RESEARCH ARTICLE Indian Journal of Agricultural Research, Volume 59 Issue 8: 1303-1311 (August 2025)

Crop Yield Prediction using Deep Learning Algorithm based on CNN-LSTM with Attention Layer and Skip Connection

Vijay H. Kalmani[1], Nagaraj V. Dharwadkar[2], Vijay Thapa[1]

10.18805/IJARe.A-6300

ABSTRACT

Background: Accurate prediction of crop production is essential for efficient agricultural resource planning. Factors such as weather, soil moistureand temperature have a direct impact on crop yields, making precise forecasting vital.

Methods: This study presents a hybrid model that enhances crop production prediction by integrating a 1D Convolutional Neural Network (CNN) with a Long Short-Term Memory (LSTM) network and an attention layer. The model is specifically applied to wheat and rice, major crops in India. The model evolves into a CNN-LSTM hybrid, designed to improve prediction accuracy by incorporating modifications, including multi-head attention and a multiplication skip connection.

Result: When compared with conventional methods like Support Vector Regressor, Decision Tree Regressor and Random Forest Regressor, the proposed hybrid model shows significantly better performance. It achieves a Root Mean Square Error (RMSE) of 0.017, indicating low prediction error, a Mean Absolute Error (MAE) of 0.09 and a strong correlation between predicted and actual yields, with an R^2 of 0.967.

Key words: Attention layer, CNN-LSTM, Decision tree regressor, Deep learning, Random forest regressor, Skip connection support vector regressor.

INTRODUCTION

Indian agriculture boasts a rich history, beginning with the Indus Valley Civilization (2600-1600 BCE), which cultivated rice, wheat, barleyand cotton using irrigation systems (Bheemabai, 2017). During the Mughal Empire (1526-1857), new crops like tobacco and potatoes were introduced, alongside advancements in farming techniques such as crop rotation (selfstudyhistory) (Batra *et al*., 2021). Colonial British policies focused on exporting raw materials, leading to decreased food production and famine in the 19th century (Shobanadevi *et al*., 2023). The Green Revolution of the 1960s and 1970s introduced high-yield crops and modern farming methods, significantly boosting food production (Howlett, 2008).

Despite its historical importance, Indian agriculture faces challenges today, particularly from climate change. Erratic rainfall, droughts and floods are reducing crop yields and productivity (Shook *et al*., 2021; Dwivedi *et al*., 2022). Dependence on monsoon rains makes agriculture vulnerable to weather variations, leading to crop failures and reduced farmer income (Mahdi *et al*., 2020). Rising temperatures are degrading soil quality and affecting crop growth (Elavarasan *et al*., 2020). While other countries adopt irrigation and modern technologies to mitigate climate impacts, many Indian farmers still use traditional methods, struggling to adapt (Majeed *et al*., 2021; Durai and Shamili, 2022).

Machine learning, a subset of artificial intelligence, enables computers to identify patterns and insights from data without explicit programming, improving over time with experience (Keerthana *et al*., 2021). Deep learning, a subset of machine learning, trains neural networks to understand complex data representations and is used for tasks like image recognition, recommender system, natural language processingand predictive analytics (Shen *et al*., 2017; Nevavuori *et al*., 2019; Baek *et al*., 2020). Recent advances include feature engineering-based LSTM models that enhance crop harvest forecasting by creating new features from existing data (Iniyan *et al*., 2023).

Related works

Jhajharia and Mathur (2023) conducted a research work in Rajasthan, India, implementing various machine learning techniques to estimate crop yield on five identified crops. The study discovered that the Random Forest, SVM

[1]Department of Computer Science and Information Technology, Rajarambapu Institute of Technology, Shivaji University Kolhapur, Sakharale-415 414, Maharashtra, India.
[2]Department of Computer Science, Central University of Karnataka, Kalaburgi-585 367, Karnataka, India.

Corresponding Authors: Vijay H. Kalmani, Nagaraj V. Dharwadkar; Department of Computer Science and Information Technology, Rajarambapu Institute of Technology, Shivaji University, Kolhapur, Sakharale-415 414, Maharashtra, India, Department of Computer Science, Central University of Karnataka, Kalaburgi-585 367, Karnataka, India.
Email: vijaykalmani@gmail.com, dharwadkarn@cuk.ac.in

How to cite this article: Kalmani, V.H., Dharwadkar, N.V. and Thapa, V. (2025). Crop Yield Prediction using Deep Learning Algorithm based on CNN-LSTM with Attention Layer and Skip Connection. *Indian Journal of Agricultural Research*. **59(8):** 1303-1311. doi: 10.18805/IJARe.A-6300.

