RICE LEAF DISEASE CLASSIFICATION USING CNN

Pallapothala Tejaswini¹, Priyanshi Singh², Monica Ramchandani³, Yogesh Kumar Rathore⁴, Rekh Ram Janghel⁵

DEPARTMENT OF INFORMATION TECHNOLOGY
NATIONAL INSTITUTE OF TECHNOLOGY, RAIPUR¹,²,³,⁴,⁵

monicaramchandani4042@gmail.com¹, pallapothalatejaswini@gmail.com², priyanshisingh1099@gmail.com³, yogeshgrathore23@gmail.com⁴, rjanglel.it@nitrr.ac.in⁵

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 Blast, Hispa are the most frequently occurring rice leaf diseases. To resolve this issue, we have studied various machine learning and deep learning approaches for detecting the diseases on their leaves by calculating their accuracy, recall, and precision to measure the performance. This study helps the farmers by detecting the 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 78.2 %, while others, such as VGG16, had a lower accuracy of 58.4%.

Keywords: rice leaf diseases; deep learning; Convolutional neural networks; machine learning; transfer learning;

1. INTRODUCTION

India is a land of Agriculture [1] 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[2]. 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 [4] 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 at 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 the majority of 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[5]. Figure 1 shown below is an example of different diseases of rice leaf.

![Brown Spot](image1.png) ![Healthy Rice Leaf](image2.png)

![Hispa](image3.png) ![Leaf Blast](image4.png)

Fig.1. Types of Rice Leaf Diseases[6]

Climate changes [7] 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 disease with the naked eye, and they must consult the 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, Hispa, LeafBlast, 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 from 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[8]. 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[9].

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.

**Fig.2. Overview of Proposed Method**

In this paper we are working on an automated system to detect the fungal disease of rice plants which is a major cause for a 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 a number of 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 of 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.

2. LITERATURE REVIEW

(HaixiaQi et al., 2021) [1], this paper dealt with the automatic identification of groundnut leaf diseases using the stack ensemble technique. The proposed research was conducted on diseased groundnut leaves to identify four groundnut leaf diseases, in this study, they merged deep learning models with traditional machine learning approaches. Deep layer networks, such as ResNet50 and DenseNet121, performed the best when it came to dataset prediction. The maximum accuracy for data augmentation was 97.59 percent. ResNet50 had the best identification performance when integrated with the LR model. (GowriShankar et al., 2020) [2], in this paper, they discussed the automatic identification of groundnut leaf diseases. A DL model was employed to increase the network's speed and accuracy in finding and classifying different disease-infected patches on groundnut leaves. To grow the efficiency of previous algorithms, they replaced the typical SVM classifier with KNN for distinguishing four different pathologies (Leaf Blight, Leaf Spot, Stem Rot, and Bud Necrosis). (Ramakrishna et al., 2015) [3], this paper discusses one of the most common illnesses that affect ground leaves in its early stages. The suggested scheme incorporates four leading phases for the detection and categorization of groundnut leaf disease. The initial process is to do a color renovation on the images that will be used as input. The plane separation would be the next step. The extraction of features is the next phase. The backpropagation algorithm is employed to detect the leaf disease as a last step. (NilamBhise et al., 2020) [4] In the suggested research, disease detection is carried out in two stages. The type of crop and the type of the disease are determined in the first step using a CNN. Tensorflow lite is used to categorize the uploaded picture numerical value to the dataset values, and Keras frameworks are used to classify the dataset values. In terms of disease diagnosis performance, the findings show that the Mobile Net Model outperforms other models. (Salini et al., 2021) [5], The focus of this research is to reduce pesticide use in agriculture while making better quality and quantity of output. For feature extraction, they use image processing techniques, and for classification, they use SVM. To improve performance and produce a better outcome, the model was combined with data augmentation. This research aims to detect three major rice plant diseases: Bacterial Leaf Blight, Brown Spot, and Leaf Blight. The input to this model is the full image for processing, and the output will be the disease that has affected the plant, as well as the model's accuracy. (Mahalakshmi et al., 2021) [6], To identify the existence of disease as well as to detect specific types of disease, the author extracted color and texture features of corn leaves then the collected features are categorized using Binary SVM and multi-class SVM. The proposed system's accuracy is 85 percent, which is its best performance. (Saleem et al., 2019) [7], this paper reviewed the detection of several plant diseases and their classification by deep learning. Alternative deep
learning models and machine learning methods for visualizing plant diseases are examined in this study with the conclusion that deep learning models are more accurate than conventional machine learning techniques. (Shruthi et al., 2019) [8], ML techniques were used to describe the steps involved in the detection of general plant disease. They have used a CNN to detect the diseases with high accuracy.

