

# Prediction and Analysis of Paddy Crops Disease in Artificial Intelligence Techniques

Tanmayee Tushar Parbat<sup>1</sup>, Rohan Benhal<sup>2</sup>, Honey Jain<sup>1</sup>, Dr. Vinayak Musale<sup>3</sup>

<sup>1</sup>B.E IT, Dr. Vishwanath Karad MIT World Peace University, Pune, Maharashtra, India

<sup>2</sup>BBA IT, Dr. Vishwanath Karad MIT World Peace University, Pune, Maharashtra, India

<sup>3</sup>Dr. Vishwanath Karad MIT World Peace University, Pune, Maharashtra, India

## ABSTRACT

### Article Info

Volume 7, Issue 6

Page Number: 261-268

### Publication Issue :

November-December-2021

### Article History

Accepted : 05 Dec 2021

Published : 15 Dec 2021

The survival of human beings is generally based on the proper productivity of agriculture. The paddy plant is considered as a major planting crop in improving the economical level of our country. Nowadays, the yield level of paddy crop might be minimized due to several diseases. Bacteria, fungi, virus and certain harmful insects are the main causative agents for such disease occurrence on the paddy crop. The diseases which affect the early stage of the paddy crops influences in the whole stage of crop cultivation. In early days of agriculture, the manual detection of diseases has been carried out by farmers. Image processing is one of the emerging techniques for identifying and classifying the different types of diseases and it overcomes the issues encountered during the manual detection of diseases. Image processing technique solves several issues involved in the cultivation of crops including, recognition and classification of plant diseases, discrimination of certain weeds and disease forecasting.

Keywords: - Paddy Crop, Machine Learning, Image Processing

## I. INTRODUCTION

Plant diseases are one of the causes in the reduction of quality and quantity of agriculture crops [1]. Reduction in both aspects can directly affect the overall production of the crop in a country [2]. The main problem is a lack of continuous monitoring of the plants. Sometimes newbie farmers are not aware of the diseases and its occurrence period. Generally, diseases can occur on any plant at any time. However, a continuous monitoring may prevent disease infection. The detection of a plant disease is one of

the important research topics in the agriculture domain. This article attempts to apply concepts of Machine Learning and Image Processing to solve the problem of automatic detection and classification of diseases of the rice plant, which is one of the important foods in India. On any plant, diseases are caused by bacteria, fungi, and virus. For rice plants, most common diseases are Bacterial leaf blight, Brown spot, Leaf smut, Leaf blast, and Sheath blight [3]. Image processing operations can be applied on external appearances of infected plants. However, the symptoms of diseases are different for different plants.

Some diseases may have brown color or some may have a yellow color. Each disease has its own unique characteristics.

Diseases differ in shape, size, and color of disease symptoms. Some of the diseases might have the same color, but different shapes; while some have different colors but same shapes. Sometimes farmers get confused and are unable to take proper decision for selection of pesticides. Capturing the images of infected leaves and finding out the information about the disease is one way to get rid of loss of crop due to disease infection. As an automated solution of this problem, cameras can be deployed at certain distances in the farm to capture images periodically. These images can be sent to a central system for analysis of diseases; the system can detect the disease and give information about the disease and pesticide selection [4].

At the core of such system would be to automatically recognize the disease that has occurred. We address this problem in this article. We briefly present our approach to solving the problem of automatic detection and classification of rice plant diseases. We collected the leaves from rice farm and prepared a dataset of images of rice plant leaves having a white background. Our system first removes the background from an image and then using K-means clustering it extracts the disease portions of the leaf image. After applying K-means clustering, some unnecessary green region is removed from disease portion using thresholding technique [5].

## II. IMAGE PROCESSING

In Image processing section, initially the image is captured from the camera and further the image is processed using k means clustering for segmenting the image. The processed image is then edge detected using three different edge detection techniques. The edge detection techniques used are sobel, prewitt and

canny algorithm. The diseased sample banana leaf has been taken for the edge detection analysis. Amongst the three edges detection methods used, canny edge detection algorithm gives the better and reliable detection. Owing to its optimality to meet with the three criteria for edge detection and the simplicity of process for implementation, it became one of the most popular algorithms for edge detection method.

