

PADDY LEAVES DISEASE DETECTION AND CLASSIFICATION USING DEEP LEARNING TECHNIQUES

MRS. SINDHU A S, DR. C K ROOPA

DEPARTMENT OF INFORMATION SCIENCE AND ENGINEERING

JSS SCIENCE AND TECHNOLOGY UNIVERSITY MYSORE- 570006

ABSTRACT

Rice is considered one the most important plants globally because it is a source of food for over half the world's population. Like other plants, rice is susceptible to diseases that may affect the quantity and quality of produce. It sometimes results in anywhere between 20–40% crop loss productions. Early detection of these diseases can positively affect the harvest, and thus farmers would have to be knowledgeable about the various diseases and how to identify them visually. Even then, it is an impossible task for farmers to survey the vast farmlands on a daily basis. Even if this is possible, it becomes a costly task that will, in turn, increase the price of rice for consumers.. The proposed model has been trained and tested on three different rice diseased datasets. The accuracy obtained on CNN model for the rice disease dataset is 96.6%. With fewer parameters, the proposed model achieves higher accuracy in comparison with the VGG16 achieved accuracy of 73.6%, VGG19 achieved accuracy of 65.7% and ResNet 50 model achieved accuracy of 70%. The system provides a user-friendly interface for uploading an image of diseased leaves, analyzing them using a trained Convolutional Neural Network (CNN) model, and offering remedies for the detected disease rice leaf.

Keywords: CNN, VHH16, VGG19

INTRODUCTION

Agriculture has always been a primary industry in any nation's economy. Various crops such as paddy, wheat, maize are grown in huge fields spanning hectares. Though factors like city expansion and urbanisation are creating obstacles, the demand for agricultural products is more than ever. Technical aspects are being brought into the conventional art of farming to increase the produce multifold even when the land for farming remains the same[1]. Hybrid seeds, Irrigation and crop care systems, and automated seed storage systems have been introduced to obtain and maintain large amounts of grains at a time. Technology is also being brought into different kinds of crop care systems to increase efficiency of care by attempting to reduce human error and loss. The effect of climate change is evident in the growth of crops over time. Unseasonal rainfall and sunlight has resulted in improper growth of plants especially for paddy, where consistent weather conditions are required. Climate change also gives rise to improper land conditions which may hamper the growth of paddy or encourage growth of unnecessary weeds. Excessive rainfall also leads to depletion of nutrients through soil erosion. These climatic influences reduce the chances of crop sustenance and must be taken care of from the start of the cropping season. The arrival of the Green Revolution in India brought along with many new industrial practices of agriculture. These include advanced practices of drip irrigation, tractors etc. The arrival of pesticides and fertilisers brought both positive and negative consequences. Even though they protect the plants from the pests, excessive dosage of these chemicals pollutes the soil which is used for the crops . Losses are also caused in biodiversity as nearby water resources are polluted. Thus it becomes

necessary to find the pest infested areas along with the intensity to make most effective use of pesticides without damaging the crops. An enhanced version of ANN is the Convolutional Neural Networks (CNN) which is greatly used in image classification and Speech recognition due to its capability of handling huge data and forming intelligent links between the neurons unlike ANN networks [3]. CNN technology can be used to create an automated crop care system by identifying the diseased paddy crops from the healthy ones by just an image. This enhancement reduces human effort along with having little scope of error due to the highly efficient computational model of neural networks.

PROBLEM DEFINITION

To build a robust machine learning model for the Detection of Rice Leaf diseases based on training and validating dataset.

CHALLENGES

- Complex background with different real condition
- Computational time complexity
- The segmentation sensitivity towards the region of interest (ROI) determination.

OBJECTIVES

- Preprocess dataset for segmentation.
- Build and train machine learning model for segmentation.
- Preprocess dataset for classification of images.
- Build and train model for classification of rice leaf disease image.
- Develop an efficient and robust model for early detection of rice leaf diseases.

