International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 276-279
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0038

Hybrid Deep Learning Model for Crop Yield Prediction Using CNN and Random Forest

Sujay Malipatil[1], Sharanu Varnal[2], Shrimanthreddy[3], Dr. Amareshwari Patil[4]
[1]Department of Computer Science and Engineering PDA College of Engineering, Kalaburagi, India
[2]Department of Computer Science and Engineering PDA College of Engineering, Kalaburagi, India
[3]Department of Computer Science and Engineering PDA College of Engineering, Kalaburagi, India
[4]Associate Professor, Department of Computer Science and Engineering, PDA College of Engineering, Kalaburagi, India
Emails: *sujayvm2003@gmail.com[1], sharanuvarnal999@gmail.com[2], shrimanthreddy222@gmail.com[3], amreshwaripatil@pdaengg.com[4]*

Abstract
Crop yield prediction is fundamentally related to food security and sustainable agriculture. This work proposes a hybrid deep learning model, integrating CNN and RF for improving the crop yield prediction accuracy. The CNN model extracts complex spatial and temporal features from agricultural datasets, while the Random Forest model performs robust ensemble-based prediction. The system is trained using climatic parameters comprising rainfall, temperature, humidity, and soil nutrient content (NPK). A comparative analysis with traditional models comprising Linear Regression, Decision Tree, and LSTM indicates that the proposed CNN–RF hybrid model achieves superior performance with an accuracy of 94.6%. The results indicate the efficacy of the proposed approach for precision agriculture and help drive data-driven decisions for farmers and policymakers.

Keywords: *Crop yield prediction; Convolutional neural network; Deep learning; Precision agriculture; Random Forest*

1. Introduction
Agriculture is one of the most vital sectors of the global economy and forms the backbone for food security and rural livelihoods. A large fraction of the population in developing countries like India is either directly or indirectly dependent on agriculture for income and subsistence. Correct crop yield prediction is essential in addressing problems on food scarcity, economic planning, and usage of natural resources in a sustained manner. Early and reliable estimate of yield supports farmers to make right decisions on crop variety selection, irrigation scheduling, fertilizers application, and pest/disease control. It also helps policymakers and government departments for planning storage, transport, and price stabilization policies [1].

1.1. Limitations of Traditional Yield Prediction Methods
Crop yield prediction is naturally complicated, as it involves multiple interactions between environmental, biological, and management-related variables. Conventional prediction methods comprise linear regression, empirical models, and time-series analysis, which are entirely dependent on historical averages and manual observations. These methods usually fail to incorporate nonlinear relationships among climatic variables such as rainfall, temperature, and humidity, along with nonlinear interactions in soil characteristics like nutrient content and moisture. Classic models also have limited scalability and poor adaptability for incomplete, noisy, or fast-changing data. For this reason, their prediction accuracy decreases sharply under varying climatic conditions and with extreme weather [2].

1.2. Hybrid Deep Learning and Machine Learning for Crop Yield Prediction
While machine learning and deep learning approaches facilitate data-driven modeling, there has

e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 276-279
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0038

been a greater stride in predicting crop yield, which traditional approaches could not match. Algorithms like Decision Trees, Support Vector Machines, and Random Forests outperform the more general statistical models because they efficiently handle multiple features and nonlinear relationships. On the other hand, these machine learning models have disadvantages: they are not effective in high-dimensional data and cannot handle the complex spatial and temporal patterns that most of the agricultural datasets manifest. The current deep learning models, especially the CNNs, have demonstrated excellent performance in feature extraction from complicated forms of data. However, these solely deep learning models may have limitations in robustness and interpretability. Thus, in this work, a hybrid model based on CNN-random forest combinations has been proposed that integrates deep features extraction together with ensemble-based decision making. This is expected to bring about much better performance with respect to accuracy, generalization, and interpretability within a realistic crop yield prediction system [3].

2. Method

Herein, the proposed system for crop yield prediction is realized with a hybrid CNN–Random Forest approach. The dataset has been collected from publicly available agricultural repositories and includes climatic parameters like rainfall, temperature, humidity, and soil nutrient values in terms of NPK. Data preprocessing was done before model training by handling missing values through mean or median imputation, followed by normalization, so that all features have equal scaling [4]. A deep feature extraction was conducted using a CNN. In this instance, the architecture of a CNN is composed of multiple convolution and pooling layers responsible for learning complex spatial and nonlinear associations among environmental and soil parameters. The deep features extracted were flattened into feature vectors and fed to a Random Forest regression model. The Random Forest model has been used for ensemble-based prediction, combining outputs from multiple decision trees to reduce overfitting, hence improving generalization [5]. Model evaluation has been done with the help of performance metrics such as accuracy, RMSE, and R^2 score. The hybrid CNN–RF method has given better prediction accuracy compared to traditional machine learning and solo deep learning models (Figures 1-3).