As discussed earlier, IoT and Image processing are combined together in agricultural field in order to increase product yield and to reduce the crop failure. We focused on plant failure due to environmental factors through IoT technology. IoT system includes sensors, Arduino and a camera that regularly captures the plant. The color, texture, shape and area of the leaf are the parameters also considered in this work. After examine the conditions of the plants we go for image processing. The initial test is done by using MATLAB software. In addition to the environmental factors, the plant with a diseased leaf can also be identified using Image processing. Based on the output and constraints the pesticides will be sprayed for the crop/plant where the disease is identified. If there is any change that corresponds to the deterioration in the plants growth, the farmer is immediately informed. Early diagnosis will thus help in taking the necessary actions to increase the produce and reduce failure of crops.

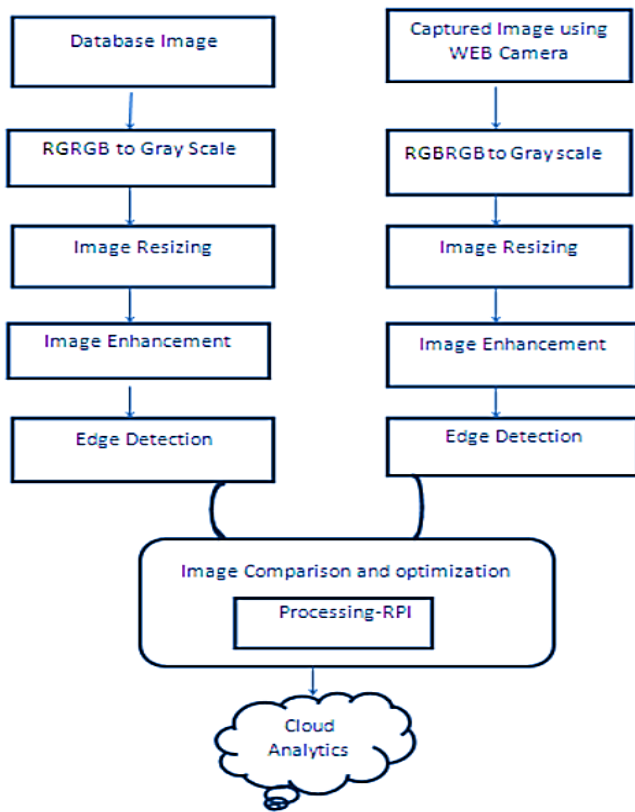


Figure 1: Block diagram of Image Processing

### III. LITERATURE REVIEW

The image processing technology is robust as well as simple to perform the detection of disease in paddy leaf, in which, some mathematical modelling techniques (Abirami et al. [1]) can be included to enhance the characteristics of image processing technology. The image processing technology is also used to compare the leaf colour chart (LCC) with the crop leaf colour (Adam et al. [2]).

A methodology was proposed by Ahmad et al. [3], to recognize the RBPH (Rice brown plant hopper) on rice stems. With the help of a smartphone, the image of rice stem was taken, then after, in order to scale the population density of RBPH, the image processing technique was applied on the captured images. For the segmentation, the steps like k-means clustering, image enhancement and median filtering were performed. The farmers can utilize this application in their android smartphones to control rice pest infection.

For recognizing the variations occurred in rice crops, Aji et al. [4], proposed a method, in which, a camera system and unmanned aerial vehicle (UAV) were employed to capture the images from the paddy leaf. From the rice field, nearly 30 samples were taken that had 1x1-meter area. The Oklahoma State University (OSU) Green seeker prototype was used to measure the Normalized Difference Vegetation Index (NDVI) values of the samples. NDVI is used to forecast the agricultural production, plant health etc. This measure is used in crop monitoring to illustrate the farmers that which part of their fields have moderate, dense vegetation. In order to get the NDVI values, the images are processed.

