



IDENTIFICATION OF PADDY LEAF DISEASES USING MACHINE LEARNING

¹K. Nagi Reddy , ²N.Yagnitha , ³M. Chandu, ⁴M. Manjula , ⁵P. Muni Rakesh

Abstract—Now a days, Farmers are facing loss in crop production due to many reasons one of the major problem for the above issue is crop diseases. This is due to lack of knowledge about the disease and pesticides or insecticides available in order to control the disease. But finding the current disease and providing best remedies requires expert opinion or prior knowledge in order to control the disease. This is a time consuming and expensive. In order to solve the above issue the Machine Learning model using Convolutional Neural Network (CNN) algorithm is developed to detect the paddy crop disease using the image.

Keywords: Paddy leaf, Deep neural network, Canny based edge detection, Image processing, DNN, convolution neural network

I. INTRODUCTION

It has been observed that agriculture is the important source of earning for the human being in several countries [1]. According to farmers requirements and their land need food plants can be harvested by them. But because of this farmers are facing many problems like various diseases in plants, natural disasters, water scarcity, etc.[2]. Using various technical facilities several issues can be solved. There is no need to search experts because such mechanism could improve the food productivity [3]. The essential topic of research is detection of plant disease [4] in agriculture domain. It is found that the demanding task is finding and classifying the diseases of plant. To avoid loss in the agricultural product yield, significant key is to detect plant diseases [5]. Diseases can be considered as a pattern in various studies of recognition of plant diseases [6]. Plant diseases are complex to monitor if it is made manually. A huge amount of work and experience is required to perform the manual processing time thus plant disease recognition as well as image processing mechanisms have been used [7]. Image processing involves recognition of diseases by fetching image. Images are pre-processed and segmented. Finally the features are extracted for classification purpose [8]. Such techniques could be implemented on external appearances of plants that are infected [9]. It is observed that in several plant leaves are significant source to find the disease. Sheath rot, leaf blast, brown spot, bacterial blight and leaf smut have been considered as diseases that are common in case of rice plants [10]. Symptoms of plant diseases are varied in case of various plants. Observations conclude that plant diseases have been found in variety of color, size and shape. There different disease associated to different features of plant leaves. Researchers have find that different plant disease have different color some have yellow in color and some have brown [11]. Many diseases have same shape but different colors but some have same color and different in shape. Normal part of characteristics related to disease might be fetched after segmentation [12].

II. PADDY DISEASES

Paddy is frequently affected by Rice Blast and Bacterial Blight thus leads to decrease in paddy crop production. The cause and the symptoms of the rice blast , bacterial blight and leaf smut is discussed below,

A. Brown Spot

Rice Blast is caused by a fungus named as *Magnaporthe oryzae* and it affects the leaf collar, collar node, neck and leaf node of the paddy crop. Rice Blast can kill the rice plants at seedling stage and leads to yield losses in case it severely infected. Initially Rice blast symptoms appear as lesions on leaf diamond shape. Later lesion became elliptical with gray center and brownish border. Then it gets enlarged and starts killing the entire leaves. In older lesions leaves became elliptical or spindle shaped and whitish to gray centers with red to brownish or necrotic border. Paddy leaf infected by Rice Blast as shown in Fig1.

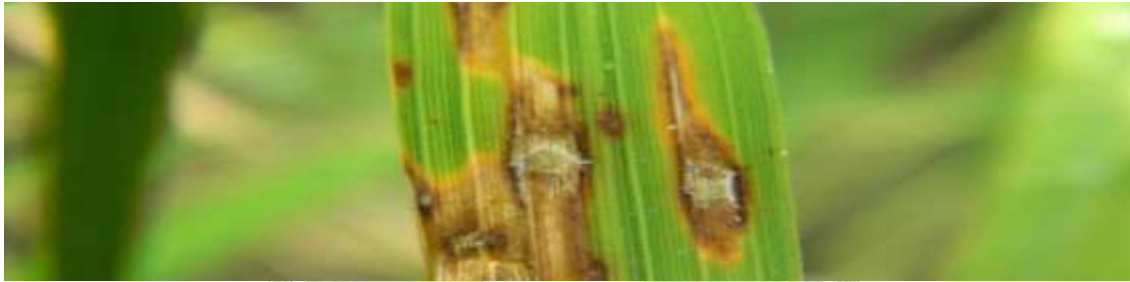


Fig. 1. Brown Spot

B. Bacterial leaf blight

Bacterial leaf blight is caused by bacteria called *Xanthomonas oryzae*. Bacterial leaf blight is one of the deadly and destructive diseases of paddy and it may leads to crop loss of 75%. Blight causes yellowing, drying of leaves and wilting of seedlings. Initially symptoms appear as water-soaked streaks in leaf tips. In later stage grayish white lesions appear on the leaves leads to yellowing and drying of leaves. Blight is an epidemic disease and it can easily spread in wind and water. The paddy leaf infected by bacterial blight is shown in Fig2.