LITERATURE SURVEY

The paper titled "Disease Identification in Paddy Leaves using CNN-based Deep Learning" by Swathika R, Srinidhi. S, Radha N, and Sowmya.K, published in IEEE on 17th May 2021[1], proposes a solution for the quick classification of whether a paddy plant is diseased or healthy, along with identifying the affected area using a CNN (Convolutional Neural Network) model, achieving an accuracy of up to nearly 70%.[1] To further enhance the study, a potential avenue for exploration would be to emphasize the acquisition of the dataset through the utilization of an unmanned aerial vehicle (UAV) for streamlined and precise data collection. The integration of UAVs would enable broader coverage, rapid data collection, and an improved understanding of crop health. By capturing aerial images from different angles and heights, the dataset would encompass a more representative sample of paddy fields. Furthermore, the use of advanced imaging technologies, such as hyperspectral or thermal cameras, could be incorporated into the UAV setup to augment the accuracy and specificity of disease identification. Real-time monitoring capabilities facilitated by UAVs would enable timely detection and management of diseases, contributing to the field of precision agriculture [1].

The research paper titled "Rice Leaf Disease Detection Using Machine Learning Techniques: A Case Study in Bangladesh" by Kawcher Ahmed, Tasmia Rahman Shahidi, Syed Md. Irfanul Alam, and Sifat Momen, published in IEEE on 24th-25th December 2019[2], investigates the application of various machine learning algorithms, such as KNN, J48 (Decision Tree), Naive Bayes, and

Logistic Regression, for detecting rice leaf diseases. The authors conducted training on their dataset using these algorithms, with the decision tree algorithm being one of the techniques utilized. As part of future work, expressed the intention to explore the effectiveness of ensemble learning methods on the same dataset [2].

The research paper titled "Rice Plant Disease Detection Using Image Processing and Probabilistic Neural Network" by Tatvan Vocational School, Bitlis Eren University, Bitlis, Turkey, published in Springer on 11th May 2022[3], focuses on utilising the Probabilistic Neural Network (PNN), an Artificial Neural Network (ANN) model, to determine the presence of diseases in rice plants based on images. The system achieved an accuracy rate of 76.8% in accurately detecting rice plant diseases.[3] For future studies, the researchers propose exploring different segmentation algorithms in the user interface program. Additionally, they suggest investigating the application of deep learning methods and different camera technologies to enhance the detection of rice plant diseases [3].

The research paper titled "Rice Plant Disease Detection and Classification Using Deep Residual Learning" by Sanjay Patidar, Aditya Pandey, Bub Aditya Shirish, and A. Sriram, published in Springer on 31st July 2020[4], proposes the utilisation of Residual Neural Network (ResNet) for classifying rice plant disease images into specific disease classes. The authors achieved an impressive accuracy of approximately 95.83% on the dataset by employing ResNet as well as other classifiers like Support Vector Machines.[4].In order to further enhance the research, a suggested framework involves implementing Internet of Things (IoT) devices for image capture in the fields. Additionally, the web interface could include a platform for farmers to engage in discussions about the prevalent trends and experiences with various diseases, facilitating knowledge sharing and collaboration within the farming community [4].

The paper titled "Rice Disease Detection using Intensity Moments and Random Forest" by Sristy Saha and Sk. Md. Masudul Ahsan, published in IEEE on 15th June 2021[5], presents a method for rice disease detection using intensity moments and the Random Forest algorithm. Random Forest, a supervised learning algorithm, utilises the CART (Classification and Regression Tree) algorithm to generate predictions through aggregation. The proposed system achieves a disease identification accuracy of 91.47%. [5] The authors suggest that by incorporating additional collaborative features, the obtained results can assist developers in rapidly identifying plant diseases. However, the specific details regarding the collaborative features and their implementation are not specified in the provided information [5].

