Journal of Soft Computing Paradigm (ISSN: 2582-2640)
www.irojournals.com/jscp/

# Deep Learning for Crop Yield Forcasting in Agriculture Using Multilayer Perceptron and Convolutional Neural Networks

**Jeevanraja S.[1], Vishnu Vardhan Reddy A.[2], Ashok Reddy S.[3], Hari Krishna Reddy V.[4]**

[1]Assistant Professor, [2-4]Student, Muthayammal Engineering College, Namakkal, Tamilnadu, India

**E-mail:** [1]jeevaraju18@gmail.com, [2]avulavishnuvardhanreddy2003@gmail.com, [3]somulaashokreddy@gmail.com, [4]harikrishnareddy983@gmail.com

**Abstract**

The primary challenge facing the agricultural sector, which is essential for ensuring global food security, is enhancing crop productivity while effectively addressing the challenges posed by plant diseases. Advanced technologies have the potential to completely transform agricultural methods, especially in the areas of computer vision and machine learning. This study uses meteorological as well as fruit and vegetables image datasets to create an integrated agricultural decision support system for crop yield estimation and disease prediction. By enabling early plant disease detection and precise crop yield estimates, the system seeks to improve precision agriculture techniques. To analyze and classify the images and predict the possibility of crop disease harming fruits and vegetables, a Convolutional Neural Network (CNN) deep learning model is used. The Multilayer Perceptron algorithm is used to train the model using a large dataset that contains historical meteorological data, allowing it to identify patterns and connections between environmental conditions. Finally, farmers receive an SMS notice with prediction specifics.

**Keywords:** Precision Agriculture, Disease Prediction, Crop Yield Values, Deep Learning, Neural Network.

## 1. Introduction

Machine learning algorithms analyze multispectral and hyperspectral imagery to enable early detection of diseases and nutritional deficiencies in crops, enabling prompt action.


*DOI: https://doi.org/10.36548/jscp.2024.4.006*
*Received: 11.12.2024, received in revised form: 03.01.2025, accepted: 24.01.2025, published: 07.02.2025*

Farmers can forecast disease outbreaks and maximize agricultural productivity by employing predictive modelling, which makes use of historical and meteorological data. Precision agriculture offers a personalized approach to fruit cultivation and transforms conventional farming methods for more abundant and sustainable crops. It does this by improving decision-making and disease management through data-driven methods. Crop yield prediction takes into consideration crop, weather, soil, and environmental factors and is essential at the local, regional, and global levels. Key qualities for prediction are extracted using decision support models. Precision farming is essential for tackling food security in the face of climate change because it incorporates technology such as variable rate, sensor systems, and management information to increase crop production and quality while lowering environmental impact.

Research on crop yield prediction considers the influence of genotype, environment, and management. Several machine learning models, including random forests, neural networks, and decision trees, have been applied to crop yield prediction, with studies often relying on environmental and management variables due to limited genetic data. Despite challenges, these models can handle large datasets and offer good accuracy. Agriculture remains a key sector for economic development and survival, particularly in India, where its importance continues to grow with increasing output demands [8-10].

The basic steps in precision farming and shown in Figure 1

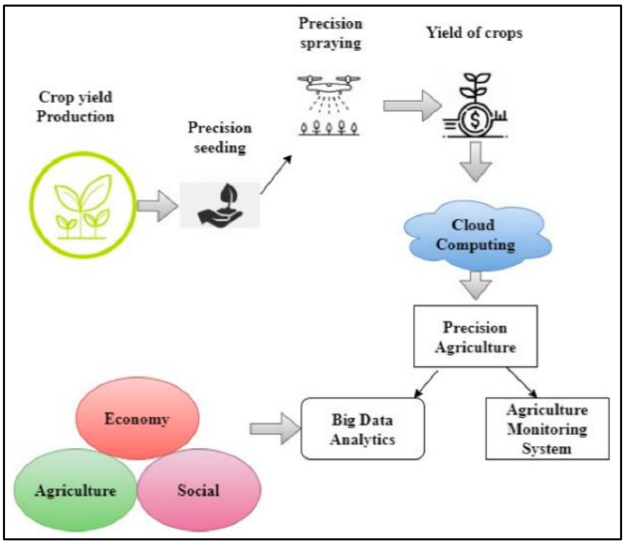


**Figure 1.** Precision Farming System [11]

## 2. Related Work

Nizom Farmon et al. [1] and colleagues used wavelet attention 2-D-CNN for crop-type classification in DESIS image classification, incorporating image dimension reduction and spectral AM. They found that a 48x48 spatial patch size and FA from 2 to 3 provided the best OA, proving the WA-CNN's ability to learn features for image categorization.

