**RICE LEAF DISEASE CLASSIFICATION USING CNN**

**ABSTRACT**

Rice is amongst the majorly cultivated crops in India and its leaf diseases can have a substantial impact on output and quality. The most important component is identifying rice leaf diseases, which have a direct impact on the economy and food security. Brown spot, Leaf Smut, Bacteria leaf blight are the most frequently occurring rice leaf diseases. To resolve this issue, we have studied various machine learning and deep learning approaches for detecting diseases on their leaves by calculating their accuracy, recall, and precision to measure the performance. This study helps farmers by detecting diseases in rice leaves in order to get a healthy crop yield. The deep learning models perform well when compared with the machine learning methods. After analyzing all of the deep learning models, we found that the 5-layer convolution model had the best accuracy of 80.2 %, while others, such as VGG16, had a lower accuracy of 69.57%.

**INTRODUCTION: -**

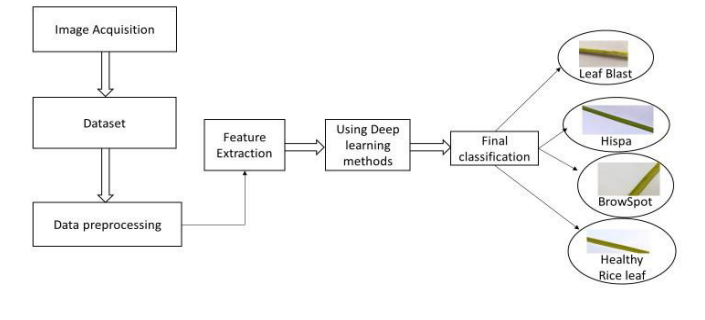
India is a land of Agriculture as it plays an essential role in our country because a lot of the people are dedicated to the agricultural industry. Crop production is amongst the major factors which are affecting domestic market conditions in our country. Agricultural firms began to search for new high-yield, cost-effective inventions as a consequence of expanding population, variable weather conditions, and political unpredictability. The health of the plant/crop is critical for achieving food security and sustainability in agriculture. However, the plants can quickly become infected with illnesses, which can cause major social and economic problems, due to a variety of factors. Crop diseases can affect its growth and development, and also crop yield and quality, and are one of the most common reasons for productivity loss. To avoid soil pollution, the illness should be detected, and certain pesticides should be used from their beginning stage.

There are a variety of methods for detecting plant diseases in their early stages. The traditional method of plant disease detection is naked eye monitoring, which is ineffective and inaccurate for large crops. The major goal of this paper is to research and diagnose rice leaf illnesses in advance, as well as to identify the disease's name so that appropriate precautions can be followed. Rice is a standout amongst the most vital food plantations in our country, as well as one of the crops having a variety of purposes and high nutritional worth, with a production volume of 104.80 million tons coming from various Indian states. Because our country is the largest producer of rice in the second position in the world, the country's rice-growing region is constantly expanding. It contains a high amount of carbohydrates and protein, as well as a significant amount of dietary fiber and minerals. Plant illnesses are caused by pathogens, fungus, bacteria, viruses, and other microbes in most cases. Rice leaves are sensitive to diseases that are caused by fungus, viruses, and the varied field environment makes it simple for pathogens to infect the leaves. Figure 1 shown below is an example of different diseases of rice leaf. Climate changes will create an ideal environment for those pathogens to thrive. The growth of crops is hampered in their initial stages due to fungi-caused illnesses. If illness strikes while the crop is still growing, it might reduce the crop's yield. Manually determining the presence of illnesses in large agriculture regions is quite challenging. Diseases, particularly in rice plants, have become a problem since farmers are unable to identify leaf diseases with the naked eye, and they must consult an expert in order to discover that specific disease, which takes more time and requires much expense. The most frequently occurring diseases in rice leaf are Brown spot, Bacteria leaf bright, Leaf Smut, Healthy. Thus, disease detection in leaves is an important topic that provides many benefits in monitoring large fields of crops. Rice leaf disease can affect yield and quality by damaging the green layer of the leaves. The way to control these rice diseases is to rapidly and precisely detect the disease type and then implement appropriate corrective actions in a timely manner. Using digital image processing techniques and deep learning networks, the detection of disease is efficient, consumes less time, and is accurate. Advances in Computer vision offer an opportunity to extend and increase plant protection. The content of the paper is organized in the following manner: In Section 2, we'll go through a few previous kinds of research that have been done using image processing techniques, ML methods, and DL models for recognizing diseases in the leaves of rice plants. Section 3 is about the different deep learning methods studied for rice leaf disease detection, and the final section concludes the experimental results and future scope.