The paper titled "Plant Disease Detection for Paddy Crop using Ensemble of CNNs" by Abhijit Acharya, Akhil Muvvala, Siddhesh Gawali, Ruhi Dhopavkar, Rutuja Kadam, and Ashish Harsola, published in IEEE on 26th May 2021[6], proposes an approach for detecting three paddy leaf diseases (Blast, Bacterial Leaf Blight - BLB, and Brown Spot) using an ensemble of pre-trained CNN architectures. The CNN architectures employed in the study include GoogLeNet, ResNet, ShuffleNet, ResNeXt, and Wide ResNet. The models are combined in an ensemble with different weights to achieve an accuracy of 95.54% for the final algorithm [6]. As part of future development, the researchers suggest testing the model on servers to ensure efficient deployment. This implies deploying the algorithm on server systems, which can handle the computational requirements of the model and make it accessible for practical use in real-world scenarios. By optimizing and deploying the model on servers, it can be readily utilized for efficient plant disease detection in paddy crops [6].

The research paper titled "Image Processing for Diagnosis Rice Plant Diseases Using the Fuzzy System" by Anwar Rifa'i and Deni Mahdiana, published in IEEE on 21st June 2021[7], presents a system for diagnosing rice plant diseases through image processing using a fuzzy system. The fuzzy system utilises various image features, including contrast, correlation, energy, homogeneity, average, variance, kurtosis, entropy, standard deviation, and skewness. The system achieved an accuracy of 94.792% in diagnosing rice plant diseases.[7] However, the authors suggest that further development is needed to create a system capable of detecting more types of rice diseases. This implies expanding the range of diseases that the system can accurately identify and diagnose. By enhancing the system's capability to detect a wider variety of rice diseases, it can provide more comprehensive support to farmers and contribute to improved disease management in rice crops[7].

The research paper titled "Rice Blast Disease Recognition Using a Deep Convolutional Neural Network" by Wan-jie Liang, Hong Zhang, Gu-feng Zhang, and Hong-xin Cao, published on 27th February 2019[8], introduces a novel method for recognizing rice blast disease using a Convolutional Neural Network (CNN). The authors establish a dataset consisting of 2906 positive samples and 2902 negative samples to train and test the CNN model. In addition to proposing the CNN-based recognition method, the paper emphasises the exploration of the extracted features by the CNN. The authors conduct comparative experiments and quantitatively analyse the performance of the CNN model in terms of accuracy, Receiver Operating Characteristic (ROC) curve, and Area Under the Curve (AUC). These evaluations provide a comprehensive understanding of the effectiveness of the proposed approach for rice blast disease recognition.[8].

The research paper titled "Rice Disease Detection by Image Analysis" by Sudarshan S. Chawathe, published in IEEE in 2020[9], presents an automatic system for detecting three main types of rice leaf diseases: Bacterial leaf blight, Leaf blast, and Brown spot. The system utilises the Random Forest decision tree classifier and achieves an accuracy of 91.47% in disease detection. Furthermore, the paper highlights the expansion of the research to more diverse datasets. The author intends to study datasets that encompass a larger collection of disease and non-disease states. This expansion aims to enhance the robustness and generalizability of the disease detection system, enabling it to accurately identify a wider range of rice leaf diseases and non-diseased states [9].

The research paper titled "Detection of Paddy Crops Diseases and Early Diagnosis Using Faster Regional Convolutional Neural Networks" by Anandhan.K and Ajay Shanker Singh, published in IEEE on 20th April 2021, introduces a system for identifying various diseases in rice plant leaf images through the utilization of mask R-CNN and Faster R-CNN algorithms. In their proposed system, the authors employ these advanced convolutional neural network architectures to detect and diagnose diseases in paddy crops. The research work includes the use of these algorithms to accurately identify and classify different types of diseases affecting rice plants. Additionally, the authors plan to extend their research by capturing video footage directly from farmers' fields. This expansion aims to increase the size and diversity of the input dataset, enabling the system to be trained and tested on a wider range of real-world scenarios. By incorporating video data, the system can potentially improve its accuracy and robustness in early diagnosis of paddy crop diseases [10].