Submitted: 21-08-2024 **Accepted:** 12-11-2024 **Online:** 26-12-2024

and Lasso Regression models performed better in predicting agricultural yield than deep learning models such as Gradient Descent and LSTM. However, the study suggested that a larger dataset and further investigation into soil and rainfall are required for practical applications of prediction models in crop production Panigrahi *et al*. (2023) formulated a forecasting model for the Indian state of Telangana from 2016 to 2018 for Bengal gram, groundnutsand maize. It utilized six supervised regression models: gradient boosting regression, random forest regression, linear regression, decision tree regression, XGBoost regression and voting regression. The research found that the XGBoost Regression and Random Forest Regression models were the most precise. A two-step approach to enhance agricultural yield, involving forecasting seasonal rainfall using modular artificial neural networks (MANNs), followed by using the rainfall data and crop-specific land area to forecast the yield of major kharif crops with support vector regression (SVR) (Khosla *et al*., 2020).

In a study by Gopal and Bhargavi (2019), a hybrid MLR-ANN model was developed that utilized the coefficients and bias from a multiple linear regression (MLR) model to initialize the weights and bias in the input layer of the artificial neural network (ANN) model, in place of random weights and bias. This approach improved the accuracy of the model over traditional methods. Nigam *et al*. (2019) investigated various machine-learning algorithms for forecasting crop yield based on variables such as temperature, rainfall, seasonand area. Simple RNN and LSTM were used to predict temperature and precipitation initially. Keerthana *et al*. (2021) discussed various machine learning algorithms, including AdaBoost regressor, Random Forest, Gradient Boosting, Decision Treesand KNN classifiers. The study found that an ensemble model consisting of Decision Tree Regressor and AdaBoost Regressor produced the most precise outcomes. Khaki *et al.* (2020) suggested a novel approach that merges CNNs and RNNs. CNN captures both the spatial relationships among soil data gathered at various depths and the temporal dependencies in meteorological data, while RNN represents the rising trend in crop production over time due to ongoing advancements in plant breeding and management techniques. Sivanantham *et al*. (2022) developed a new method called QRECF-DFFMPC to improve prediction accuracy while minimizing time consumption. This approach comprises an input layer, hidden layersand an output layer. The empirical orthogonal function in hidden layer 1 is used to select appropriate features. Quantile regression is then applied in hidden layer 2 to evaluate the features and produce the regression result for each data point. Satpathi *et al*. (2023) conducted a comparative analysis for Chhattisgarh using ANN, LASSO, ELNET and ridge regression with 21 years of historical rice data from three districts: Raipur, Surgujaand Bastar. The study found that ANN performed better with Raipur and Surguja data, while ELNET performed better with Bastar data. Additionally, different ensemble models were used, with performance being comparable for Raipur and Surguja, while Bastar performed better with Random Forest.

MATERIALS AND METHODS

Dataset

The crop prediction model is trained on a dataset featuring information on rice and wheat, including soil, climate conditions and yield potential (Fig 1). It also incorporates historical weather and soil data from various regions. (Kaggle, 2024) this dataset, which covers nearly all Indian states, is essential for training and assessing the model's accuracy in predicting crop yields based on current conditions.

Theoretical background

CNN

Convolutional Neural Networks (CNNs) are a class of deep neural networks widely used for image and signal processing tasks, such as image classification, object detection and speech recognition (Kiranyaz *et al*., 2021). A 1D CNN, specifically, is designed to process one-dimensional signals like audio or time series data. It can be viewed as a special case of 2D CNNs, where the input is a one-dimensional sequence and the filters are one-dimensional vectors (Srivastava *et al*., 2022). The main advantage of 1D CNNs is their ability to learn local features from input signals, such as sound wave shapes or time series patterns, making them effective for tasks like speech recognition or anomaly detection. The core of CNNs is the convolution operation, a mathematical process where a

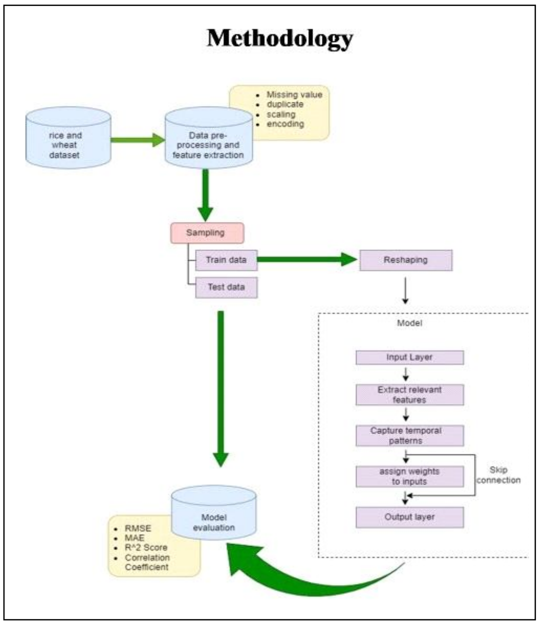

Fig 1: Architecture of the hybrid model integrating a 1D CNN, LSTM and attention layer for crop yield prediction.

filter slides over an input signal to produce a set of feature maps, crucial for extracting relevant features. To analyse the proposed crop yield prediction, we used the CNN architecture as shown in Fig 2.

$$yij = \sum_m \sum_n x(i - m)(j - n)W_{mn} + b \quad ...(1)$$

x = Input image.
w = Filter.
b = Bias term.

