

## Identification of Paddy Leaf Diseases using Evolutionary and Machine Learning Methods

Nilam Sachin Patil<sup>1</sup>, E. Kannan<sup>2</sup>

<sup>1</sup>VelTech Rangarajan Dr. Sagunthala R & D Institute of Science and Technology, Chennai.

<sup>2</sup>VelTech Rangarajan Dr. Sagunthala R & D Institute of Science and Technology, Chennai.

**Article History:** Received: 10 November 2020; Revised: 12 January 2021; Accepted: 27 January 2021; Published online: 05 April 2021

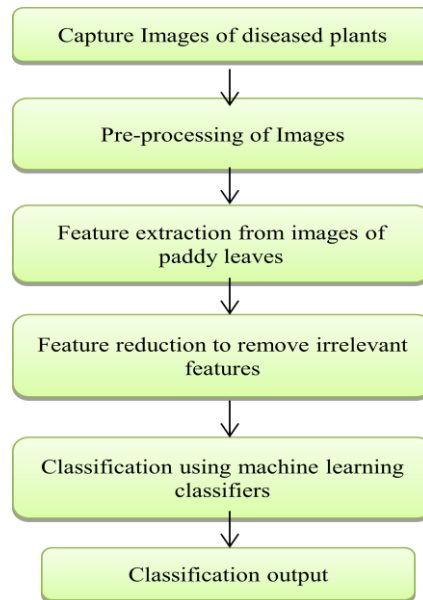
**Abstract:** In the field of agriculture, especially paddy plants, there is a demand for research to classify the paddy diseases at early stages. This is feasible if there are automated systems that can assist the farmers to recognize the paddy diseases from the paddy leaf images of the plants. The recognition of agricultural plant diseases by utilizing the image-processing and machine learning techniques can certainly minimize the reliance on the farmers to protect the yield of paddy crops. In this paper, an attempt has been made to pre-process the images to prepare the feature-set for Classifiers and then feature extraction algorithms are used to extract the relevant features from the processed images. The feature-set is then supplied to the classifiers for identification of Paddy Leaf diseases. The usage of cascaded classifiers has been explored to detect the diseases of paddy leaves. An attempt has also been made to use genetic algorithm with nearest neighbour algorithm to identify the diseases of paddy leaves. The proposed automated system can be used on Android , Windows platform and Apple platform for quickly identifying the paddy leaf diseases as the entire implementation has been performed using MATLAB. The proposed automated system can certainly help the farmers to classify the diseased paddy leaves at early stage to protect the crops from further damage.

**Keywords:** Paddy disease detection, Classification, Adaboost algorithm, Bagging classifier.

### 1. Introduction

The machine learning based automated systems are the need of the hour for the Indian fields that produce great amount of rice to identify the diseases as soon as initial symptoms of the diseases appear on the paddy leaves and to save the yield from further damage (Qing & Zexin, 2009). Machine learning techniques help to extract the features from the images and then redundant features can also be reduced using machine learning methods (Pugoy & Mariano 2011). Once the feature set is ready, again machine learning and/or deep learning methods can be used to classify the images in appropriate classes. The plant images can be easily identified using machine learning as well as deep learning approaches with high accuracy (Phadikar & Sil, 2008). In Figure. 1, a generic approach to identify the plant diseases using machine learning approaches is demonstrated in a pictorial representation.

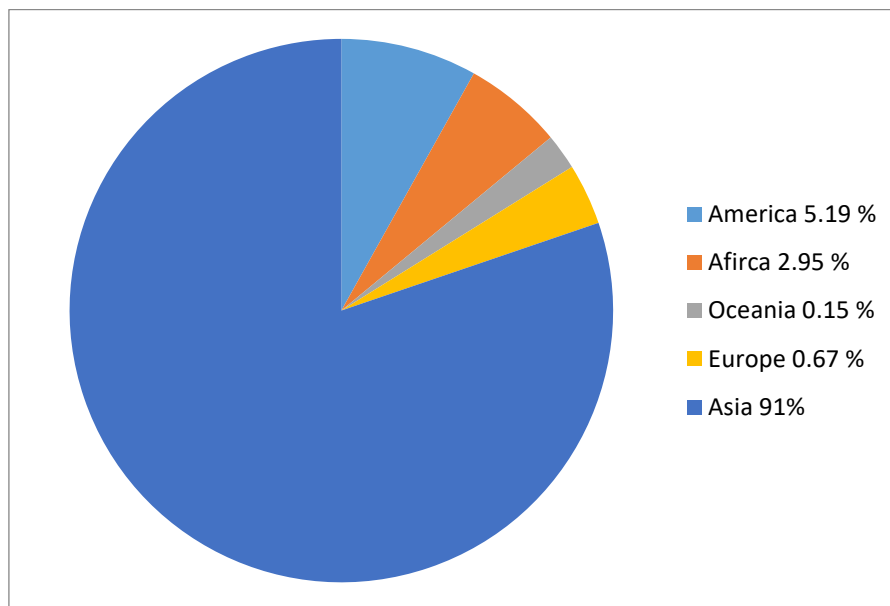
In generic approach, the images of diseased plants are captured and processed, so that the features from the images can be extracted for further processing. After pre-processing of images, the features of the images are extracted to prepare a feature set. The parameters of feature set can directly be supplied to the classifiers or feature reduction algorithms can be used if feature set contains redundant features that do not play any role in classification of the images. The classifiers are trained using the feature set and tested after attaining an acceptable accuracy. The model gets approval after gaining good accuracy in both trained as well as tested data/ feature sets. However, machine learning approaches are providing acceptable accuracy in classification of images of diseased plants but deep learning techniques are gaining more attention because of delivering better accuracy in identification of plant diseases from the images of the diseased plants.



**Figure 1.** Generic scheme of classification of paddy leaf diseases using ML

### 1.1. Background of the research

Agriculture is India's largest source of revenue. The vast majority of the people in India rely on agriculture for their income. Agriculture is the most important sector of the Indian economy (Agriculture Sector in India,2015). Over 58% of the rural residents rely on farming for their primary source of income relates (Verma, 2017). Rice is a basic food for the majority of rural populations, and is the second most-produced cereal. Rice is a nutritional staple in India. Rice disease causes 10 to 15% of crop losses in Asia (Papademetriou, 2000). An agricultural crop is grown in five continents, namely Asia, Africa, America, Europe, and Oceania. Deliver and Devour 91.05% of the world's rice according to the Food and Agriculture Organisation of the United Nations (FAOSTAT) The rice production as per the population in different areas across the world like Africa-2.95% of , America-5.19% ,Asia-91%, Europe-0.67% and Oceania-0.15% of the population as seen in Figure. 2



**Figure 2.** Rice production across the World (Source FAOSTAT)

Rice consumption is projected to increase more rapidly than supply in most countries. In this case, some kind of harm to the crop is unacceptable (Khoenkaw ,2016). Prevalence of rice disease has often been difficult to ascertain. a naked-eye study was the only way to diagnose the rice disease For disease detection to be effective, constant observation of the field is needed (Prakashet & Saraswathy, 2017). Visual research relies on the use of continuous human interpretation, which greatly increases its expense, effort, and (Jaskaran & Harpreet, 2018).