Fig. 2. Bacterial Blight

C. Leaf Smut

Leaf Smut, caused by the fungus *Entyloma oryzae*, is a widely distributed, but somewhat minor, disease of rice. The fungus produces slightly raised, angular, black spots (sori) on both sides of the leaves as shown in the Fig3. Although rare, it also can produce spots on leaf sheaths.



Fig. 3. Leaf Smut

III. LITERATURE SURVEY

Shruthi U, Nagaveni V, Raghavendra B K proposed “A review on machine learning classification techniques for plant disease detection” shows different machine learning classification are used to detect the crop disease. The different classification methods are Artificial Neural Network (ANN) classification technique, K-Nearest Neighbor (KNN) classification technique, Convolutional Neural Network

classification, fuzzy classifier and Support Vector Machine (SVM) classification methods. From above mentioned techniques Convolution Neural Network provides high accuracy compared to other methods and detect more number of diseases in multiple crops [1].

V.Vanitha proposed “Rice disease detection using deep learning” which proposes an automatic disease detection using deep neural network. The developed model is capable of detecting 3 different disease of paddy and also detect the healthy leaf image. The dataset was trained with three CNN architectures and achieved a high accuracy of 99.53% [2].

Shamim Mahbub, Md. Abu Nasim, Md. Jahid Hasan, Md. Shahin Alom proposed “Rice disease identification and classification by integrating support vector machine with deep convolutional neural network” represents a system to identify the rice disease and help the farmers to take proper decision to control the disease and also help them to increase production. They have built a AI model by integration of Support Vector Machine(SVM) with Deep Convolutional Neural Network(DCNN). Model identifies and classifies nine types of rice diseases with an accuracy of 97.5% [3].

Anuradha badge proposed “Crop disease detection using Machine learning: Indian Agriculture” how diseases affecting the less yield and how machine learning technique will help us to detect the disease and help the farmers to take necessary action. They used canny’s edge detection algorithm for the efficient detection of crop disease by taking image of crop They take wheat crop for their research paper [4].

The survey help us to find out the best machine learning algorithm, for the proposed system as described in [1] CNN provides highest accuracy and processing speed compared to other algorithms.

Table1. Methodologies

S. No	Author/Year	Objective	Methodology
1	Ramesh/2020	Finding and Classifying diseases in paddy leaf	Optimized Deep Neural Network(DNN)
2	X.E. Pantaz/2019	To Perform Leaf disease detection in different crop species	Image Processing
3	Kim/2018	Considering recent developments in precision agriculture	Nanotechnology

IV. PROPOSED METHODOLOGY

In this system for detecting the Brown Spot, Leaf Smut and Bacterial leaf blight disease of paddy, 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. Both healthy and disease leaf image dataset are collected. Here we have collected 2000 images of Rice blast, 2000 images of Bacterial Blight and 2000 healthy paddy leaf images. These images are trained with the help of Convolutional Neural Network (CNN) Algorithm.

- ▶ The model that is proposed by us to detect and classify the infected plant leaves consists of 4 phases.

Those phases are

- ▶ Dataset Collection
- ▶ Image Preprocessing
- ▶ Segmentation
- ▶ Selection of Classifier

Data set Collection: Firstly, the images of leaves were collected from online sources such as GitHub, Kaggle and also some of the image’s dataset consists of 20,000 images divided into 19 different classes. The dataset consists of both healthy and infected leaves which covers diseases like brown spot, leaf smut, bacterial blight etc.,

Image preprocessing: In this step images are resized to smaller pixel size in order to speed up the computations. The acquired images contain some noise. This noise is removed using some filtering techniques like feature extraction.

Segmentation: In this step, segmentation of images is done in order to separate the leaves from the background. Segmentation is performed using K-means clustering with 2 cluster centers, one for

background and one for foreground. K-means clustering is unsupervised learning technique that is used to segregate the datapoints in the predefined number (k) of clusters or groups on the basis of their similarities.