Activation function

After each convolution operation, an activation function is applied to the output of the layer. The network may learn more complicated patterns in the input thanks to the activation function, a non-linear function that incorporates non-linearity into the network. Some commonly used activation functions in CNNs include ReLU, sigmoid and tanh.

$$y = f(x) \quad ...(2)$$

x= Input to the activation function.

Pooling operation

Pooling is a down sampling operation that reduces the spatial dimensions of the feature maps while preserving their important features. The two most commonly used pooling operations in CNNs are max pooling and average pooling.

$$y_{ij} = max_{m,n} x\ (i*s+m)\ (j*s+n) \quad ...(3)$$

x = Input feature map.
y = Output of the pooling operation.
s = Stride (the distance between adjacent pooling regions).
m and n= Indices of the pooling region.

Fully connected layer

Each neuron in it is a conventional neural network is linked to every other neuron in the layer above it. A matrix multiplication operation is used to calculate the result of a fully connected layer, which continues by a bias term and an activation function.

$$y = f(w_x + b) \quad ...(4)$$

x = Input.
W = Weight matrix.
b = Bias term.

LSTM

It is a form of RNN (Saini and Nagpal, 2022) which incorporates gating mechanisms to selectively retain or discard information from previous time steps. This makes it well-suited for processing long sequences of data and modelling long-term dependencies. The LSTM cell contains a forget gate, an input gate, a candidate activation, a cell state update, an output gateand a hidden state output as shown in Fig 3. The forget gate in an LSTM network decides the extent to which the prior cell state should be preserved, while the input gate regulates the amount of the candidate activation that must be appended in the cell state. The output gate determines the amount of the cell state that should be emitted as the hidden state. The equations for an LSTM cell can be written as:

$$f_t = \sigma(W_f[x_t, h_{t-1}] + b_f) \quad ...(5)$$

$$i_t = \sigma(W_i[x_t, h_{t-1}] + b_i) \quad ...(6)$$

$$c_t = \tan \tan h(W_c[x_t, h_{t-1}] + b_c) \quad ...(7)$$

$$c_t = f_t \times C_{t-1} + i_t \times g_t \quad ...(8)$$

$$O_t = \sigma(W_o[x_t, h_{t-1}] + b_o) \quad ...(9)$$

$$h_t = O_t \times \tan \tan h(C_t) \quad ...(10)$$

$i_t \rightarrow$ Input gate.
$f_t \rightarrow$ Forget gate.
$O_t \rightarrow$ Output gate.
$h_t \rightarrow$ Hidden gate.
$C_t \rightarrow$ Cell state at timestamp (t).
$C_t \rightarrow$ Candidate for cell state at timestamp (t).

Attention layer

An attention layer in neural networks selectively focuses on relevant parts of the input data by computing a weight vector based on the similarity between a query vector and key vectors (Ahmad *et al*., 2018). This weight vector is used to compute a weighted sum of value vectors, producing the

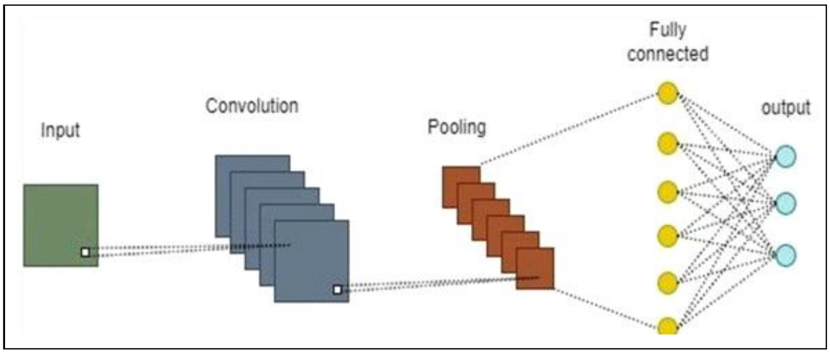

Fig 2: Architecture of the CNN component, detailing the layers involved in feature extraction from input data.

output of the attention layer. Multi-head attention allows the model to capture multiple aspects of the input data (Filippi *et al*., 2019), commonly used in natural language processing and computer vision.

Skip connection

Skip connections, also known as residual connections, bypass one or more layers in a neural network to mitigate vanishing gradient problems, making it easier for the network to learn complex representations (Zhang *et al*., 2020).

Addition-based skip connection

In this method, the output of a layer is added to the output of a previous layer, preserving original information and aiding in gradient flow during training.

Multiplication-based skip connection

This approach multiplies the output of a layer elementwise with the output of a previous layer, selectively emphasizing important features. Both methods of skip connections improve deep neural network performance by enabling the simultaneous learning of shallow and deep features.

Proposed model

A crop yield prediction model that incorporates a 1D CNN, LSTMand attention layer is depicted in Fig 4. The LSTM records temporal patterns, the CNN extracts features from environmental inputand the attention layer improves accuracy by concentrating on important details. By combining LSTM and attention outputs, skip connections enhance pattern recognition.