When the unprecedented growth of the number of people occurs, so does the steady increase in the market for food products. The strain places the whole population under the necessity of utilizing new technologies for early detection and successful care, so there is little space for error. Image analysis techniques are one of the cost-effective and reliable approaches used for differentiating plant diseases (Jitesh & Harshad, 2016); (Iswarya & Maheswari, 2019). Plants are susceptible to infection when fungi and bacteria cause illness in them. leaf blight, brown blot, sheath blight, and leaf scorch (Rice Production , 2015). These diseases will place severe economic pressure on the rice farmers across the board. farmers may ignore diseases or struggle to recognize them, which lead to the loss of the crop Per illness has its own solution. when a disease appears on a plant, farmers have to monitor its spread (Basavaraj & Surendra 2020). This disease detection process needs some due diligence during the selection of pesticides. seemingly-infected cattle and letting themselves be captured by an automated device may be a potential alternative for farmers. With such mechanism the farmers can be kept updated on diseases immediately, many of them can save money and time from the major economic losses (Usha & Priyadharshini, 2019).

## **1.2. Need of the research**

One of most popular problems from agricultural science is the identification of diseases from the images of the plant leaves and to suggest the remedies to get rid of these diseases. The remedies can be suggested by the automated systems if the identification of the plant diseases can be made accurately. The usage of the image processing techniques along with machine learning based image identification techniques minimize the need of manual ways for the farmers to safeguard their crops from the diseases. The automated strategies are very useful for classifying paddy leaf disease from the images of the diseased plants. Rice diseases are identified from the images of the diseased leaves which enables accurate detection and classification of the paddy leaf diseases without investing time on manual inspection. Machine learning based methods assist in detecting the paddy leaf diseases in a faster and accurate manner. The outcome of these automated systems can eventually improve the agricultural production and reduces economic losses. Hence, this paper proposes image processing, feature extraction and classification techniques to identify the paddy leaf diseases through an automated system.

## **1.3. Contributions of the paper**

An attempt has been made to use machine learning algorithms for identification of the paddy leaf disease.

1. The images of paddy leaves have been processed to reduce the unwanted distortions, enhances image features for further processing. It involves resizing of images, brightness correction, filtering, illumination corrections, noise removal, geometric transformations, and grey scale transformations.
2. After pre-processing of images, feature extraction algorithms are applied, so that the relevant feature dataset can be attained from the processed images.
3. Next, cascaded classifiers with modifications with respect to the problem statement are applied to obtain better accuracy.
4. The proposed automated system can be used on Android and Windows platform for quickly identifying the paddy leaf diseases.

## **1.4. Organization of the paper**

In this paper is structured mainly into five sections. The paper begins with background study, need for the work and contribution of the research work, followed by literature survey in section II, this section discusses the existing methods used in identification of paddy leaf diseases. Section III describes about the proposed methodology for identification of the paddy leaf diseases. Next section discusses the obtained results. The last section explains conclusion of the research study presented in this paper.

## **2. Literature review**

Geraldin B. Dela Cruz(2019) introduced a smartphone application that assisted the rice farmers to determine nitrogen deficiency depending on the plant coloration. The tool may be used instead of or in conjunction with conventional nitrogen use. An easy to use technology for farmers was proposed without any training to the model. This paper introduced automated image processing methods using numbers to obtain high accuracy results. The target outcomes were computed by using the z-score statistical approach.

Anthon G. and Wickarchi N.(2009) discussed the crucial role of Paddy disease classification and recognition in terms of financial growth of the agricultural sector. To find a reliable image-based diagnostic device for paddy diseases, digital cameras were used to take images under experimental conditions for this research study. Three

tropical diseases were chosen for this study: rice blast (*Magnaphelgrisea*), rice sheath blight (*Rhizoctoniasolani*) and brown dots (*Cochiobotriamiyabeanus*). Digital picture processing began with a green paddy bacterial infection leaf. These images were partitioned into the segments using the tools of geometry and trigonometry. There was a great degree of precision was achieved for classifying the images of diseased paddy leaves.

Jitesh P. Shah et al. (2016) had provided a survey report on the infected rice samples by using the image analysis and machine learning techniques. This paper surveyed several image processing and machine learning methods that were applied for plant disease diagnosis and classification. This study provided detailed insights into 19 research works based on rice plant diseases. This research also offered a survey on critical parameters including dataset size, classifier number, preprocessing, classifier forms, and accuracy etc. J. Yang et al. (2016) suggested an accurate and efficient Nitrogen dosing with a laser-induced fluorescence (LIF). The LIF approach was used to determine the vitamin B2 levels in rice using ultraviolet light(355 nm) (excitation light source). This described the differences in the fluorescence spectra observed in Nitrogen-supplemented and Nitrogen-free rice leaves. The numbers of relevant features were then extracted from the fluorescence spectra, and tailored for N-fertilizer dose. Binary SVM method has been used for classification. The accuracy of their system was around 95%.

Archana K.S. & Sahayadhas A. (2018) discussed a disease detection technique in the agriculture sector by using advanced picture analysis to extract features and classification algorithms. The major challenge in detection of plant leave diseases are feature recognition/feature extraction and accurate classification. This paper proposed an algorithm for accurately predicting the bacterial infection(*Oryza sativa*)in rice plants at early stages. Numerous picture segmentation and classification techniques were attempted for identifying the paddy leaf diseases. The detailed analysis was made to assess the performance of the proposed work.

Latte M V & Shidna S. (2016) described the methods to detect and analyze the Paddy leaf deficiencies including nitrogen, phosphorus, and potassium with pattern detection techniques. Using the nutrient elements technology, the aim was to identify several leaf elements that could become potential reason behind deficiency in paddy leaves. In this paper, a pattern detection method was proposed to analyze the color pattern of the paddy leaves such as red, green, blue, and gold light brown color. The rice plant colors could also depict the dead leaves. A database of phosphorus, potassium, and nitrogen-deficient paddy leaves was prepared before applying the pattern detection techniques. A comparative study was made to analyze the performance of proposed study over the existing methods. The color patterns were used to detect the deficiency in paddy leaves more accurately in comparison to other nutritional components.

Prabira K.S et al. (2020) had examined the use of image processing methods for diagnosing the rice plant diseases. The authors discussed image processing methods, disease recognition methods, and feature extraction methods on the basis of state-of-the-art work. This paper also provided highlights on the recent developments in automated rice plant disease detection techniques, and also discussed the potential challenges and opportunities in the field of paddy leaf diseases. Prab Kumar et al. (2017) discussed the techniques to identify the stages of the rice plant diseases. The color abnormalities and dead zones on leaves were visualized. The experimental results proved that the procedure for identification of faulty components on the leaves of the rice crop was highly accurate. S.Phadke and J. Coval (2008) demonstrated the simple prototype for identifying the rice plant diseases along with the identification of stages of therice plant diseases based on the images of the rice plants. Digital cameras were used to identify the soiled parts of the rice plants with high accuracy.