To identify and detect the white-backed plant hoppers (WBPHs, *Sogatella furcifera* (Horvth)), a novel three-layer detection method was applied on the images of rice plants present in paddy fields via image processing. Both Ada Boost classifier and SVM classifier were employed, in which, SVM classifier was trained on local binary pattern (LBP) and Gabor features to remove impurities and identify the white-backed plant hoppers (WBPHs). Likewise in histogram of oriented gradient (HOG) features, the Ada Boost classifier is trained. In final layer, to determine the WBPHs the SVM classifier was trained on the HOG features (Anami et al. [5]).

Nidhis et al. (2019) had proposed a method with the image processing to recognize the type of disease affected over the paddy leaves. The impact of disease infection was explained according to the portion of the diseased area. The pesticides were applied on plant leaves respect the severity for eliminating the rice blast, bacterial blight and brown spot because which are affects the productivity.

With the assistance of image processing, Devi and Archana et al. [6], had developed an approach to conduct disease determination in paddy leaves automatically. For the purpose of feature extraction, the SIFT, gray scale co-occurrence matrix and DWT were used. Then, the extracted features were sent to

various classifiers such as back propagation neural network, multiclass SVM, KNN and Naive Bayesian to differentiate both normal and diseased plants. Furthermore, by using image processing, various difficulties and disease degrading the quality of Malaysias type of rice were also observed. For investigating and differentiating the level of rice, like RGB colour model, edge detection and histogram were mostly preferred (Bashir et al. [7]).

The image pre-processing is required to boost the values of pixel by removing the various noise in the image, which results in quality enhancement in images. As a result of image enhancement, the required information is manipulated in the image, so that, the user can utilize the required information. The image pre-processing enhance the visual understanding regarding that image. The image processing techniques are categorized into the frequency and spatial domain techniques. In spatial domain techniques, the image pixels were directly manipulated, whilst, in the indirect method manipulation of pixels was performed at the frequency domain. In the frequency domain enhance-ment technique, the kernels or convolution was used to make the transformation. Poor contrast or blur are generally considered to be the weakness in images. Image pre- processing overcomes the above issues and improves the quality of the image. The pixel of the upper and lower threshold value is mapped to the new pixel value for im- proving the quality of the image. For performing additional analysis and processing task, the visual impact of pixel brightness could be increased (Batoa et al. [8]).

To differentiate the disease affected and healthy leaves, eight features are extracted from the paddy leaf images. The HSV image is derived at the pre-processing phase from RGB image (Chen et al. [9]), which could be used for disease detection.

#### IV. METHODOLOGY

In this section we will explain about the methodologies which we have used.

A **decision tree** is a graphical representation of possible solutions to a decision based on certain conditions. It's called a decision tree because it starts with a single box (or root), which then branches off into a number of solutions, just like a tree. Decision trees are helpful, not only because they are graphics that help you 'see' what you are thinking, but also because making a decision tree requires a systematic, documented thought process. Often, the biggest limitation of our decision making is that we can only select from the known alternatives. Decision trees help formalize the brainstorming process so we can identify more potential solutions.

**Random forests** or random decision forests are an ensemble learning method for classification, regression and other tasks, that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. Random decision forests correct for decision trees' habit of over fitting to their training set. The first algorithm for random decision forests was created by using the random subspace method which, in Ho's formulation, is a way to implement the "stochastic discrimination" approach to classification proposed by Eugene Kleinberg.

An **Artificial Neural Network**, often just called a neural network, is a mathematical model inspired by biological neural networks. A neural network consists of an interconnected group of artificial neurons, and it processes information using a connectionist approach to computation. In most cases a neural network is an adaptive system that changes its structure during a learning phase. Neural networks are used to model complex relationships between inputs and outputs or to find patterns in data. The inspiration for neural networks came from examination of central nervous systems. In an artificial neural network, simple artificial nodes, called "neurons", "neuroses", "processing elements" or

“units”, are connected together to form a network which mimics a biological neural network.

In machine learning, **support vector machines** (SVMs, also support vector networks) are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. Given a set of training examples, each marked as belonging to one or the other of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier (although methods such as Platt scaling exist to use SVM in a probabilistic classification setting).

