ accelerates Fuzzy C-Means (FCM) segmentation algorithms in medical image processing by leveraging GPU capabilities. This approach significantly enhances processing speed, achieving

up to 8.9x faster execution without reducing segmentation accuracy, which is crucial for medical diagnostics. Alsmirat M. A. et al.,¹⁸ explore using digital cameras as an alternative to specialized fingerprint devices, explicitly addressing the challenge of image compression. It investigates the optimal compression ratio for maintaining high accuracy in fingerprint identification, establishing that a 30-40% compression ratio is effective without significant accuracy loss.

Nhi, N. T. U. et al.¹⁹ introduces a semantic-based image retrieval system that combines a custom C-Tree structure with a neighbor graph (Graph-C-Tree) to enhance retrieval accuracy. The system employs k-nearest Neighbor (k-NN) for visual word creation and uses an ontology framework with SPARQL queries to improve semantic retrieval. The proposed method demonstrated high precision across several datasets, outperforming comparable approaches.

Dikshit, A et al.²⁰ proposed an attention-based model to forecast meteorological droughts at a short-term forecast range for five sites situated in Eastern Australia. However, the study's reliance on large-scale climatic indices and historical data may limit its ability to account for sudden or unpredicted climate anomalies, potentially affecting the accuracy of short-term drought forecasts. Additionally, the model's performance might vary with different geographical regions due to the specific climatic and environmental characteristics not fully represented in the dataset.

Qian, W. et al.²¹ proposes an enhanced GAN-based image style transfer (IST) method that incorporates a circular local binary pattern (LBP) as a texture before improving image detail. A dense connection residual block and attention mechanism are integrated to enhance high-frequency feature extraction. At the same time, a total variation (TV) regularizer is added to the loss function to reduce noise and smooth results. The method outperforms existing approaches in generating more detailed and higher-quality styled images.

Wang, H et al.²² introduce a visual saliency-guided image retrieval model that integrates the Itti visual saliency model with a multi-feature fusion approach. The model uses a two-stage complexity classification (cognitive load and cognitive level) and a group sparse logistic regression model to enhance retrieval accuracy, especially in complex multimedia scenarios.

Agana, N. An et al.²³ proposed a deep belief network consisting of two Restricted Boltzmann Machines for long-term drought prediction using lagged values of SSI as inputs. The proposed approach may limit its ability to capture unexpected climate shifts or anomalies, potentially impacting the accuracy of long-term drought predictions. Additionally, the model's effectiveness might vary across different regions, influenced by the unique hydrological and climatic conditions not fully captured in the training dataset.

Li, J. et al.²⁴ introduce a new color image watermarking method using Quaternion Hadamard Transform (QHT) and Schur decomposition. The technique leverages the correlation between color channels to embed watermarks resiliently. Additionally, a quaternion Zernike moment-based detection method is employed to handle geometric distortions, making the watermark extraction process robust against such attacks.

Chopra, M. et al.²⁵ provides a comprehensive review of advancements in generative models for text-to-image synthesis, focusing on GAN-based architectures such as DCGAN, StackGAN, StackGAN++, and AttnGAN. By examining these models, the paper highlights how iterative refinement, hierarchical structuring, and attention mechanisms contribute to generating semantically coherent and realistic images from text descriptions.

Proposed approach

Dataset representation and preprocessing

Assume a multivariate time series dataset with n observations and m features. Each observation at time t is represented as a vector:

$$\mathbf{x}_t = [x_{t,1} \quad x_{t,2} \quad \cdots \quad x_{t,m}]^T \quad (1)$$

The dataset X is represented as a matrix:

$$X = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \cdots \quad \mathbf{x}_n]^T = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,m} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,m} \end{bmatrix} \quad (2)$$

Feature range and differences

To understand the variability and range of features, the following processes are undertaken:

Range of features

The range of a feature j is given by:

$$\text{Range}(x_{:,j}) = \max(x_{:,j}) - \min(x_{:,j}) \quad (3)$$

Calculating the range helps identify the spread of the data, which is crucial for understanding its variability and scaling the model appropriately.

Feature differences

Differencing is used to stabilize variance and make the series stationary:

$$\Delta x_{t,j} = x_{t,j} - x_{t-1,j} \quad (4)$$

The differenced dataset ΔX is:

$$\Delta X = \begin{bmatrix} \Delta x_{2,1} & \Delta x_{2,2} & \cdots & \Delta x_{2,m} \\ \Delta x_{3,1} & \Delta x_{3,2} & \cdots & \Delta x_{3,m} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta x_{n,1} & \Delta x_{n,2} & \cdots & \Delta x_{n,m} \end{bmatrix} \quad (5)$$

Differencing removes trends and seasonality, making the data suitable for time series modeling, particularly with neural networks like BiLSTM and LSTM.

Granger causality test

The Granger Causality Test determines causal relationships between variables in a time series dataset. It assesses whether one variable's past values help predict another's future values.

