A CONVOLUTIONAL NEURAL NETWORK-BASED AUTOMATED METHOD FOR DETECTING PADDY LEAF DISEASE

*Pramod K** [1], *Dr. N.V.Balaji* [2]

ABSTRACT

India's main crop for production is paddy. Over the past eleven years, the agricultural sector has contributed around 18.08 % of India's GDP. Unfortunately, the farmers who put forth a lot of effort to raise this crop have to deal with significant losses due to crop damage brought on by many paddy illnesses. About seven or eight of the more than thirty paddy leaf diseases that exist are fairly prevalent in India. Among the various paddy leaf diseases, the most common and detrimental ones are Brown Spot Disease, Blast Disease, Bacterial Leaf Blight, and others. Paddy plant growth and productivity are being hampered by these diseases, which can result in significant financial and environmental losses. Crop damage can be significantly decreased and farmer losses can be avoided if these diseases can be identified early on with high precision and speed. Four disease categories and one paddy leaf class that is healthy have been covered in this paper. The primary goal of this paper is to provide the best results for paddy leaf disease detection by using deep learning CNN models for automated detection, which can achieve the highest accuracy, as opposed to the time-consuming manual disease detection process, which is also of questionable accuracy. After analyzing four models (VGG-19, Inception-Resnet-V2, ResNet-101, and Xception), it was discovered that Inception-Resnet-V2 had a greater accuracy of 92.68%.

Keywords: Paddy leaf disease, deep convolutional neural network (DNN), transfer learning, VGG-19, ResNet-101, Inception-ResNet-V2

Department of Computer Science[1,2],
Karpagam Academy of Higher Education, Coimbatore, India[1,2]

* Corresponding Author

I. INTRODUCTION

The primary source of energy and a vital component of the proteins that almost three billion people consume, paddy (Oryzasativa species) remains the primary crop in human food systems at the beginning of the twenty-first century [1]. Asia-Pacific produces almost 90% of the world's paddy [2]. Rice is planted in approximately 75% of the edited region and more than 80% of the fully irrigated zone in India, making it the primary crop used to provide food. When growing paddy, farmers frequently face a number of challenges, including diseases, pests, climate change, increased population, and arable land damage [3]. These days, farmers are losing interest in paddy cultivation as a result of these numerous issues. This essay has only addressed the diseases and pests that contribute to the different issues that arise when growing rice. The three main groups of paddy illnesses are bacterial, fungal, and miscellaneous. Bacterial blight, bacterial leaf streak, brown spot, leaf smut, leaf scald, panicle blight, bronzing, and other diseases are among them [4].Keep in mind that the negative effects of climate change, particularly the rise in temperature, have recently caused the incidence of diseases to reach extreme levels (IPCC, 2007). According to estimates, a variety of pests and diseases cause 4–14% of India's annual rice yield to be lost. Brown spot and bacterial leaf curse (BLB) are two real rice infections at the moment. However, the new technologies for diseases and pests are still limited [5].

Paddy disease is typically identified manually by experts using their unaided eyes, which takes more time and is expensive on large homesteads [6]. It is difficult to quantify and occasionally makes a mistake when identifying the type of disease [7]. Paddy production has recently declined due to a lack of knowledge about the proper way to treat paddy plant leaf disease [8]. The diagnosis of paddy leaf diseases must be done quickly and appropriately in order to get around this. A 1979–1981 survey [9] found that there were 20 paddy leaf diseases in India, with 13 of those diseases deemed to be the most important. Bacterial leaf blast is one of the most dangerous diseases in 2019, according to the Indians Rice Knowledge Bank. The Deep Convolutional Neural Network was chosen for this study, and the dataset was trained using four pre-trained DNN-based models