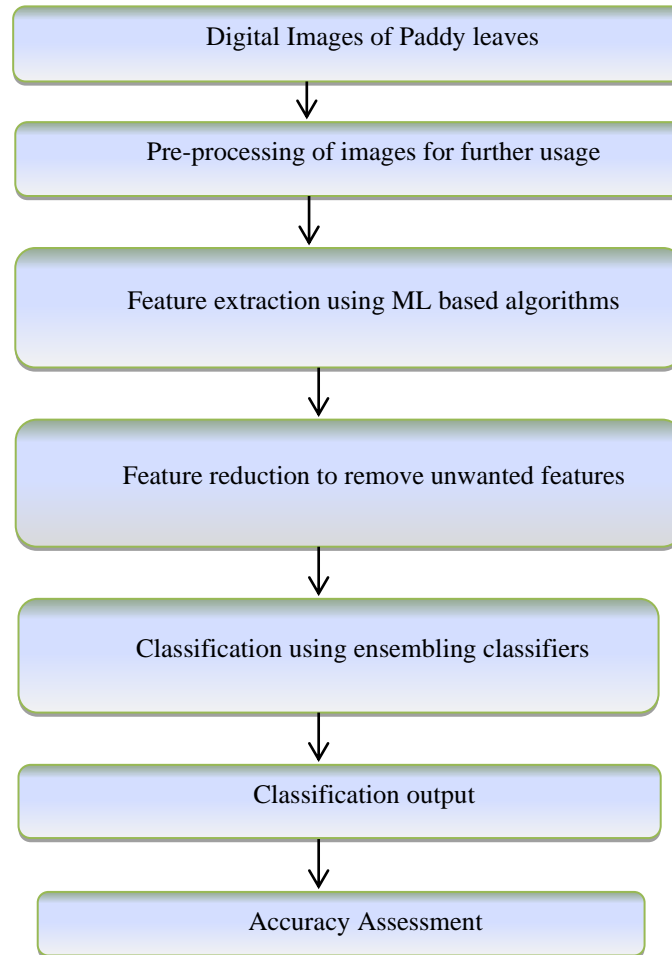
Ramesh S. & Vydeki D. (2018) discussed a disease detection technique for rice plants. The images of healthy and infected leaves were collected for supplying the inputs to the system. For the training process, the simulation produced a reliable estimation of 90% of the infected images and an additional 86% for the healthy ones. S. Sengupta & A. K. Das (2017) proposed rule-based classification method using the principles of Particle Swarm Optimization (PSO) and Association Theory (AT) to present an incremental classification scheme for classifying the diseases. The characteristics of rice diseases are constantly shifting with temperature, biological, and geographic influences, and thus the incremental classifier is appropriate for the prediction of diseases. Accuracy of the suggested technique has been assessed using all rice disease models and many times and using state-of-art classification methods. Statistical methods are often used to examine the results of the process which are used to determine their meaning and performance. Y. Zhou et al. (2020) conducted a systematic study on over 1883 rice varieties in China in order to classify blast-resistance alleles. This research work included the range of viable blast resistance alleles to include thirteen novel alleles (resistance strains).

However, many research approaches have been presented by the researchers but still there is a need of simple and viable automated system that can quickly recognize the rice plant diseases and help the farmers to take corrective measures to protect their crops at early stages,

### 3. Proposed Methodology

The proposed work involves four major steps:

- a) Pre-processing of diseased paddy leaves images for further processing;
- b) Extraction of features from the images of paddy leaves using different machine learning algorithms;
- c) Filtration of unwanted features using feature reduction techniques;
- d) Classification of paddy leaf diseases using cascaded classifiers based on machine learning.



**Figure 3.** Proposed approach

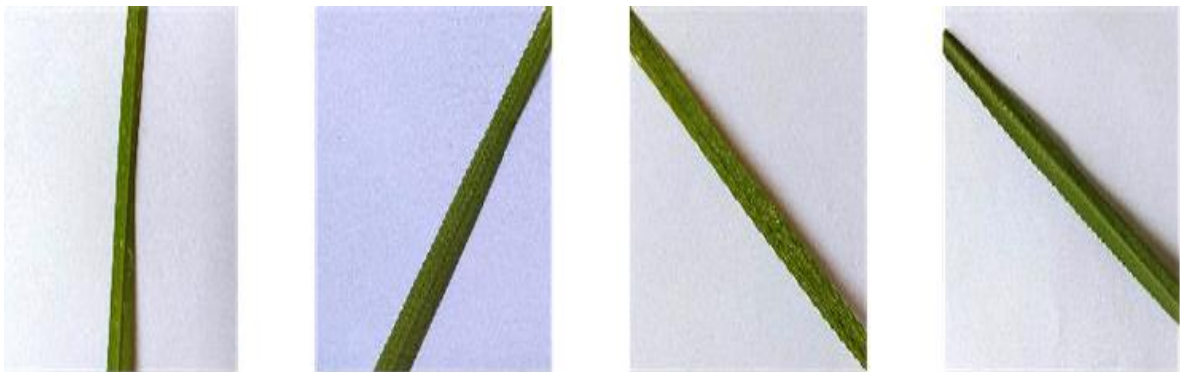
**Dataset:** The collection of data and preparation of dataset is very important for developing any machine learning based application, so for developing the Paddy leaf diseases diagnostic system. We have collected the images/videos for paddy leaves manually from the fields of Maharashtra, and we have also used online repositories of images such as Kaggle for the training of machine learning algorithms in initial phases.

**Images Pre-processing:** Before using the dataset to train our model, the series of pre-processing steps are to be applied to the data to enhance the images for further usage. During pre-processing, the image data reduces the unwanted distortions, enhances image features for further processing. It involves resizing of images into uniform sizes, brightness correction, filtering, illumination corrections, focus corrections, noise removal, thresholding, geometric transformations, and grey scale transformations. This process prepares the images of diseased paddy leaves for further analysis.

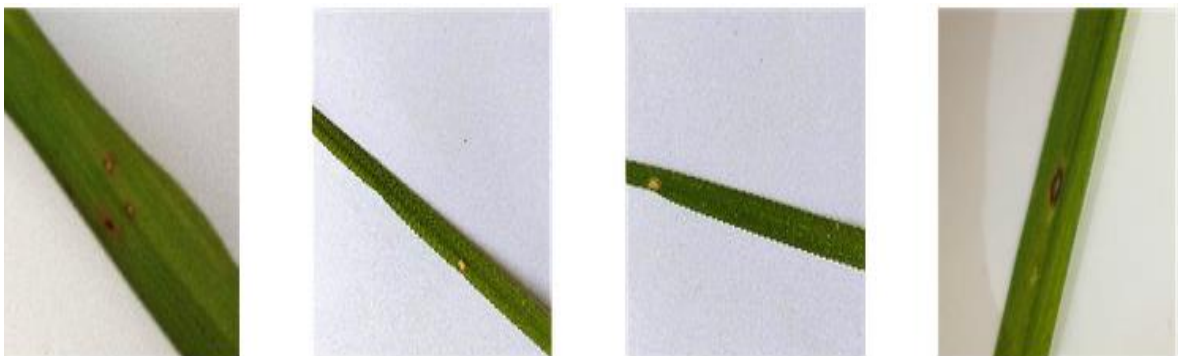
Following are the raw images from different categories like Hispa, Brownspot, Leafblast and Healthy leaves.



