RICE PLANT (ORYZA SATIVA) DISEASE CLASSIFICATION USING MACHINE LEARNING ALGORITHMS

J. Arockia Jackuline Joni[1]*, Dr. M. Mary Shanthi Rani[2]

Abstract
To prevent reductions in agricultural product production and quantity, it is crucial to identify plant diseases. The application of various machine learning and image processing techniques reduces the issues in the agricultural industry. With the aid of various ML and image processing approaches, the major objective of this review is to identify rice plant diseases using picture inputs of Infectious rice plants. Moreover, the key Machine Learning (ML) and image processing concepts for identifying and categorizing plant diseases are covered. k-Nearest Neighbor Classifier (KNN), Support Vector Machine (SVM), Genetic Algorithms (GA), and Probabilistic Neural Networks (PNN) are a few of the classification methods utilized in agricultural research. The quality of a conclusion can vary depending on the input data, so choosing a categorization method is an important responsibility. The categories for plant leaf diseases are used in a variety of sectors, including biology, agriculture, etc.
This research presents a thorough analysis of rice plant illnesses, picture dataset size, pre-processing, segmentation methods, and classifiers.

Keywords: K-Nearest Neighbor Classifier (KNN), Support Vector Machine (SVM), Genetic Algorithms (GA), Probabilistic Neural Networks (PNN), Oryza, Discrete Wavelet Transform (DWT), Blast Disease (BD), Brown Spot Disease (BSD), and Narrow Brown Spot Disease (NBSD)

[1]*M.Sc., M.Phil.,(Ph.D.,) Assistant Professor, Department Of Computer Applications, Fatima College, Mary Land, Madurai. Email: Arojoni12@Gmail.Com
[2]Research Supervisor, Associate Professor, Department of Computer Science & Applications, The Gandhigram Rural Institute (Deemed to Be University), Dindigul. Email: Drmaryshanthi@Gmail.Com

***Correspondence Author:** J. Arockia Jackuline Joni
*M.Sc., M.Phil.,(Ph.D.,) Assistant Professor, Department Of Computer Applications, Fatima College, Mary Land, Madurai. Email: Arojoni12@Gmail.Com

DOI: - 10.48047/ecb/2023.12.si5a.0150

I. INTRODUCTION

In many nations around the world, agriculture is the main industry for generating revenue. Based on the importance of agriculture, farmers choose their crops, paddies, and relevant pesticides to promote plant growth in the constrained amount of time [1]. In many nations, rice is the main crop grown for food [2]. The agriculture industry is currently dealing with serious issues with rice plants as a result of illnesses that reduce the quality and output of the harvests. The varied causes of the reduced production rate, including insufficient expert availability on the farm, ignorance of fertiliser management, and ignorance of illnesses and pests, are listed below[3].

Plant diseases both directly and indirectly cause some environmental harms. As these diseases spread around the globe, they harm a plant's ability to function as a whole and also harm the economy by drastically reducing the amount of crops grown [4].

Many bacterial and fungal diseases harm plants [5]. The several diseases that affect rice plants are Sheath blight, NBSD Leaf blast, and Brown spot [6]. The destruction left behind by the Blast demonstrates how serious the illness is. Bypolaris Oryza (a fungus) is responsible for another prominent rice disease known as brown spot, which is present throughout the growing season. When rice is produced on soils lacking in silicon, brown patches typically get worse [7]. In Asia, 10 to 15 percent of the yield is lost to rice illnesses [8]. An specialist in agriculture has recently supported manually the analysis and monitoring of plant diseases, which takes more work and processing time [9]. Farmers can have trouble identifying the infections, which results in crop loss. One efficient method for farmers is to use an automated system to process photos of "seem to seem" infected leaves.

To identify illnesses, one looks at the plant leaf that displays the symptoms. The automation technology enables farmers to quickly detect infections. The failure of certain plants to yield was caused by the slow disease detection. Early illness detection is therefore crucial.

A number of ML and image processing techniques have been developed to identify diseases in rice plants. Accuracy in the diagnosis of plant diseases using ML algorithms depends on three procedures: feature segmentation, feature extraction, and classification method. Deep learning techniques have produced favorable results for picture classification. The illnesses of the mango [11], apple [12], tomato [13], rice [14], and wheat [15] have recently been the focus of research [15]. In numerous instances, they have used fruits or leaves to identify diseases based on photographs. In other situations, they have also used photographs with uniform backgrounds.

1.1 GENERALSTRUCTURE

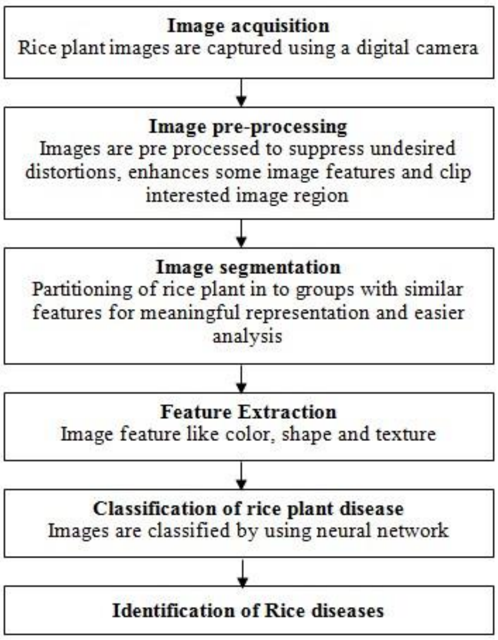


Figure 1: General architecture for rice disease identification

The paddy fields are seen in the pictures. The diseased leaf segments in those photos are then segmented during pre-processing. The features are then extracted from the segmented images, and finally, the diseases are classified using ML algorithms. The effectiveness of such a system depends on how well it executes the ML and image processing tasks [16]. The following processes are described in relation to the classification and detection of illnesses affecting rice plants: 1) Picture acquisition with a digital camera to record the illnesses of rice plants, 2) Pre-processing improves the features of the taken image, which improves the data for further processing, The preprocessed photos are divided into groups with similar properties for simpler analysis in step three. Step four, feature extraction, extracts the features from the segmented based on forms, colours, and textures, and Step five, classification, categories various rice plant diseases. Figure 1 illustrates the procedures used to spot the rice illness.