II. RELATED WORK

- There have previously been a lot of studies done on various rice illnesses. There is currently research being done on rice diseases and how to treat them. Utilizing machine learning techniques created by Kawcher Ahmed et al. to identify three paddy leaf diseases [10]. With a primary focus on three major paddy leaf diseases, they used 10-fold cross-validation techniques and four machine learning models to complete their job and increase accuracy.
- MilonBiswas et al. [11] used a single classifier and focused on just three paddy diseases. Image segmentation, grayscale conversion, SVM classifier application, and outcome prediction come first.
- Wen-Liang Chen et al. [12] state that bacterial blast leaf disease is one of the most prevalent paddy illnesses. Using the Internet of Things and artificial intelligence technologies, they mainly focus on agricultural sensors that generate non-image data that can be automatically trained and evaluated in real-time by the AI process. They can identify plant diseases almost accurately.
- Using KNN and ANN algorithms, S. Ramesh et al. [18] suggested a method for detecting rice blast leaf disease. They primarily addressed one rice leaf disease, Indian rice crops, and early disease detection techniques. They obtained the highest ANN accuracy of 99%.
- A mechanism that uses deep learning techniques to identify rice leaf disease from real-time video was proposed by Dengshan Li et al. [15]. They employed a variety of deep CNN models, including VGG16, ResNet-50, ResNet-101, and YOLOv3, in addition to faster-RCNN for image detection from video.
- Using image processing techniques, GittalyDhingra et al. [08] conducted a thorough investigation into a variety of paddy diseases. They talked about two aspects of classifying and detecting rice diseases.
- Junde Chen et al. [11] examine five paddy leaf diseases using a deep learning methodology with transfer learning. Two deep learning models, such as the Dense-Net and Inception module, were used, and they achieved an accuracy of 98.63%.

III. PROPOSED SYSTEM

The identification of paddy disease has been thoroughly investigated in the past utilizing a variety of techniques and machine learning and deep learning ideas. This study used a customized deep learning model, as depicted in the flowchart (Fig. 1), and a benchmarked methodology. Following the collection of images of infected paddy leaves, the image pre-processing phase began. The pre-processed photos are processed by a deep convolutional neural network. The convolutional blocks of the models extract the salient features from the input images. The DNN model determines each node's weights based on the image features. The activation function, like soft max, helps classify the given data in the final dense layer of the model, which consists of five neural nodes. The machine pre-processes the images by rotating, zooming, flipping, shuffle, and resizing them after receiving the image from the dataset. There will be four deep CNN models used

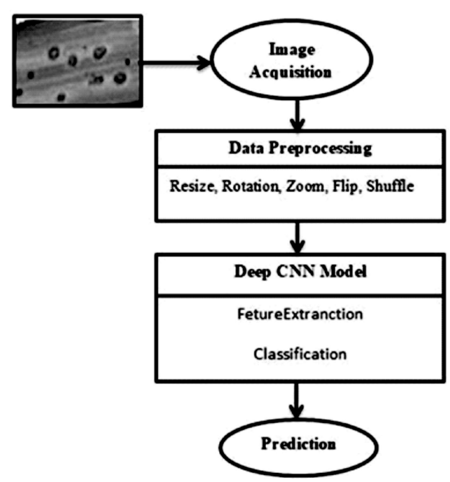

Figure 1: Proposed Model

Classification based on Transfer Learning

In the field of machine learning, transfer learning refers to the idea of applying learned information to another model in order to address a related issue [17]. The ImageNet dataset is used to train Keras applications that are based on deep CNN. The ImageNet project is a large visual database designed for visual object recognition research. Millions of pictures and thousands of classes are used to build deep convolutional neural network-based models [10].

Several convolutional, pooling, and dense layers are present in Keras deep learning applications. As seen in Figures 3, 4, and 5, the architectures can be divided into several blocks of layers. The network's convolutional blocks contain several convolutional layers that use the input data to extract features. Fig. 4 showed viable convolutional layer architecture, a residual inception block [32]. In order to enhance the dimensionality of the filter before concatenation, a filter expansion layer (1x1 Conv Linear) was applied after each inception block, as shown in Figure 4.

Model's acquired feature parameters are transferable. Related issues can be resolved more successfully with a new model that uses the pre-trained weights than with a general model [12]. Dense layers for classification tasks are found in the last block of Keras applications.