METHODOLOGY

- The image collected through various means are resized to 256*256*3 dimensions and used as input to the system.
- Preprocess the dataset for image classification with each image of dimension 256*256*3.
- Divide the dataset into a train, test and validation set in the ratio of 8:1:1 that consists of 4 classes of diseases.
- Build the CNN architecture for classification.
- Train the model using a training set.
- Analyze the results
- Deploy the model into a Graphical User Interface.

In this system for detecting the Brown Spot, Bacterial leaf blight, Leaf blast diseases of paddy leaf, it involves major two phases one is training the model and the other part is detecting the given image of the disease. The first phase deals with training the model using the image dataset. Images of diseased dataset are collected. Here we have collected 1946 images of Brown Spot, 1440 images of Leaf blast, 1674 images of Bacterial Blight. The images are downloaded from the Kaggle, UCI Machine Learning Repository, and github repository. These images are trained with the help of Convolutional Neural Network (CNN) Algorithm.

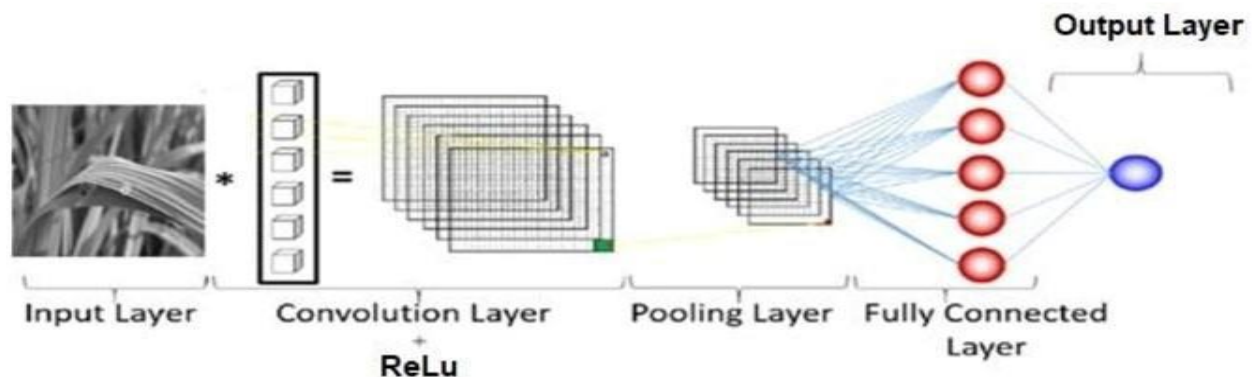


Fig 1 CNN Architecture

We adopted a CNN architecture consisting of an input layer, convolutional layers, rectified linear unit (ReLU) activation function, pooling layers, and fully connected layers. The input layer takes the resized image as input, and the subsequent layers perform operations such as feature extraction, non-linearity modeling, dimensionality reduction, and classification. Finally, the output layer produces the disease detection result for the input leaf image. This CNN architecture enables accurate and efficient analysis of paddy diseases based on the input image features in our project.

1. CONVOLUTIONAL LAYER

The convolutional layer helps to extract features from the input images. The convolutional layer scans the images using filters to create an activation map. This process captures local patterns and spatial relationships between pixels, enabling the model to learn meaningful features for disease detection. By applying convolutional operations, the model can automatically detect relevant patterns and structures in the input images, leading to accurate identification of paddy diseases.