First, we reshape the input data to have the shape (n,t,m), where n is the no. of samples, t is the no. of timesteps and m is the no. of features.

Then, we define the model architecture as follows:

Input layer: $x \in R^{n \times t \times m}$

CNN layer: C = CNN $(X) \in R^{n \times t \times} c_{units}$

LSTM layer: (q) = LSTM $(C) \in R^{n \times t \times} q_{units}$

Attention layer: (a) = Attention $(q) \in R^{n \times} a_{qunits}$

Skip Connection: $(S) = (S) = q_{units} * C_{units}$

Flatten layer: (r) = Dense $(S) \in R^{n \times t \times} r_{units}$

Output layer: (y) = Dense $(r) \in R^{n \times 1}$

Where,

CNN: Convolutional Neural Network layer with h_{units} units.

LSTM: Long Short-Term Memory layer with q_{units} units.

Attention: Attention layer that takes as input the LSTM layers and outputs an attention vector of a_{units} dimensions.

Dense: Fully connected layer with r_{units} units.

The model is trained to minimize the mean squared error loss function:

$$\text{MSE}(Y, y) = \frac{1}{n} \sum_{i=1}^{n} (Y_i - y_i)^2 \qquad ...(11)$$

with 500 epochs using the Adam optimizer with a batch size of 32.

Algorithm 1: CNN-LSTM with Attention.

1: Preprocess and reshape (X, Y).

2: Input: Train_X, Train_Y.

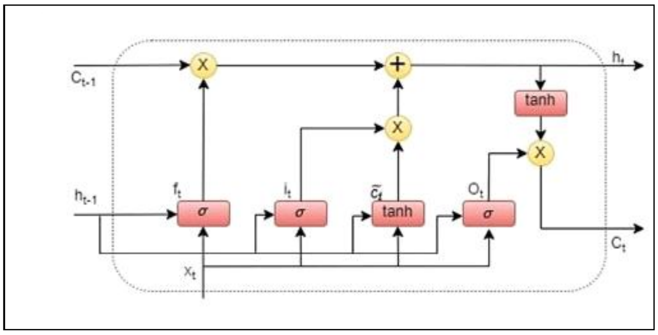

Fig 3: Structure of an LSTM cell, highlighting its gating mechanisms that manage information flow over time.

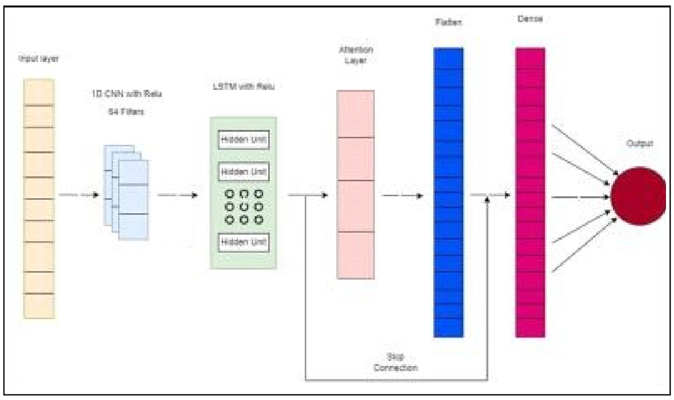

Fig 4: Hybrid 1D CNN-LSTM model with attention layer and skip connections.

3: Hyper-Parameters: optimizer, rate, pool size, batch size.
4: Initialize ().
5: Model=sequential ().
6: Convolution (filters, kernel size, activation).
7: LSTM (units, activation).
8: Attention (weights, activation)
9: Add skip connections:
9.1: LSTM_Skip = LSTM (units, activation).
9.2: Attention_Skip = Attention (weights, activation).
10: Model. Compile (Train_X, Train_Y, epochs, batch size).

RESULTS AND DISCUSSION

The dataset we used for this study had 12 characteristics and was filtered for rice and wheat, yielding about 5000 rows. Both training and testing datasets were separated and the CNN layer was in charge of feature selection. Metrics such as the correlation coefficient, R^2 score, RMSE and MAE were used to assess the performance of the various models that were tested. The CNN-LSTM model was modifiedand the best attention layer was chosen to be combined with skip connections after it was compared to other attention layers. Table 1 demonstrates that the CNN-LSTM with skip connections and multi-head attention performed better than the other models.

Also, we have implemented other machine-learning algorithms like decision tree regressor, random forest regressor and support vector regressor using Jupiter. Since, this is a regression model we don't have accuracy as an evaluation metric therefore for accuracy we have multiplied the correlation coefficient by 100.

$$RMSE = \frac{1}{n}\sqrt{\sum_{i=1}^{n} (y_i - y_i)^2} \quad ...(12)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} \left| y_i - y_i \right| \quad ...(13)$$

n = No. of observations.
y_i = Actual value.
y_i = Predicted value.

$$R^2 = \frac{1 - SS_{res}}{SS_{tot}} \quad ...(14)$$

SS_{res} = Sum of squared residuals.
SS_{tot} = Total sum of squares.