S. P. Raja et.al.[2] used ensemble feature selection and classification techniques to predict plant cultivation yield size. The ensemble technique offers better prediction accuracy than existing classification techniques. Modern forecasting techniques can provide financial benefits. Efficient feature selection methods and relevance-based data features are crucial for machine learning model precision.

The research proposes a two-model approach for county-level crop yield prediction: a first model combining 2D-CNN and LSTM attention, and a second model using 3D-CNN and ConvLSTM. The first model uses multi-2D-CNNs for feature extraction, while the second model uses 3D-CNN and ConvLSTM for simultaneous spectral and spatial extraction. The approach shows significant improvements compared to recent models for crop yield prediction [3].

The study by Sisheber et al.[4] improved crop yield estimation in Ethiopian smallholder agricultural systems by assimilating data fusion retrieved from Landsat and MODIS data. The improved yield estimation performance could be valuable for crop production monitoring and food security analysis in fragmented agricultural landscapes.

Arnab Muhuri et al [5] and colleagues proposed a geodesic distance-based scattering power decomposition for compact polarimetric SAR data over agricultural landscapes. The technique decomposes the polarized portion of the total backscattered power proportional to target similarity measures. A compensation strategy using the CP radar vegetation index (CpRVI) is proposed to compensate for pseudo-power components.

Crop rotation is essential for increasing crop yield, improving soil health, and reducing plant disease. However, mapping crop rotation is challenging due to the lack of accurate data. A new crop rotation classification scheme is proposed that integrates temporal information into static crop maps. The scheme defines three main rotation systems, divides them into nine subsystems, and uses Sentinel-1 and Sentinel-2 images to identify them. The 2020 map showed 81% producer, 79% user, and 81% overall accuracies, demonstrating the potential of direct mapping for crop rotation systems [6].

This research proposes an innovative within-season emergence (WISE) phenology normalized deep learning model for scalable within-season crop mapping. The model normalizes crop time-series data using WISE crop emergence dates, allowing it to accommodate spatiotemporal variations in crop phenological dynamics. Results show that the model outperforms calendar-based approaches in Illinois, yielding over 90% accuracy for classifying corn and soybeans at the end of the season. It also provides satisfactory performance 85% of the time, one to four weeks earlier than calendar-based approaches [7].

Precision agriculture is transforming conventional farming methods by incorporating technologies like GPS, drones, and IoT. GPS allows farmers to map and manage their farms, while variable rate technology allows for efficient resource use. Drones and satellites provide real-time imagery for crop health tracking. IoT sensors gather information about ambient factors, soil conditions, and moisture content, enabling better irrigation techniques. Advanced farm management software helps farmers access and analyze data streams, leading to more productive, economical, and ecologically responsible farming.

## 3. Proposed System

The integrated precision agriculture management system is a comprehensive solution for modern agriculture, utilizing a deep learning-based data management system to protect environmental and agricultural data. The system uses edge computing technologies to improve real-time processing, reducing latency and reliance on centralized processing. It also incorporates an AI-driven decision support system, analysing complex datasets to provide farmers with insights into crop management. The system uses convolutional neural network algorithms to categorize diseases from fruit and vegetable images, and MLP to predict crop outcomes using soil and weather datasets. Farmers receive these insights through the text message. Figure 2 shows the proposed flow diagram for disease prediction and crop prediction.