Purpose of the Granger causality test

Granger causality is crucial for identifying dependencies and selecting relevant features for the model. If a variable x_i Granger-causes x_j , including x_i can improve the predictive power for x_j .

Conducting the test

For testing causality between two variables x_i and x_j , construct lagged vectors:

$$y_t = \begin{bmatrix} x_{j,t-1} \\ x_{j,t-2} \\ \vdots \\ x_{j,t-p} \end{bmatrix}, \quad z_t = \begin{bmatrix} x_{i,t-1} \\ x_{i,t-2} \\ \vdots \\ x_{i,t-p} \end{bmatrix} \quad (6)$$

Fit the following models:

Autoregressive model

$$x_{j,t} = \mathbf{a}^T y_t + \epsilon_t \quad (7)$$

Extended model with x_i

$$x_{j,t} = \mathbf{a}^T y_t + \mathbf{b}^T z_t + \epsilon_t \quad (8)$$

Where \mathbf{a} and \mathbf{b} are coefficient vectors. The significance of coefficients \mathbf{b} indicates whether x_i Granger-causes x_j .

Normalization

Normalization is a critical preprocessing step that ensures each feature contributes equally to the model. It prevents features with more extensive ranges from dominating the learning process.

Min-max scaling

Min-Max scaling transforms each feature into the range $[0, 1]$:

$$\tilde{x}_{t,j} = \frac{x_{t,j} - \min(x_{:,j})}{\max(x_{:,j}) - \min(x_{:,j})} \quad (9)$$

This scaling is useful when the data does not have significant outliers and the features are approximately uniformly distributed.

Z-score normalization

Z-score normalization transforms each feature to have a mean of 0 and a standard deviation of 1:

$$\tilde{x}_{t,j} = \frac{x_{t,j} - \mu_j}{\sigma_j} \quad (10)$$

Where μ_j and σ_j are the mean and standard deviation of feature j . This method is beneficial when data follows a Gaussian distribution.

Data splitting

Data splitting involves dividing the dataset into training and testing sets to ensure that the model is evaluated on unseen data. The split is typically performed using a ratio r .

Training set

The training set is used to fit the model. It comprises the first k observations, where:

$$k = \lfloor n \times r \rfloor \quad (11)$$

Training set:

$$X_{\text{train}} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,m} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{k,1} & x_{k,2} & \cdots & x_{k,m} \end{bmatrix} \quad (12)$$

Testing set

The testing set is used to evaluate the model's performance:

$$X_{\text{test}} = \begin{bmatrix} x_{k+1,1} & x_{k+1,2} & \cdots & x_{k+1,m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,m} \end{bmatrix} \quad (13)$$

Splitting ensures that the model's predictions are unbiased and that its performance can be generalized to new data.

Bi-directional LSTM model

A Bi-directional LSTM processes the input sequence in both forward and backward directions. Given an input sequence $X = [x_1, x_2, \dots, x_T]$, where each x_t is a feature vector:

- Forward LSTM:

$$\vec{h}_t = \text{LSTM}(x_t, \vec{h}_{t-1}) \quad (14)$$

- Backward LSTM:

$$\overleftarrow{h}_t = \text{LSTM}(x_t, \overleftarrow{h}_{t+1}) \quad (15)$$

- Output of bi-directional LSTM:

$$h_t = \begin{bmatrix} \vec{h}_t \\ \overleftarrow{h}_t \end{bmatrix} \quad (16)$$

For this model, the bi-directional LSTM has an output dimension of $2 \times 64 = 128$.

LSTM layers

Each LSTM layer applies the following transformations:

- Forget gate:

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (17)$$

- Input gate:

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (18)$$

$$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (19)$$

· Cell state:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (20)$$

· Output gate:

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (21)$$

· Hidden state:

$$h_t = o_t \odot \tanh(c_t) \quad (22)$$

The LSTM layers in this model are configured as follows:

- First LSTM: 32 units, returns sequences
- Second LSTM: 16 units, returns sequences
- Third LSTM: 8 units, returns a single sequence

Dropout layer

The Dropout layer randomly sets a fraction of input units to zero during training:

$$h_t^{\text{drop}} = h_t \odot \text{Mask}(p) \quad (23)$$

where $\text{Mask}(p)$ is a binary mask with dropout probability $p = 0.25$.

Dense layer

The Dense layer computes the final output using:

$$\hat{y} = Wh + b \quad (24)$$

This layer produces an output vector with dimensions matching $\text{trainY.shape}[1]$.

Loss function

The model is compiled with the Mean Squared Error (MSE) loss function:

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2 \quad (25)$$

where \hat{y}_i is the predicted value, y_i is the true value, and N is the number of samples.

Optimizer

The Adam optimizer updates model weights using the following:

$$\theta_{t+1} = \theta_t - \alpha \frac{m_t}{\sqrt{v_t} + \epsilon} \quad (26)$$

Where:

- α is the learning rate.
- m_t and v_t are the first and second moment estimates.
- ϵ is a small constant to prevent division by zero.