Thousands of image categories are used to train Keras deep learning network architectures, and thousands of nodes are included in the final dense layer to classify every category. Removing the top layer of the model and adding a tailored fully connected layer to identify the required classes of images is an innovative approach when the dataset has insufficient data [14].

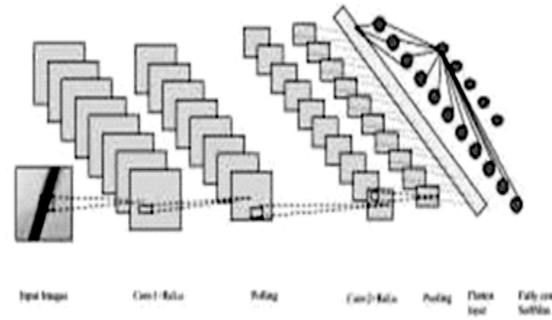

Figure 2: shows the model's schematic based on a CNN.

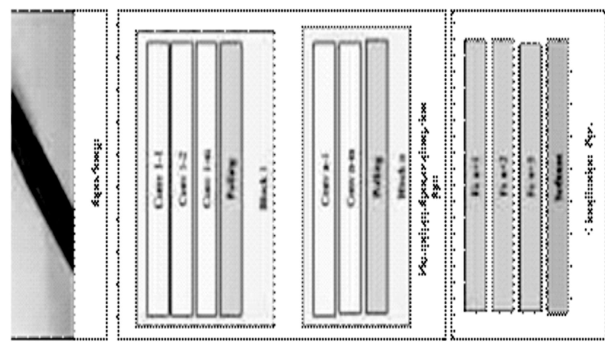

Figure 3: shows the VGG block schematic, which consists of a series of convolutional layers.

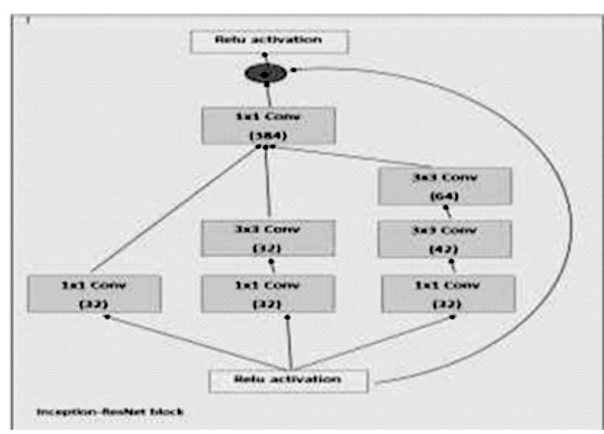

Figure 4: Inception ResNet Block Schema

Inception-ResNet Block Schema for Inception-ResNe-V2 Network [18] (Fig. 4).This is Fig. 5's Inception-Resnet block.

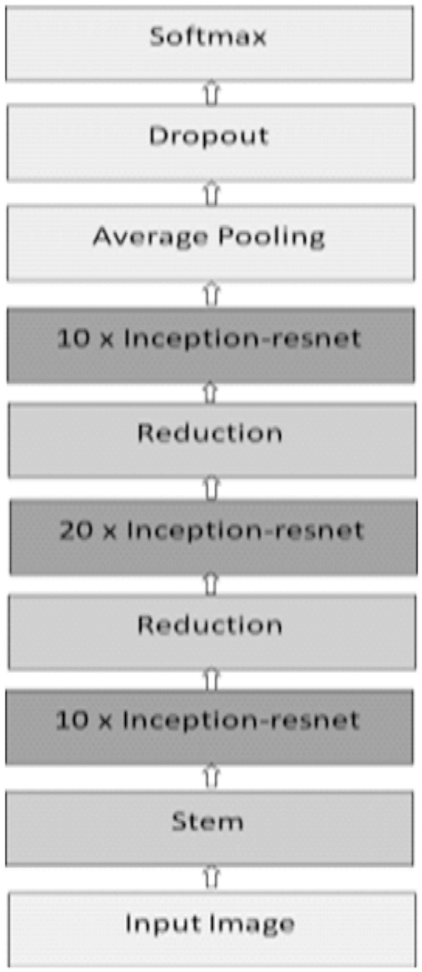

Figure 5 : Inception-ResNet-V2 Network Schema [18]