(Yang Lu et al., 2018) [9], this article mainly focuses on the identification of diseases on rice crops. To detect diseases in the rice crop, it employs deep Convolutional neural networks. One of the backdrops of this study is that they used less data for training. (Azathet et al., 2021) [10], this research paper presents the work on the cotton leaf for the detection of diseases and pest diagnosis using image processing and Deep Learning. The researchers used CNN to detect diseases in cotton leaves. In terms of recognizing specific diseases, the model has a 96.4 percent accuracy rate. Recently, different deep learning methods [11-17] have been applied with Convolutional neural networks and computer vision to detect plant disease, recognition of plant leaves for medicine purposes, pest detection, counting of the wheat head, etc. In real-time robots need to detect the plant leaf disease and provide proper pesticide over it in large farms and weed detection with a bounding box helps to remove the weed with herbicides.

The following section includes a tabular column that gives a quick overview of the recent works related to the different plant leaf disease detection. The various machine learning and deep learning algorithms used for identifying the diseases, which we will look at in the next section.

Table 1. – A quick summary of studied research works on the detection of leaf diseases using different classifiers.

| Article | Crop   | Dataset                                | Classifier                          | Accuracy | Advantages                  | Drawbacks                      |
|---------|--------|----------------------------------------|-------------------------------------|----------|-----------------------------|--------------------------------|
| [1]     | Peanut | 6029 Lab images                        | ResNet50, DenseNet121               | 97.59%   | Best prediction performance | It is complicated in field environment |
| [5]     | Rice   | Kaggle dataset with 1000 images        | Image processing techniques and SVM | 80%      | Easy to train and flexible model | Low performance               |
| [6]     | Corn   | 1292 lab images                        | Binary SVM and Multi-Class SVM      | 85%      | Exhibits best performance on the dataset created | Only suitable to cornfield and it is not suitable for other fields |
| [3]     | Groundnut | 100 lab images                     | Back Propagation                   | 97.41%   | Best Performance            | Works with only one groundnut disease i.e., |
Findings from the Literature: From the above literature, we found that plant diseases are region-specific due to different environmental conditions and geographic locations. Here many authors proposed a deep learning-based model to detect plant diseases more accurately. We also observed that the deep learning models are suggested only when a large dataset is available. Here, in our paper, we are working on the fungal disease of rice plants and having a dataset of 1600 images. So, we are proposing an automated deep learning model in the next section for the prediction of such diseases.

3. PROPOSED METHODOLOGY

3.1 DATASET DESCRIPTION

In the present study, the leaf dataset consists of four types of diseased rice leaf images; these are Hispa, Brownspot, LeafBlast, and Healthy. This dataset consists of 1600 images of rice leaves with the various symptoms of the diseases, which consists of 400 images of Brownspot, 500 images of Hispa, 300 images of LeafBlast, and 400 images of Healthy. 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 [32]. This leaf dataset contains a combination of Hispa, Brownspot, LeafBlast and Healthy. The sample images are shown in Fig.1.

| Disease Type | Train Images | Test Images | Validation Images |
|--------------|--------------|-------------|-------------------|
| Hispa        | 200          | 150         | 150               |
| Brownspot    | 200          | 100         | 100               |
| LeafBlast    | 150          | 60          | 90                |

Table 2. Information about the division of dataset
3.2 Data Preprocessing:

In our present world scenario[11], 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.

3.3 DEEP LEARNING METHODS FOR CLASSIFICATION

3.3.1 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[15].

Our architecture mainly contains the following layers:

- Convolution layer
- Pooling layer
- Fully connected layer

![Fig.3. Layers of CNN](image-url)
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.

### 3.3.2 VGG-19:

VGG-19 is a 19-layer deep CNN. You may use the ImageNet database [17], that has been trained on over a million photographs, to import a pre-trained version of the network. The network can sort photographs into 1000 distinct object categories.

A fixed-size (224 * 224) RGB image was supplied as input to this network, suggesting that matrix was of form (224,224,3). The one and only preprocessing was to eliminate the mean RGB value obtained across the whole training set from each pixel. To cover the complete visual notion, they employed kernels with the size of (3 * 3) as well as a stride size of 1 pixel. Spatial padding was used to maintain the image's spatial resolution. Maximum pooling was obtained with stride 2 across a 2 × 2-pixel frame. Since prior models used tanh or sigmoid functions, the Rectified linear unit (ReLu) was developed to add non-linearity to improve model classification and processing time, and it proved to be considerably better to them.