**Figure 4.** Image is depicting paddy leaves with BrownSpot disease



**Figure 5.** Image is depicting paddy leaves with Hispa disease



**Figure 6.** Image is depicting paddy leaves with LeafBlast disease



**Figure 7.** Image is depicting healthy paddy leaves

The dimension of the image is 1051 x 1051 and the format of the image is JPEG. We have reduced a dimension when the images are transformed from the BGR color space to gray scaled images. Grayscale is a range of monochromatic shades from black to white. Many image editing programs allow us to convert a color

image to black and white, or grayscale. This process removes all color information, leaving only the luminance of each pixel. The luminance of a pixel value of a grayscale image ranges from 0 to 255. The conversion of a color image into a grayscale image is converting the RGB values (24 bit) into grayscale value (8 bit). Gray scale is the most preferred pre-processing technique in image processing as it is one layer image from 0-255 whereas the RGB technique has three different layered images. Therefore we have preferred grey scale image instead of RGB. RGB color increases the complexity of the model and hence converting the images to a grayscale is beneficial in order to reduce the complexity of the ML based classification model.

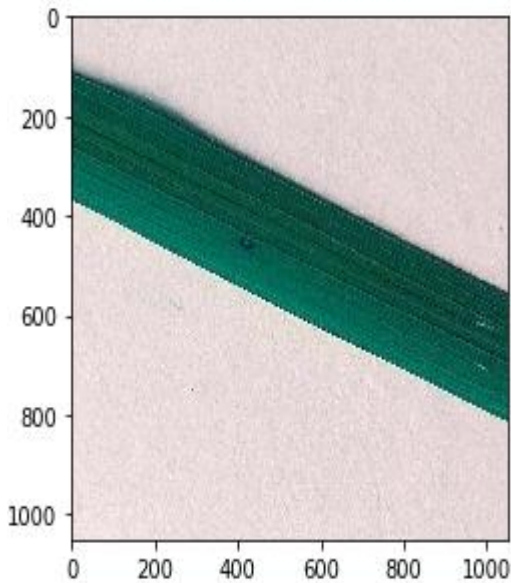


Figure 8. Original Image Shape (1051, 1051, 3)

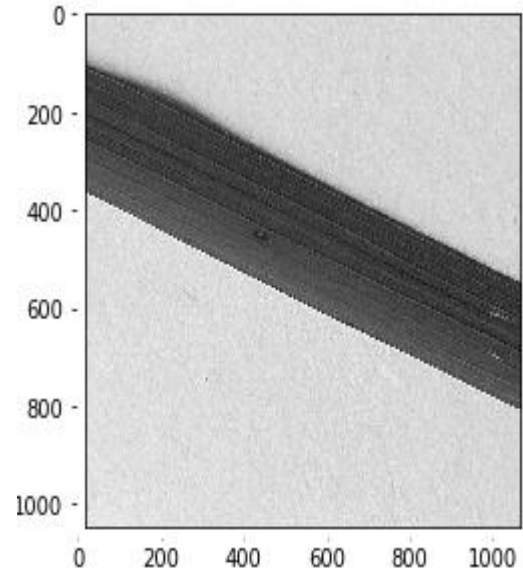


Figure 9. Grey Scaled image (1051, 1051)

Then we have taken analyzed the distribution of the intensity values of all the pixels presented in the images. It was required to flatten the image matrix and put the intensity values of all the pixels in one single row vector to find the prominent peaks.

