Crop Disease Detection Using the Machine Learning Techniques

L.V. Suresh Kumar, P. Dilleshwara Rao, N. Hari Priya, P. Dinesh, V. Pravallika, M. Avinash, K. Sai praneeth, L. Charath Chandu

Department of Electrical and electronics Engineering, GMR Institute of Technology, Rajam, india.

ABSTRACT

Agriculture, fundamental to human civilization, confronts ongoing challenges from crop diseases that hinder plant growth and jeopardize food security. Swift and precise disease detection is crucial to curb their spread and limit crop damage. However, conventional methods of disease prediction and categorization are often cumbersome and inaccurate, resulting in significant yield reductions if diseases remain undetected. Recent research has concentrated on utilizing machine learning methods like Random Forest (RF), Support Vector Machine (SVM), Decision Tree (DT), alongside deep learning techniques such as Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), Deep Convolutional Neural Network (DCNN), and Deep Belief Network (DBN) for identifying crop pests. By integrating computer-aided imaging technologies into farming practices, farmers can promptly identify and classify diseases, thereby reducing losses and increasing productivity. This article presents an overview of the latest progress in machine learning and deep learning strategies for crop pest detection, emphasizing their efficacy and challenges. It also discusses the potential impact of these techniques on agricultural sustainability and suggests future research directions to improve crop health monitoring and management.

Keywords**:** Image processing, feature extraction ,Remote sensing, Random Forest (RF), Support Vector Machine (SVM), Deep Convolutional Neural Network (DCNN), and Deep Belief Network (DBN).

INTRODUCTION

Plant pests and diseases are a significant threat to almost all major types of plants and global food security. This study presents a detailed systematic literature review on the latest approaches, datasets, and challenges in plant disease detection[1]. The agricultural production is greatly affected by various plant diseases. Classifying the severity of crop diseases is the requirement for formulating disease prevention and control strategies[2]. Precision agriculture is a rapidly developing field aimed at addressing current concerns about agricultural sustainability. Machine learning is the cutting edge technology underpinning precision agriculture, enabling the development of advanced disease detection and classification methods[3]. The comprehensive study underscores the critical threat plant pests and diseases pose to global food security, prompting a systematic exploration of advanced detection methods and disease severity classification to formulate effective prevention strategies. Precision agriculture, propelled by cutting-edge technologies like machine learning, emerges as a beacon of hope for sustainable farming practices, offering tailored resource management and optimized yields while mitigating environmental impact. Machine learning's prowess in analyzing vast datasets facilitates early disease detection and precise classification, revolutionizing agriculture by empowering farmers with real-time insights and decision support. Despite strides made, challenges persist, including data standardization and accessibility, but ongoing research promises to leverage emerging technologies to bolster traceability and collaboration across the agricultural value chain, fostering a future where food security and sustainability converge seamlessly.

LITERATURE REVIEW

An in-depth analysis of artificial intelligence (AI), machine learning (ML), and deep learning (DL) techniques for plant disease detection is presented in a recent paper. The study evaluates existing literature through a systematic assessment of publicly available datasets and data collection methodologies. Findings highlight the use of vision-based AI methods, the efficacy of classifiers like logistic regression (LR) and support vector machines (SVMs), challenges in disease localization, and the increasing adoption of cognitive convolutional neural networks (CNNs) with attention mechanisms and transfer learning. Notably, the study emphasizes the necessity for developing compact models suitable for smaller devices, capable of handling diverse crops and diseases with fewer parameters.[1] In our study, we delve into the categorization of crop diseases using a modified light-weight convolutional neural network (CNN) with an attention mechanism. Our main focus is on accurately diagnosing agricultural illnesses, especially in their early stages when disease severity variations are subtle. We enhance the fine-grained categorization by incorporating multi-scale convolution kernels and a coordinate attention mechanism into the Squeeze next architecture. To evaluate our model's performance, we utilize the AI Challenger 2018 plant disease detection dataset. Our results reveal an impressive accuracy of 91.94%, outperforming the original Squeeze next model by 3.02 percentage points. In comparative tests with mobilenetv2, exception, and ReseNet50, our approach demonstrates the ability to maintain a compact model size while achieving slightly higher accuracy than exception.[2] Machine Learning and Deep Learning for Plant Disease Classification and Detection, it utilize machine learning and deep learning approaches within precision agriculture to classify and identify plant diseases. Through your research, you're exploring how these technologies can enhance disease detection methods and address concerns related to agricultural sustainability. This study provides accessible datasets for plant disease detection and classification, introducing a novel approach to categorizing research based on the employed technique, whether classification or object detection. As part of your investigation, you're conducting a thorough computer analysis using state-of-the-art algorithms to identify plant diseases and predict their occurrence. This findings indicate that YOLOv5 demonstrates remarkable object detection accuracy, while ResNet50 and MobileNetv2 strike the best balance between accuracy and training time for image classification tasks.[3]