II. LITERATURE REVIEW

This section deals with several studies on rice disease detection. Table lists the main procedures used to identify illnesses of rice plants.

Table 1: Basic image processing steps in the detection of rice plant diseases

Author Name	Key Technique	Description	Limitations
Santanu Phadikar *et al* [17]	Segmentation	To spot the infected leaf part of the rice plant images, a region identification methodology centred on Fermi energy was recommended. NGLDM (Neighboring gray-leveled dependence matrix) centered textural features were extorted to categorize the various diseases of rice plants.	Segmentation accuracy is low.
C.Kumar Charliepaul [18]	Feature Extraction	Symptoms were characterized using features like, color and shape of the infected part of the rice plant and the extorted feature was utilized for recognizing the diseases.	The risk factors in the recognized diseases were not found
Amit Kumar Singh *et al* [19]	Classification	A methodology to recognize the utmost occurring disease in rice plant say, Rice Leaf blast (RLB) utilizing SVM classifier was propounded.	Segmentation process was not ameliorated. Employed the conventional framework for segmentation.

In order to prevent the effects of rice plant diseases, Manoj Mukherjee et al. [20] established a framework for processing the images of paddy leaves via the histogram. By using this system, one might identify diseases early on and take prompt action to reduce productivity loss. The leaf's image was first captured, after which it was processed. The image was afterwards converted from RGB to grayscale, and then a histogram was extracted using MATLAB methods. The obtained pictures were provided as the data for classifying and rating the illnesses. A consultation treatment unit for the disease was developed once the disease and its stage had been identified with the help of agricultural experts.

In order to provide a user interface for the technologically illiterate users, in particular farmers, Namita Mittal et al. [21] proposed an icon-centric information retrieval architecture. Internet. In addition, it made it easier for farmers to diagnose crop diseases instantaneously using pattern recognition and digital image processing rather than having to wait for specialists to visit the farms and diagnose the diseases. Results from experiments demonstrated this methodology's efficacy. The image processing part demonstrated the results based on the training of 25 photos for each disease category. By increasing the quantity of photos used to train the system, the results could be much better.

Using image processing techniques, Gayathri and Neelamegam [22] proposed a framework to automatically detect leaf diseases. According to the methods, pre-processing, picture acquisition, segmentation, and paddy leaf disease classification were carried out in order to identify those diseases. The features are extracted using the hybridised DWT (Discrete Wavelet Transform), GLCM (Gray Level Co-occurrence Matrix), and scale-invariant feature transform techniques. Finally, to distinguish between healthy and diseased plants, the extracted features were fed into a variety of classifiers, including K-Nearest Neighbors (KNN), neural networks (NN), backpropagation, multiclass SVM, and Naive Bayesian. In order to classify leaf disease, many classification techniques were looked at. The results showed that the multi-class SVM performed better than other classifiers, with an accuracy of 98.63%.

For classifying the four types of paddy diseases, including leaf streak, bacterial leaf blight (BLB), brown patches on leaves, and RLB diseases caused by fungi and bacteria, Khaing and Chit Su [23] suggested an automated system. This system included feature extortion, feature extraction, leaf checking, image collection, and pre-processing. Standard deviation, mean, contrast, energy, entropy, and correlation of the image's colour, statistical, and textural aspects were extracted using PCA (Principal Component Analysis), GLCM, and Color Grid-based Moment. Finally, SVM was used to classify it based on the disorders. With the original grayscale conversion and the updated grayscale conversion, this methodology achieved performance with 72.70% and 90%, respectively.

III. OUTLINE OF THE RICE PLANT DISEASE

A. Paddy

Many elements have a negative impact on paddy output. Disease is one of the key causes. The different varieties are listed in the table below along with their corresponding symptoms, subsequent treatment, and post-detection.

B. Current Techniques

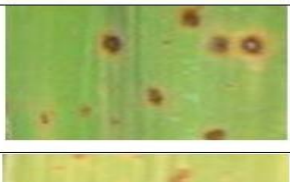
An image processing technique is being used in a study by Kurniawati et al. to create a prototype system to automatically and accurately detect and classify the paddy diseases with Blast Disease (BD), Brown Spot Disease (BSD), and Narrow Brown Spot Disease (NBSD).

The procedure entails a number of steps, including picture collection and acquisition, segmentation through pre-processing, shape feature and colour feature extraction, paddy disease classification review with an eye towards lesion type, boundary colour, spot colour, and broken paddy leaf colour [1] [2].

Grape Leaf Disease Identification via Color Imagery Using Hybrid System Intelligence by A. Meunkaewjinda et al. The procedure demonstrates segmentations based on grape leaf colour, disease segmentation, analysis, and disease pattern categorization. The segmentation attribute for grape leaves emphasises the pre-processing module that separates out the background data that isn't relevant. Self-organizing feature modification is used in the segmentation process. Genetic algorithms were then used to attach and map the features during the engagement optimization process. A support vector machine method is run from up to down turnovers for classification after being equipped with an optimization process. A subsequent Gabor wavelet platform analysis removes the segmented picture [3][4].

Using a framework for the detection and classification of plant leaf and stem diseases, Dheeb Al Bashish et al. conducted a study [5]. The study demonstrates that it would be extremely expensive for developing nations to rely just on the observation of the naked eye to detect such diseases. It might be extremely realistically significant to provide quick, automatic, affordable, and accurate image processing-based solutions for this task. The suggested framework makes use of image processing and is made up of a segmentation procedure that makes use of the K-Means technique before going through a pre-trained neural network strategy [6].

Table1. Different Paddy Diseases

	Paddy Blast Disease	Disease can infect paddy at all growth stages and all aerial parts of plant (Leaf, neck and node).
	Paddy Brown Spot Disease	Initial lesions are water -soaked to greenish gray and later become grayish white with brown margin
	Narrow Brown Spot Disease	Short, narrow, elliptical to linear brown lesions usually on leaf blades, sheaths, pedicels, and glumes or rice hulls.