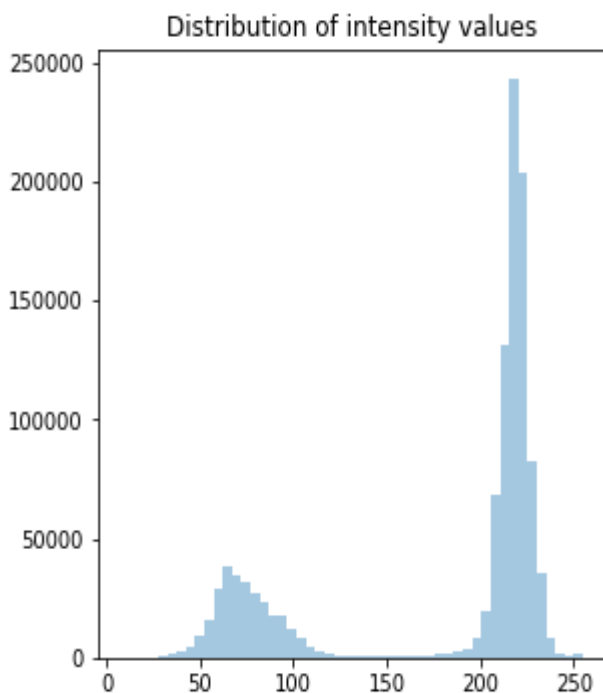


Figure 10. Presenting the distribution of intensity values

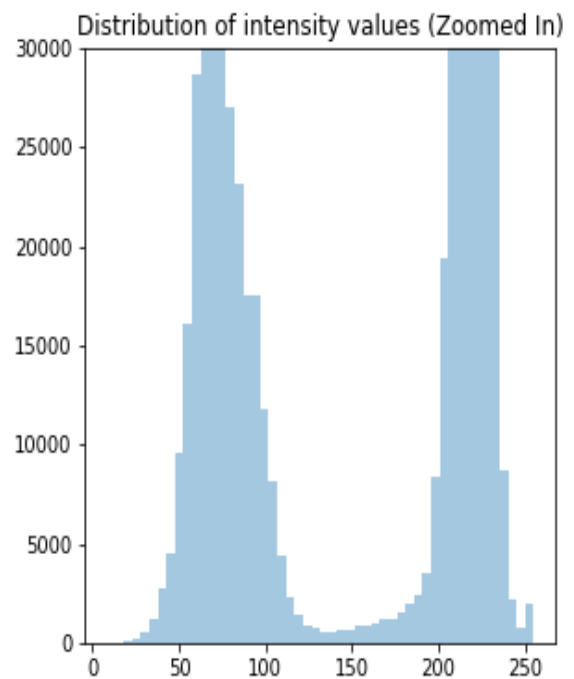


Figure11. Presenting the distribution of Zoomed intensity values



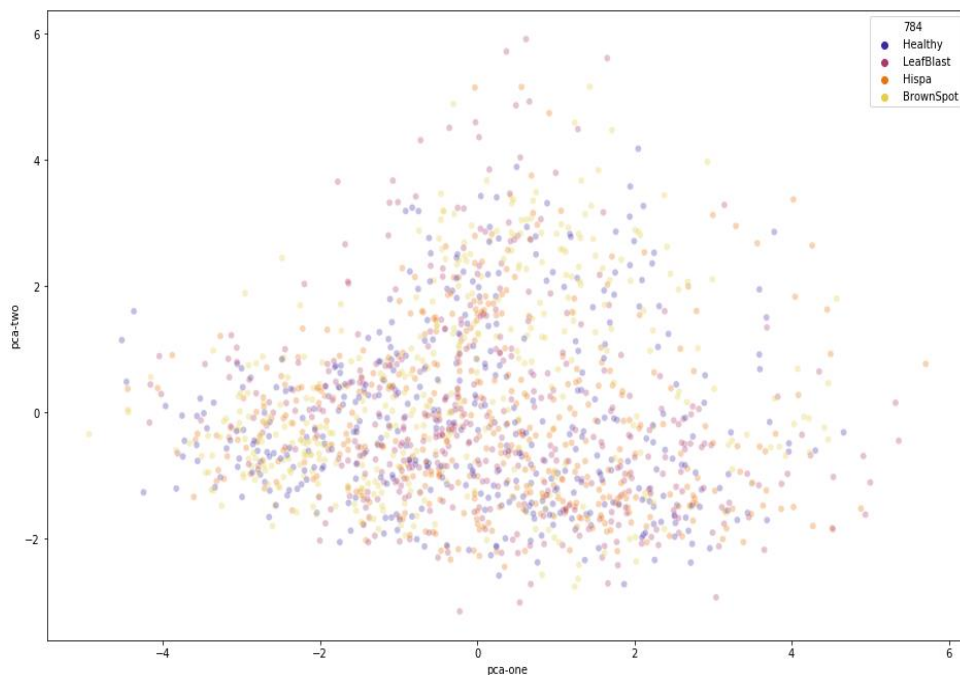
It can be observed that there are 2 prominent peaks. The count of pixels with intensity values around 0 is extremely high (30000). It is expected that the leaf should cover a smaller portion of the picture as compared to the background color which is primarily black. The next thing to do is to separate the two, that is, the leaf from the background. The optimal separation value is somewhere around 20 but rather than relying on such descriptive statistics, we have used a more formal approach known as Otsu's method. Otsu's method assumes that the image contains two classes of pixels following bi-modal histogram (foreground pixels and background pixels), it then calculates the optimum threshold separating the two classes so that their combined spread (intra-class variance) is minimal, or equivalently, so that their inter-class variance is maximal. Otsu's method exhibits relatively good performance if the histogram can be assumed to have bimodal distribution and assumed to possess a deep and sharp valley between two peaks. Next masking has been done. Masking helps in locating the whole leaf from the image.

We have also applied K-means for segmentation of the leaf from the background. The comparison has been made between the labels of Otsu and K-Means at a pixel level, summing over the booleans and dividing them by the total number of pixels in the image. If the result is 1, it means there is no difference at all.

**Feature Extraction:** For a given image of diseased paddy leaves, the feature extraction begins from an initial set of data and builds features, so that the more informative feature dataset can be prepared. Feature extraction is a kind of process, where an initial set of raw variables is processed to more manageable groups.

In the proposed research methodology, we have used HOG (Histogram of oriented gradients) feature extraction method. HOG is a feature descriptor used to detect objects in computer vision and image processing. The HOG descriptor technique counts occurrences of gradient orientation in localized portions of an image - detection window, or region of interest (ROI). It divides the image into small connected regions called cells, and for each cell compute a histogram of gradient directions. It then discretizes each cell into angular bins according to the gradient orientation. Each cell's pixel contributes weighted gradient to its corresponding angular bin. The grouping of cells into a block is the basis for grouping and normalization of histograms. Normalized group of histograms represents the block histogram. The set of these block histograms represents the descriptor. HOG is a proven feature descriptor used to detect objects in computer vision and image processing. Hence, we have used HOG for feature extraction.

**Feature Reduction:** The feature reduction technique known as PCA (Principal component analysis) has been applied to remove redundant features which do not impact the decision making of classifiers. The reduced feature-set has been prepared and used further by the classifiers to identify the diseases of paddy leaves.



**Figure 12.** PCA is presenting extracted features from images of diseased plants, where each paddy leaf image is represented by a distinguished colour.



**Classification using Cascaded Classifiers:** We have then made use of cascaded classifiers (such as AdaBoost, and Bagging classifiers) with modifications in the algorithms for identification of diseases of paddy leaves by using the reduced feature set to achieve better accuracy. Cascading is a particular case of ensemble learning based on the concatenation of several classifiers, using all information collected from the output from a given classifier as additional information for the next classifier in the cascade.

#### **a. Adaboost Algorithm**

Boosting is a general ensemble method that creates a strong classifier from a number of weak classifiers. This is done by building a model from the training data, then creating a second model that attempts to correct the errors from the first model. Models are added until the training set is predicted perfectly or a maximum number of models are added. AdaBoost is a kind of gradient boosting with built-in functionality of cross-validation. It allows the user to run a cross-validation at each iteration of the boosting process and thus it makes the process of getting the exact optimum number of boosting iterations in a single run quite easy. Hence, we are using AdaBoost in our research work for promising results.

The Adaboost algorithm has been applied into three steps:

1. A loss function is being optimized, such as cross entropy for our classification problem.
2. A weak learner is being allowed to make predictions.
3. An additive model has been used to add weak learners to minimize the loss function.
4. Newer weak learners have been added to the proposed machine learning model to correct the residual errors of all the previous trees.

#### **b. Bagging Classifier**

Bagging Classifier is one of the most powerful and popular machine learning algorithms. It is an ensemble algorithm that is used for classification. Bagging stands for bootstrap aggregation and is a powerful statistical method for estimating a quantity from the given dataset. Bagging tries to employ similar learners on small samples of the datasets and then takes a mean of all the results. Bagging uses bootstrap sampling to obtain the data subsets for training the base learners. An ensemble method in Bagging classifier combines the results from multiple machine learning algorithms altogether to make better and accurate predictions than any single algorithm. For aggregating, the outputs of base learners, bagging uses voting for classification. We have made use of bagging with diverse algorithms for voting and we have attempted to improve the output by experimenting with multiple algorithms using bagging concept. The results obtained are shown in the next section.

#### **c. Genetic Algorithm(GA) based Classifier**

The GA based classifier combines evolutionary technique and dimensionality reduction technique by representing the extracted features as a chromosome in GA. The pre-processing includes feature set (e.g., fractal dimension and texture). The training of GA based classifier makes use of a weighted nearest neighbour algorithm (NNA). There is a need to consider an appropriate representation that could be used by a genetic algorithm. In order to consider suitable representation in GA, the chromosome is defined. To distinguish among the rice plant diseases, the relevant weights are assigned to the features during training phase (TP). These weights are used in the validation stage and the weights of the best chromosome are used further to evaluate the performance of the GA based classifier.