Accurate diagnosis of plant diseases is crucial for sustainable agriculture, according to the paper "Pathogen-Based Classification of Plant Diseases: A Deep Transfer Learning Approach for Intelligent Support Systems". The study finds that the most accurate models are EfficientNetV2B2 and EfficientNetV2B3, using Keras transfer learning models with a variety of datasets, such as Agri-ImageNet and natural photos of sunflower and cauliflower. Developing intelligent support systems for plant disease detection is the goal of the project. Standardization and real-time processing are two issues that should be the focus of future study in order to increase agricultural output.[4]

The paper "Plant Disease Detection and Classification by Deep Learning—A Review" delves into the use of deep learning, highlighting its feature extraction and autonomous learning capabilities, in the diagnosis of crop leaf diseases. It discusses developments, difficulties, and new approaches in deep learning for plant disease diagnosis through the use of sophisticated imaging techniques. In addition to highlighting the need of early disease identification for efficient agricultural management, the review attempts to support scientists in creating effective, impartial methods for crop preservation. It demonstrates how the replacement of human approaches with deep learning and image processing is transforming the detection of plant diseases.[5]

CONVOLUTION NEURALNETWORKS

A new two-stage neural network design was suggested for classifying plant diseases. It employed both traditional augmentation methods and state-of-the-art generative adversarial networks to enhance the image dataset. Convolutional Neural Nets (CNNs) are widely used for picture identification and classification in paddy crop disease detection. Using drones or cellphones, high-resolution photos of paddy fields are first taken. The photos are then labeled with certain illnesses, such as blast and bacterial blight. Image quality is improved by preprocessing methods like scaling and normalization. The input layers, pooling layers to minimize spatial dimensions, convolutional layers for feature extraction, fully connected layers for classification, and output layers that provide illness probabilities are all part of the CNN architecture. Using backpropagation and gradient descent to reduce classification errors, the model is trained on a labeled dataset that is divided into training, validation, and test sets. Metrics including recall, accuracy, precision, and F1-score are used to assess performance.

CNN-based diagnosis makes early illness identification easier. identification, scalability, and precision farming, enabling effective supervision of vast paddy fields. Case studies have shown that certain systems can diagnose illnesses like bacterial blast and blight with high accuracy, and some of these systems can be incorporated into mobile applications for on-the-spot diagnosis. The datasets were gathered from different paddy fields and used in order to efficiently train and evaluate the models. Obtaining high-quality datasets, generalizing models, integrating them seamlessly with farm management systems, and having real-time processing capabilities are among the challenges. All things considered, CNNs are quite beneficial for precision agriculture, and future developments should bring about even greater accuracy and application.