Model prediction and evaluation

Prediction

Given an input sequence X_{input} , the model predicts output \hat{y}_t :

$$\hat{y}_t = \text{Model}(X_{\text{input}}) \quad (27)$$

The BiLSTM and LSTM models leverage sequential data to predict future rainfall patterns accurately.

Comparison of actual and predicted data

Evaluate predictions using Mean Absolute Error (MAE) and Root Mean Square Error (RMSE):

$$\text{MAE} = \frac{1}{N} \sum_{t=1}^N |y_t - \hat{y}_t| \quad (28)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{t=1}^N (y_t - \hat{y}_t)^2} \quad (29)$$

Where N is the number of observations in the test set. These metrics provide insight into the accuracy of the model's predictions by comparing them against actual values.

Results and discussion

Dataset representation

In order to test our proposed model, we used data collected from NASA's GMAO MERRA-2 assimilation model²⁶. The data was collected from 2000 to 2020 in Mumbai, India. The data contains precipitation, specific humidity, relative humidity, and temperature. Hence, in this paper, we predict the precipitation using information on particular moisture, relative humidity, and temperature. In our study, we utilized daily time intervals for both training and prediction purposes. Daily data granularity allows our model to capture short-term rainfall fluctuations, which is particularly valuable for identifying immediate drought conditions.

Data preprocessing

In our study, we addressed missing data points through linear interpolation, which fills gaps by estimating values based on neighboring data. This method is effective in maintaining continuity within the time series without introducing significant bias. For anomalous values, we applied a z-score thresholding approach to detect outliers. Any data points that fell beyond three standard deviations from the mean were flagged as potential anomalies. These values were subsequently replaced using the rolling median imputation technique, which is less sensitive to extreme variations and preserves the dataset's underlying trends.

Granger causality test

Figure 3, presents the correlation information about specific humidity, relative humidity, and temperature concerning precipitation. The figure clearly shows that the correlation between specific humidity and precipitation is 0.724540. Also, the correlation between relative humidity and precipitation is 0.743112. Moreover, the correlation between temperature and precipitation is -0.779399. Hence, from Figure 3 it is clear that all the features are highly correlated to precipitation.

We plot the scatter plot to get more information about the correlation between specific humidity, relative humidity, and temperature, as represented in Fig. 4. The figure shows that particular humidity and relative humidity positively correlate with precipitation because precipitation also increases with specific humidity and relative humidity. However, temperature is negatively correlated to precipitation because the precipitation increases with a decrease in temperature.

The correlation matrix and scatter plot confirm that the features (specific humidity, relative humidity, and temperature) are related to precipitation. But to get more information about the relationship, we plot the Fig. 5. With a mean of around 20 units, the histogram for specific humidity shows a distinct right-skewed distribution, showing a concentration of greater specific humidity values. By contrast, the relative humidity distribution is bimodal, with maxima about at 50% and 90%, implying two quite common groups in the sample about atmospheric moisture content. The temperature data is likewise very bimodal, with most observations falling around 10°C and 25°C. Crucially for modeling their effect on precipitation patterns, these histograms-overlaid with kernel density estimates-offer a visual grasp of the variability and average values for every meteorological variable.

Performance analysis

To analyze the performance of our proposed model, we plot the model training and validation Loss in Fig. 6. From the figure, it is clear that at the start of the epochs, the training loss is higher than the validation loss. However, as the training increases the value of training loss also decreases, this shows that our purposed model is training efficiently. Finally, the variation in training and validation loss is almost similar after the fifth epoch. Hence, it is evident that our proposed model gets trained very quickly (in 5 epochs).

To get more information about prediction accuracy, we compare the predicted values from our model with the actual values, as represented in Fig. 7. The figure clearly shows that the forecasting values follow the actual values with a high accuracy value. However, there is some mismatch value for the prediction at the pike values