**K nearest neighbors** is a simple algorithm that stores all available cases and classifies new cases based on a similarity measure (e.g., distance functions). KNN has been used in statistical estimation and pattern recognition already in the beginning of 1970’s as a non-parametric technique [14-15].

### V. RESEARCH METHODOLOGY

The flow diagram shown in Fig. 2 indicates the process steps used for Blast disease recognition using KNN and ANN classifier algorithms.

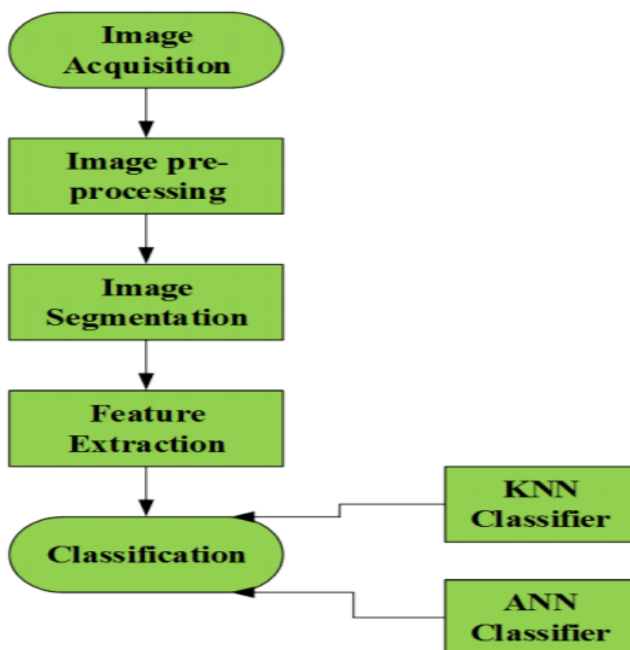


Figure 2: Flow Diagram of Research Methodology

This process consists of five various stages such as Image Acquisition, Image pre-processing, Image Segmentation, Feature Extraction and Classification.

In the image acquisition step, the digital camera is employed to capture the image 512×512 pixels of rice leaf. The pixel of images is reduced to 256×256 pixels for increasing the accuracy. The gray scale image is created from the resized colour image for carrying out subsequent process. The HSV image is extracted from the RGB image during the phase of image pre-processing.

DNN based Classification:- In Artificial Intelligence, a machine learning subset is called deep learning, which has the potential of learning unstructured or unsupervised data. Input layer, output layer and hidden layers are the essential components of the DNN framework. The proposed DNN framework is specified in Figure 3.

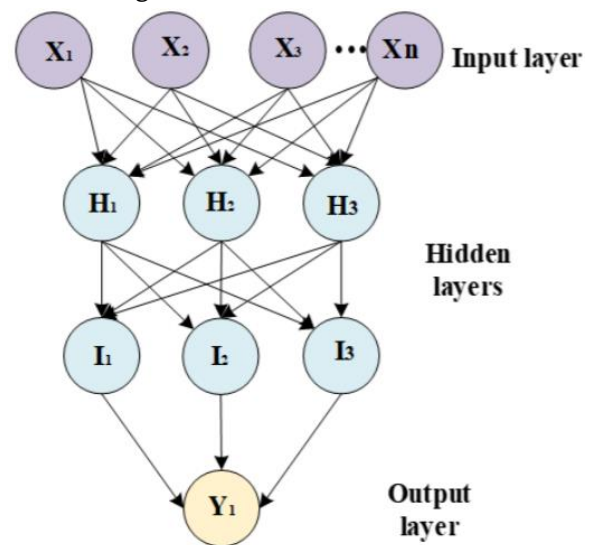


Figure 3: Artificial Intelligence Architecture

DNN indicates the type of ML (Machine Learning) while the system utilizes many layers of nodes to derive high-level functions from input information. It means transforming the data into a more creative and abstract component. Nodes are little parts of the system, and they are like neurons of the human brain. When a stimulus hits them, a process takes place in

these nodes. Some of them are connected and marked, and some are not, but in general, nodes are grouped into layers. The system must process layers of data between the input and output to solve a task. Creative and analytical components of information are analyzed and grouped to ensure that the object is identified correctly. The creation of neural network is inspired by the working of human brain and its functions. Artificial intelligence and machine learning, which is a subset of AI, play an essential part in its functionality. It starts working when a developer enters data and builds a machine learning algorithm, mostly using the “if ... else ...” principle of building a program. The deep neural network does not only work according to the algorithm but also can predict a solution for a task and make conclusions using its previous experience.