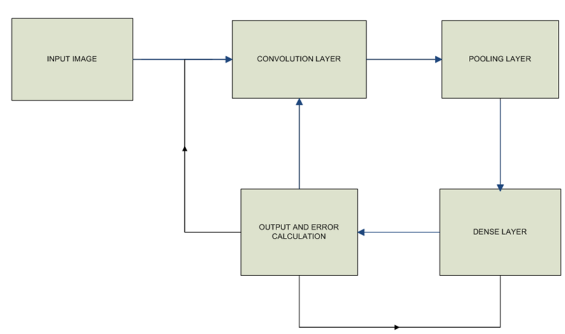


Fig1. Processing steps of CNN network

2.1 Convolution Layer

In order to extract features from input photos and use Convolutional Neural Networks (CNNs) for paddy crop disease diagnosis, the convolution layer is crucial. To create feature maps that emphasize certain attributes like edges, textures, and disease patterns, it uses tiny, learnable filters called kernels that glide over the pictures and multiply them element-by-element. Different features are detected by each filter, and to add non-linearity, non-linear activation functions—typically ReLU—process the resultant feature maps. While deeper layers recognize more complicated patterns like disease spots and lesions, first convolution layers capture fundamental information.

2.2 Pooling Layer

The pooling layer reduces computational complexity by downsizing images. It comes in two types: max pooling and average pooling. In max pooling, each NXN submatrix outputs the maximum value, while in average pooling, the average value of each submatrix is computed. The choice of convolutional and pooling layers depends on specific requirements.

2.3 Dense layer

The convolutional or pooling layer's output feeds into the dense layer, which requires flattened input. Flattening occurs in the flatten layer. One or more dense layers are employed to forecast plant diseases, with the final dense layer providing the network's output. This entire process is termed forward propagation, where the input undergoes multiplication with weights and summation of all inputs.

DATA SETS

An overview of datasets that are accessible to the public is given in this section. These datasets are utilized for various reasons; some are used for object detection to identify plant illnesses, while others are used for classification to ascertain if a picture of a plant is healthy or diseased.

A. Classification

The 14000 leaf photos in the Plant Village dataset—both healthy and diseased—are divided into 38 groups according to species and illnesses. Images of six crop species, including apple, cashew, mango, paddy, wheat, and tomato, are included in the dataset. It encompasses illnesses caused by fungi, bacteria, viruses, and mites. It also shows pictures of six crop kinds' healthy, disease-free leaves.

B. OBJECT DETECTION

In order to improve accuracy for our object identification algorithms, we combined two feature extractors (conv2D and MaxPooling2D) with CNN to identify pests and illnesses of rice plants. Using a variety of techniques to increase the dataset size, they used the CNN algorithms on a bespoke dataset of around 13867 photos, of which 11101 are utilized as the training set and 2775 as the validation set. According to the findings, the model has an accuracy of 85.2%.

C. DATA SETS USED

These are a few of the data sets we utilized for both training and validation.