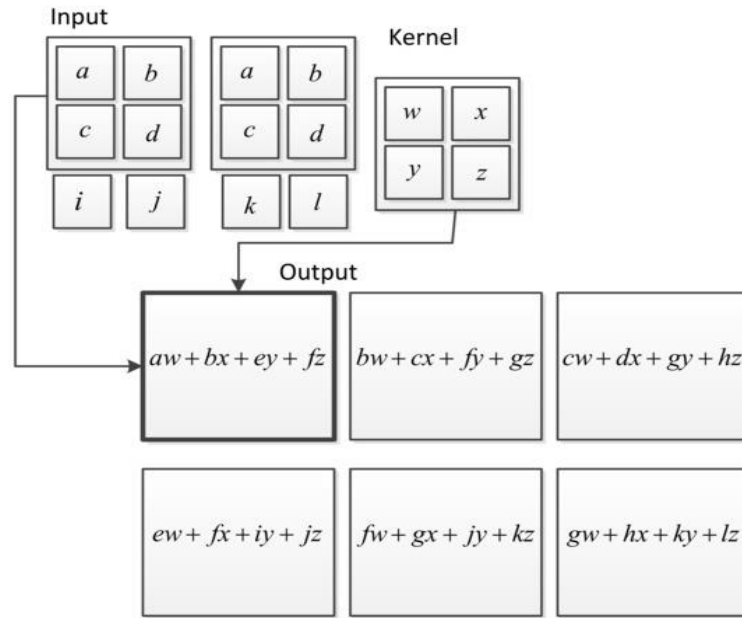


Fig 2 Convolution layer

2. RECTIFIED LINEAR UNIT (ReLU)

The utilization of the ReLU (Rectified Linear Unit) activation function in the hidden layers of our model. ReLU is a commonly used activation function known for its simplicity and effectiveness. It replaces negative pixel values in the image with 0, while retaining the positive pixel values unchanged. This non-linear transformation helps introduce non-linearity into the model, allowing it to learn complex patterns and improve the accuracy of disease detection.

3. POOLING LAYER

We incorporate a pooling layer within the Convolutional Neural Network (CNN) architecture. The pooling layer reduces the data size generated by the preceding convolutional layer, making it more efficient to store. By reducing the dimensionality of the data, the pooling layer helps retain important features while discarding redundant information, contributing to efficient and effective disease detection. The internal working of the pooling layer is visualized in Figure below, showcasing its role in the overall CNN model.

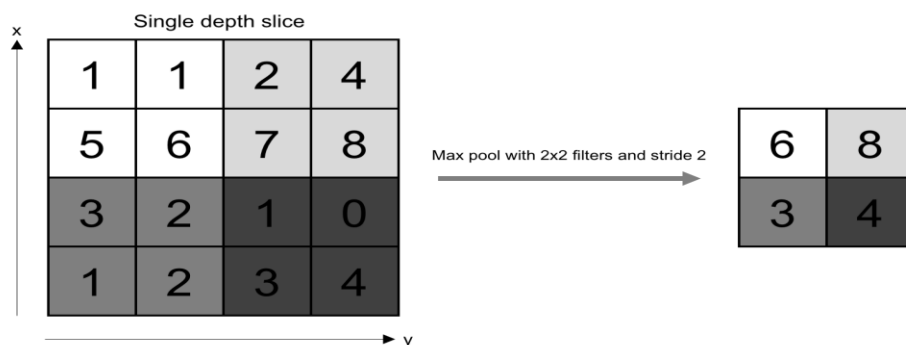


Fig 3 Pooling layer

4. FULLY CONNECTED LAYER

Fully connected input layer – The preceding layers' output is "flattened" and turned into a single vector through a set of which is used as an input for the next stage. The first fully connected layer – adds weights to the inputs from the feature analysis to anticipate the proper label. Fully connected output layer – offers the probability for each label in the end. Fig 4 shows the internal working of fully connected layer.