Selection of Classifier: This is the classification problem as we have to classify the type of disease on the leaf of the paddy. So, we have plenty of machine learning algorithms that can be applied on this dataset.

- ✚ **Algorithms used:** Support vector machine(SVM),
- ✚ K-Nearest Neighbors(KNN), Linear Regression.

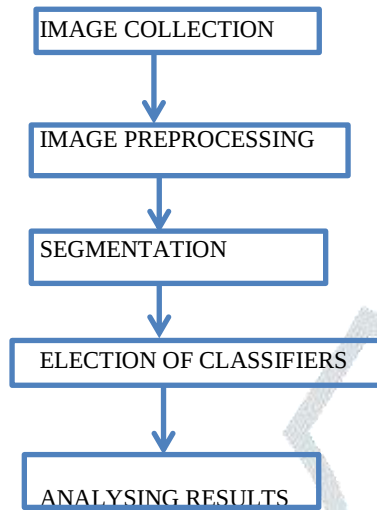


Fig. 4: Block Diagram

A. CNN Architecture

A CNN consists of an input layer, multiple hidden layer and an output layer. In hidden layer consist of Convolution layer, Rectified Linear Unit, pooling layer and fully connected layer. The CNN architecture for the proposed model is shown in figure 5. The input layer takes the resized, gray scaled image and output layer produces the detection of the disease and provides remedies. The detailed explanation of the remaining layers as follows

1. Convolutional Layer

The training data (images of the diseased and healthy rice plant) was sent to input layer of CNN. The convolution operation is then performed on input samples; the input is convolved with filters called kernels, that is, a number of filters slide over the feature map of the previous layer, to produce output feature maps.

2. Rectified Linear Unit (ReLu)

In this layer is usually called as activation function layer, ReLu is one of the activation function. There are various types of activation function available such as sigmoid, Tanh, ReLU, Softmax, etc. In our model ReLu activation function is used in hidden layers. It is the most widely used activation function. In ReLu layer the image with negative pixel values are replaced with pixel value 0 and remaining pixel retain as it is. The ReLU function can be written in the mathematical form in equation 1.

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } x \geq 0 \end{cases} \quad (1)$$

where x is a pixel value

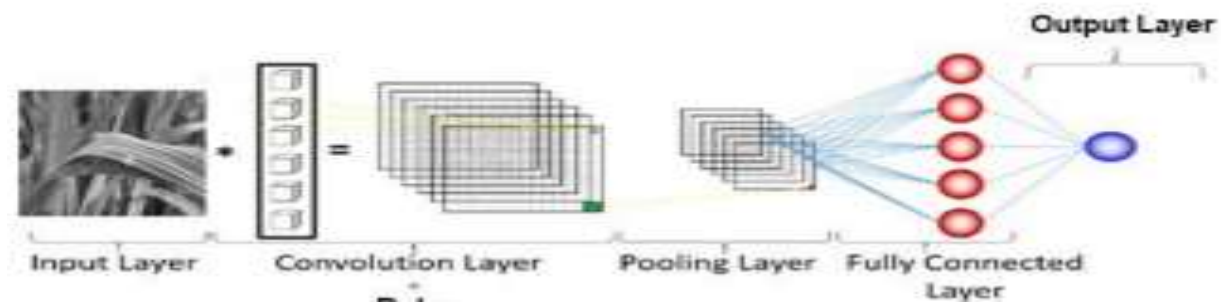


Fig. 5. Convolution Neural Network Architecture

3. Pooling Layer

A pooling layer performs reduction operation along the dimensions of image (Width, Height), resulting dimensionality reduction. The primary aim of pooling operation is to reduce the size of the images as much as possible. This scans across the image using a window and compresses the image extracting features. Average pooling and Max pooling are the most commonly used methods in pooling layers. In max pooling largest value of the pixel is taken from the selected window of the image, while average pooling takes the average of all pixel values within the window.

4. Fully Connected Layer

After the convolution + RELU+ Pooling layers, these layers stack up many times until the image is reduced to a vector. In this layer actual classification will going to happen. In this layer all the neurons are interconnected; this layer produces an N-dimensional vector, where every neuron in this layer contains the vectors of the features extracted from the image. The proposed system has concentrated on detecting the paddy diseases and, thus leads to increase in paddy crop production. In this system it detects the most common and frequently occurring paddy diseases (Rice blast , Brown Spot and bacterial Blight) and provides pesticides or insecticides as a remedy to control the disease. The type of paddy disease is detected by CNN algorithm. The proposed system architecture is shown in Fig 6.