IV. DISEASE DETECTION IN RICE PLANTS USING MACHINE LEARNING TECHNIQUE

A convolutional NN-based RLB recognition algorithm was suggested by Wan-jie Liang [24]. (CNNs). For testing and training the CNN, a data set with 2902 negative and 2906 positive samples was taken into account. To implement this methodology, various rice disease identification systems were created. Automatic rice disease diagnosis was accomplished by fusing identity prototypes with rice disease domain knowledge. In addition, comparisons between quantitative and qualitative analyses were done. The results of this automated diagnosis methodology were satisfactory. These results demonstrated that high-level features extracted by Convolution Neural Network are much more effective and discriminative than the conventional handmade features including Haar-WT (Wavelet Transform) and LBP histogram (local-binary pattern). Also, quantitative evaluation results showed that CNN with SVM and CNN with Softmax both performed and were accurate in the same ways. As a result, CNN was a top-performing methodology for RLB detection and could be used in real-world settings.

A way to improve the convolution neural network deep learning ability was proposed by Yang Lu and Shujuan Yi [25]. 500 photos of healthy and diseased leaves and stems were taken from a rice farming area and are included in the collection. Ten prevalent illnesses of rice plants were taught to the convolution neural network by picture recognition. Cross validation is done across 10 times, and the CNN centric approach outperforms the other frameworks in terms of accuracy, convergence rate, and recognition competency.

Eusebio L. Mique, Jr. [26] examined a programme that helped rice farmers identify pests and diseases by using image processing and CNN. It examined the prevalence of pests.

Controlling, educating farmers about various rice pests and illnesses, identifying the types of pests that attack rice fields, and providing information on how they might be controlled and managed. An application for identifying rice pests and even illnesses was proposed using image processing and CNN. This model's primary goal is to achieve a training accuracy of 90.90%. A lower cross-entropy value suggests that the trained design could make predictions earlier or classify images more accurately. With the aid of this cutting-edge application, farmers can simply control and manage rice pests.

V. SEGMENTATION METHODS FOR DETECTION OF RICE PLANT DISEASES

Table 2 Enumerates the segmentation methodologies used in rice disease diagnosis.

Table2: Comparison of various segmentation methods in recognition of rice diseases

Author Name	Algorithm Used	Merits	Demerits
Archana and Arun Sahayadhas [27]	K-meansclustering (KMC)algorithm	Automatically recognized the disease and separated the leaf color, illumination and sign color as of the diverse color channel.	The pigment of the plant was not properly extracted and types of diseases were not differentiated.
Libo Liu [28]	OTSU	Segmentation efficiency and accuracy was high	Isolated noises and certain holes that were existent in the image after segmentation
Xiong Xiong *et al* [29]	Panicle SEG	Attained better segmentation accuracy execution speed was augmented	Could not be applied in disparate field environments, indoor rice images, and disparate rice accessions

VII. CLASSIFICATION METHODS FOR DISEASES OF RICE PLANTS

Applications for machine learning include the categorization of distributed denial of service (DDOS) attacks [51], criminal networks [52], Apriori based probability tree classification [53], general disease prediction [54][55], facial recognition [56], and forgery detection [57][58], among others. In order to effectively detect and identify the three common microorganisms that caused diseases in the Philippines' rice fields—(a) Xanthomonasoryzae, (b) Thanatephoruscucumeris, and (c) Magnaportheoryzae—John William Orillo et al. [30] examined a framework of using sound signal processing scheme. Using an electret condenser microphone and a simulated anechoic chamber, sound signals from samples of rice leaves damaged by the aforementioned bacteria were captured. To remove noise, spectral subtractions were used to analyse the sound signals. Every input's features were demanded by the Mel Frequency Cepstral Coefficients for the ANFIS learning method. On the basis of 450 sound data (recorded), the fuzzy NN was used to train the system, with 80% of the data being used for training and 20% for testing. In order to prevent additional infection, a tool was also built that would generate a report in PDF format demonstrating the diagnosis and treatment approaches for the contaminated sample.

A prototype for identifying and categorizing rice illnesses based on images of sick rice plants was studied by Harshad kumar et al. [31]. The three diseases affecting rice plants are BLB, Leaf smut, and Brown spot. The aforementioned images of rice farmland were taken with a digital camera. Three segmentation techniques and four back ground removal strategies were evaluated. Centroid feeding centric KMC was advised for segmenting the sick area as of a leaf image in order to permit correct feature exploitation. Eliminating the green pixels in the sick area improved KMC's output. On the basis of colour, texture, and shape, several features were demanded. In SVM, a technique for multi-class classification. On a test dataset, this approach had an accuracy of 73.33%, and on a training dataset, it had an accuracy of 93.33%. The 10 and 5-folded cross validations were also used, as a last step.

To find rice plant illnesses, Santanu Phadikar et al. [32] looked at a classification approach with an automated scheme. It sought to categorise various rice diseases by coercing features from the area that was diseased in the photos under consideration. A technique known as fermi energy-centric segmentation was suggested to separate the disease infected area of the image from its background. Using variables like colour, position, and shape, the symptoms of the diseased area were identified and then put into practise by algorithms. Notable characteristics are selected using Rough Set Theory to reduce classifier complexity and information loss (RST).

H. Al-Hiary et alautomated .'s software-based recognition and classification of plant leaf diseases [33] was proposed, and it provided a quicker and more accurate solution. The four main phases of this processing framework are as follows. After segmentation, the next two phases were introduced. The primary stage identified primarily green-colored pixels. Next, based on predetermined threshold values that were assessed using the Otsu approach, those green pixels were muted. The other additional stage was completely removing all of the pixels with "0" RGB values as well as those on the object's exterior areas (the infected cluster). The practical results showed that this method was effective for identifying illnesses in plant leaves. The created algorithms effectively could identify and classify the diagnosed diseases with accuracy ranging from 83% to 94% and could achieve a 20% speedup.