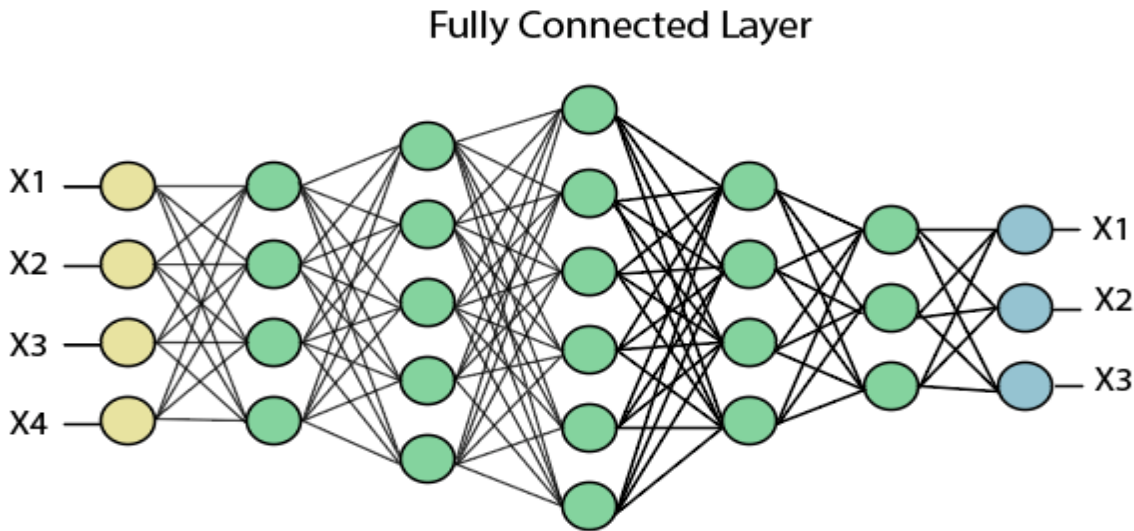


Fig 4 Fully Connected layer

In our proposed system for paddy disease detection, we utilize fully connected layers within the Convolutional Neural Network (CNN) architecture. The fully connected input layer takes the output from preceding layers and converts it into a flattened vector, which serves as input for the next stage. The first fully connected layer applies weights to the inputs obtained from the feature analysis to predict the correct label. Finally, the fully connected output layer provides the probability distribution for each possible disease label. Our system focuses on detecting the most common paddy diseases, including Rice Blast, Brown Spot, Bacterial Blight and healthy. The dataset is loaded and preprocessed by resizing the images. The features are then extracted using the convolutional layer of the CNN, which aids in training the model. After training, the model becomes capable of disease detection. During the testing phase, the same preprocessing steps are applied to the input image. The image data is passed through the CNN algorithm, and the extracted features are compared with the training model's features. This enables the model to accurately detect the paddy leaf disease. Utilizing a larger and diverse dataset during training enhances the model's accuracy in disease detection.

RESULT AND DISCUSSION

In this section, the models constructed in the project to develop CNN and machine learning models that would detect and categorize rice leaf illnesses based on photographs of healthy and diseased leaves into three categories: brown spot, bacterial leaf blight, leaf blast, along with healthy are detected. In this study, Python version 3.11 is used for the implementation of the project work. The CNN model takes a long time to train the model. So, system configuration and GPU should be good and there should be proper installation of all the libraries before training the model and to execute the model smoothly. Various libraries were installed such as TensorFlow, Keras, numpy.

```

Epoch 1/10
116/116 [=====] - 383s 3s/step - loss: 0.9676 - accuracy: 0.5234 - val_loss: 0.6805 - val_accuracy: 0.6579
Epoch 2/10
116/116 [=====] - 272s 2s/step - loss: 0.6410 - accuracy: 0.7102 - val_loss: 0.5872 - val_accuracy: 0.7249
Epoch 3/10
116/116 [=====] - 253s 2s/step - loss: 0.4969 - accuracy: 0.7971 - val_loss: 0.3748 - val_accuracy: 0.8349
Epoch 4/10
116/116 [=====] - 259s 2s/step - loss: 0.3500 - accuracy: 0.8612 - val_loss: 0.2635 - val_accuracy: 0.9163
Epoch 5/10
116/116 [=====] - 263s 2s/step - loss: 0.3020 - accuracy: 0.8813 - val_loss: 0.2050 - val_accuracy: 0.9354
Epoch 6/10
116/116 [=====] - 342s 3s/step - loss: 0.1662 - accuracy: 0.9343 - val_loss: 0.1490 - val_accuracy: 0.9402
Epoch 7/10
116/116 [=====] - 239s 2s/step - loss: 0.1803 - accuracy: 0.9335 - val_loss: 0.1393 - val_accuracy: 0.9426
Epoch 8/10
116/116 [=====] - 221s 2s/step - loss: 0.1152 - accuracy: 0.9609 - val_loss: 0.0547 - val_accuracy: 0.9785
Epoch 9/10
116/116 [=====] - 244s 2s/step - loss: 0.0856 - accuracy: 0.9715 - val_loss: 0.2756 - val_accuracy: 0.9187
Epoch 10/10
116/116 [=====] - 220s 2s/step - loss: 0.0951 - accuracy: 0.9663 - val_loss: 0.0274 - val_accuracy: 0.9928