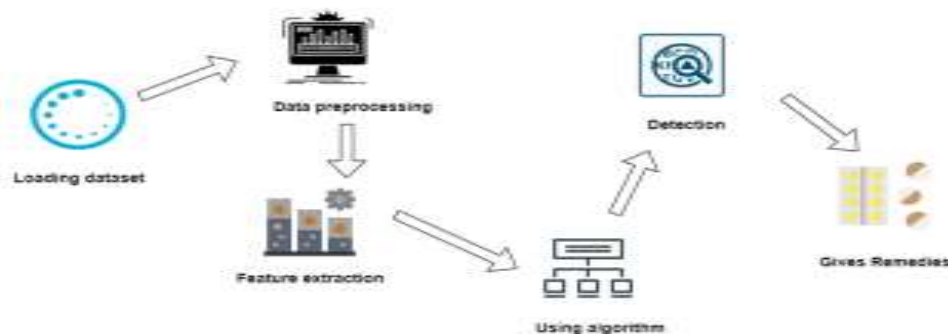


Fig.6. Proposed Architecture

As shown in Fig 6, first the data set is loaded and the data set is preprocessed by reducing the size and converting it to the gray scale, then the features are extracted using the convolutional layer of CNN these features will help the model in training. After training, the model is able to detect the disease. In the testing phase same data preprocessing is done and the image data is uploaded to CNN algorithm. The given image is considered as vector and compared with the training model vector then the model is able to detect the disease and provides the remedies. If more data set is used during the training the model provides better accuracy while detecting the paddy leaf disease.

B. Brown Spot

Performance parameters:

Fig. 7 show that Different diseases of paddy leaf at different levels, from these levels, the components of the confusion matrix can be obtained.

The confusion matrix is shown in Table 2 derived from the image of Brown spot. The confusion matrix will have following parameters , True Negative (TN), True Positive (TP), False Negative (FN)and False Positive (FP). The parameters utilized for verifying results are accuracy, precision, recall and f1 score that are defined by:

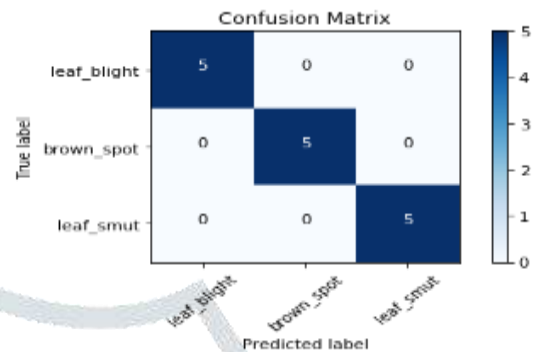
$$\text{Accuracy} = \frac{TN+TP}{TN+FP+TP+FN}$$

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

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

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

Table 2: Confusion Matrix



It has been stated from the Confusion Matrix of table 2 that the parameters of Brown Spot, here '0' is considered as "Negative value" and '5' is considered as "Positive Value".

Table3: Results of proposed method (CNN)

Type of diseases	Accuracy	Precession	Recall	F1 Score
Brown spot	99%	98.95%	98%	98.4%
Bacterial Leaf Blight	98.5%	99.9%	100%	99.9%
Leaf Smut	98%	100%	99.99%	99.99%
Average	98.5%	99.61%	99.33%	99.43%