VII. TECHNIQUES FOR IDENTIFYING RICE'S MAIN DISEASES

Using an image processing framework, M.N. Abu Bakar et al. [34] established a comprehensive methodology for identifying the illnesses on leaves known as RLB. It consists of three steps: preprocessing, image segmentation, and image analysis using the Hue Saturation Value (HSV) colour space.

Image segmentation, the most important task in image processing, was used to isolate the illness region, and pattern recognition built on the Multi Level Thresholding methodology was implement-ted. As a result, three categories—infection stage, worst stage, and spreading stage—were created to describe the severity of the RLB disease.

A strategy for identifying sheath blight was suggested by Faranak Ghobadifar et al. [35] using SPOT pictures as the primary data. This approach to detecting the pest infestation by remote sensing (RS) would reduce the cost of food production, limit environmental risks, and improve natural pest control before this problem grows. Analysis of precision farming procedures was done using ENVI4.8 and SPSS software (Environment for Visualizing

Images). As a result, the photographs' early and late growing seasons were different. Certain image indices, such as RVI14, SDI24, and SDI14, supported a sound method for differentiating between healthy and unhealthy plants. These indexes were used to develop strategies for identifying the sheath blight disease using RS.

The inclusion of Basmati cultivars was suggested by Shahzad Amir Naveed et al. [36] to identify the Bacterial Blight in "Pakistani rice germplasm". Several types of seeds are collected from various research institutions and then sown in pots. Using a DNA marker connected to the xa5 gene, DNAs were extracted and checked for polymorphism. 45 lines out of 88 germplasm lines at the time of the polymorphism survey demonstrated the presence of xa5 genes such MB66, MB33, MB2, and MB57. In comparison to the parent line, all of the afore mentioned lines displayed 240bp amplification, but 43 germplasm lines did not. Ten Pakistani basmati cultivars were also assessed for xa5 gene recognition, although this failed to reveal the presence of a specific gene.

Takashi Kobayashi [37] evaluated the use of aerial hyper-spectral photography to determine how serious a disease like panicle blast is in field crops. In northern Japan, the band ratio between low P and high R2 is at its most constant (R498 to 515) to provide hyper-spectral RS imaging at the dough phase of rice grain growth (R700 to 717). Due to percentage-based visual estimations of illnesses, there was a considerable increase in this band ratio (R2=0.83).

A triplex PCR framework was proposed by Eun-Sung Song [38] for the efficient race-associated detection of a BB pathogen, such as the rice disease Xanthomonas oryzaepv Oryzae (XOO). To do this, a model based on "2 genes" was created (XorII and hpaA). A genomic locus and a shorter-patch-repair endonuclease are used in place of AFLP to perform the K5 and K3 races (Amplified Fragment Length Polymorphism). Without any pre-processing, amplicons were obtained using an assay to detect the pathogen in solution, which was also used to create the template (for example, by preparing infected leaf DNA and isolating lesion bacterial cells).

Thirty-five wild rice plants were screened as part of an experiment conducted by Anil Kumar et al. [39] during the 2013 monsoon season to identify the presence of BB resistance genes and accessions against the X00's BX043 dye. A progress curve with a "area under disease" that generates resistant controls was consequently created. After measuring the severity of the diseases, it was determined that none of the accessions were resistant. Eleven accessions had moderate levels of resistance, twenty one displayed little harm, and three displayed harmful responses. Five resistance genes also produced a variable genetic frequency ranging from 00.000 to 45.710%. The NKSWR-25 accession's resistance genes are xa5, xa2, and xa4, whereas the NKSWR16, NKSWR32, NKSWR36, NKSWR41, NKSWR42, NKSWR53, NKSWR64, NKSWR97, and NKSWR99 accessions each contained two resistance genes (from xa5, Xa2, and Xa4). Hence, these accessions were used to introduce particular BB resistance genes into widely used, high-yield rice cultivars.

VIII PROPOSED METHODOLOGY

A. Planning:

Image processing is increasingly needed in a variety of application areas, including the ones mentioned above. The suggested methodology includes paddy disease pre-processing, segmentation, and disease classification. The technique that is applied for the classification of paddy disease in our system process can be categorised as the feed- forward neural network technique. Below is a flowchart depicting in-depth function fundamentally of the proposed method.

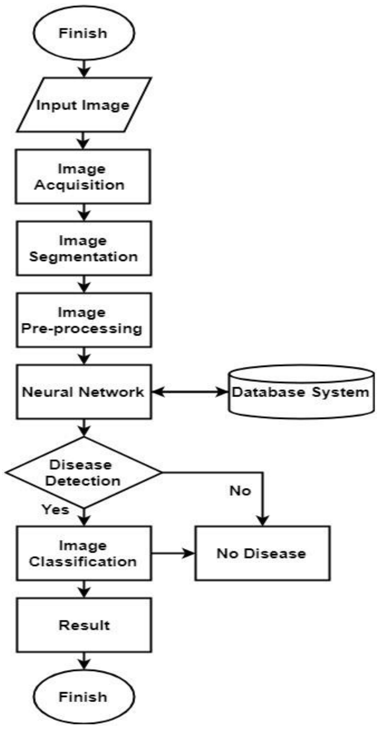


Fig.1: Depiction of system flow

B. Image Acquisition:

Firstly, we have to prepare the paddy images themselves. The RGB colour images of paddy leaf are captured using Olympus fe-4050 digital camera with pixel resolution \s2048x1024. Those images are cropped into smaller image and stored in BMP format \sC. Image Segmentation