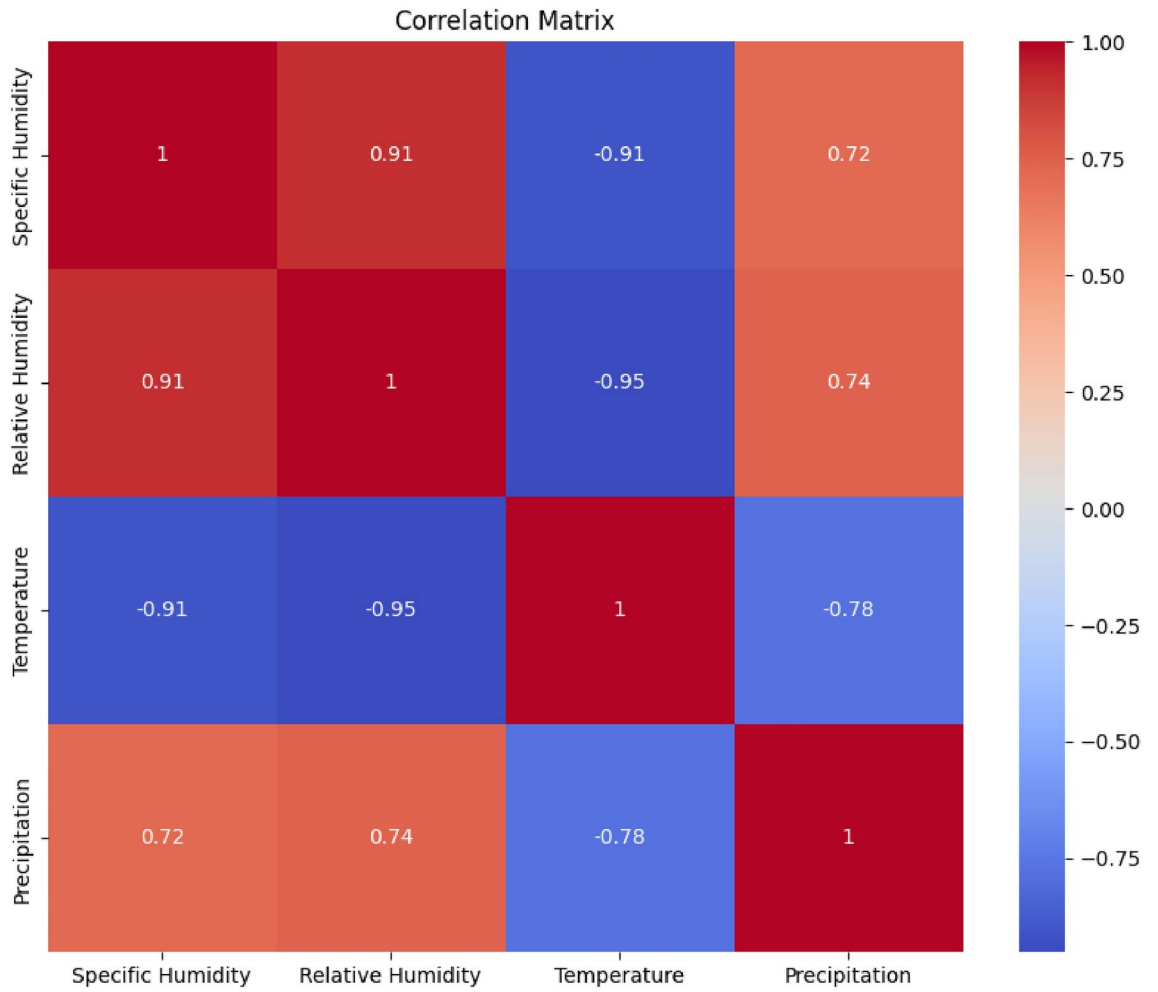


Fig. 3. Correlation matrix of selected features.

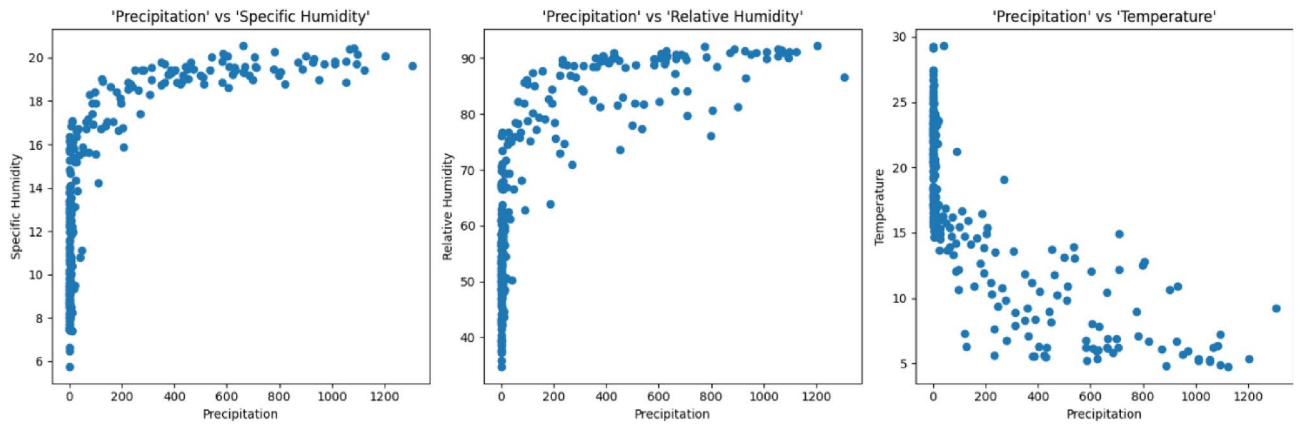


Fig. 4. Scatter plot of selected features.

of the actual precipitation. In this context, we will improve our model for better accuracy in the future and test it on more real-world models.

Comparative analysis

In this subsection, we compare our proposed model qualitatively and quantitatively.