In the Table 2, we evaluated the performance of different models and compared their accuracy. Our modified model CNN-LSTM multi-head attention with multiplication skip connection outperformed other models with 98% accuracy.

Root mean square error (RMSE) measures the difference between actual and predicted values, with a lower RMSE indicating higher model accuracy. In Fig 5, the support vector regressor has the highest RMSE, while the CNN-LSTM multi-head attention with multiplication skip connection has the lowest, showing minimal error. Mean Absolute Error (MAE) similarly gauges prediction accuracy, with the CNN-LSTM model having the smallest MAE in Fig 6.

Table 1: Performance metrics comparison of crop yield prediction models.

Models	RMSE	MAE	R^2 Score	Correlation coefficient
CNN-LSTM	0.020	0.010	0.947	0.973
CNN-LSTM self attention	0.023	0.012	0.937	0.979
CNN-LSTM scaled dot-product attention	0.022	0.012	0.944	0.978
CNN-LSTM multi-head attention	0.019	0.010	0.958	0.983
CNN-LSTM multi-head attention with addition skip connection	0.019	0.010	0.958	0.980
CNN-LSTM multi-head attention with concatenation skip connection	0.018	0.011	0.963	0.982
CNN-LSTM multi-head attention with multiplication skip connection	0.017	0.009	0.967	0.984

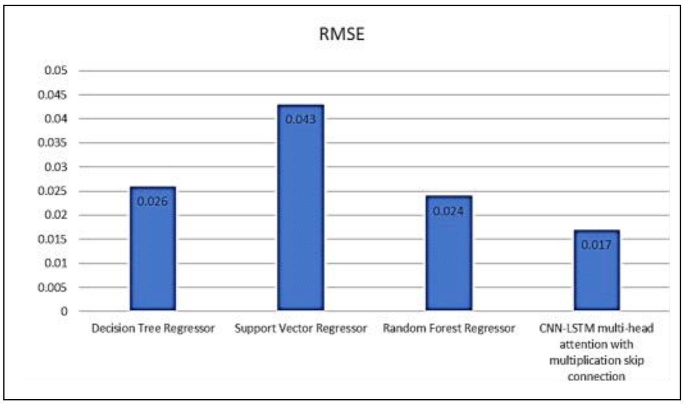

Fig 5: Please provide self-explanatory figure titles root mean square error (RMSE) comparison of prediction models.

The R^2 score, ranging from 0 to 1, indicates how well the model explains the variance in the target variable. The CNN-LSTM model in Fig 7 achieves the highest R^2 score of 0.959, reflecting excellent performance.

In Fig 8, we can see the line graph of actual data and predicted data. We have taken 100 data points for the comparison. The value in the fig is between 0 to 1 because scaling was performed in the original data. In the graph, the predicted data at some points like 42,49 and 60 have some high spike than the actual data which means still the model is not 100% accurate.

Table 2: Performance evaluation of various crop yield prediction models.

Model	Accuracy
Decision tree regressor	95%
Support vector regressor	90%
Random forest regressor	96%
CNN LSTM	97%
CNN-LSTM multi-head attention with multiplication skip connection	98%

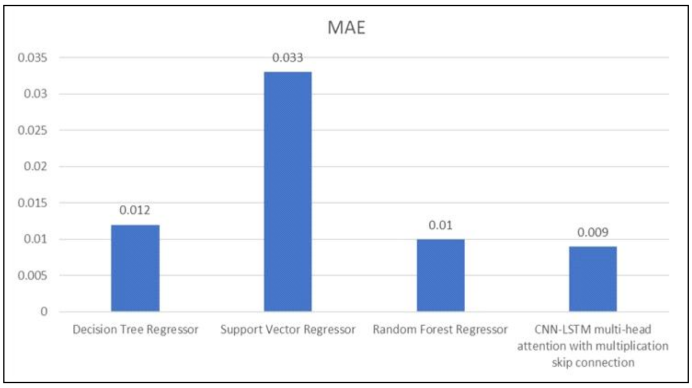

Fig 6: Mean absolute error (MAE) of different models.

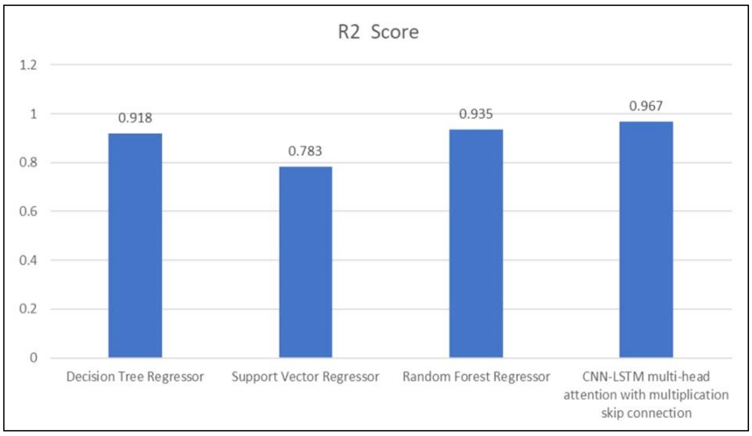

Fig 7: R^2 score analysis for crop yield prediction models.