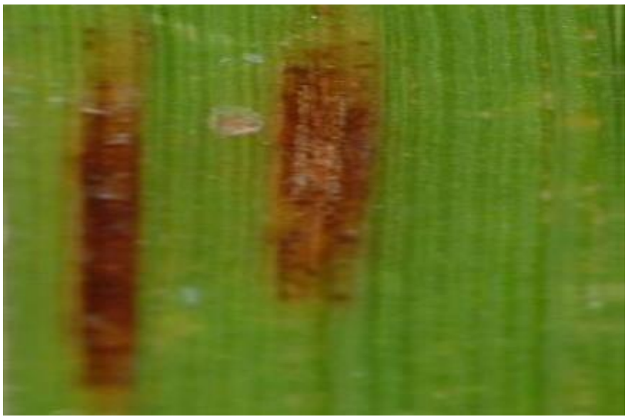


Fig.2 Narrow Brown Spot in Paddy Leaf

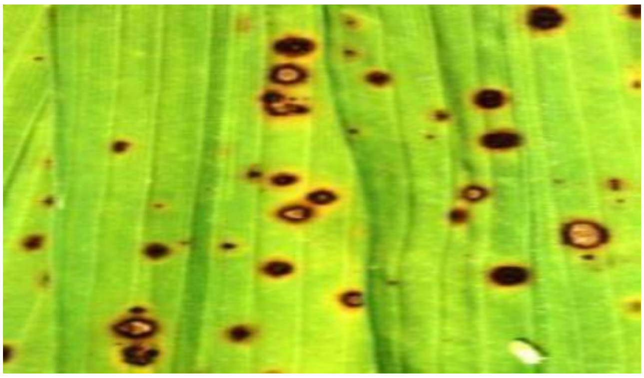


Fig.3 Brown Spot in Paddy Leaf

C. Image segmentation

Primarily, the RGB image has to be converted to a binary image for segmentation. In case of plants, vein of the leaves is different considering intensity and disease spot colour. If segmentation is applied through grayscale image mechanism, vein will present as a binary image pattern engaged with the diseased spot. The interested region to be considered is only disease spots, not the vein. To minimise the effect of the presence of vein, RGB image needed to be colour transformed before segmentation. First, the extra noise in the image needs to be removed. An average filter mechanism is employed for this purpose. Within a rectangular filter window around each pixel, the average filter calculates the mean (average) of the gray-scale values. This results in the image being smoothed (eliminating noise). The filtered pixel's value is determined by:

$$r = (a_1 + a_2 + \cdots + a_9)/9 \quad (1)$$

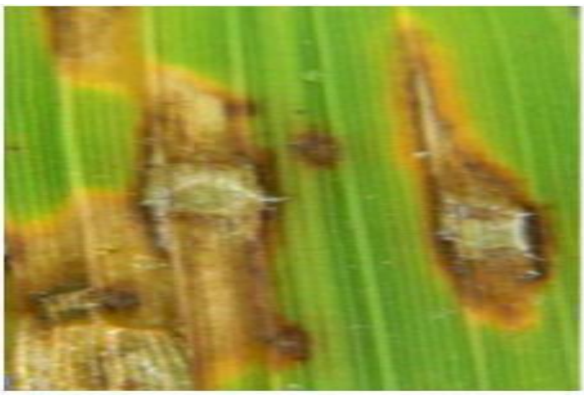


Fig.4 Spot Area in Paddy Leaf

Fig.5 Image Segmentation Generated From Paddy Leaf

We can transform the image into binary form after the noise has been removed. The Otsu method [7], which automatically executes clustering- based picture thresholding by finding the lowest point. between two classes of the histogram, taking the between- class variation into consideration, is used to carry out the operation. The formula for Otsu's approach is as follows:

$$\sigma_w^2(t) = q_1(t)\sigma_1^2(t) + q_2(t)\sigma_2^2(t)$$

Table 3 provides an overview of several image where the class x mean is denoted by. As demonstrated below, the resulting image can then be used to distinguish the sick areas of the leaf:

Author Name	**Technique Used**	**Disease Identified**	**Accuracy**	**Merits**	**Demerits**
Mohd adzhar abdul kahar *et al* [10]	Neuro-Fuzzy expert system	LBD, BSD, BLB	74.21%.	Recognized the diseases at their early stages.	Issues in tackling the noises and other lighting problems due to external forces.
Khaing War Htun, and Chit Su Htwe [23]	PCA, Color Grid-centered Moment and GLCM for feature extortion and SVM for classification	Leaf blast, leaf streak, BLB, leaf brown spot	90%	Attained Highest accuracy	This methodology was not applicable for categorization of crop diseases
Chowdhury Rafeed Rahman *et al* [42]	Deep CNN- centered classification	Pests and diseases in rice plants were recognized	95%	Accurately and timely detect the diseases	Deep learning methodology contained several layers for classification. So it took more time to spot the diseases contrasted with others
M. Akila and P. Deepan [43]	R-FNN, R- CNN, SSD	Diseases and pests of various plants were identified	88%	Ability to compete with complex scenarios and effectively identifies disparate diseases.	Time Consuming
Suman T1, Dhruvakumar T2 [44]	SVM classifier	Rice blast diseases, narrow brown spot, BLB, brown spot,	70%	Efficiently classified 4 kinds of diseases in rice	Lowest accuracy when contrasted with others
Fahrul Agus *et al* [45]	ESforRPD2 application, Unified Modelling Language and Waterfall Paradigm	8 sorts of diseases and 48 symptoms of the rice plants were recognized	87.5%	Showed Good Reliability	Performance of this method was low compared with other expert systems

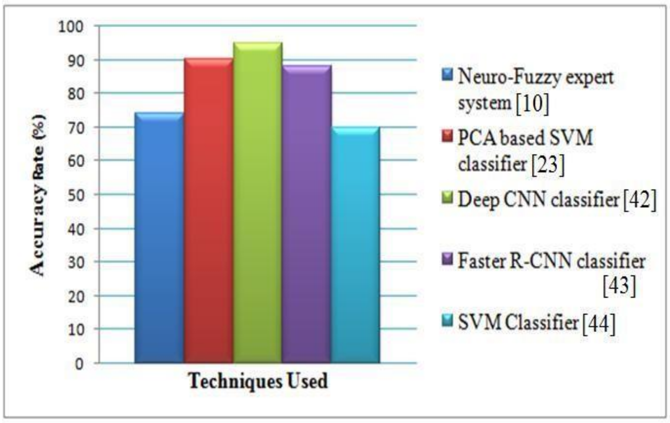
D. Classification

The neural network is used for the processes of pattern recognition, system identification, and system control, where information only moves in one direction, with some of it passing through input nodes and towards hidden nodes and the remainder being sent in the opposite direction to output nodes that need to be redirected. As training pictures for the network, 10 samples each of images with Blast disease, Brown spot disease, and narrow-brown spot disease are employed. The photos will be classified as either having paddy blast, brown spot disease, or narrow brown spot illness based on the training and testing procedure of the feed forward neural network described above [8].