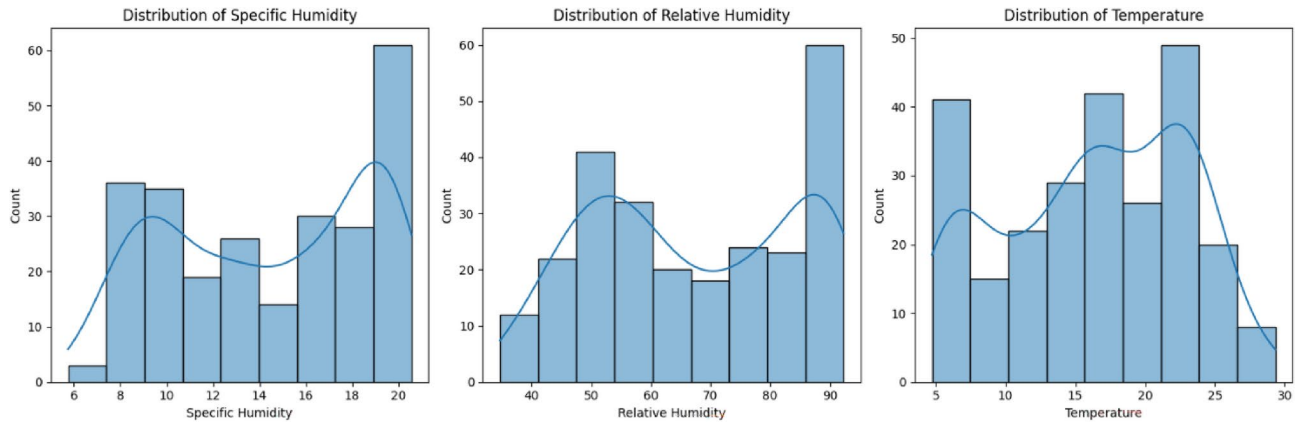


Fig. 5. Histogram of selected features.

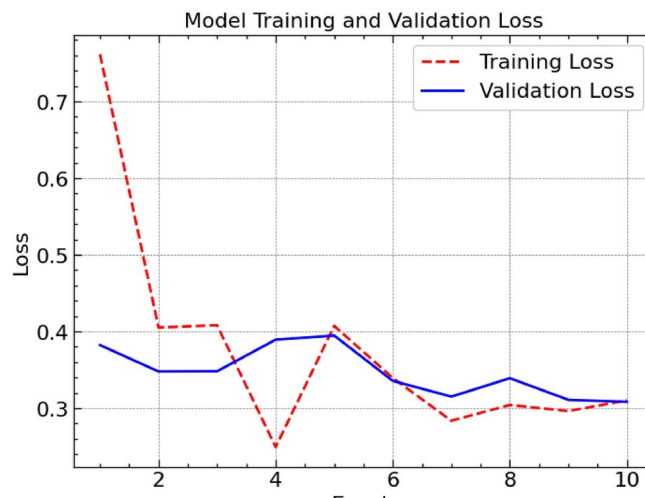


Fig. 6. Variation of model training and validation loss over epochs.

Quantitative analysis

We compared the Bi-LSTM and GRU models in the quantitative analysis with our proposed model. Figure 8 presents this quantitative analysis. Figure 8a presents the comparison of accuracy, from the figure it is clear that the proposed approach has the lowest accuracy than the rest of the models. Similarly, Fig. 8b, c present the comparison of Mean Absolute Error (MAE) and Mean Square Error (MSE). The figures show that our proposed model outperformed the other two models.

Qualitative analysis

This section will present the qualitative analysis of current work with our proposed approach. The comparison details are presented in Table 1.

The approach proposed by Wang, X. et al.⁷ is complex because The complexity of this study arises from integrating multiple deep-learning modules within a committee machine framework, which requires sophisticated techniques for series decomposition and feature extraction. This approach demands advanced computational methods to ensure that the models effectively leverage soil moisture memory for accurate sub-seasonal drought forecasting.

The complexity of the model proposed by Dikshit, A. et al.¹ stems from utilizing the Long Short-Term Memory (LSTM) model to predict drought characteristics across multiple scales and dimensions. This approach requires sophisticated techniques for data preprocessing, model training, and evaluation, ensuring that the LSTM effectively captures temporal and spatial variations in drought dynamics for improved forecasting accuracy.

Anshuka, A. et al.⁸ is complex due to using the LSTM model to capture spatio-temporal patterns in hydrological extremes. This involves sophisticated methods for processing multivariate data and incorporating eigenvector values of SST, necessitating advanced computational techniques to accurately model the dynamic interactions within the hydrological and climatological variables.