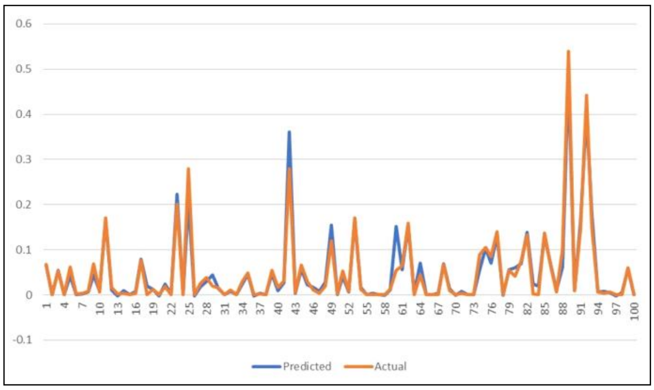

Fig 8: Actual vs. predicted crop yields.

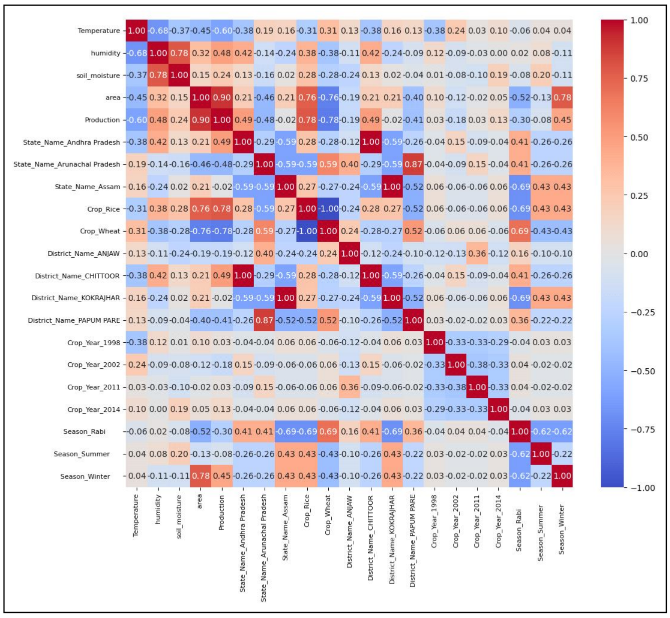

Fig 9: Correlation matrix of variables.

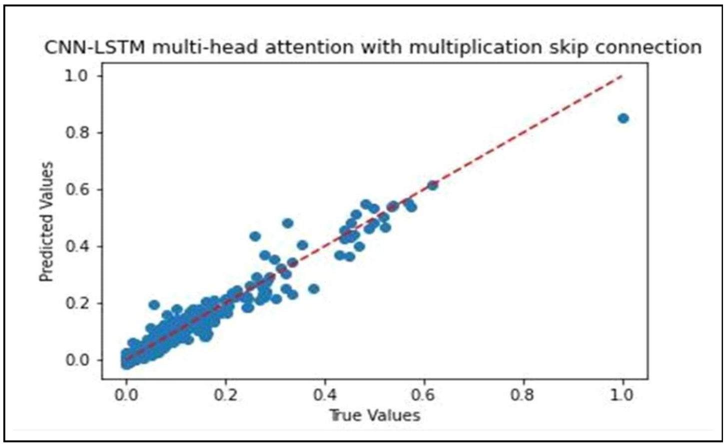

Fig 10: CNN-LSTM mutli-head attention with multiplication skip connection.

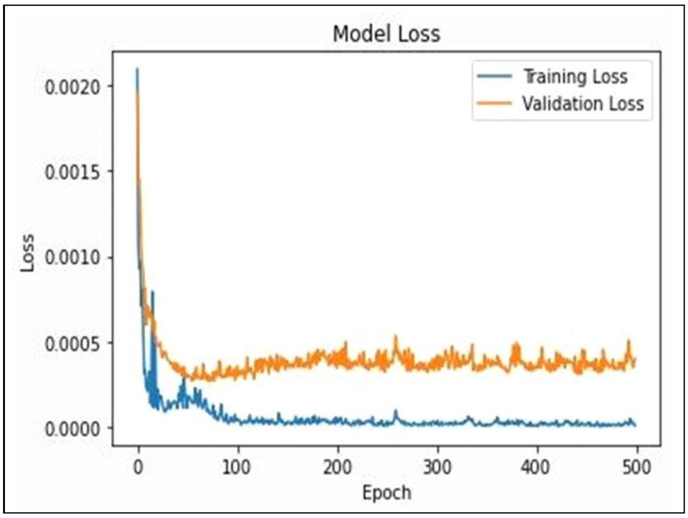

Fig 11: Training and validation loss over epochs.