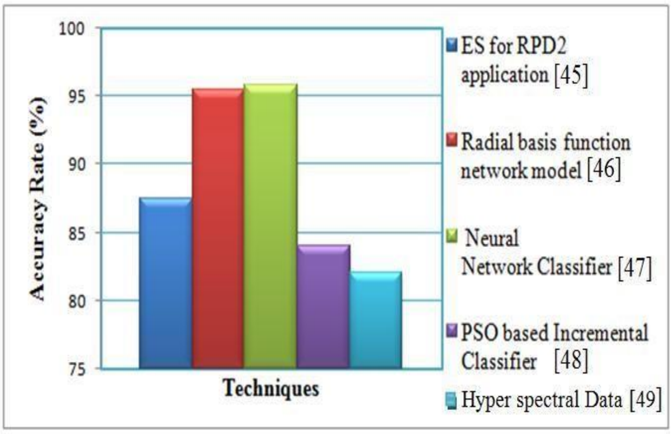
IX COMPARATIVE ANALYSIS

Table 3 below compares and outlines various ML and image processing techniques used in the identification and classification of rice illnesses.
Table 3 compares various methods for diagnosing rice illnesses.

processing and machine learning (ML) approaches used in the recognition of rice illnesses. Various classifiers are used to identify different types of rice illnesses such rice blast, leaf brown spot, sheath rot, and BB. Every technique has a benefit. The effectiveness of several classification approaches and other image processing methodologies used in various studies is compared in terms of accuracy. Using figure 2, the performance comparison graph of the various methods examined in Table 3 is explained in depth.



(a)



(b)

Figure 2: Accuracy Level comparison of different Techniques(a) Neuro-Fuzzy expert system, PCA based SVM, Deep CNN, Faster R-CNN, and SVM (b)

ESforRPD2,Radial basis function, NN classifier, PSO based incremental classifier

The accuracy level of various classifiers and other types of approaches employed in the detection of rice disease is explained in Figure 2. In order to identify different types of rice diseases, a number of ML techniques were used. When compared to other methods, the NN-based classifiers produced better results. In contrast to the RBFN model used in [46], which reached 95.5% accuracy level for disease identification, the NN classifier used in [47] achieved the highest degree of accuracy (95.8%). Additionally, when compared to all other methodologies, the SVM classifier used in [44] had the lowest accuracy rate.

X. IMPLEMENTATION

A. ACQUISITION AND PRE-PROCESSING

Using a CCD colour camera (Nikon D80, zoom lens 18-200mm, F3.5-5.6, and 10.2 Megapixels) in the paddy field of the China National Rice Research Institute, located in Fuyang, Zhejiang, China, photographs of three rice illnesses (RBLB, RSB, and RB) were captured (N, E). Images of each disease are taken from 72 samples. The JPEG format used to store these photographs is uncompressed. The resolution of all images was decreased to 800 by 600 pixels in order to ease computing effort and standardise image resolution. Since the photographs were taken outside, namely in a rice field, noise from insects, dust, and dewdrops was unavoidable. The aforementioned average or median filter was applied in order to lessen their impact during image analysis.

B. SEGMENTATION

We created an algorithm that examines the spot colour and delineates the segmented disease spot region from rice leaf in order to acquire the integrated rice disease spots. Here is how the algorithm was explained: The photos are first the following when converted from a red, green, and blue (RGB) colour representation to a y1 and y2 representation:

$$\begin{aligned} y_1 &= 2g - r - b \\ y_2 &= 2r - g - b \end{aligned} \qquad (3)$$

The maximum class square method of error is related using the transformation that was previously specified. Since the Otsu technique [7] was employed to separate disease spots from rice leaves where two threshold values and are calculated automatically. Initial research revealed that Otsu's approach can erase the edge of some illness areas. To determine the illness category, it is crucial to determine the disease spot's edge. The threshold values needed to be changed in order to maintain the edge of the illness spot during segmentation. The following guideline was followed in doing this:

$$p_s = \begin{cases} (0,0,0), & if\ y_1 > T_1 + 10\ and\ y_2 < T_2 - 10 \\ p_r, & if\ y_1 > T_1 + 10\ and\ y_2 > T_2 - 10 \end{cases}$$

where, is the original image's pixel, and the disease affected patch region will be identified by ( ) By comparing the noise region with the disease spot area, any remaining noisy parts in the image can be eliminated. Finally, compressing horizontal, vertical, and diagonal segments completes the area's edge detection [9]. Following that, using the edge of the disease spot as a guide, we may acquire the binary picture of the disease spot. A duplicate of the original image can be used to determine the pixels of illness spots using the following formula:

$$p_d = \begin{cases} p_r, & p_r \in A_2 \\ (0,0,0), & otherwise \end{cases}$$

C. FEATURE EXTRACTION

Color, shape, and texture characteristics are the most common types of image features in image processing. Because the colour qualities that are greatly impacted by its mechanism are applicable to outside light. In order to classify the illness spots using characteristic values, a selection is made to compare the colour, shape, and texture properties of the disease spots. The area (A) and perimeter (P) of illness spots are estimated from a binary image. By rotating the image at the same angle, the Minimum Enclosing Rectangle (MER) of a disease location can be determined [3]. The length (l) and breadth (w) of a disease spot are represented by the long axis length and short axis length of the MER. Using the area, perimeter, and MER of the illness spot, shape properties including rectangularity, compactness, elongation, and roundness are determined. In contrast to this work, three paddy diseases are classified using an adaption of textural features from the grey level co-occurrence matrix [10]. Initially, using the set of equations stated in, we converted the image from a red, green, and blue (RGB) format to a hue, saturation, and value (HSV) colour representation (6).