The basic framework is as follows:

- 1) Initialize vectors to be considered for NNA.
- 2) Train the GA using a fitness function which is based on the previously generated vectors and the weighted NNA.
- 3) Validate the output of the best individual/chromosome

##### **A. Initialization of NNA**

A portion of the TP is used to initialize the vectors that are used by NNA. During this procedure, weighted distances to a given individual are determined between the image vectors and initialized vectors.

##### **B. Training of GA**

The training phase of the GA needs the problem formulation as an individual or a chromosome, and the fine tuning of evolutionary operators such as mutation, crossover, and elitism. The replacement mutation and alternating-position crossover are used with the rate of 0.03 and 0.8 respectively.

#### 4. Results and Discussions

This section provides insights into the results obtained by using the three machine learning algorithms.

**Assessment of Accuracy:** The results obtained from machine learning based classifiers will be assessed to measure the performance in terms of accuracy of the detected paddy leaf diseases. We will be using Confusion matrix, Area under Curve, F1 score, Precision and Recall matrices to compare the performance of the cascaded algorithms used for classification.

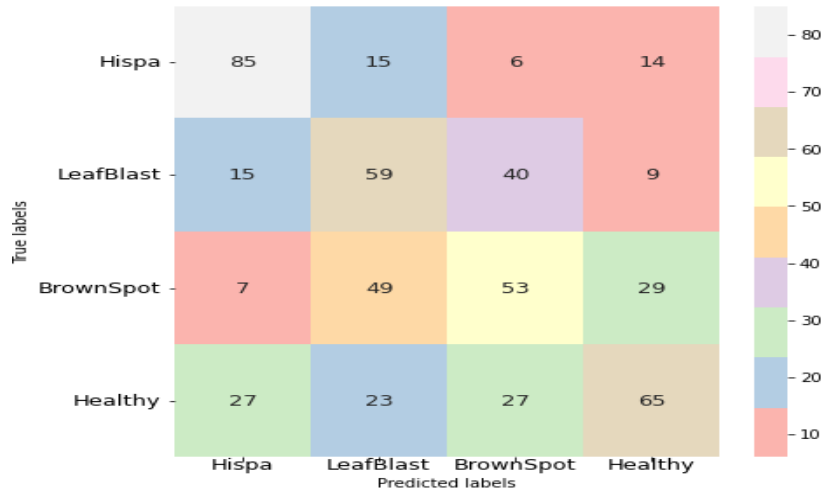


Figure 13. Confusion Matrix by AdaBoost classifier

Following are the observations for the above confusion matrix as shown in Figure 13.

- In the testset, 120 images belong to "Hispa" disease. 85 images are correctly identified and 35 are misidentified.
- In the testset, 123 images belong to "LeafBlast" rice plant disease. 59 images are correctly identified and 64 are misidentified.
- In the testset, 138 images belong to "BrownSpot" rice plant disease. 53 images are correctly identified and 85 images are misidentified.
- In the testset, 142 images belong to "Healthy" rice plants. 65 images are correctly identified and 77 images are misidentified.
- 

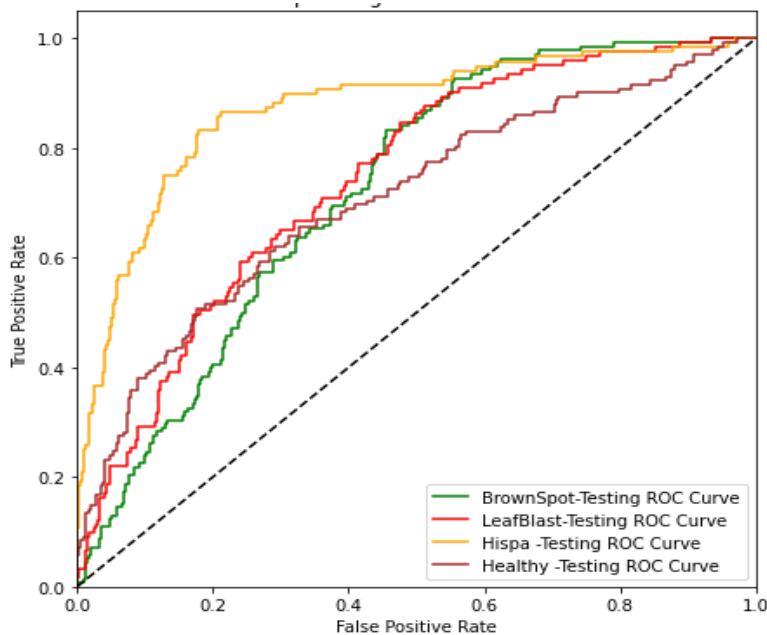
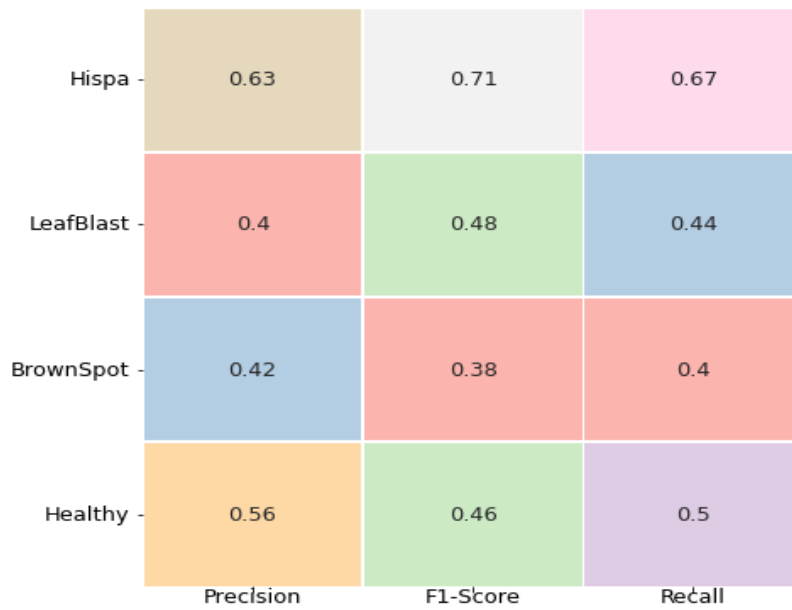