**Bacteria leaf bright Brown spot**  **Leaf Smut**

In this paper we are working on an automated system to detect the fungal disease of rice plants which is a major cause for the loss of a rice plant. This type of disease may occur and spread due to climate change and moisture on the leaves. The above-mentioned figure (Fig1) represents the entire process of the method which we have proposed. The first step is acquiring the image. The act of obtaining a picture from a source is known as image acquisition. The input can be taken from different resources, especially hardware such as sensors or cameras etc. This is a crucial step in the entire process because processing is impossible without an image. This step is always the initial phase in the process. The second step is dataset collection. The dataset in our method consists of several images of the four diseases of rice leaf which we are predicting. We train the dataset using a wide variety of images to get better accuracy. The next step is preprocessing data. Collected data should be preprocessed involving data cleaning and removing all the inconsistencies from the data. We do this process to obtain a clean dataset for achieving better accuracy. Then we perform feature extraction which involves extraction of only those features which are important, and which are most required. Then we use various DL methods which we are employing in our method and train and test the dataset accordingly. Finally, the images are classified according to their respective diseases. In our method, we have tested four diseases that affect the rice plant.



**DATASET** **DESCRIPTION: -**

In the present study, the leaf dataset consists of four types of diseased rice leaf images; these are Bacteria leaf bright, Brown spot, Leaf Smut, and Healthy. This dataset consists of 160 images of rice leaves with the various symptoms of the diseases, which consists of 40 images of Brown spot, 40 images of Bacteria leaf bright, 40 images of Leaf Smut. These images are all in JPEG format and have good resolution with a width and height of 1449\*1449. The image background is also white so there is no need to apply any background subtraction method. This leaf dataset contains a combination of Bacteria leaf bright, Brown spot, Leaf Smut.

Information about the division of dataset

|  |  |  |  |
| --- | --- | --- | --- |
| Disease Type | Train image | Test image | Validation image |
| Bacteria leaf bright | 26 | 10 | 4 |
| Brown Spot | 26 | 10 | 4 |
| Leaf smut | 26 | 10 | 4 |

**DATA** **PREPROCESSING: -**

In our present world scenario, input data involves a lot of noise, has missing values, outliers, and is inconsistent. Data preprocessing involves removal of the noise, missing data, and organizing data in a proper format so that accuracy is increased. It enhances the quality of the data. This step involves data cleaning, data transformation, and data reduction (data compression) Data cleaning involves cleaning the data. It removes the noise present in the data. Data transformation is transforming high-level data into low-level data for easier calculations. Data reduction involves reducing the data dimensions so that the data is not high dimensional, but the quality of data remains the same. Data cube aggregation which is summarizing the data.

**DEEP LEARNING METHODS FOR** **CLASSIFICATION: -**

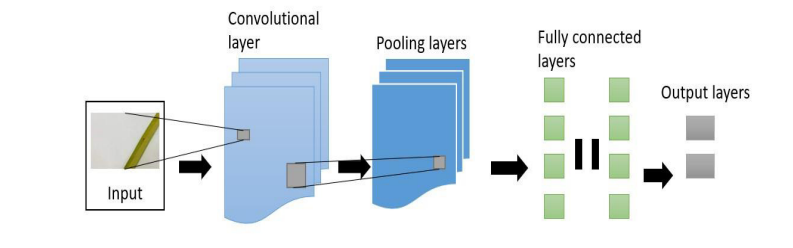
**CNN: -**

Here, we applied a Convolutional neural network (CNN) based approach which is a method of DL that takes input as an image and gives importance to many other objects in the image, as well as differentiates between them. The amount of pre-processing needed by a CNN is substantially less than that required by other classification methods. While simple techniques need the hand-engineering of filters, with enough training, CNN learns these filters/characteristics. Our architecture mainly contains the following layers:

● Convolution layer

● Pooling layer

● Fully connected layer

 **Layers of CNN**

The above figure represents the working of CNN. The input in the form of the image after preprocessing the data and extracting the required features when passed through CNN passes through 3 layers of CNN and it is precisely represented. The final output is then displayed.

● **Input Layer**: The input layer of CNN consists of the dataset. The input data will be represented as a 3X3 matrix.

● **Convolution Layer:** A layer that uses filters to learn from smaller sections of input data to obtain features from an image.

● **Pooling Layer:** This layer is used to shrink the image's dimensionality, lowering the processing power required for subsequent layers. There are two variations of pooling. They are:

● **Max pooling:** The pixel with the maximum value as input is selected and transferred to the output while parsing input. It is the most used approach compared to average pooling.

● **Fully Connected Layer (Dense):** This is one of CNN's last layers, and it can recognize features that are significantly linked with the output class. The result is a one-dimensional vector created by flattening the pooling layer results.