IV. DATASET DESCRIPTIONS AND PADDY DISEASE TYPES

A. Types of Diseases

Diseases and insects alone cause a lot of food grains to be wasted. To remove these rice illnesses, research is being done all around the world. Although there were more than 30 rice illnesses in Bangladesh overall, 20 of them were recorded in the 1979–81 study [9].However, thirteen diseases are prevalent in all three seasons (Boro, Aus, and Aman). The primary diseases were bacterial blight, bacterial leaf streak, sheath blight, sheath rot, leaf blast, brown spot, grain spot, stem rot, and leaf scald, while the lesser ones were Zinc-deficiency, Tungro, Bakanae, and Cercospora leaf spot. Only the disorders that have been discussed are briefly discussed in section B.

B. Descriptions of the Datasets

Four paddy leaf diseases were selected for this study. There are 984 images in the dataset. gathered the information from a number of online sources, including Kaggle [15] and the UCI machine learning repository [16]. Table II provides information on the dataset's size and categorizes it into train, validation, and test groups. A thorough explanation of the dataset's classes may be found in Figures 6 and 7.

1) Brown spot: Brown spot is a common fungal disease that affects paddy leaves. Round, tiny, dark brown to purple-brown marks is visible at the beginning stage (Fig. 6(a)). Large patches on the leaves will grow over time and have the potential to kill the entire leaf.
2) Leaf blast: Magnaportheoryzae is a type of fungus that causes this paddy leaf disease. The main signs of this illness are spindle-shaped, white to grey-green spots with dark red to brownish borders (Fig. 6(b)).
3) Leaf blight: One type of bacteria called Xanthomonasoryzae is the cause of this blight disease. As seen in Fig. 6(c), infected leaves first turn greyish green, then yellow, then straw in color, and finally die.
4) Leaf smut: The fungus Entylomaoryzae is the cause of leaf smut, a prevalent paddy leaf disease. Both sides of the infected leaf will exhibit angular, black patches (sori), as shown in Fig. 6(d). Black dots that are between 0.5 and 5.0 millimeters long and 0.5 and 1.5 millimeters wide are present on the leaves.
5) Healthy leaf: The leaf should be green in color and free of any indications of disease.

C. Data Pre-processing

The Keras Image Data Generator function participated in the pre-processing of the data. 823 photos with three color channels and varying pixel values make up the dataset.Next, resize each image to 256 by 256 pixels. A different perspective of the visual object is obtained by irregularly rotating the training images within a 15-degree range. Shear range, zoom, and width and height shift range are all fixed at 0.1. The only popular preprocessing method used on both training and testing datasets is rescaling images. There is a batch size of eight for the training dataset and one for the test set.

Table 1. Count of each Disease's Test and Train Images

Paddy leaf disease class name	Number of Images	Number of Images Used for Train and Validation	Number of Images used for Test
BrownSpot	157	130	26
LeafBlast	152	127	27
LeafBlight	210	180	27
Leaf Smut	212	179	26
HealthyLeaf	225	189	26
Total	956	805	132

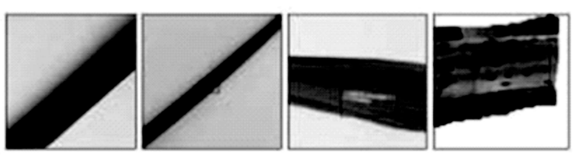

Figure 6: Paddy Leaf Disease (a) Brown Spot, (b) Leaf Blast, (c) Leaf Blight, and (d) Leaf Smut.

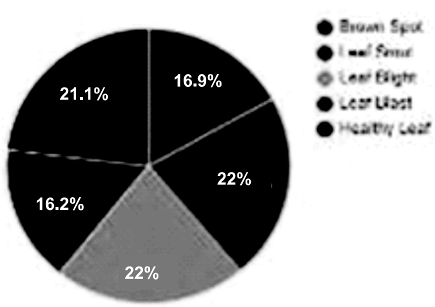

Figure 7. Class Data Share Percentages in the Dataset.