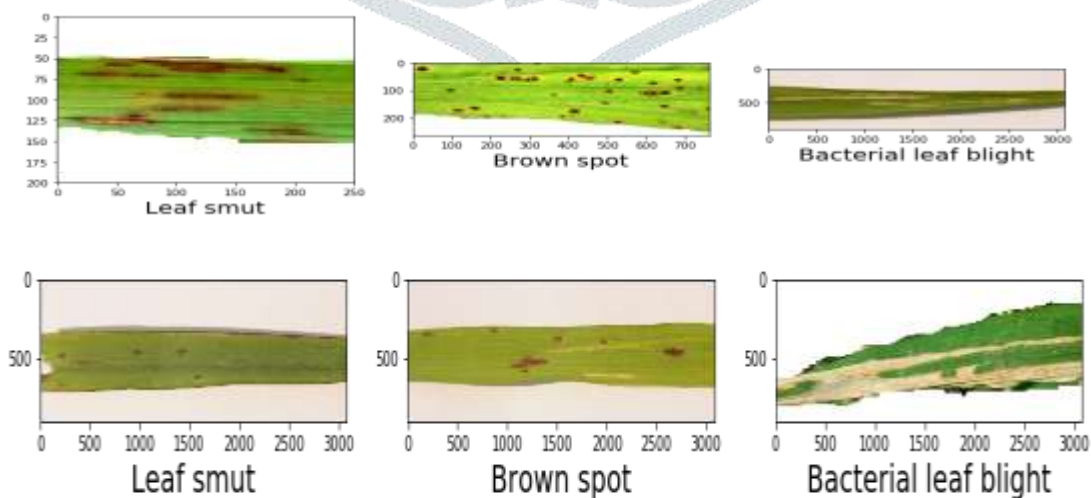


Fig. 7 Different diseases of paddy leaf at different levels

V. EXISTING ANALYSIS

Table4: Existing Results (ANN)

Type of Diseases	Accuracy	Precession	Recall	F1 Score
Brown Spot	93%	94%	93%	93%
Bacterial Leaf blight	93.2%	94%	93.2%	93%
Leaf Smut	93.2%	94%	93%	93.2%
Average	93.14%	94%	93.06%	93.06%

VI. RESULTS AND DISCUSSION

The Indian farmers are facing many problems like various diseases in plants, natural disasters, water scarcity, etc.. Using various technical facilities several issues can be solved. This paper uses CNN for identifying the diseases as one of the machine learning techniques and compared this with the traditional ANN algorithm. By comparing these algorithms the following are drawn from table 3 and 4.

The accuracy, Precession, Recall and F1 Score of the Brown Spot, Bacterial Leaf blight and Leaf Smut have calculated and tabulated in table 3 and 4 using the existing ANN and proposed CNN. The average accuracy, Precession, Recall and F1 Score also calculated in the both the cases.

Based on the both the tables 3 and 4 it is inferred that the average accuracy of CNN and ANN is 98.5% and 93.14% respectively. So there is 5.36 % of improvement of accuracy is achieved by CNN compared with ANN.

The average precession of ANN 94% and CNN is 100%, there is an improvement of 6% using CNN when compared with the ANN.

The Recall and F1 Score of ANN is 93.06% whereas for CNN these values are 99.33% and 99.43% respectively. Hence a there is an improvement of 6.27 and 6.37 respectively using CNN compared with the ANN.

VII. CONCLUSION

Farmers are facing the issues with paddy crop disease identification and unable to find effective pesticide or insecticide to control the infected disease. To combat this issue a machine learning model using the Convolutional Neural Network (CNN) algorithm is proposed to detect the Brown Spot, Bacterial Blight and Healthy paddy leaf images. By comparing the results obtained using ANN and CNN it is concluded that CNN is detecting the diseases of the paddy leaves with high accuracy, Precession, Recall and F1 Score. Hence CNN can be better option for detection of diseases when compared to ANN.

VIII. REFERENCES

- [1] Ramesh, S., &Vydeki, D. (2020). Recognition and classification of paddy leaf diseases using Optimized Deep Neural network with Jaya algorithm. Information processing in agriculture, 7(2), 249-260.
- [2] Vo-Tong Xuan (2018). Rice production, agricultural research, and the environment. Routledge, In Vietnam's rural transformation (2018), pp. 185-200
- [3] X.E. Pantazi, D. Moshou, A.A. Tamouridou (2019). Automated leaf disease detection in different crop species through image features analysis and one class classifiers. Comput Electron Agric, 156 (2019), pp. 96-104
- [4] M.K. Elkazzaz, E.A. Salem, K.E. Ghoneim, M.M. Elsharkawy, G.A. El-Kot, Z.A. Kalboush (2015). Integrated control of rice kernel smut disease using plant extracts and salicylic acid. Arch Phytopathol Plant Protect, 48 (8) (2015), pp. 664-675
- [5] M. Yusof, N.F. Mohd, M. Rosli, R. Othman, M.H.A A. Mohamed (2018). M-DCocoa: M-agriculture expert system for diagnosing cocoa plant diseases. Proc. International Conference on Soft Computing and Data Mining. 2018 (2018), pp.