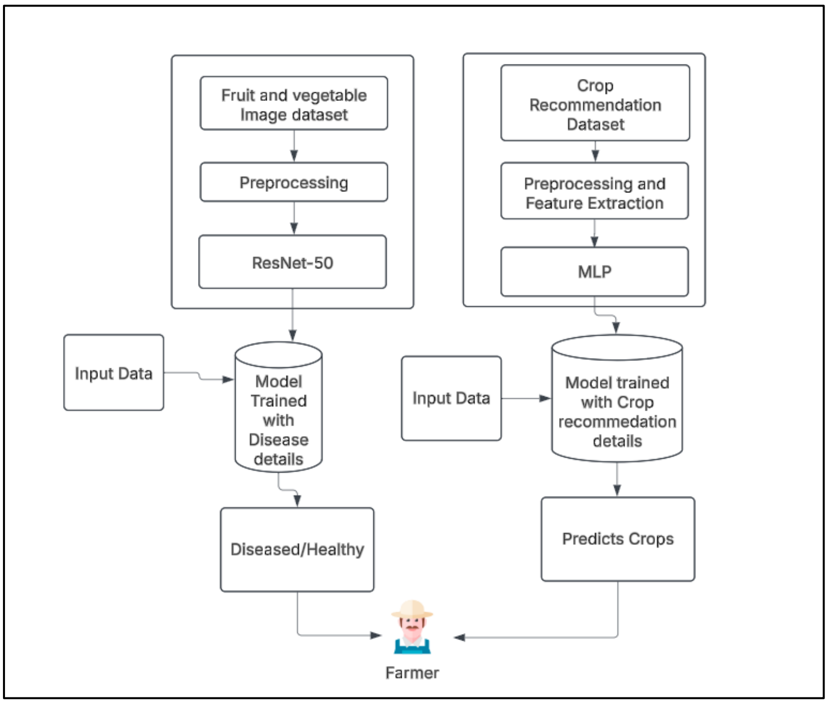


**Figure 2.** Proposed Flowchart

The proposed precision agriculture management system begins with data collection, followed by preprocessing and feature extraction for the machine learning model. The models are trained on 70% of the dataset and tested on the remaining 30%, which was split from the total dataset.

### 3.1 Dataset Description

The first dataset used in the study is a collection of fruit and vegetable images obtained from Kaggle (https://www.kaggle.com/datasets/muhammad0subhan/fruit-and-vegetable-disease-healthy-vs-rotten). It consists of 28 directories, each representing a combination of healthy and rotten images for 14 different types of fruits and vegetables. A total of 1,000 images (500 healthy and 500 diseased) were considered for the study. These images were resized to uniform dimensions of 224x224 pixels. The pixel values were normalized to the range [0, 1] to speed up training and improve convergence. Data augmentation techniques were applied to enhance the model's robustness and increase the dataset size. The images were converted into numerical arrays and labeled using one-hot encoding.

The second dataset for crop prediction was sourced from Kaggle (https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset). This dataset contains 2,200 records, with information on 22 different crops, including rice, maize, chickpea, kidney beans, pigeon peas, moth beans, mung bean, black gram, lentil, pomegranate, banana, mango, grapes, watermelon, muskmelon, apple, orange, papaya, coconut, cotton, jute, and coffee. Each crop is characterized by 7 features, which are crucial for crop recommendation. Table 1 below illustrates the 7 features and their corresponding unique values.

**Table 1.** Dataset Features

| Features | Unique Values |
|---|---|
| N (Nitrogen levels) | 137 unique values |
| P (Phosphorus levels) | 117 unique values |
| K (Potassium levels) | 73 unique values |
| Temperature (°C) | 2,200 unique values |
| Humidity (%) | 2,200 unique values |
| pH | 2,200 unique values |
| Humidity (%) | 2,200 unique values |
| Rainfall (mm) | 2,200 unique values |

Regarding feature variability per crop, each crop has 100 data points with distinct values for temperature, humidity, pH, and rainfall. However, nutrient levels (N, P, K) show slight variations across different crops. The dataset was initially pre-processed to handle missing values. Relevant features such as N, P, K, temperature, humidity, pH, and rainfall were extracted. Label encoding was performed to assign a unique number to each crop, as the MLP could not process categorical data. The features were then scaled to ensure balanced learning across all inputs.

The ResNet-50 was trained on the fruit and vegetable image dataset to identify the disease in the fruits and vegetables and MLP (Multilayer Perceptron) was used for the crop prediction. The MLP was trained on the metrological dataset to predict the crops.