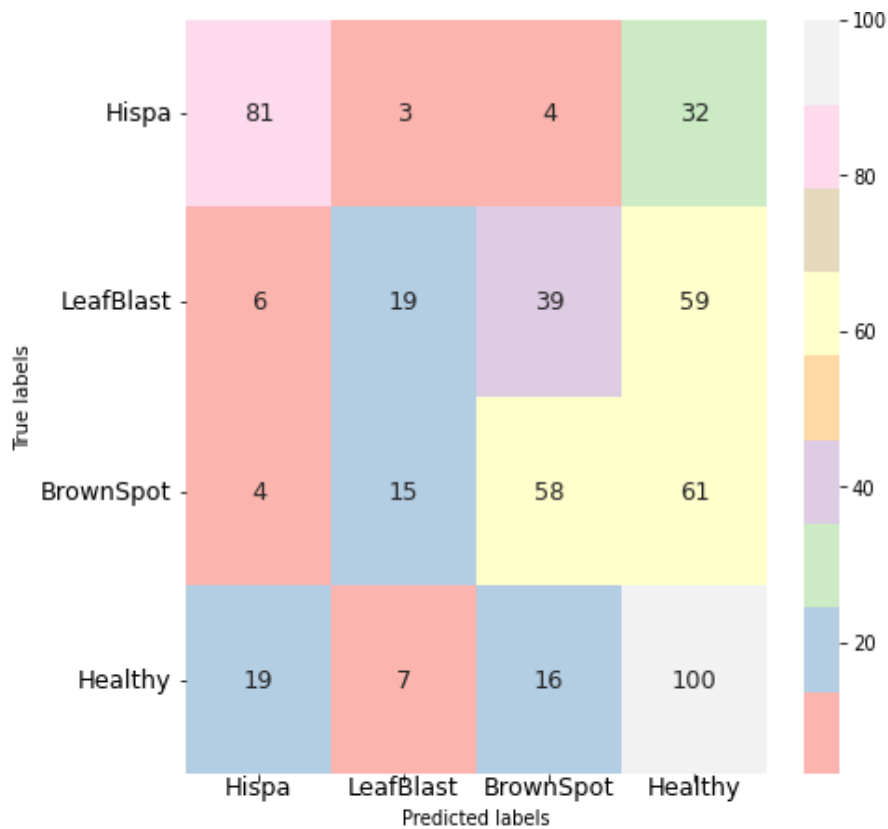
Figure 14. RoC curve obtained from the AdaBoost classifier

**Figure 14.** depicts area under curve or RoC for AdaBoost classifier. ROC curve measures the performance for the classification of paddy leaf diseases s at threshold settings. It plots the True Positive Rate against False Positive Rate at different classification thresholds. AdaBoost is showing good performance for detection of rice plant diseases.



**Figure 15.** Performance evaluation of AdaBoost using Precision, F1-Score and Recall

Figure15. depicts the performance evaluation parameters such as Precision, Recall and F1-Score for AdaBoost classifier and the results can be further improved by fine tuning of parameters in the proposed technique.



**Figure 16.** Confusion Matrix -Bagging classifier

Following are the observations for the above confusion matrix as shown in Figure 16.

- In the testset, 120 images belong to "Hispa" disease. 81 images are correctly identified and 39 are misidentified.
- In the testset, 123 images belong to "LeafBlast" rice plant disease. 19 images are correctly identified and 104 are misidentified.
- In the testset, 138 images belong to "BrownSpot" rice plant disease. 58 images are correctly identified and 80 images are misidentified.
- In the testset, 142 images belong to "Healthy" rice plants. 100 images are correctly identified and 42 images are misidentified.

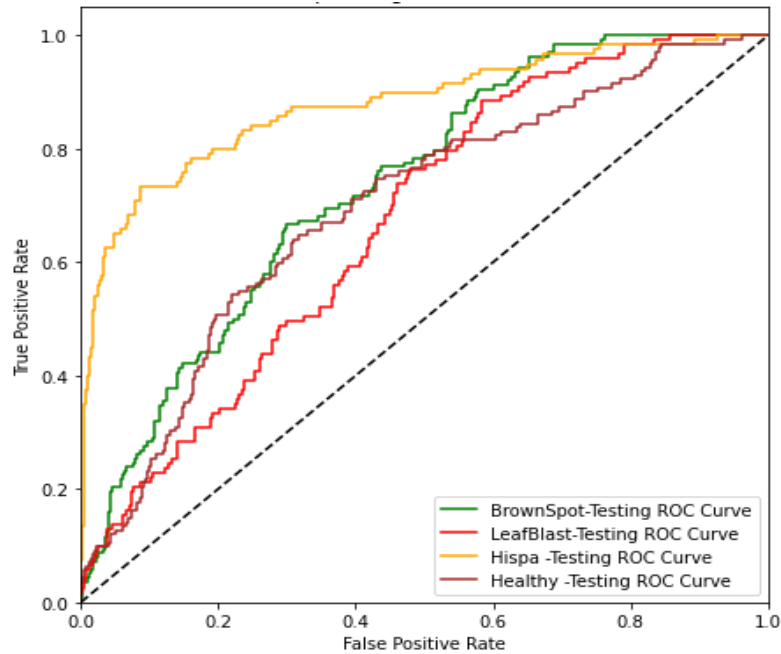


Figure 17. RoC curve obtained from the Bagging classifier

Figure 17. provides AUC for Bagging classifier and shows the aggregate measure of performance across all possible classification thresholds.

Hispa	0.74	0.68	0.7
LeafBlast	0.43	0.15	0.23
BrownSpot	0.5	0.42	0.45
Healthy	0.4	0.7	0.51
	Precision	F1-Score	Recall

Figure 18. Performance evaluation parameters of Bagging classifier

Figure 18 shows precision score, recall score and F1-score with respect to Bagging Classifier.

Hispa	90	10	6	14
Leafblast	10	<b>69</b>	<b>20</b>	<b>24</b>
BrownSpot	15	<b>12</b>	<b>78</b>	<b>33</b>
Healthy	7	<b>7</b>	<b>14</b>	<b>114</b>
	Hispa	Leafblast	BrownSpot	Healthy

**Figure 19.** Confusion matrix for GA based Classifier

Following are the observations for the above confusion matrix as shown in Figure 19.

- In the testset, 120 images belong to "Hispa" disease. 90 images are correctly identified and 30 images are misidentified.
- In the testset, 123 mages belong to "LeafBlast" rice plant disease. 69 images are correctly identified and 54 are misidentified.
- In the testset, 138 images belong to "BrownSpot" rice plant disease. 78 images are correctly identified and 50 images are misidentified.
- In the testset, 142 images belong to "Healthy" rice plants. 114 images are correctly identified and 28 images are misidentified.

Hispa	0.75	0.73	0.75
Leafblast	0.56	<b>0.70</b>	<b>0.58</b>
BrownSpot	0.56	<b>0.66</b>	<b>0.60</b>
Healthy	0.80	<b>0.61</b>	<b>0.69</b>
	Precision	Recall	F1-Score

**Figure 20.** Performance evaluation of GA classifier using Precision, F1-Score and Recall

Figure 20 shows performance of GA with respect to evaluation parameters such as F1-score, precision and recall. The results show that GA based classifier outperform other two classifiers. Next is AdaBoost which performs more accurately than Bagging Classifier and then comes Bagging Classifier. All the classifiers can be

used to prepare an automated classifier for the detection of paddy leaf diseases. None of the discussed methods give 100% accuracy but still these methods can be used for saving the time of farmers for identifying the diseases of paddy leaves. The accuracy attained by GA classifier is 96% at training feature-set and 91% at testing feature-set. The accuracy attained by AdaBoost is 88% at training feature-set and 84% at testing feature-set whereas the accuracy attained by Bagging algorithm is 86% at training feature-set and 81% at testing feature-set.