363-371

- [6] Kim, DaeYoung, A. Kadam, S. Shinde, R.G. Saratale, J. Patra, G. Ghodake (2018). Recent developments in nanotechnology transforming the agricultural sector: a transition replete with opportunities. *J Sci Food Agric*, 98 (3) (2018), pp. 849-864.
- [7] Astonkar, R. Shweta, V.K. Shandilya (2018). Detection and Analysis of Plant Diseases Using Image Processing. *Int Res J Eng Technol*, 5 (4) (2018), pp. 3191-3193
- [8] Singh, B. Kumar, S.S. Ganapathy subramanian, A. Singh (2018). Deep learning for plant stress phenotyping: trends and future perspectives. *Trends Plant Sci*, 23 (10) (2018), pp. 883-898
- [9] M. Kamal, A.N.I. Mahanijah, F.A.R. Masazhar (2018). Classification of leaf disease from image processing technique. *Indonesian J Elect Eng Comput Science*, 10 (1) (2018), pp. 191-200
- [10] D. Patrício, R.R. Inácio (2018). Computer vision and artificial intelligence in precision agriculture for grain crops: a systematic review. *Comput Electron Agric*, 153 (2018), pp. 69-81
- [11] B.H. Prajapati, J.P. Shah, V.K. Dabhi (2017). Detection and classification of rice plant diseases. *IntellDecisTechnol*, 11 (3) (2017), pp. 357-373
- [12] J.G. Barbedo, L.V. Arnal, T.T.S. Koenigkan (2016). Identifying multiple plant diseases using digital image processing. *BiosystEng*, 147 (2016), pp. 104-116
- [13] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, D. Stefanovic (2016). Deep neural networks based recognition of plant diseases by leaf image classification. *Comput Intel Neurosci* (2016), pp. 1-11
- [14] P. Mohanty, D.P. Sharada, M.S. Hughes (2016). Using deep learning for image-based plant disease detection. *Front Plant Sci*, 7 (1419) (2016), pp. 1-10
- [15] A.-K. Mahlein (2016). Plant disease detection by imaging sensors—parallels and specific demands for precision agriculture and plant phenotyping. *Plant Dis*, 100 (2) (2016), pp. 241-251
- [16] F. Pinki, N. Tazmim, S.M.M Islam Khatun (2017). Content based paddy leaf disease recognition and remedy prediction using support vector machine. In: *Proc. In Computer and Information Technology (ICCIT)*, 20th International Conference (2017), pp. 1-5
- [17] Y. Lu, S. Yi, N. Zeng, Y. Liu, Y. Zhang (2017). Identification of rice diseases using deep convolutional neural networks. *Neurocomputing*, 267 (2017), pp. 378-384
- [18] G. Dhingra, V. Kumar, H.D. Joshi (2019). A novel computer vision based neutrosophic approach for leaf disease identification and classification. *Measurement*, 135 (2019), pp. 782-794
- [19] Lipsa Barik, A Survey on Region Identification of Rice Diseases Using Image Processing, *International Journal of Research and Scientific Innovation (IJRSI) | Volume V, Issue I, January 2018 | ISSN 2321-2705*
- [20] G. Jayanthi, K.S. Archana, A. Saritha, Analysis of Automatic Rice Disease Classification Using Image Processing Techniques, *International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-8, Issue-3S, February 2019.*
- [21] Saradhambal.G1, Dhivya.R2, Latha.S3, R.Rajesh4, Plant Disease detection and its solution using image classification, *International Journal of Pure and Applied Mathematics Volume 119 No. 14 2018, 879-884 ISSN: 1314-3395 (on-line version) url: http://www.ijpam.eu Special Issue*
- [22] S.Ramesh, Recognition and Classification of paddy leaf diseases using Optimized Deep Neural Network with Jaya Algorithm, 6 sep 2019, In Press
- [23] De. Neha Mangla, Priyanka B Raj, Soumya G Hegde, Pooja R, Paddy Leaf Disease using Image Processing and Machine Learning, publishes on *IJ IREEICE* on 2 feb 2019, vol 7
- [24] K.S. Archana1, ArunSahayadhas, Automatic Rice Leaf Disease Segmentation Using Image Processing Techniques, published in *International Journal of Engineering & Technology*, 7 (3.27) (2018) 182-185.
- [25] Komal Bashir, Mariam Rehman, And Mehwish Bari, Detection and Classification of Rice Diseases: An Automated Approach Using Textural Features, in Vol. 38, No. 1, 239-250, January 2019.