$$\begin{cases} v = \max(r, g, b) \\ s = \begin{cases} (v - \min(r, g.b) * 255)/v & v \neq 0 \\ 0 & v = 0 \end{cases} \\ h = \begin{cases} (g - b) * 60/s & v = r \\ 180 + (b - r) * 60/s & v = g \\ -240 + (r - g) * 60/s & v = b \\ if\ h < 0\ then\ h = h + 3 = 60 \end{cases} \end{cases} \qquad (4)$$

Where,, and stand for the respective red, green, and blue (RGB) pixel values for the image. The number of distinct grey levels in a picture was then used to construct a GLCM with square matrix from an HSV spatial system. An element of a GLCM of an image represented the probability of two grey level pixels, one of which is at location ( ) and the other of which is at a distance and an orientation angle of ( ) from place ( ). In our investigation, was chosen with orientation angles of. Each image's GLCM is used to extract five texture properties, including contrast, uniformity, entropy, inverse difference, and linearity correlation in all orientation angles. Each image contains a total of 60 texture feature values from 3 spatial (HSV) systems and 4 orientation angles.

D. IMAGE COLOR TRANSFORM

The photos are converted from RGB to one of three other colour spaces, including YCbCr, HIS, and CIELAB. To remove extraneous spots, the color transformed images are run through a median filter. Otsu thresholds are applied to the RGB image in the final stage, converting the RGB colour component A into the CIELAB colour component after the RGB image is first turned into CIEXYZ. The LAB colour model's brightness and colour information are independent of one another. This colour model also specifies brightness with L; A represents colours from green to red; and B guides colours from blue to yellow [16].

E. SMOOTHING & SEGMENTATION

In our work, we apply an image smoothing technique with a median filter to reduce unneeded noise. A numerical collection's median is the value at which half of the values are less than or equal to it and the other half are larger than or equal to it. The initial step in median filtering involves moving windows and shorting all of the pixels inside each window. The median metrics value was then computed and set to the centre pixel. If the number of elements in the window is odd, the middle value is assigned as the median value; otherwise, the median value is the average of the two middle values.

Detecting the illness location requires a method after image smoothing. While removing the diseased spot from a plant leaf, it's crucial to choose a threshold of grey. The threshold can be selected as the bottom of the bar if there is a deep and sharp length between the two peaks of the histogram. Yet, an issue arises if the valley is flat and wide. In that situation, separating items from background is not possible using this technique. As a result, the Otsu method is a useful approach that may be applied to this kind of segmentation. In our research, we employed the Otsu approach to automatically attain and choose the right ideal threshold.

F. OTSU'S APPROACH

The OTSU method uses a threshold at level K to separate pixels belonging to the two classes' background and object. Then, the calculated class means and class variances. The next step is to find a threshold K that maximises one of the object functions (l, k, or n), which are specified as:

$$l = \frac{\sigma_B^2}{\sigma_W^2} \quad k = \frac{\sigma_T^2}{\sigma_W^2} \quad n = \frac{\sigma_B^2}{\sigma_T^2}$$

The illness spot is identified using the CIELAB colour system, the H component of the HIS colour space, and a component of the YCbCr colour space. Images that have been segmented with illness spots as the The three methods are compared to determine which is the most effective for locating illness spots.

Point of interest

1) The YCbCr colour model is a popular one for digital video. In this colour model, Y stands for the luminance component, whereas and stand for the colour component. is the difference between the components that are blue and that which is between the components that are red [13] [14].
2) HSI: Based on human colour perception, HSI is a device-dependent colour model. In this colour scheme, the letter H stands for hue, which depicts a pure colour and is typically correlated with the light's wavelength. S stands for Saturation, which in the HSI colour model represents how vibrant things are. I write "Intensity," which displays the light's amplitude [14] [15].
3) CIELAB: The CIE established the device independent CIELAB system to categorise colour according to the human perception. When an image is being converted from First, an RGB image is converted into CIEXYZ, then RGB colour component to CIELAB colour component. radiance and the colour information in the LAB colour model is distinct from one another. This colour model also specifies brightness with L; A represents colours from green to red; and B guides colours from blue to yellow [16].

XI. CONCLUSION

The main focus of the rice plant disease recognition system is on quick and precise disease prediction in crops. Management of plant diseases. Recognizing diseases in rice plants at an early stage helps paddy researchers and farmers respond quickly to safeguard the plant. The primary goal of this research is to review and compile information on approaches for applying image processing and machine learning to detect diseases in rice plants. Many segmentation algorithms were used to retrieve the rice plant's diseased leaf image. The procedures the issues discussed in this paper aid scholars in resolving a number of issues that either directly or indirectly affect society. A strategy will be recommended for future study to address the research issue highlighted above as well as to investigate ML and segmentation techniques that may facilitate the detection of plant diseases. Future performance comparisons with classic algorithms' computational demands could be made.