### VI. SIMULATION RESULTS

The image dataset is processed using ANN and KNN Classifiers to detect the blast disease in rice crops. The performance of the classifier algorithms is analysed using the common QoS parameters as described in this section. Four outcomes such as false negative, true negative, false positive and true positive are used to derive the performance measure of ANN and KNN via confusion matrix.

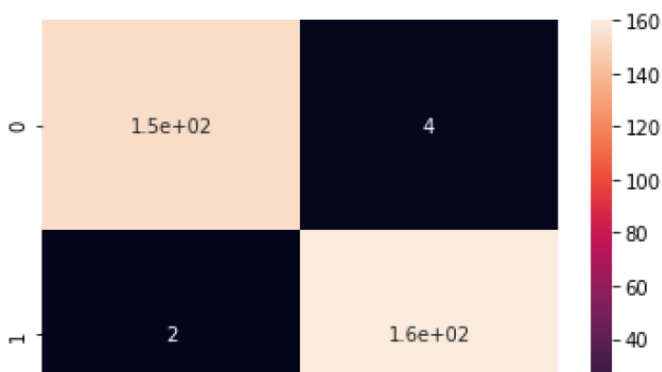


Figure 4: Confusion Matrix

	precision	recall	f1-score	support
0	0.99	0.97	0.98	158
1	0.98	0.99	0.98	162
accuracy			0.98	320
macro avg	0.98	0.98	0.98	320
weighted avg	0.98	0.98	0.98	320



Figure 5: Fruit Leaf

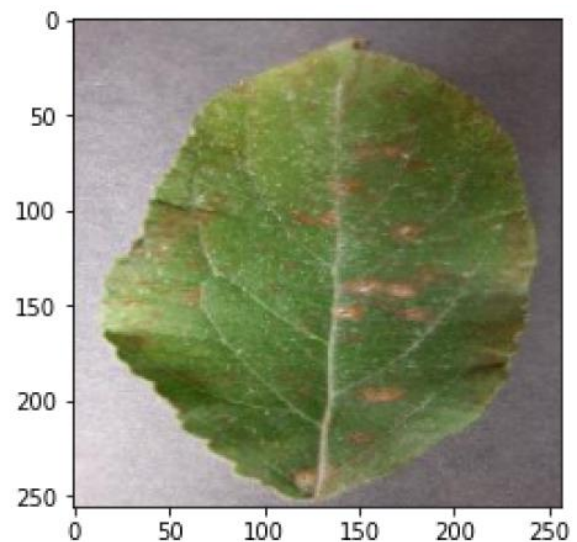


Figure 6: Color Fruit Leaf

### VII. CONCLUSION

The research work of this paper addressed the detection of diseases of fruit plant through image processing techniques with the help of optimized machine learning algorithms. In this, k-means clustering is involved for clustering in the segmentation stage. Two features such as texture and colour are adapted for feature extraction. The application of machine learning algorithms namely, ANN, KNN, DNN and JAYA optimization algorithm

facilitates in the classification process to gain the supreme performance. The future trends that could evolve in this reported thesis is presented here. There is a lot of scope for extending this work by assisting several image processing techniques such as deformation, transformation and inclusion of images in addition to the pre-processing, segmentation and other related proposed stages. In addition to the colour and texture features the shape features could also be considered for the feature extraction stage of image processing.