## 4. Experimental Results

The implementation uses a combination of libraries and techniques for both ResNet-50 (CNN) and MLP models. In the ResNet-50 implementation, the tensorflow.keras library is used for model construction and transfer learning. Image preprocessing techniques like resizing are applied using cv2.resize() to ensure uniform input dimensions, followed by normalization by dividing pixel values by 255.0 to scale them between 0 and 1. Data augmentation is performed using ImageDataGenerator to improve model generalization by applying random transformations such as rotation, zooming, and flipping. Transfer learning with pre-trained weights is leveraged through tensorflow.keras.applications.ResNet50, allowing the model to benefit from previously learned features. For the MLP model, data preprocessing begins with handling missing values through fillna() to avoid errors during training. Feature selection is done by narrowing down to relevant meteorological data, including parameters like nitrogen (N), phosphorus (P), potassium (K), temperature, humidity, pH, and rainfall. For categorical variables, LabelEncoder() is used for label encoding. Feature scaling is applied using MinMaxScaler() to normalize the input data between 0 and 1, which helps improve the training stability and model performance. Libraries such as sklearn are used for preprocessing tasks, while pandas, numpy, opencv, and PIL facilitate data manipulation and image handling throughout the process. Different performance measures such as accuracy, sensitivity, specificity, error rate and precision can be derived for analyzing the performance of the system. The evaluation results presented in Table 2 depicts the accuracy of different machine learning algorithms for comparison, with the Multi-Layer Perceptron (MLP) achieving the highest performance at 93%. K-NN achieved 64%, Naive Bayes 52%, and Random Forest 75%. K-NN, based on classifying by nearest neighbors, performs moderately but struggles with noisy or high-dimensional data. Naive Bayes, assuming feature independence, showed the lowest performance at 52%, particularly when data assumptions are not met. Random Forest, using an ensemble of decision trees, delivered better results with 75%. Despite the performance of these models, the MLP stands out as the most effective for the task.

**Table 2.** Accuracy of Crop Recommendation

| Algorithm | Accuracy (%) |
|---|---|
| K-NN algorithm | 64 |
| Navies Bayes | 52 |
| Random Forest | 75 |
| Multi-layer perceptron algorithm | 93 |

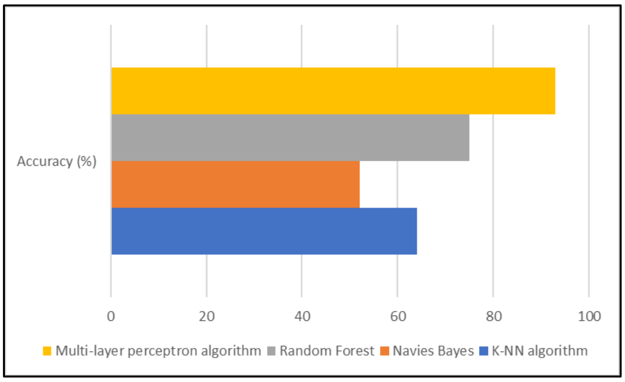


**Figure 3.** Crop prediction- Performance Chart

From Figure 3, proposed MLP algorithm provides improved accuracy in crop datasets analysis

### b. Fruit and Vegetable based Disease Prediction

The graph demonstrates that performance-wise, deep learning algorithms surpass more traditional machine learning methods. The accuracy rate is improved when compared to machine learning algorithms.

**Table 3.** Fruit and Vegetable Disease Detection Accuracy

| Algorithm | Accuracy |
| --- | --- |
| K-NN algorithm | 60% |
| Navies Bayes | 56% |
| Random Forest | 77% |
| ResNet-50 | 92% |

The Table 3 presents the accuracy of different machine learning algorithms, highlighting ResNet-50 as the proposed model for image classification. The K-NN algorithm achieved 60%, showing moderate performance but limited by its reliance on distance metrics, which may not capture the complexities in image data. Naive Bayes performed slightly lower at 56%, struggling due to its assumption of feature independence, which doesn't hold well for image data with correlated pixel values. Random Forest, an ensemble model, achieved 77%, demonstrating better performance by combining multiple decision trees, but still falls short compared to deep learning models. ResNet-50, with an accuracy of 92%, outperforms all other models, because of its deep architecture and residual learning mechanism, which allows it to learn complex features and patterns from images efficiently. This makes ResNet-50 the most effective model for the task.