In Fig 9, the diagonals are all 1/dark red, representing a perfect correlation since each variable is compared to itself. As correlations strengthen between different variables, the numbers increase and the colors darken accordingly.

A scatter plot visually represents the relationship between two variables by plotting data points on a two-dimensional graph. A trend line (diagonal) indicates the strength of the relationship: tightly clustered points suggest a strong relationship, while scattered points indicate a weaker one. Fig 10 show that data points in the Hybrid 1D CNN-LSTM with attention layer have points more closely grouped, though outliers are present.

We trained the Hybrid 1D CNN-LSTM model for 200 epochs using the MSE loss function and Adam optimizer. As shown in Fig 11, the training loss drops quickly during the first 100 epochs and then continues to decrease more gradually. The validation loss follows a similar pattern. The loss function's downward trend indicates that the model improves over time, but it hasn't yet reached its optimal performance, as the loss is still above zero.

CONCLUSION

We developed a modified CNN-LSTM model for predicting crop yields, specifically for rice and wheat, using a dataset containing crop, soiland climate information. We trained and evaluated several models, including decision tree regressor, random forest regressor, support vector regressor and various CNN-LSTM variations (*e.g.*, CNN-LSTM with multi-head attention and different skip connections). Among these, the CNN-LSTM multi-head attention with multiplication skip connection outperformed the others based on metrics like RMSE, MAE, R^2 score and correlation coefficient. Although the results were promising, the study's small dataset size is a limitation. Future work with larger datasets may yield even better results. Table 1 highlights that the top two models-Hybrid 1D CNN-LSTM with attention and Random Forest-could potentially deliver higher performance if combined in an ensemble.

Funding

The authors did not receive support from any organization for the submitted work. No funding was received to assist with the preparation of this manuscript. No funding was received for conducting this study. No funds, grants, or other support was received.

Conflict of interest

The authors have no relevant financial or non-financial interests to disclose. The authors have no conflicts of interest to declare that are relevant to the content of this article. All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The authors have no financial or proprietary interests in any material discussed in this article.

REFERENCES

Ahmad, I., Saeed, U., Fahad, M., Ullah, A., Rahman, M.H.U., Ahmad, A. and Judge, J. (2018). Yield forecasting of spring maize using remote sensing and crop modeling in Faisalabad-Punjab Pakistan. *Journal of the Indian Society of Remote Sensing.* **46(10):** 1701-1711. https://doi.org/10.1007/s12524-018-0825-8.

Baek, S.S., Pyo, J. and Chun, J.A. (2020). Prediction of water level and water quality using a CNN-LSTM combined deep learning approach. *Water.* **12(12):** 3399. https://doi.org/10.3390/w12123399.

Batra, K. and Gandhi, P. (2021). Prediction of nematode population dynamics using weather variables in leguminous crops. *Indian Journal of Agricultural Research.* **55(3):** 383-386. doi: 10.18805/IJARe.LR-4572.

Bheemabai S. Mulage. (2017). History of agriculture system in India: A legal perspective. *International Journal of Humanities Social Sciences and Education.* **4(7):** 25-30.

Dwivedi, A.K., Upreti, H. and Ojha, C.S.P. (2022). Wheat yield modelling using infocrop and DSSAT crop simulation models. *Indian Journal of Agricultural Research.* **56(6):** 646-652. doi: 10.18805/IJARe.A-5981.

Durai, S.K.S. and Shamili, M.D. (2022). Smart farming using Machine Learning and Deep Learning techniques. *Decision Analytics Journal.* **3:** 100-041. https://doi.org/10.1016/j.dajour.2022.100041.

Elavarasan, D. and Vincent, P.M.D. (2020). Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications. *IEEE Access*. **8:** 86886-86901. https://doi.org/10.1109/access.2020.2992480.

Filippi, P., Jones, E.J., Wimalathunge, N.S., Somarathna, P.D., Pozza, L.E., Ugbaje, S.U. and Bishop, T.F. (2019). An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning. *Precision Agriculture*. **20:** 1015-1029.

Gopal, P.M. and Bhargavi, R. (2019). A novel approach for efficient crop yield prediction. Computers and Electronics in Agriculture. 165: 104968. https://doi.org/10.1016/j.compag.2019.104968.

Howlett, Peter. (2008). Travelling in the social science community: assessing the impact of the Indian Green Revolution across disciplines, Economic History Working Papers 22513, London School of Economics and Political Science. Department of Economic History.

Iniyan, S., Varma, V.A. and Naidu, C.T. (2023). Crop yield prediction using machine learning techniques. *Advances in Engineering Software.* **175:** 103-326. https://doi.org/10.1016/j.advengsoft.2022.103326.

Jhajharia, K. and Pratistha, M. (2022). Machine learning approaches to predict crop yield using integrated satellite and climate data. *International Journal of Ambient Computing and Intelligence.* **13(1):** 1-17.