### VIII. REFERENCES

- [1] Abirami, M. T. S. (2017), 'Application of image processing in diagnosing guava leaf diseases', *International Journal of Scientific Research and Management* 5(7), 5927– 5933.
- [2] Adam, E., Deng, H., Odindi, J., Abdel-Rahman, E. M. and Mutanga, O. (2017), 'Detecting the early stage of phaeosphaeria leaf spot infestations in maize crop using in situ hyperspectral data and guided regularized random forest algorithm', *Journal of Spectroscopy* 2017.
- [3] Ahmad, F., Ku-Mahamud, K. R., Sainin, M. S. and Airuddin, A. (2017), 'Segmentation method based on artificial bee colony for recognizing leaf lesion', *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)* 9(2-3), 103–107.
- [4] Aji, A. F., Munajat, Q., Pratama, A. P., Kalamullah, H., Setiyawan, J. and Arymurthy, A. M. (2013), 'Detection of palm oil leaf disease with image processing and neural network classification on mobile device', *International Journal of Computer Theory and Engineering* 5(3), 528.
- [5] Anami, B. S., Pujari, J. and Yakkundimath, R. (2011), 'Identification and classification of normal and affected agriculture/horticulture produce based on combined color and texture feature extraction', *International Journal of Computer Applications in Engineering Sciences* 1(3), 356–360.
- [6] Archana, K. and Sahayadhas, A. (2018), 'Automatic rice leaf disease segmentation using image processing techniques', *International Journal of Engineering and Technology* 7(3.27), 182–185. Athanikar, G. and Badar, P. (2016), 'Potato leaf diseases detection and classification system', *International Journal of Computer Science and Mobile Computing* 5(2), 76– 88.
- [7] Bashir, K., Rehman, M. and Bari, M. (2019), 'Detection and classification of rice diseases: An automated approach using textural features', *Mehran University Research Journal of Engineering and Technology* 38(1), 239–250.
- [8] Batoa, H., Limi, M. A., Hamzah, A., Cahyono, E. D., Arimbawa, P., Yusria, W. O. and Gafaruddin, A. (2019), 'External factors affecting lowland rice farmers' use of chemical pesticides in Welala Village, Kolaka Timur Regency, Indonesia', *Journal of Agricultural Extension* 23(2), 80–89.
- [9] Chen, J., Liu, Q. and Gao, L. (2019), 'Visual Tea Leaf Disease Recognition Using a Convolutional Neural Network Model', *Symmetry* 11(3), 343.
- [10] Dhingra, G., Kumar, V. and Joshi, H. D. (2019), 'A novel computer vision based neutrosophic approach for leaf disease identification and classification', *Measurement* 135, 782–794.
- [11] El Massi, I., Es-saady, Y., El Yassa, M., Mammass, D. and Benazoun, A. (2017), 'Automatic recognition of vegetable crops diseases based on neural network classifier', *International Journal of Computer Applications* 975, 8887.
- [12] Ellur, R. K., Khanna, A., Yadav, A., Pathania, S., Rajashekara, H., Singh, V. K., Krishnan, S. G., Bhowmick, P. K., Nagarajan, M., Vinod, K. et al. (2016), 'Improvement of basmati rice varieties for resistance to blast and bacterial blight

- diseases using marker assisted backcross breeding', *Plant Science* 242, 330–341.
- [13] Habib, M. T., Majumder, A., Jakaria, A., Akter, M., Uddin, M. S. and Ahmed, F. (2018), 'Machine vision based papaya disease recognition', *Journal of King Saud University Computer and Information Sciences*.
- [14] Hamuda, E., Mc Ginley, B., Glavin, M. and Jones, E. (2017), 'Automatic crop detection under field conditions using the hsv colour space and morphological operations', *Computers and electronics in agriculture* 133, 97–107.
- [15] Hassanien, A. E., Gaber, T., Mokhtar, U. and Hefny, H. (2017), 'An improved moth flame optimization algorithm based on rough sets for tomato diseases detection', *Computers and electronics in agriculture* 136

**Cite this article as :**

Tanmayee Tushar Parbat, Rohan Benhal, Honey Jain, Dr. Vinayak Musale, "Prediction and Analysis of Paddy Crops Disease in Artificial Intelligence Techniques", *International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT)*, ISSN : 2456-3307, Volume 7 Issue 6, pp. 261-268, November-December 2021.

Journal URL : <https://ijsrcseit.com/CSEIT217673>