### 3.3.3 VGG-16:

We have used Visual Geometry Group (VGG) 16 layers model. The input image size has set to 224 × 224 × 3, and it is connected to next convolution layer Conv1. Then combination of convolutional and max pooling layers is applied for feature extraction and feature reduction respectively. In a few settings, it includes 11 convolutional filters which are considered to extract the features from input image and send these features to subsequent layers for further
processing[18]. To extract the features from corners of the image we used “similar padding” of one row and column each side and stride is set to “1”. Spatial pooling is done via five max-pooling layers that follow part of the Conv layers.

3.3.4 RESNET:

An Artificial Neural Network Based model ResNet (Residual Network) has been used here to solve the problem of the vanishing/exploding gradient. In this network, we use a technique called skip connections. The skip connection skips a few training steps and connects directly to the output. The shortcut link is introduced to a VGG-19-inspired 34-layer basic network architecture. As a result of these shortcut connections, the architecture becomes a residual network.

A residual neural network (ResNet) is a form of artificial neural network (ANN) based on cerebral cortex pyramidal cell structures. Residual neural networks employ skip connections, also known as shortcuts, to go around some layers. Double- or triple-layer skips with nonlinearities (ReLU) and batch normalization are used in the bulk of ResNet models. We have used the above network in its standard form[1].

3.3.5 Xception:

Xception is a deep CNN with 71 layers. You may use ImageNet to import a pre-trained network that is already trained on a number of input images. The network arranges all the input images into 1000 different categories, such as pencils, pens, books, and many more. The network has a massive library of various representations of features for a long-range of input datasets. The input data size for this pre-trained network is 299 * 299 pixels.

Xception is a convolutional neural network with just convolution layers according to their depth. The Xception architecture's feature extraction building is constructed up of 36 convolutional layers. In our experimental evaluation, we will just look at picture classification, therefore our results will be constrained. Just after the convolutional base layer, regression will be utilized. Fully-connected layers, as illustrated in the example, can be introduced before the logistic regression layer. The experimental assessment portion. Six categories are used to categorize the 36 convolutional layers. There are a total of 14 modules, each having its own linear residual connection. Everything else is constructed around them, with the exception of the first and last modules.

3.3.6 5-Layer CNN:

A five-layer CNN[19] is similar to a three-layer CNN in the exception that it has two additional layers apart from the five Layer CNN at the end. As usual, the input image went through three layers of the CNN and the output is passed through 2 additional layers namely:

- Dropout
- Activation Functions

**Dropout:** This is done when the output after the FC layer causes overfitting of the model. Overfitting generally occurs when the model works very well such that it gives a negative impact on the performance of the model.

In dropout, due to overfitting of the model, a few neurons are removed randomly from the
NN of the training process with reduced model size.

Activation Functions: It is amongst the most necessary parameters of the CNN Model. It defines the complex relationship between the variables in the model[20]. It decides the start of the network and the end of the network in the NL. The common activation functions used are ReLU, tanh etc.

4. EXPERIMENTAL RESULTS

We have studied various deep learning algorithms such as VGG 16, VGG 19, ResNet, Xception, and 5D CNN on the same rice dataset to measure the accuracy of each method. The below screenshot (Fig 3) presented the training process. Whereas training and validation accuracy were shown in Fig 4, where the y-axis shows the accuracy obtained after each iteration represented in the x-axis. Similarly, Fig 5 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.

Fig.4. Test accuracy of Brownspot, Healthy, Hispa and Leafblast

Fig.5. Graph of training and validation accuracy
**Fig.6.** Graph of training and validation loss

**Table.3.** Accuracy of different DL Methods

| Based Model | Accuracy   |
|-------------|-----------|
| VGG 16      | 58.4%     |
| VGG 16 (with first three block of frozen) | 72.2%     |
| VGG 19      | 72.4%     |
| XCeption    | 72.2%     |
| ResNet50    | 72.2%     |
| 5-layer convolution | 78.2% |
CONCLUSION AND FUTURE SCOPE

In this study, we have performed the classification of various rice leaf diseases using a few DL methods for four rice leaf diseases. We used a dataset of diverse rice leaves with illnesses, which we subsequently processed using several standard deep learning methods like VGG19, VGG16, Xception, Resnet, along with a handmade 5-layer convolutional network. We discovered that the 5-layer convolutional network performs the best at identifying rice leaves out of all of them. From Fig 6 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 less 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.