The model proposed by Maity, R. et al.⁹ is complex due to the utilization of a one-dimensional convolutional neural network to extract meaningful insights from diverse hydrometeorological precursors. This approach

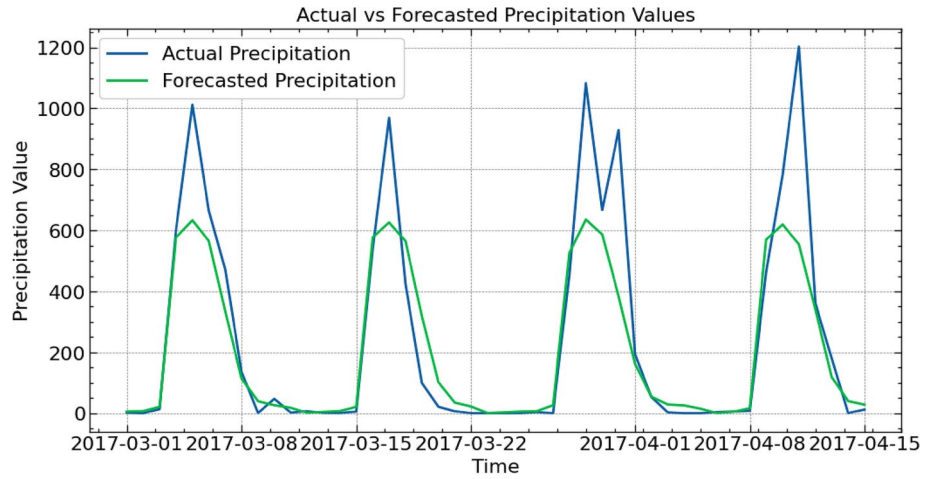


Fig. 7. Prediction of precipitation by proposed model.

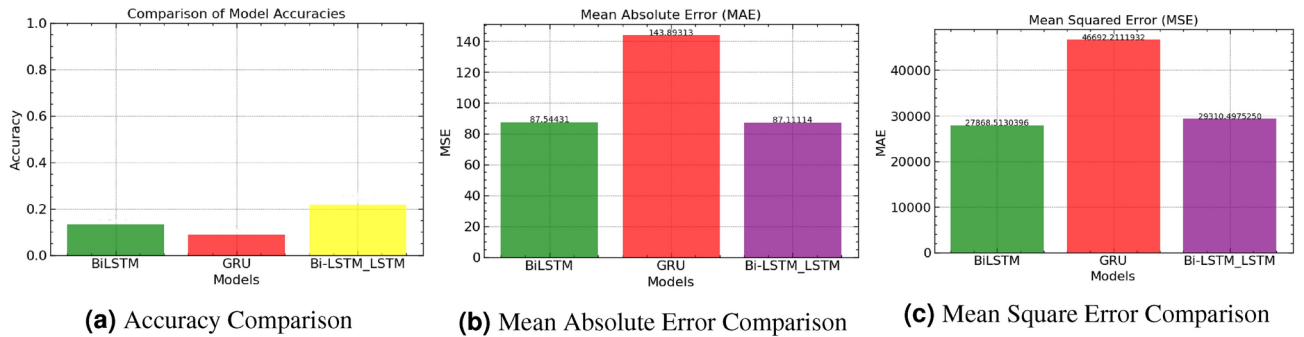


Fig. 8. Quantitative analysis.

Model	Hypothesis testing	Technique	Complexity
Wang, X. et al. ⁷	✗	KRNN, CNN, GCN, MLP	High
Dikshit, A. et al. ¹	✗	LSTM	High
Anshuka, A. et al. ⁸	✗	LSTM	High
Maity, R. et al. ⁹	✗	CNN	High
Abbes, A. et al. ¹⁰	✗	LSTM	High
Vo Q. T. et al. ¹¹	✗	LSTM-CM	High
Moteri, M. A. A. et al. ¹³	✗	SA-CLSTM	High
Dikshit, A. et al. ¹⁴	✗	DL, XAI	High
Poornima S. et al. ¹⁵	✗	LSTM	High
Kaur, A. et al. ¹⁶	✗	ANN, ANN-GA, DNN, SVM	High
Dikshit, A et al. ²⁰	✗	Attention-based model	High
Agana, N. A et al. ²³	✗	Deep Belief Network	High
Proposed approach	Granger causality test	Bi-LSTM	Low

Table 1. Comparative analysis.

requires advanced data integration and modeling techniques to accurately capture the intricate relationships between meteorological variables and drought conditions, highlighting the challenges of representing complex environmental processes in a computational framework.

The complexity of the model proposed by Abbes, A. et al.¹⁰ lies in integrating Multi-Resolution Analysis Wavelet Transform (MRA-WT) with Long Short-Term Memory (LSTM) networks to handle non-stationary time-series data. This approach requires advanced techniques for data preprocessing and model training to

analyze and forecast complex drought patterns effectively, demonstrating the intricate relationship between temporal and spectral characteristics of drought indices.

The model proposed by Vo Q. T. et al.¹¹ is complex due to integrating the LSTM model with a climate prediction model (GS5) to leverage physical process simulation and low prediction bias. This hybrid approach requires sophisticated techniques to balance and optimize the strengths of both models, ensuring accurate and reliable drought forecasts while minimizing uncertainties and biases.

The complexity of the model proposed by Moteri, M. A. A. et al.¹³ stems from using a convolutional extended short-term memory model with self-attention (SA-CLSTM) to capture the intricate interactions between drought factors in coastal arid regions. This approach requires advanced techniques for integrating multiple data sources and optimizing model parameters to ensure accurate and reliable drought severity, category, and geographic variation forecasts.