XII. REFERENCES

1. Kurniawati, N.; Abdullah, S.; Abdullah, S. Investigation on Image Processing Techniques for Diagnosing Paddy Diseases. International Conference of Soft Computing and Pattern Recognition 2009, pp. 273–277.
2. Cunha, J. Application of Image Processing Techniques in the Characterization of Plant Leafs. IEEE Intl Symposium on Industrial Electronics. 2003.
3. Meunkaewjinda, A.; Kumsawat, P.; Attakit mongcol, K.; Srikaew, A. Grape Leaf Disease Detection from Color Imagery using Hybrid Intelligent System. Proceeding of ECTI-CON 2008, pp. 513–516.
4. Zhihao, Q.; Minghua, Z. Detection of rice sheath blight for in-season disease management using multispectral remote sensing. International Journal of Applied Earth Observation and Geo information 7 2005, pp.115–128.
5. Shen, W.; Wu, Y.; Chen, Z.; Wei, H. Grading Method of Leaf Spot Disease Based on Image Processing. CSSE 2008.
6. Moshou, D.; Bravo, C.; West, J.; Wahlen, S.; McCartney, A.; Ramon, H. Automatic detection of ‗yellow rust' in wheat using reflectance measurements and neural networks. Comput. Electron. Agric., vol. 44, no.3 2004, pp.73– 188.
7. Otsu, N. A threshold selection method from gray-level histogram. IEEE Trans. Syst., Man and Cybern, vol.91979, pp. 66–68.
8. Ju-Won, L.; Han-Wook, L.; Jong-Hoe, L.; Ick Tae, K.; Gun-Ki, L. A Study on Lung Nodule Detection using Neural Networks. IEEE TENCON 1999, pp. 1150–1153.
9. Bradski, G.; Kaehler, A. Learning OPENCV: computer vision with the OpenCV library. California: O'Reilly Media, Inc. 2008, pp. 238 – 250.
10. Perona, P.; Malik, J. Scale-space and edge detection using anisotropic diffusion. IEEE Transaction on pattern Analysis and Machine Intelligence, vol. 12 1990, pp. 629–639.
11. Yue, S.; Li, 218 P.; Hao, P. SVM classification: Its contents and challenges. Appl. Math. J. Chinese Univ. Ser. B, vol.18, no.3 2003, pp. 332–334.
12. Yao, Q.; et al.. Application of support vector machine for detecting rice diseases using shape and color texture features. Engineering Computation ICEC'09. International Confere-nce on. IEEE, 2009.
13. Song, K.; Liu, Z.; Su, H.; Guo, C. A Research of Maize Disease Image Recognition of Corn Based on BP Networks. Third International Conference on Measuring Technology and Mechatronics Automation IEEE 2011, pp. 246–249.
14. Ying, G.; Miao, L.; Yuan, Y.; Zelin, H. A Study on the Method of Image Pre-Processing for Recognition of Crop Diseases. International Conference on Advanced Computer Control, IEEE 2008, pp. 202–206.
15. Al-Hiary, H.; Bani-Ahmad, S.; Reyalat, M.; Braik, M.; Al Rahamneh, Z. ―Fast and Accurate Detection and Classification of Plant Diseases. International Journal of Computer Applications Volume 17 No.1 2011.
16. luo, L.; Zohu, G. Extraction of the Rice Leaf Disease Image Based on BP Neural Network. International Conference on Computational Intelligence and Software Engineering, IEEE 2009.
17. Gayathri Devi T and Neelamegam P, ― Paddy leaf disease detection using SVM with RBFN classifier‖, International Journal of Pure and Applied Mathematics, vol. 117, no. 15, pp. 699-710, 2017.
18. Vanitha V, ―Rice disease detection using deep learning‖, International Journal of Recent Technology and Engineering, vol. 7, pp. 2277-3878, no. 5S3, 2019.
19. Hemanta Kalita, Shikhar Kr Sarma and Ri dip Dev Choudhury, ―Expert system for diagnosis of diseases of rice plants: prototype design and implementation‖, International Conference on Automatic Control and Dynamic Optimization Techniques, pp. 723-730, 2016.
20. Sujeet Varshney and Tarun Dalal,―A novel approach for the detection of plant diseases‖, IJCSMC, vol. 5, no. 7, pp. 44-54, 2016.
21. Mohan K J, Balasubramanian M and Palanivel S, ―Detection and recognition of diseases from paddy plant leaf images‖, International Journal of Computer Applications, vol. 144, no. 12, 2016.
22. Shaik Anzala Noor, Nimmala Anitha, Sreenivasulu P, ―Identification of fungal diseases in paddy fields using image processing techniques‖, International Journal of Advance Research, Ideas and Innovations in Technology, vol. 4, no. 3, 2018.
23. Santanu Phadikar, Jaya Sil and Asit Kumar Das,―Classification of rice leaf diseases based on morphological changes‖, Internatio- nal Journal of Information and Electronics Engineering, vol. 2, no. 3, pp. 460-463, 2012.
24. Leonard Gianessi P,―Importance of pesticides for growing rice in south and southeast asia‖, International Pesticide Benefit Case Study, vol. 108, 2014.

25. Sachin Khirade D and Patil A B, ―Plant disease detection using image processing‖, International conference on computing communication control and automation, pp. 768-771, 2015.
26. Kahar M A, Sofianita Mutalib and Shuzlina Abdul-Rahman, ―Early detection and classification of paddy diseases with neural networks and fuzzy logic‖, in 17th International Conference on Mathematical and Computational Methods in Science and Engineering, pp. 248-257, 2015.
27. Arivazhagan S and Ligi S V, ―Mango leaf diseases identification using Convolutional Neural Network‖, International Journal of Pure and Applied Mathematics, vol. 120, no. 6, pp.11067- 11079, 2018.
28. LiuB, Zhang Y, He D and Li, Y, ―Identification of apple leaf diseases based on deep convolutional neural networks‖, Symmetry, vol. 10, no. 1, pp.11, 2017.
29. Fuentes A, Yoon S, Kim S and Park D, ―A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition‖, Sensors, vol. 17, no. 9, 2017.
30. Lu Y, Yi, S, Zeng N, Liu, Y and Zhang Y,―Identification of rice diseases using deep convolutional neural networks‖, Neuro computing, vol. 267, pp.378-384, 2017.
31. Jin, X, Jie L, Wang S, Qi, H and Li S, ―Classifying wheat hyperspectral pixels of healthy heads and Fusarium head blight disease using a deep neural network in the wild field‖, Remote Sensing, vol. 10, no. 3, pp.395, 2018.
32. Jitesh Shah P, Harshad kumar B Prajapati and Vipul K Dabhi, ―A survey on detection and classification of rice plant diseases‖, In IEEE International Conference on Current Trends in Advanced Computing, pp. 1-8, 2016
33. Phadikar S, Sil J and Kumar Das A, ―Region identification of infected rice images using the concept of Fermi energy‖, Advances in Computing and Information Technology, pp. 805-811, 2013.