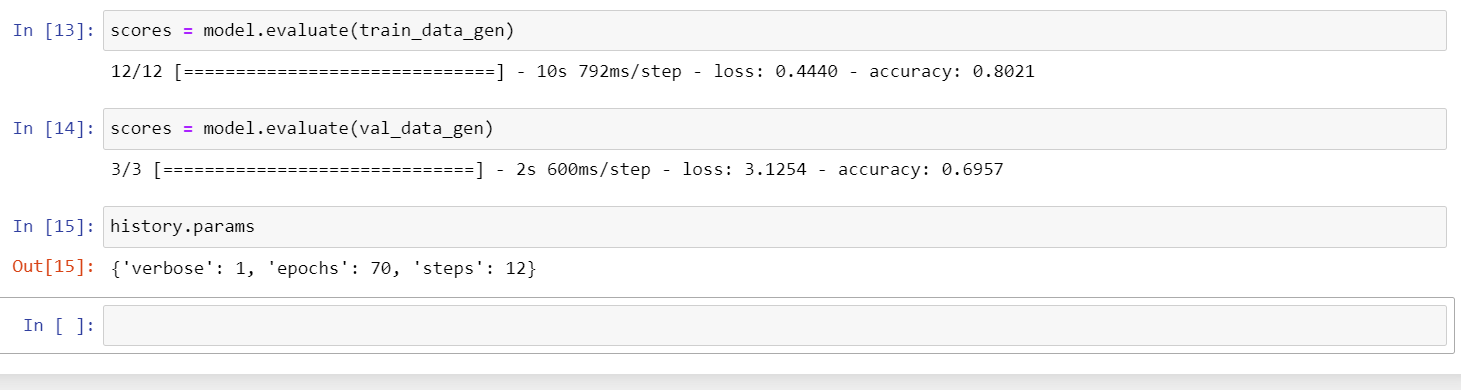
● **Dropout Layer:** Used to reduce model overfitting problem by removing a random set of neurons in that layer. It is connected with the FC layer.

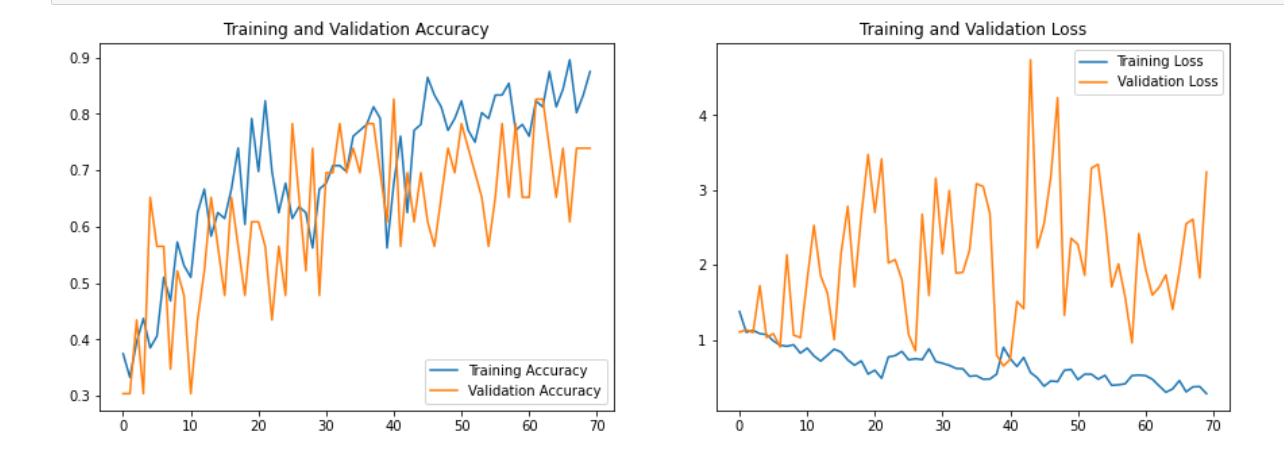
● **SoftMax Layer:** This is the network's last layer that assists in classifying individual input images of the dataset into several classes depending on the learned properties from the network.

● **Output Layer:** The output layer holds the final classification result.

**EXPERIMENTAL RESULTS:-**

We have studied various deep learning algorithms such as CNN on the same rice dataset to measure the accuracy of each method. The screenshot below presents the training process. Whereas training and validation accuracy were shown, where the y-axis shows the accuracy obtained after each iteration represented in the x-axis. Similarly, contains the training and validation loss, where the y-axis shows the loss (in percentage) when training started and thereafter increase or decrease in loss after each iteration represented in the x-axis.



**CONCLUSION:-**

In this study, we have performed the classification of various rice leaf diseases using a few DL methods for rice leaf diseases. We used a dataset of diverse rice leaves with illnesses, which we subsequently processed using several standard deep learning methods like 5-layer convolutional network. We discovered that the 5-layer convolutional network performs the best at identifying rice leaves out of all of them. We can conclude that the accuracy of our proposed 5-layer CNN model is approximately 6 percent more than the other standard deep learning models. Also, we observed that by adjusting the training parameters like learning rate, number of epochs, and optimizer methods, we can get significant accuracy with a handmade model having a smaller number of layers than the other traditional models. The better we can detect infections, the simpler it will be for farmers to protect their crops. In the future, we will broaden the scope to include more diseases and algorithms, making disease detection vast, easier and faster.