V. MODEL DETAILS

Deep Convolutional Neural Network: Deep CNN models are a kind of feed-forward neural network that uses network parameter adjustments to reduce the cost function value. These models have made significant contributions to the fields of object detection, natural language processing, image or video classification, and visual imagery analysis. Convolutional neural networks, which have more layers than a standard CNN model, are the foundation of the deep CNN model. Convolutional, activation, pooling, flatten, dropout, batch normalization, and dense layers are typically included in the model. To prevent overfitting, the dropout is used. To assist in identifying complex patterns in the data, activation functions such as ReLu and Softmax are added to network layers. Table V displays the various deep CNN models and demonstrates how well the deep CNN model performs in categorization and identification. The four models were chosen on the basis of their architectural design variation and depth. The network's topological depth is indicated by depth, and the validation accuracy using the ImageNet validation dataset is indicated by the Top 5 accuracy in Table III.

Table 2. Description of Keras-Based Deep Learning CNN Models [13]

ModelName	Top5Accuracy	Parameters	Depth
VGG-19	0.898	153,567,231	25
ResNet-101	0.837	39,607,153	98
Xception	0.867	20,810,371	124
Inception-ResNet-V2[32]	0.834	49,737,879	567

VI. RESULT ANALYSIS

Four CNN deep learning Keras pre-trained algorithms, listed in Table IV, were used in this investigation to categorize and identify the leaf illnesses. The study's evaluation of various algorithms showed that, with an accuracy of 0.837, Inception-ResNet-V2 was the most accurate of all. Additionally, Inception-ResNet-V2 fared better than the others in terms of F1 score, recall, and precision. Resnet-101 has outperformed Inception-ResNet-V2 with an accuracy of 0.898. The Xception model has the highest accuracy of 0.8942, while VGG-19 has the lowest accuracy of 0.834. In terms of accuracy, precision, recall, and

F1 Score, it has performed the worst.. For all training procedures, 100 epochs are taken into consideration. Table III, which includes a statistical analysis of these different models, presents all of this data. The following definitions apply to the evaluation metrics (accuracy, precision, recall, and f1 score) [18].

1) Accuracy: Confusion matrix data is used to calculate accuracy. Accuracy, which is the total ratio of properly predicted data to all data in the dataset, is the most logical performance statistic. The formula for accuracy is shown in Eq. (4). More accuracy can only be attained when the false positive and false negative values in the dataset are almost equal.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

2) Precision: The ratio of correctly predicted positive values to all positive predicted values is the algorithm's precision measurement [18]. Equation displays the precision formula.

$$\text{Precision} = \frac{TP}{TP+FP}$$

3) Recall : Recall is defined as the ratio of accurately predicted positive values to the actual positive class in the confusion matrix. The recall formula can be found in Equation.

$$\text{Recall} = \frac{TP}{TP+FN}$$

4) F1 Score: The F1 Score is the weighted average of Precision and Recall. Because of this, this score includes both false positive and false negative values. The formula for the F1 score is displayed in equation (7). Even though precision is hard to comprehend on its own, F1 is typically more helpful than accuracy.

$$\text{F1 Score} = \frac{2\times Recall \times Precision}{Recall + Precision}$$

Because it has the greatest accuracy of the three models, the model Inception-ResNet-V2 has generated a performance table for this particular algorithm for each class of the dataset, mentioning the precision, recall, and F1 score of each class. As shown in Table IV, the precision for Brown Spot, Leaf Blast, Leaf Blight, and Leaf Smut has been over 0.90, while the precision for images of healthy leaves is 0.78. Recall was above 0.90 for the other classes and 0.71 for the class leaf blast. Lastly, regarding the F1

Table 4. Statistical Analysis of Different Pre-trained Keras Model

Models	Accuracy	Precision	Recall	F1Score
VGG19	0.8143	0.8176	0.8035	0.8041
ResNet-101	0.9152	0.9215	0.9056	0.9056
Inception-ResNet-V2	0.9286	0.9371	0.9262	0.9286
Xception	0.8942	0.8963	0.8865	0.8823