Fig.2 Anthracnose in Cashews

Fig.3 Die Black or Pink Disease in Cashews

Fig. 4 Inflorescence blight in Cashews

Fig.5 Red Rust In Mango

Fig.6 Mango Malformation

Fig.7 Phoma Blight in Mango

Fig.8 Powdery Mild dew in Mango

Fig.9 Root Rot And Damping In Mango

Fig.10 Scab on Mango

Fig.11 Sooty mold in Mango

Fig12. Black Sheath Rot in paddy

Fig13. Blast in paddy

Fig.14 Brown Spot in paddy

Fig.15 False Smut in Paddy

Fig.16 Dead Heart in Paddy

Fig.17 Downy Mild Dew In Paddy

Fig.18 Hispa In Paddy

Fig.19 Tungro in Paddy

Fig.20 Bacterial Spot in Tomato

Fig.21 Bacterial Wilt In Tomato

Fig.22 Buck Eye Rot in Tomato

Fig.23 Damping off In Tomato

Fig.24 Furasium In Tomato

Fig.25 Late Blight In Tomato

Fig.26 Leaf Mold in Tomato

Fig.27 Leaf Roll in Tomato

Fig.28 Yellow Leaf Curl in Tomato

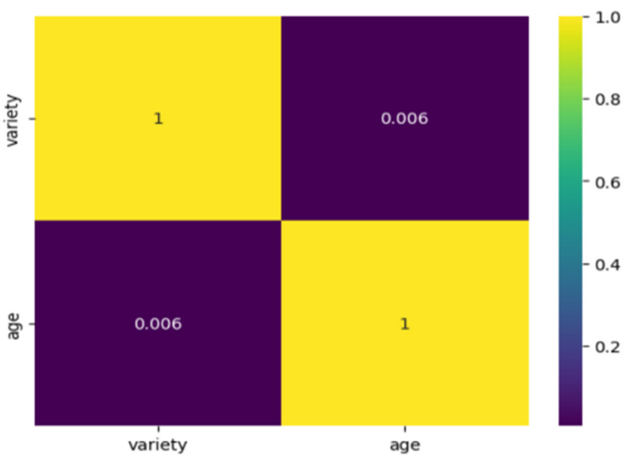
Fig.29 Brown Rust In Wheat

Fig.30 Head Scab In Wheat

For the purpose of training and verifying the diseases, a few other disease data sets are also obtained. When it comes to cereal crops like paddy, wheat, etc., the most often occurring diseases are brown rust and blast. Anthracnose is typically seen in fruits.

Results

A range of outputs are included in the study's conclusions, such as extensive data sets that specify the input variables and circumstances used to train and evaluate the model. Understanding the biological and environmental components that affect illness prediction requires access to these data sets. The study also produced heat maps, which show where in the paddy fields the illness is most likely to occur. By offering insightful



information on the spatial distribution of possible disease outbreaks, these heat maps facilitate more effective resource management and targeted treatments. The model's capacity to correctly forecast certain crop diseases, allowing for early identification and prompt control, was one of the study's key findings. For farmers and other agricultural stakeholders, this capacity is especially important since it enables the adoption of preemptive interventions that may lessen the negative effects of disease on crop quality and productivity.