REFERENCES

[1] Qi, H., Liang, Y., Ding, Q., & Zou, J. (2021). Automatic Identification of Peanut-Leaf Diseases Based on Stack Ensemble. Applied Sciences, 11(4), 1950.
[2] Rahman, C. R., Arko, P. S., Ali, M. E., Khan, M. A. I., Apon, S. H., Nowrin, F., & Wasif, A. (2020). Identification and recognition of rice diseases and pests using convolutional neural networks. Biosystems Engineering, 194, 112-120.
[3] Ramakrishnan, M. (2015, April). Groundnut leaf disease detection and classification by using back propagation algorithms. In 2015 international conference on communications and signal processing (ICCSP) (pp. 0964-0968). IEEE.

[4] Ramesh, S., Hebbar, R., Niveditha, M., Pooja, R., Shashank, N., & Vinod, P. V. (2018, April). Plant disease detection using machine learning. In 2018 International conference on design innovations for 3Cs compute communicate control (ICDI3C) (pp. 41-45). IEEE.

[5] R Salini, A.J Farzana, B Yamini, (2021). Pesticide Suggestion and Crop Disease classification using Machine Learning , 63(5), 9015-9023.

[6] Mahalakshmi, S. D., & Vijayalakshmi, K. (2021). Agro Suraksha: pest and disease detection for corn field using image analysis. Journal of Ambient Intelligence and Humanized Computing, 12(7), 7375-7389.

[7] Sony, A. (2019). Prediction of Rice Diseases Using Convolutional Neural Network (in Rstudio). Int. J. Innov. Sci. Res. Technol, 4(12), 595-602.

[8] Lu, Y., Yi, S., Zeng, N., Liu, Y., & Zhang, Y. (2017). Identification of rice diseases using deep convolutional neural networks. Neurocomputing, 267, 378-384.

[9] Lu, Y., Yi, S., Zeng, N., Liu, Y., & Zhang, Y. (2017). Identification of rice diseases using deep convolutional neural networks. Neurocomputing, 267, 378-384.

[10] Zekiwos, M., & Bruck, A. (2021). Deep Learning-Based Image Processing for Cotton Leaf Disease and Pest Diagnosis. Journal of Electrical and Computer Engineering, 2021.

[11] A. Jayakumar, “Detection and Classification of Leaf Diseases in Maize Plant using Machine Learning MSc Research Project Data Analytics.” [Online]. Available: https://agriculturegoods.com/why-is-agriculture-important/

[12] Vaishnave, M. P., Devi, K. S., Srinivasaan, P., & Jothi, G. A. P. (2019, March). Detection and classification of groundnut leaf diseases using KNN classifier. In 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN) (pp. 1-5). IEEE.

[13] Shrivastava, V. K., Pradhan, M. K., Minz, S., & Thakur, M. P. (2019). Rice plant disease classification using transfer learning of deep convolution neural network. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences.

[14] Hasan, M. J., Mahbub, S., Alom, M. S., & Nasim, M. A. (2019, May). Rice disease identification and classification by integrating support vector machine with deep convolutional neural network. In 2019 1st international conference on advances in science, engineering and robotics technology (ICASERT) (pp. 1-6). IEEE.

[15] Shrivastava, V. K., & Pradhan, M. K. (2021). Rice plant disease classification using color features: a machine learning paradigm. Journal of Plant Pathology, 103(1), 17-26.

[16] Ahmed, K., Shahidi, T. R., Alam, S. M. I., &Momen, S. (2019, December). Rice leaf disease detection using machine learning techniques. In 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI) (pp. 1-5). IEEE.

[17] Ramesh, S., &Vydeki, D. (2018, September). Rice blast disease detection and classification using machine learning algorithm. In 2018 2nd International Conference on Micro-Electronics and Telecommunication Engineering (ICMETE) (pp. 255-259). IEEE.

[18] Matin, M. M. H., Khatun, A., Moazzam, M. G., & Uddin, M. S. (2020). An Efficient Disease Detection Technique of Rice Leaf Using AlexNet. Journal of Computer and Communications, 8(12), 49.
[19] Liang, W. J., Zhang, H., Zhang, G. F., & Cao, H. X. (2019). Rice blast disease recognition using a deep convolutional neural network. Scientific reports, 9(1), 1-10.

[20] Sethy, P. K., Barpanda, N. K., Rath, A. K., & Behera, S. K. (2020). Deep feature based rice leaf disease identification using support vector machine. Computers and Electronics in Agriculture, 175, 105527.

[21] Dataset available at: https://www.kaggle.com/minhhuy2810/rice-diseases-image-dataset, seen on 12 January.