Table V. The Performance Score For Each Class Of Inception- Resnet-v2 Model

Classes	Brown Spot	Healthy Leaf	Leaf Blast	LeafB light	Leaf Smut
Precision	1.00	0.78	0.95	0.91	1.00
Recall	0.93	1.00	0.71	1.00	1.00
F1Score	0.96	0.88	0.82	0.98	1

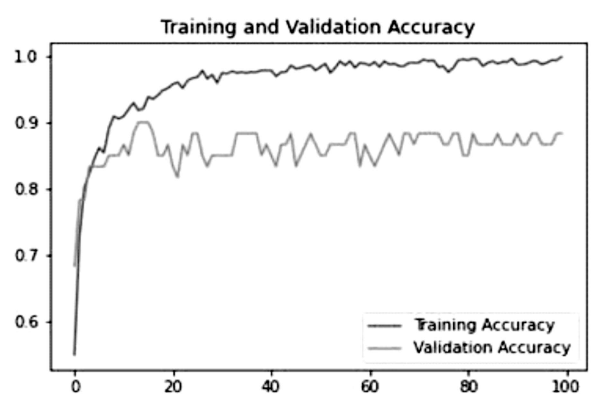

Figure 8: Training and Validation Accuracy Graph.

Figure 8 displays the model's training and validation accuracy graph. It is clear that the orange line, which represents the validation accuracy, is varying between roughly 0.80 and 0.90, while the blue line, which represents the training accuracy, has reached about 1.0.

Error Analysis

Even though disease detection by hand is challenging, technology has greatly simplified the process for us. Even so, there are still certain limitations that prevent technology from producing flawless outcomes like humans. When it comes to detecting diseases, the machine occasionally becomes confused. There are some mistakes after selecting the best modelFig. 9 displays two data Brown Spot with Leaf Blast and Leaf Blight, as well as eight data conflicts between Leaf Blast and Healthy Leaf. It's a modest number, though.

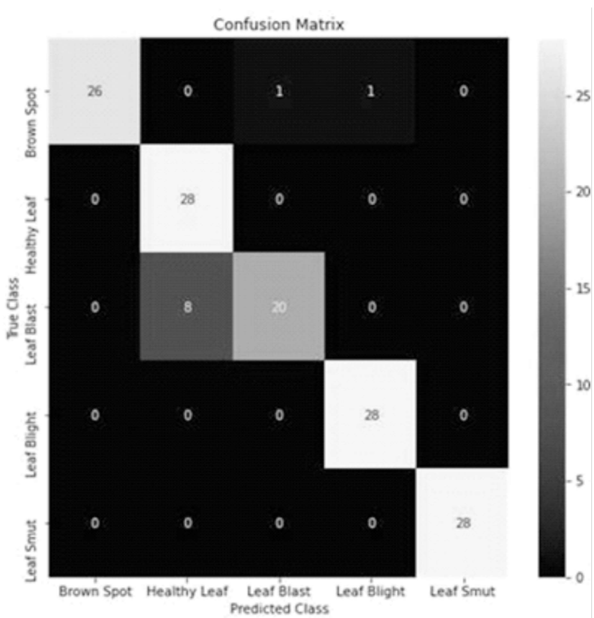

Figure 9: Inception-ResNet-V2 Model Confusion Matrix

VII. CONCLUSION

For this study, four benchmark deep learning network designs were investigated using different statistical metrics and their performance was assessed. The Inception-ResNet-V2 network design yielded the best test accuracy, 92.68%, when the algorithm's accuracy, precision, recall, and F1 score were evaluated. The model training and testing data used in this paper came from local paddy businesses and various online sources. Five classes make up the dataset; four of these classes include images of paddy leaf diseases that are commonly prevalent, and one class includes images of healthy leaves. The Inception-ResNet-V2's distinct architecture, which consists of a stem, reduction, and inception-resnet blocks with a depth of 571, has a greater impact on dataset adaptation than other networks. With a testing accuracy of 91.52%, the Res Net-101 network came in second. Transfer learning techniques were applied in order to produce a more precise prediction of the paddy leaf diseases. This transfer learning modification improved accuracy while decreasing the complexity of model training time.