The model proposed by Dikshit, A. et al.¹⁴ is complex due to integrating XAI with deep learning models to understand local interactions among predictors for different drought conditions. This approach requires sophisticated techniques to analyze and interpret model outputs using Shapley additive explanations (SHAP), ensuring that predictions align with physical model interpretations across various spatio-temporal scales.

The approach by Poornima S. et al.¹⁵ is complex because it uses LSTM networks to handle real-time nonlinear data. This approach requires advanced techniques for data preprocessing, model training, and evaluation to effectively capture temporal patterns and improve long-term drought predictions, challenging the conventional statistical methods' capabilities.

The model proposed by Kaur, A. et al.¹⁶ is complex due to the integration of IoT with dimensionality reduction techniques and advanced machine learning models, such as DNN and ANN-GA, to predict drought conditions. This framework requires sophisticated data processing and analysis at both the Fog and Cloud layers, necessitating advanced techniques to handle large-scale data and optimize model performance for accurate drought assessment and management.

The complexity of the model proposed by Dikshit, A et al.²⁰ stems from using an attention-based deep learning model to capture the nonlinear relationships and dependencies between hydrometeorological and climatic factors. This approach requires sophisticated data integration and interpretation techniques to understand how the model forecasts droughts across various lead times, enhancing transparency and trust in the predictions.

The complexity of the model proposed by Agana, N. An et al.²³ study arises from employing a Deep Belief Network with two Restricted Boltzmann Machines to address the nonlinear and nonconvex optimization challenges in drought prediction. This approach requires advanced techniques for data preprocessing and model training to effectively capture complex temporal patterns, ensuring superior performance over traditional methods like Multilayer Perceptron (MLP) and Support Vector Regression (SVR).

Conclusion

The paper used a hybrid model to predict rainfall, an essential factor in drought prediction. The proposed hybrid model is the stacked models of Bi-LSTM and LSTM. The stacked model takes the input from NASA's GMAO MERRA-2 assimilation model and uses time series analysis to predict the rainfall. As the rainfall is directly related to drought, our proposed models help predict drought. We also compared our proposed model with standard deep learning models and previous literature to present the effectiveness of our proposed model. Our proposed hybrid mode outperformed the standard deep learning models in accuracy, mean square error, and mean absolute error. In addition to that, our proposed model is less complex than the current literature. However, our model still needs improvement. Therefore, in future work, we explore additional meteorological variables and transfer learning to improve accuracy and false negatives in our model. Specifically, we plan to explore the integration of satellite data, such as remote sensing data on soil moisture and vegetation indices, to improve the model's predictive accuracy by incorporating additional environmental factors. This data could offer valuable insights into drought conditions and enhance the model's performance by providing finer spatial resolution. Additionally, we aim to test the model across diverse regions with varying climatic conditions to assess its generalizability. For instance, applying the model to semi-arid and humid regions would allow us to evaluate its adaptability and robustness in different drought contexts. We also plan to implement transfer learning techniques to leverage pre-trained models on similar datasets, which could improve prediction accuracy while reducing computational requirements.

Data availability

The datasets generated and analyzed during the current study are available in the Kaggle repository, <https://www.kaggle.com/datasets/poojag718/rainfall-timeseries-data>

Received: 5 September 2024; Accepted: 14 November 2024

Published online: 13 December 2024

References

1. Dikshit, A., Pradhan, B. & Huete, A. An improved spei drought forecasting approach using the long short-term memory neural network. *J. Environ. Manag.* <https://doi.org/10.1016/J.JENVMAN.2021.111979> (2021).
2. AghaKouchak, A. et al. Status and prospects for drought forecasting: opportunities in artificial intelligence and hybrid physical-statistical forecasting. *Philos. Trans. R. Soc. A* <https://doi.org/10.1098/RSTA.2021.0288> (2022).
3. Hao, Z., Singh, V. P. & Xia, Y. Seasonal drought prediction: Advances, challenges, and future prospects. *Rev. Geophys.* **56**, 108–141 (2018).
4. Statistica. Countries most exposed to droughts worldwide in 2024, by risk index score. <https://www.statista.com/statistics/1395543/countries-most-exposed-to-droughts-by-risk-index-score/> (2024). Accessed 30 June 2024.