```

Fig 5 The model is trained for 10 epochs with the learning rate of 0.01 and the accuracy obtained is 96.6%.

The above figure shows the CNN model was trained for 10 epochs using a learning rate of 0.01. The trained model achieved an impressive accuracy of 96.6%. This indicates that the model has effectively learned the patterns and features in the dataset, enabling accurate classification of rice diseases.

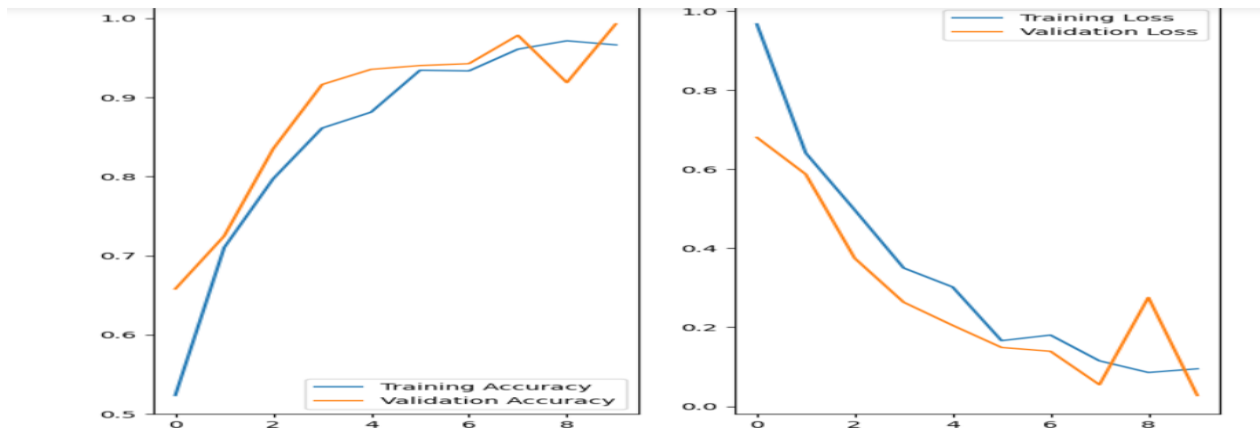


Fig 6 Training and Validation learning graph for 10 epochs

Fig6 represents the training and validation learning graph for 10 epochs. This graph provides valuable insights into the CNN model's performance over the training period, showing the trends in both training and validation accuracy or loss. It helps in understanding the model's convergence, overfitting, and generalization capabilities, aiding in the evaluation and improvement of the model's training process.

COMPARISON GRAPH

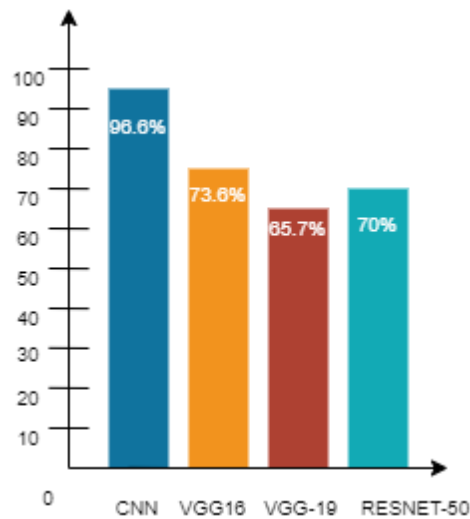


Fig7 Comparison graph

The dataset used in the project consists of 946 images of Brown Spot, 1440 images of Leaf blast, and 1674 images of Bacterial Blight for training the models. The evaluation results, shown in the graph, indicate the accuracy achieved by different models. The CNN model achieved an accuracy of 96.6%, while the VGG16, VGG19, and ResNet-50 models achieved accuracies of 73.6%, 65.7%, and 70% respectively. These results demonstrate the effectiveness of the CNN model in accurately detecting and classifying rice leaf diseases compared to the other models.