REFERENCE

[1] Trébuil, G., 2011. Rice production systems in Asia: The constant presence of an essential cereal on a continent in mutation.

[2] Papademetriou, M.K., 2000. Rice production in the Asia-Pacific region: issues and perspectives. Bridging the rice yield gap in the Asia-Pacific region, 220.

[3] Bangladesh Rice Knowledge Bank, available at<<http://www.knowledgebank-brri.org/riceinban.php>>, last accessed on 25-11-2020, 08.40 am.

[4] Wikipedia, list of rice leaf diseases, available at <<https://en.wikipedia.org/wiki/List_of_rice_diseases>>, last accessed on 28-11-2020, 05.10 am.

[5] Mondal, M. H. (2010). Crop agriculture of Bangladesh: Challenges and opportunities. Bangladesh Journal of Agricultural Research, 35(2), 235-245.

[6] Mohanty, Sharada P., David P. Hughes, and Marcel Salathé. "Using deep learning for image-based plant disease detection." Frontiers in plant science 7 (2016): 1419.

[7] Sethy, P.K., Negi, B., Barpanda, N.K., Behera, S.K. and Rath, A.K., 2018. Measurement of disease severity of rice crop using machine learning and computational intelligence. In Cognitive Science and Artificial Intelligence (pp. 1-11). Springer, Singapore.

[8] Ramesh, S. and Vydeki, D., 2019. Application of machine learning in detection of blast disease in South Indian rice crops. Journal of Phytology, pp.31-37.

[9] Shrivastava, Vimal K., et al. "Rice Plant Disease Classification using Transfer Learning of Deep Convolution Neural Network." International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences (2019).

[10] Li, Dengshan, Rujing Wang, Chengjun Xie, Liu Liu, Jie Zhang, Rui Li, Fangyuan Wang, Man Zhou, and Wancai Liu. "A recognition method for rice plant diseases and pests video detection based on deep convolutional neural network." Sensors 20, no. 3 (2020): 578.

[11] Dhingra, G., Kumar, V. and Joshi, H.D., 2018. Study of digital image processing techniques for leaf disease detection and classification. Multimedia Tools and Applications, 77(15), pp.19951-20000.

[12] G. Bebis and M. Georgiopoulos, "Feed-forward neural networks," in IEEE Potentials, vol. 13, no. 4, pp. 27-31, Oct.-Nov. 1994, doi: 10.1109/45.329294.

[13] T. Kwiatkowski and H. Bölcskei, "A Mathematical Theory of Deep Convolutional Neural Networks for Feature Extraction," in IEEE Transactions on Information Theory, vol. 64, no. 3, pp. 1845-1866, March 2018, doi: 10.1109/TIT.2017.2776228.

[14] Zeiler M.D., Fergus R. (2014) Visualizing and Understanding Convolutional Networks. In: Fleet D., Pajdla T., Schiele B., Tuytelaars T. (eds) Computer Vision – ECCV 2014. ECCV 2014. Lecture Notes in Computer Science, vol 8689. Springer, Cham. https://doi.org/10. 1007/978-3-319-10590-1_53.

[15] S Agatonovic-Kustrin, R Beresford,Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research,

[16] M. Oquab, L. Bottou, I. Laptev and J. Sivic, "Learning and Transferring Mid-level Image Representations Using Convolutional Neural Networks," 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, 2014, pp. 1717-1724, doi: 10.1109/CVPR.2014.222.

[17] S. Ghosal and K. Sarkar, "Rice Leaf Diseases Classification Using CNN With Transfer Learning," 2020 IEEE Calcutta Conference (CALCON), Kolkata, India, 2020, pp. 230-236, doi: 10.1109/CALCON49167.2020.9106423.

[18] UCI Machine Learning Repository, available at <<https://archive.ics.uci.edu/ml/datasets/Rice+Leaf+Diseases>>