Fig.31 Heat Map From of the Predicted crop

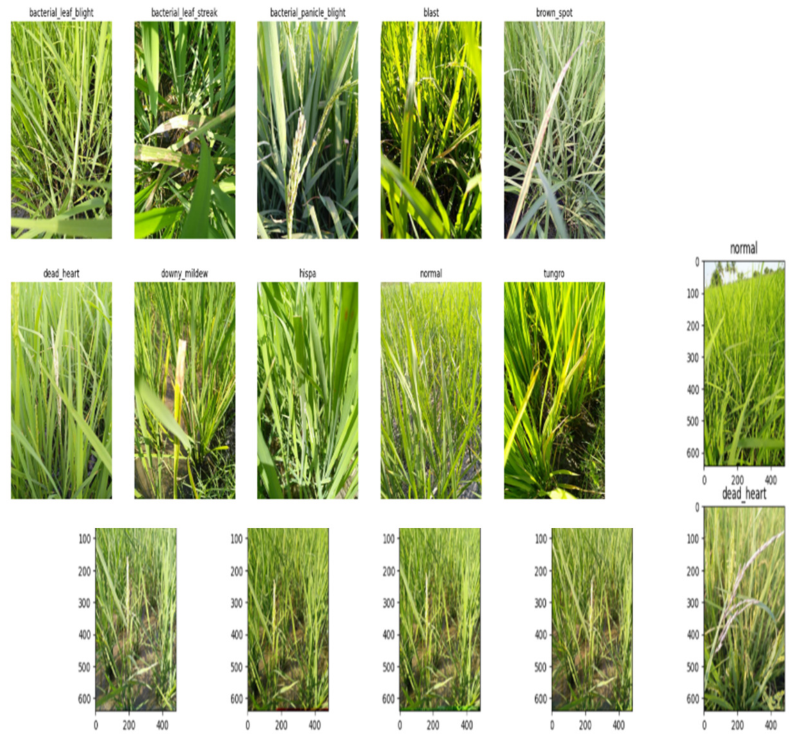


Fig.32 Data From The Trained Data Sets

Fig.33 Output For the Predicted Crop

Fig.34 Matlab Plot Between Train And Validation Set

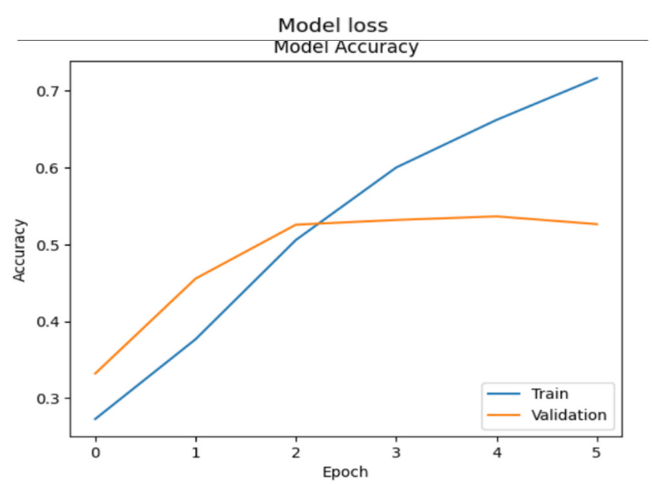


Fig.35 Matlab Plot Accuracy Between Trained And Validation Sets

Conclusions

The usefulness of many machine learning (ML) and deep learning (DL) models, including convolutional neural networks (CNN), for the identification and categorization of agricultural illnesses was assessed in our study using publically available datasets. Between the models that were evaluated, the CNN algorithm performed better than the others. On the AI Challenger 2018 dataset, it achieved an accuracy rate of 91.94%, outperforming models like MobileNetV2 and ResNet50.

A number of crop diseases, including as blast, bacterial blight, and anthracnose, were easier to diagnose because to the high accuracy of the CNN model. Utilizing high-resolution photos of crops including rice, mango, and tomato, the training and validation procedures highlighted the model's adaptability to a variety of crop kinds. The results contained extensive statistics and heat maps that illustrated the regions most vulnerable to illness, assisting in early identification and prompt action.

Even with the encouraging outcomes, there are still issues. To achieve the full potential of these models, concerns like standardization, real-time processing, and data quality must be resolved. To increase agricultural production and sustainability, future research should concentrate on improving these models' real-time application capabilities and optimizing them for a wider variety of crops.

Finally, our research demonstrates how CNN algorithms may be used in precision agriculture to provide a formidable instrument for disease control. These models have the potential to greatly advance more productive and sustainable farming methods by tackling current issues.

References

A Systematic Literature Review on Plant Disease Detection: Motivations, Classification Techniques, Datasets, Challenges, and Future Trends Shafik, W., Tufail, A., Namoun, A., De Silva, L. C., & Apong, R. a. a. H. M. (2023). A Systematic literature review on plant disease detection: motivations, classification techniques, datasets, challenges, and future trends. *IEEE Access*, *11*, 59174–59203. https://doi.org/10.1109/access.2023.3284760[1]

Crop Disease Recognition Based on Modified Light-Weight CNN With Attention Mechanism Liu, Y., Gao, G., & Zhang, Z. (2022). Crop disease recognition based on Modified Light-Weight CNN with attention Mechanism. *IEEE Access*, *10*, 112066–112075. https://doi.org/10.1109/access.2022.3216285[2]

Machine Learning and Deep Learning for Plant Disease Classification and Detection Balafas, V., Karantoumanis, E., Louta, M., & Ploskas, N. (2023). Machine learning and deep learning for plant disease classification and detection. IEEE Access, 11, 114352–114377. https://doi.org/10.1109/access.2023.3324722[3]

Pathogen-Based Classification of Plant Diseases: A Deep Transfer Learning Approach for Intelligent Support Systems

Rani, K. P. A., & Gowrishankar, S. (2023b). Pathogen-Based Classification of Plant Diseases: A deep transfer learning approach for intelligent support systems. IEEE Access, 11,64476–64493. https://doi.org/10.1109/access.2023.3284680 [4]

Plant Disease Detection and Classification by Deep Learning A Review ,Saleem, M. H., Potgieter, J., & Arif, K. M. (2019). Plant disease detection and classification by deep learning. Plants, 8(11), 468. https://doi.org/10.3390/plants8110468 [5]