Kaggle. (2024). Crop Production Prediction Kaggle, www.kaggle.com/code/milan400/crop-production-prediction/input?select=crop_production.csv.

Keerthana, T., Kaviya, K., Deepthi Priya, S. and Suresh Kumar, A. (2021). AI Enabled Smart Surveillance System. Journal of Physics: Conference Series, vol. 1916, , IOP Publishing Ltd. International Conference on Computing, Communication Electrical and Biomedical Systems (ICCCEBS). 25-26 Mar. Coimbatore, India.

Khaki, S., Wang, L. and Archontoulis, S.V. (2020). A CNN-RNN framework for crop yield prediction. *Frontiers in Plant Science.* **10:** 1750.

Khosla, E., Dharavath, R. and Priya, R. (2020). Crop yield prediction using aggregated rainfall-based modular artificial neural networks and support vector regression. *Environment Development and Sustainability.* **22(6):** 5687-5708. https://doi.org/10.1007/s10668-019-00445-x.

Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M. and Inman, D.J. (2021). 1D convolutional neural networks and applications: A survey. *Mechanical Systems and Signal Processing.* **151:** 107398. https://doi.org/10.1016/j.ymssp.2020.107398.

Majeed, A., Niaz, A., Sameen, A., Ahmad, H.B., Younus, M., Aftab, M. and Arif, M. (2021). Bed planting techniques improved crop yield by efficient use of added nitrogen fertilizer. *Indian Journal of Agricultural Research.* **55(5):** 542-548. doi: 10.18805/IJARe.A-599.

Mahdi, M.D., Mrittika, N.J., Shams, M., Chowdhury, L. and S. Siddique. (2020). A deep gaussian process for forecasting crop yield and time series analysis of precipitation based in Munshiganj, Bangladesh, IGARSS-IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA. pp. 1331-1334. doi: 10.1109 /IGARSS39084.2020.9323423.

Nigam, A., Garg, S., Agrawal, A. and Agrawal, P. (2019). Crop yield prediction using machine learning algorithms. Fifth International Conference on Image Information Processing (ICIIP), Shimla, India. pp. 125-130. doi: 10.1109/ICIIP47207.2019.8985951.

Nevavuori, P., Narra, N. and Lipping, T. (2019). Crop yield prediction with deep convolutional neural networks. *Computers and Electronics in Agriculture.* **163:** 104-859. https://doi.org/10.1016/j.compag.2019.104859.

Panigrahi, Bharati, Kathalaand, K.C.R., Sujatha, M. (2023). A machine learning-based comparative approach to predict the crop yield using supervised learning with regression models. *Procedia Computer Science.* **218:** 2684-2693. Science Direct. https://doi.org/10.1016/j.procs.2023.01.241.

Saini, P. and Nagpal, B. (2022). Deep-LSTM model for wheat crop yield prediction in India. Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT). https://doi.org/10.1109/ccict 56684.2022.00025.

Satpathi, A., Setiya, P., Das, B., Nain, A.S., Jha, P.K., Singh, S. and Singh, S. (2023). Comparative analysis of statistical and machine learning techniques for rice yield forecasting for Chhattisgarh, India. *Sustainability.* **15(3):** 2786. https://doi.org/10.3390/su15032786.

Shen, D., Wu, G. and Suk, H.I. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering.* **19(1):** 221-248. https://doi.org/10.1146/annurev-bioeng-071516-044442.

Shobanadevi, C., Elangaimannan, R. and Vadivel, K. (2023). Stability analysis for yield and Its component characters in blackgram [*Vigna mungo* (L.) Hepper]. *Indian Journal of Agricultural Research.* **57(6):** 737-739. doi: 10.18805/IJARe.A-5638

Shook, J., Gangopadhyay, T., Wu, L., Ganapathysubramanian, B., Sarkar, S. and Singh, A.K. (2021). Crop yield prediction integrating genotype and weather variables using deep learning. *PLoS ONE.* **16(6):** e0252402. https://doi.org/10.1371/journal.pone.0252402.

Sivanantham, V., Sangeetha, V., Alnuaim, A.A., Hatamleh, W.A., Anilkumar, C., Hatamleh, A.A. and Sweidan, D. (2022). Quantile correlative deep feedforward multilayer perceptron for crop yield prediction. *Computers and Electrical Engineering.* **98.** Published. https://doi.org/10.1016/j.compeleceng.2022.107696.

Srivastava, A.K., Safaei, N., Khaki, S., Lopez, G., Zeng, W., Ewert, F., Gaiser, T. and Rahimi, J. (2022). Winter wheat yield prediction using convolutional neural networks from environmental and phenological data. *Scientific Reports.* **12(1)**. https://doi.org/10.1038/s41598-022-06249-w.

Zhang, W., Quan, H., Gandhi, O., Rajagopal, R., Tan, C.W. and Srinivasan, D. (2020). Improving probabilistic load forecasting using quantile regression NN with skip connections. *IEEE Transactions on Smart Grid.* **11(6):** 5442-5450. https://doi.org/10.1109/tsg.2020.2995777.