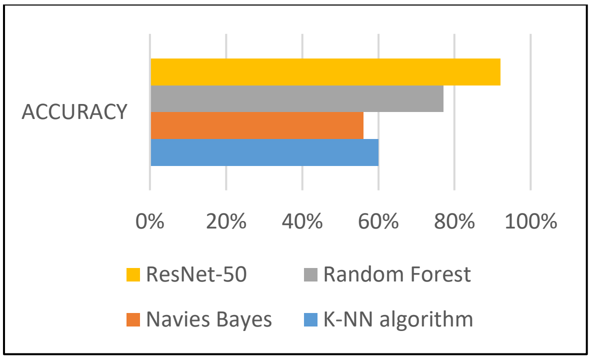


**Figure 4.** Fruit-based Disease Prediction - Chart

From the above graph in Figure 4, ResNet-50 algorithm provides high level accuracy rate than the existing algorithm

## 5. Conclusion

In summary, the “Integrated Precision Agriculture Management System” is an innovative strategy that uses cutting-edge technologies to transform contemporary agriculture. This method offers several advantages that can enhance farming techniques' production, sustainability, and efficiency while addressing several issues in the field. In the future by streamlining resource use, enhancing crop production forecasts, and providing intuitive user interfaces, the system would equip the farmers with the knowledge and resources they need to make wise decisions on a daily basis.

## References

[1] Farmonov, Nizom, Khilola Amankulova, József Szatmári, Alireza Sharifi, Dariush Abbasi-Moghadam, Seyed Mahdi Mirhoseini Nejad, and László Mucsi. "Crop type classification by DESIS hyperspectral imagery and machine learning algorithms." IEEE Journal of selected topics in applied earth observations and remote sensing 16 (2023): 1576-1588.

[2] Raja, S. P., Barbara Sawicka, Zoran Stamenkovic, and G. Mariammal. "Crop prediction based on characteristics of the agricultural environment using various feature selection techniques and classifiers." IEEE Access 10 (2022): 23625-23641.

[3] Nejad, Seyed Mahdi Mirhoseini, et al. "Multispectral crop yield prediction using 3D-convolutional neural networks and attention convolutional LSTM approaches." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 16 (2022): 254-266.

[4] Sisheber, Biniam, et al. "Assimilation of Earth Observation data for crop yield estimation in smallholder agricultural systems." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (2023).

[5] Muhuri, Arnab, et al. "Geodesic distance based scattering power decomposition for compact polarimetric SAR data." IEEE Transactions on Geoscience and Remote Sensing (2023).

[6] Liu, Yuan, Qiangyi Yu, Qingbo Zhou, Cong Wang, Sonoko Dorothea Bellingrath-Kimura, and Wenbin Wu. "Mapping the complex crop rotation systems in Southern China considering cropping intensity, crop diversity, and their seasonal dynamics." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 15 (2022): 9584-9598.

[7] Yang, Zijun, Chunyuan Diao, and Feng Gao. "Towards Scalable Within-Season Crop Mapping With Phenology Normalization and Deep Learning." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 16 (2023): 1390-1402.

[8] Wang, Yiqun, Hui Huang, and Radu State. "Cross Domain Early Crop Mapping using CropSTGAN." IEEE Access (2024).

[9] Qu, Xuzhou, Dong Shi, Xiaohe Gu, Qian Sun, Xueqian Hu, Xin Yang, and Yuchun Pan. "Monitoring lodging extents of maize crop using multitemporal GF-1 images." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 15 (2022): 3800-3814.

[10] Khaki, Bahare Delsous, Mansour Chatrenour, Mir Naser Navidi, Masoud Soleimani, Saham Mirzaei, and Stefano Pignatti. "Land Suitability Assessment Based on Feature-

Level Fusion of Sentinel-1 and Sentinel-2 Imagery: A Case Study of the Honam Region of Iran." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (2024).

[11] Song, Chuanhong, Wenbo Ma, Junjie Li, Baoshan Qi, and Bangfan Liu. "Development trends in precision agriculture and its management in china based on data visualization." Agronomy 12, no. 11 (2022): 2905.