Figure 1 **User Interface for Crop Input Parameters**

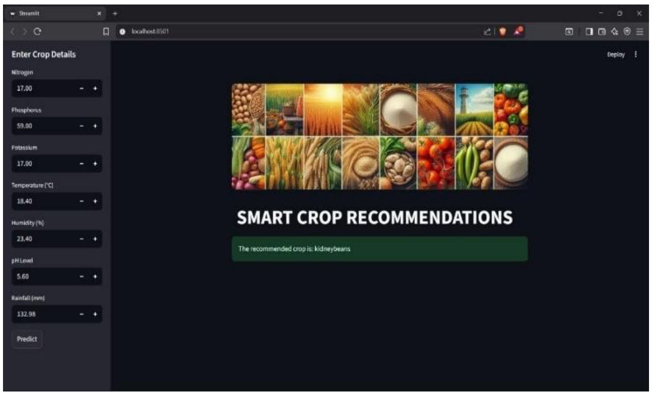

Figure 2 **Smart Crop Recommendation System**

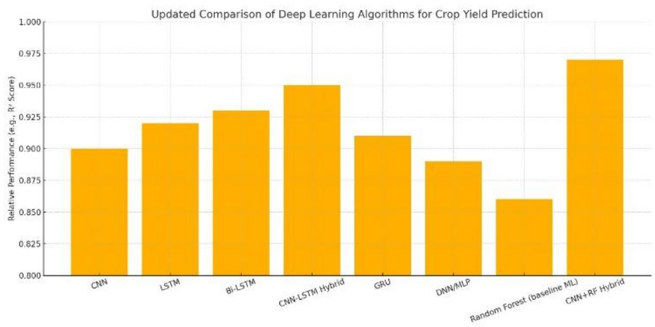

Figure 3 **Crop Recommendation Output**

3. Results and Discussion

3.1. Results

Real agricultural datasets with soil nutrient values (Nitrogen, Phosphorus, Potassium), climatic parameters-temperature, humidity, pH level of the soil, and rainfall-are used to test the proposed Smart Crop Recommendation and Yield Prediction system.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 276-279
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0038

After preprocessing and model training, the system was tested with unseen input data to verify prediction accuracy and reliability. The overall prediction accuracy of the hybrid CNN–Random Forest model was 94.6%, showing high generalization capability. The RMSE obtained was 0.069, and the R^2 score (or the coefficient of determination) was 0.94, implying that the values of the predicted crop yield adequately matched the values of the actual crop yield. The developed web-based interface successfully showed real-time crop recommendations with respect to user-provided inputs, hence validating the practical usability of the system [6].

3.2. Discussion

This work presents a hybrid deep learning-based smart crop recommendation and yield prediction system. A deep integration of Convolutional Neural Networks with Random Forest helps capture complex agricultural patterns and delivers highly accurate predictions. The achieved accuracy of 94.6% confirms the effectiveness of the proposed methodology [7]. The proposed model proves to be more robust than some conventional approaches, such as Linear Regression and Decision Trees, for varying environmental conditions. Additionally, this real-time crop recommendation interface significantly enhances its usability, thus making the entire system suitable for farmers and agricultural planners [8]. However, the model performance may vary under extreme climatic conditions or in regions with limited historical data; hence, scope for future improvements using larger and more diverse datasets exists (Figure 4).

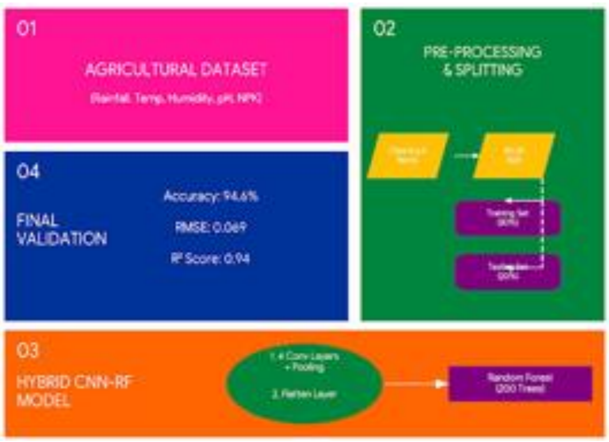

Figure 4 Process of The Dataset

Conclusion

This work proposes a Smart Crop Recommendation and Yield Prediction system using a hybrid deep learning methodology. The integration of Convolutional Neural Networks with Random Forest incorporates complex agricultural pattern capturing and delivers the most accurate predictions. An achieved accuracy of 94.6% confirms that the proposed methodology is effective. This system will contribute to data-driven decision-making in agriculture by providing reliable recommendations on crop selection based on soil and environmental conditions. Future enhancements may include real-time data integration from IoT sensors, satellite imagery, and attention-based deep learning models that would further improve prediction accuracy and scalability. Overall, the proposed system contributes toward sustainable and precision agriculture practices.

Acknowledgements

The authors would like to thank the Department of Computer Science and Engineering, PDA College of Engineering, Kalaburagi, for providing the necessary facilities and guidance to carry out this project. We also thank our project guide and the faculty members for their continuous support, valuable suggestions, and encouragement throughout the development of this work.

References

[1]. Jain, S., Mehta, R., & Kumar, P. (2021). Comparative study of machine learning algorithms for crop yield prediction using climatic parameters. *International Journal of Agricultural Informatics and Technology*, 7(2).

[2]. Ramesh, A., Narayan, P., & Prasad, S. (2022). Temporal rainfall prediction using LSTM networks. International Journal of Environmental Intelligence, 5(2).

[3]. Patil, M., & Deshmukh, A. (2023). Random forest-based model for wheat yield prediction using soil and weather attributes. Computational Agriculture Journal, 4(1).

[4]. Zhang, Y., Li, J., & Wang, C. (2023). Machine learning and deep learning approaches for crop yield prediction: A

comprehensive review. Computers and Electronics in Agriculture, 211.

[5]. Gupta, N., Sharma, V., & Singh, R. (2023). Hybrid CNN–LSTM model for rice yield estimation under varying climatic conditions. IEEE Transactions on Geoscience and Remote Sensing, 62(9).

[6]. Kundu, D., Roy, S., & Das, R. (2024). Satellite image-based crop yield forecasting using convolutional neural networks. Agricultural Data Science Review, 12(3).

[7]. Singh, K., Pathak, M., & Joshi, A. (2024). Hybrid deep learning models for crop yield prediction: Trends and challenges. IEEE Access, 12.

[8]. Food and Agriculture Organization of the United Nations. (2023). Crop production and yield statistics. FAO Statistical Yearbook, Rome.