5. Anshuka, A., Ogtrop, F. v. & Vervoort, R. W. Drought forecasting through statistical models using standardised precipitation index: A systematic review and meta-regression analysis. *Nat. Hazards* <https://doi.org/10.1007/S11069-019-03665-6> (2019).
6. Houmma, I. H., Mansouri, L. E., Gadal, S., Garba, M. & Hadria, R. Modelling agricultural drought: A review of latest advances in big data technologies. *Geomat. Nat. Haz. Risk* <https://doi.org/10.1080/19475705.2022.2131471> (2022).
7. Wang, X. et al. Sub-seasonal soil moisture anomaly forecasting using combinations of deep learning, based on the reanalysis soil moisture records. *Agric. Water Manag.* <https://doi.org/10.1016/J.AGWAT.2024.108772> (2024).
8. Anshuka, A., Chandra, R., Buzacott, A. J., Sanderson, D. & van Ogtrop, F. F. Spatio temporal hydrological extreme forecasting framework using lstm deep learning model. *Stoch. Environ. Res. Risk Assess.* **36**, 3467–3485 (2022).
9. Maity, R. et al. Potential of deep learning in drought assessment by extracting information from hydrometeorological precursors. *J. Water Clim. Change* **12**, 2774–2796 (2021).
10. Abbes, A. B., Inoubli, R., Rhif, M. & Farah, I. Combining deep learning methods and multi-resolution analysis for drought forecasting modeling. *Earth Sci. Inform.* <https://doi.org/10.1007/S12145-023-01009-4> (2023).
11. Vo, T. Q., Kim, S.-H., Nguyen, D. & Bae, D. Lstm-cm: a hybrid approach for natural drought prediction based on deep learning and climate models. *Stoch. Environ. Res. Risk Assess. (Print)* <https://doi.org/10.1007/S00477-022-02378-W> (2023).
12. Al-Ayyoub, M., AlZu'bi, S., Jararweh, Y., Shehab, M. A. & Gupta, B. B. Accelerating 3d medical volume segmentation using gpus. *Multimed. Tools Appl.* **77**, 4939–4958 (2018).
13. Moteri, M. A. A. et al. An enhanced drought forecasting in coastal arid regions using deep learning approach with evaporation index. *Environ. Res.* <https://doi.org/10.1016/J.ENVRES.2024.118171> (2024).
14. Dikshit, A. & Pradhan, B. Interpretable and explainable ai (xai) model for spatial drought prediction. *Sci. Total Environ.* <https://doi.org/10.1016/J.SCITOTENV.2021.149797> (2021).
15. Poornima, S. & Pushpalatha, M. Drought prediction based on spi and spei with varying timescales using lstm recurrent neural network. *Soft Comput.* <https://doi.org/10.1007/S00500-019-04120-1> (2019).
16. Kaur, A. & Sood, S. K. Deep learning based drought assessment and prediction framework. *Ecol. Inform.* <https://doi.org/10.1016/J.ECOINE.2020.101067> (2020).
17. Alsmirat, M. A., Jararweh, Y., Al-Ayyoub, M., Shehab, M. A. & Gupta, B. B. Accelerating compute intensive medical imaging segmentation algorithms using hybrid cpu-gpu implementations. *Multimed. Tools Appl.* **76**, 3537–3555 (2017).
18. Alsmirat, M. A., Al-Alem, F., Al-Ayyoub, M., Jararweh, Y. & Gupta, B. B. Impact of digital fingerprint image quality on the fingerprint recognition accuracy. *Multimed. Tools Appl.* **78**, 3649–3688 (2019).
19. Nhi, N. T. U. et al. A model of semantic-based image retrieval using c-tree and neighbor graph. *Int. J. Semantic Web Inf. Syst. (IJSWIS)* **18**, 1–23 (2022).
20. Dikshit, A., Pradhan, B., Assiri, M. E., Almazroui, M. & Park, H.-J. Solving transparency in drought forecasting using attention models. *Sci. Total Environ.* **837**, 155856 (2022).
21. Qian, W., Li, H. & Mu, H. Circular lbp prior-based enhanced gan for image style transfer. *Int. J. Semantic Web Inf. Syst. (IJSWIS)* **18**, 1–15 (2022).
22. Wang, H., Li, Z., Li, Y., Gupta, B. B. & Choi, C. Visual saliency guided complex image retrieval. *Pattern Recognit. Lett.* **130**, 64–72 (2020).
23. Agana, N. A. & Homaifar, A. A deep learning based approach for long-term drought prediction. *SoutheastCon 2017.* <https://doi.org/10.1109/SECON.2017.7925314> (2017).
24. Li, J., Yu, C., Gupta, B. B. & Ren, X. Color image watermarking scheme based on quaternion Hadamard transform and Schur decomposition. *Multimed. Tools Appl.* **77**, 4545–4561 (2018).
25. Chopra, M., Singh, S. K., Sharma, A. & Gill, S. S. A comparative study of generative adversarial networks for text-to-image synthesis. *Int. J. Softw. Sci. Comput. Intell. (IJSSCI)* **14**, 1–12 (2022).
26. Gupta, P. Rainfall timeseries data. <https://www.kaggle.com/datasets/poojag718/rainfall-timeseries-data> (2023). Accessed 30 Jan 2024.

Acknowledgements

Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2024R 343), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. The authors extend their appreciation to the Deanship of Scientific Research at Northern Border University, Arar, KSA for funding this research work through the project number NBU-FFR-2024-1092-13.

Author contributions

Final Manuscript Revision, funding, Supervision: B.B.G., K.T.C.; study conception and design, analysis, and interpretation of results, methodology development: A.G., V.A., draft manuscript preparation, figure and tables: S.B., A.A., R.W.A.

All authors reviewed the results and approved the final version of the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to B.B.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024