CONCLUSION

A web-based system has been successfully implemented for Rice Leaf disease detection for Rice leaves using Convolutional Neural Network along with VGG16, and is successfully being detected by a system with training accuracy 96.6% using a training dataset of 1946 images of Brown Spot, 1440 images of Leaf blast, 1674 images of Bacterial Blight. Convolutional neural network has been developed with three hidden layers to classify the leaf disease images. The system allows users to input an image and accurately detects the specific disease affecting the rice leaf, providing valuable information for diagnosis and treatment purposes.

FUTURE SCOPE

Furthermore, the application of machine learning will enhance disease classification and prediction models. By training these models on large datasets of annotated rice leaf images, they can learn to recognize disease patterns and distinguish between different types of leaf diseases with high accuracy. Additionally we are planning to develop an Android app that focuses on rice leaf disease detection, leveraging the advancements in technology and machine learning algorithms. This app will empower farmers and agronomists to easily and conveniently diagnose and monitor rice leaf diseases using their smartphones. By utilizing the device's camera and

integrating disease detection models, the app will provide real-time analysis and accurate identification of diseases affecting rice crops. Additionally, it will offer valuable insights on disease management strategies, including recommended treatments and preventive measures. Our aim is to provide a user-friendly tool that enhances the efficiency with high accuracy and effectiveness of disease detection in rice cultivation, ultimately contributing to higher crop yields and sustainable farming practices.

REFERENCES

- [1] Swathika R, Srinidhi. S, Radha N, and Sowmya.K, Disease Identification in Paddy Leaves using CNN-based Deep Learning published in IEEE on 17th May 2021.
- [2] Kawcher Ahmed, Tasmia Rahman Shahidi, Syed Md. Irfanul Alam, and Sifat Momen, "Rice Leaf Disease Detection Using Machine Learning Techniques: A Case Study in Bangladesh", published in IEEE on 24th-25th December 2019.
- [3] Tatvan Vocational School, Bitlis Eren University, Bitlis, Turkey, "Rice Plant Disease Detection Using Image Processing and Probabilistic Neural Network", published in Springer on 11th May 2022.
- [4] Sanjay Patidar, Aditya Pandey, Bub Aditya Shirish, and A. Sriram, "Rice Plant Disease Detection and Classification Using Deep Residual Learning" published in Springer on 31st July 2020.
- [5] Sristy Saha and Sk. Md. Masudul Ahsan, "Rice Disease Detection using Intensity Moments and Random Forest" published in IEEE on 15th June 2021
- [6] Abhijit Acharya, Akhil Muvvala, Siddhesh Gawali, Ruhi Dhopavkar, Rutuja Kadam, and Ashish Harsola, "Plant Disease Detection for Paddy Crop using Ensemble of CNNs" published in IEEE on 26th May 2021.
- [7] Anwar Rifa'i and Deni Mahdiana, "Image Processing for Diagnosis Rice Plant Diseases Using the Fuzzy System" published in IEEE on 21st June 2021.
- [8] Wan-jie Liang, Hong Zhang, Gu-feng Zhang, and Hong-xin Cao, "Rice Blast Disease Recognition Using a Deep Convolutional Neural Network" published on 27th February 2019.
- [9] Sudarshan S. Chawathe, "Rice Disease Detection by Image Analysis" published in IEEE in 2020.
- [10] Anandhan.K and Ajay Shanker Singh, "Detection of Paddy Crops Diseases and Early Diagnosis Using Faster Regional Convolutional Neural Networks", published in IEEE on 20th April 2021.