## 5. Conclusion

An attempt has been made to use machine learning algorithms such as AdaBoost and Bagging Classifier along with evolutionary algorithm such as GA for the identification of the paddy leaf diseases. The images of paddy leaves have been processed to reduce the unwanted distortions, to enhance the image features for further processing. It involves resizing of images, brightness correction, filtering, illumination corrections, noise removal, geometric transformations, and grey scale transformations. After pre-processing of images, feature extraction algorithms are applied, so that the relevant feature dataset can be obtained from the processed images. Next, the cascaded classifiers with modifications with respect to the problem statement are applied to obtain better accuracy. Genetic algorithm has also been tried with NNA for identifying the images of paddy leaf diseases. The proposed automated system can be used on Android and Windows platform for quickly identifying the paddy leaf diseases. It can certainly help the farmers to classify the diseased paddy leaves at an early stage to protect the crops from further damage.

## References

1. Agriculture Sector in India (2015) [Online]. Available: <http://www.ibef.org/industry/agriculture-india.aspx>.
2. Anthonys, G. & Wickramarachchi, N. (2009). An image recognition system for crop disease identification of paddy fields in Sri Lanka, *International Conference on Industrial and Information Systems (ICIIS)*, *IEEE*, 403-407, doi.org/10.1109/iciinfs.2009.5429828
3. Archana, K S. & Arun, S. (2018). Automatic Rice Leaf Disease Segmentation Using Image Processing Techniques, *International Journal of Engineering & Technology*, 7(3.27), 182, doi.org/10.14419/ijet.v7i3.27.17756
4. Basavaraj, S. A, Naveen N. M. & Surendra, P.,(2020) Deep learning approach for recognition and classification of yield affecting paddy crop stresses using field images, *Artificial Intelligence in Agriculture*, 4,1- 20
5. FAOSTAT, Food and Agriculture Organization of the United Nations: Home, [www.fao.org](http://www.fao.org)
6. Geraldin B. Dela, C., (2019). Nitrogen Deficiency Mobile Application for Rice Plant through Image Processing Techniques, *International Journal of Engineering and Advanced Technology*, 8(6), 2950-2955, doi.org/10.35940/ijeat.f8721.088619
7. Gianessi, L. P. (2014). Importance of Pesticides for Growing Rice in South and South East Asia,” pp. 30–33,
8. Iswarya, P. & Maheswari, D. (2019) Paddy Leaf Disease Identification and Classification System A Review *International Journal of Computer Sciences and Engineering*, 7 (5) , 976- 979
9. Jaskaran, S. & Harpreet, K (2018). Plant Disease Detection Based on Region-Based Segmentation and KNN Classifier, *International Conference on ISMAC in Computational Vision and Bio-Engineering (ISMAC-CVB)*.
10. Jitesh, P. S, Harshadkumar, B. P & Vipul, K. D., (2016). A survey on detection and classification of rice plant diseases, *IEEE International Conference on Current Trends in Advanced Computing (ICCTAC)*, doi.org/10.1109/icctac.2016.7567333
11. Jaganmohan, K., Balasubramanian, M. & S.Palanivel(2016) Detection and recognition of diseases from paddy plant leaf images, *International Journal of Computer Applications*, 144(12) , 34-41
12. Khoenkaw, P.,(2016). An image-processing based algorithm for rice seed germination rate evaluation, 2016 *International Computer Science and Engineering Conference (ICSEC)*, Chiang Mai, 1-5.
13. Latte, M V & Sushila, S.(2016). Multiple nutrient deficiency detection in paddy leaf images using color and pattern analysis, *International Conference on Communication and Signal Processing (ICCSP)*,*IEEE*, doi.org/10.1109/iccsp.2016.7754352
14. Papademetriou M.K., (2000). Rice production in the Asia-Pacific Region: Issues and Agriculture Economics and Importance of Agriculture in National Economy (2015) [Online]. Available: <http://agriinfo.in/?page=topic&superid=9 &topicid=185>.
15. Phadikar.S. & Sil, J., (2008). Rice disease identification using pattern recognition techniques, *11th International Conference on Computer and Information Technology*, Khulna, 420-423.

16. Prabira K.S, Nalini K.B, Amiya, K.R& Santi, K.B. (2020).Image Processing Techniques for Diagnosing Rice Plant Disease: A Survey, *Procedia Computer Science*, 167, 516-530,doi.org/10.1016/j.procs.2020.03.308
17. Prabira, K., Baishalee N. & Nilamani, B (2017). Detection of Healthy and Defected Diseased Leaf of Rice Crop using K-Means Clustering Technique, *International Journal of Computer Applications*, 157(1), 24 -27, doi.org/10.5120/ijca2017912601
18. Prakash, R. M., Saraswathy, G. P., Ramalakshmi, G. Mangaleswari K. H. & Kaviya, T. (2017), "Detection of leaf diseases and classification using digital image processing," *International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, pp. 1-4
19. Pugoy R, & Mariano V.Y., (2011). Automated rice leaf disease detection using color image analysis. *3rd international conference on digital image processing*, 8009. Chengdu: SPIE; F1-F7.
20. Qing, Y., Zexin,G., (2009). Application of support vector machine for detecting rice diseases using shape and color texture features," *International Conference on Engineering Computation*, pp. 79-83.
21. Ramesh, S. & Vydeki, D. (2018). Rice Blast Disease Detection and Classification Using Machine Learning Algorithm, *2nd International Conference on Micro-Electronics and Telecommunication Engineering (ICMETE)*, IEEE, doi.org/10.1109/icmete.2018.00063
22. Rice Production (Peace Corps): Chapter 14 - Diseases of rice (2015),[Online]. Available: <http://www.nzdl.org>.
23. Santanu P. & Jaya S. (2008). Rice disease identification using pattern recognition techniques, *11th International Conference on Computer and Information Technology, IEEE*, doi.org/10.1109/iccitechn.2008.4803079
24. Shampa, S. &Asit K. D (2017). Particle Swarm Optimization based incremental classifier design for rice disease prediction, *Computers and Electronics in Agriculture*, 140, 443 – 451, doi.org/10.1016/j.compag.2017.06.024
25. Usha, K., Kanagasuba,R. S, Jaichandran R &Priyadharshini C (2019) Detection and Classification of Paddy Crop Disease using Deep Learning Techniques, *International Journal of Recent Technology and Engineering* ,8(3), 4353 -4359
26. Verma, P., (2017). Rice Productivity and Food Security in India, Centre for Management in Agriculture, Indian Institute of Management Ahmedabad.
27. Yang, J., Gong, W. & Shi, S. (2016). Accurate identification of nitrogen fertilizer application of paddy rice using laser-induced fluorescence combined with support vector machine, *Plant, Soil and Environment*, 61(11), 501-506,doi.org/10.17221/496/2015-pse
28. Ying, Z., Fang, L., Qiong ,W., Weicong He, Bin Yuan &Wenya Yuan. (2020). Identification of Novel Alleles of the Rice Blast-Resistance Gene Pi9 through Sequence-Based Allele Mining, *Rice, Springer*, 13(1), doi.org/10.1186/s12284-020-00442-z