An IoT based System with Edge Intelligence for Rice Leaf Disease Detection using Machine Learning

S. M. Shahidur Harun Rumy
Department of Computer Science & Engineering
Southeast University
Dhaka, Bangladesh
rumywithy@gmail.com

Md. Ishan Arefin Hossain
Department of Computer Science & Engineering
Southeast University
Dhaka, Bangladesh
ishanarefin@gmail.com

Forji Jahan
Department of Computer Science & Engineering
Southeast University
Dhaka, Bangladesh
swarna.4g@gmail.com

Tanjina Tanvin
Department of Computer Science & Engineering
Southeast University
Dhaka, Bangladesh
tanvintonni06@gmail.com

2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS) | 978-1-6654-4067-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/IEMTRONICS52119.2021.9422499

***Abstract*—Bangladesh is one of the top five rice-producing and consuming countries in the world. Its economy dramatically depends on rice-producing. Rice leaf disease is the biggest problem in the agriculture sector. This is the main reason for the reduction of the quality and quantity of the crops. The spread of the disease can be avoided by continuous monitoring. However, manual monitoring of diseases will cost a large amount of time and labor. So, it is a good idea to have an automated system. This paper presents a rice leaf disease detection system using a lightweight Artificial Intelligent technique. We are applying the edge computing concept here. Our edge device is Raspberry Pi. We have processed all our data in Raspberry Pi. We consider three rice plant diseases, namely Brown Spot, Hispa, and Leaf Blast. They are the most common type of rice leaf disease in Bangladesh. We have used clear images of healthy and infected rice leaves with white background. After applying the necessary preprocessing, we have extracted the necessary features from the images. Then we have made an image classification model with various machine learning algorithms by feeding these features. We have learned that the Random Forest algorithm performed the best. By using our image classification model, we have achieved 97.50% accuracy on our edge device.**

Keywords—Edge Intelligence, IoT, Machine Learning, Rice Leaf Disease Detection, Image Classification

I. Introduction

As an agricultural country, about 135 million people of Bangladesh consider rice as their staple food. In this region, rice considers as low cost and most nutrient food. More or less, 48% of rural employment is the blessings of it. Besides, half of the agricultural GDP comes from the rice sector. The Rice sector contributes to one-sixth of the national income of Bangladesh. Around 80% of the total irrigated area is under this sector [1]. The livelihood of the people in rural areas depends mostly on rice cultivation. To ensure stable economic growth and maintain desired goals, disease-free rice cultivation can play an important role. So, it is evident that the proper cultivation of rice is a primary concern for Bangladesh.

Therefore, we have come up with an automated system, which will contribute to the development of Bangladesh's agriculture sector. Our proposed system will automatically detect if a leaf is healthy or infected from the leaf's image. Our system is based on the edge computing concept. Edge computing places networked computing resources as close as possible to where data is created. Edge computing is associated with the Internet of Things and the application of small computing devices like Raspberry Pi [2].

We are using Raspberry Pi as our edge device. We process all of our data in Raspberry Pi. Raspberry Pi is a credit-card-sized single-board computer. It is very low in price. It has some GPIO pins to communicate with external components. Its default operating system name is Raspberry Pi, which is based on Debian [3].

We are detecting three rice leaf diseases, namely Brown Spot, Hispa, and Leaf Blast, because these are the most common rice leaf disease in Bangladesh [4].

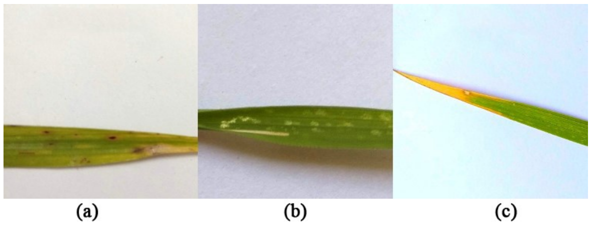

Fig. 1. Infected Leaves (a) Brown Spot, (b) Hispa, (c) Leaf Blast

Brown Spot is caused by Cochliobolus miyabeanus fungus. It causes a prominent spot on the leaves [5]. A species of leaf beetle named Dicladispa armigera is the reason for Hispa. It scrapes the upper surface of leaves and causes white patches [6]. Magnaporthe oryzae is a fungus that causes Leaf Blast disease in rice leaves. It causes grey-green or white spots with dark green or reddish-brown border [7].

We have taken images of healthy and infected leaves with white backgrounds [8]. After preprocessing the images, we have applied machine learning algorithms to make our image classification model. We have tried different machine learning algorithms, but random forest turns out to be the most efficient.

A. Internet of Things:

Internet of Things (IoT) refers to devices embedded with sensors and technologies that are connected to the internet for exchanging information. It can be a self-driving car, an automatic door, a coffee maker, a wearable device, or a washing machine connected to the internet for collecting and sharing information. Most of the time, this information exchanging happens without the interference of humans. In some cases, these devices make use of artificial intelligence [10].

B. Edge Computing:

Edge computing is a distributed computing system that processes data on an edge device instead of a data center. It makes real-time data processing effective. This reduces problems with latency and connectivity. It ensures that only critical data will be sent over the network, and no-critical data will be processed on the edge device. This reduces latency for time-sensitive devices [11]. Since edge devices process data themselves, the latency is almost zero.

C. Machine Learning:

Machine learning is a method that enables systems to learn automatically from data and improve their accuracy over time from experience without being programmed to do so. It is a branch of artificial intelligence. In machine learning, training data is given to the learning algorithm. Then learning algorithm generates a set of rules from the pattern and feature of the data. Based on these rules, it makes a prediction on new data. The quality of the algorithm defines the accuracy of the prediction. [12].

D. Random Forest Algorithm:

Random forest algorithm is a supervised learning technique. This is a popular machine learning algorithm for classification. However, it can be used for regression problems also. It consists of a combination of multiple decision trees. It uses the bagging method for training. The prediction of random forest is based on the majority of all the decision trees. Random forest eliminates the risk of overfitting and it takes less time for training [9].

E. Motivation:

Bangladesh is the fourth-largest rice producer. Rice is grown on about 10.5 million hectares. Almost all of the 13 million farm families of Bangladesh grow rice [1]. So, it is essential to produce good quality and quantity of rice. But the main problem is the disease of rice crops. Detection of rice plant disease has always been challenging. It requires constant human observations. But it is much more time-consuming and requires more labor.

This research presents an automatic rice leaf disease detection system. It detects if any of the crops are infected by disease or not. This will make the life of farmers easier, reduce the cost of labor and save time. Our proposed system is cost-friendly also, as we are using Raspberry Pi, a cheap microcomputer.

F. Research Paper Outline:

Our research paper has six sections. The sections and their contents are as follows:

- Section 1: Introduction. This section gives the overview, objectives, and motivation of our research report.
- Section 2: Literature Review. This section presents a concise survey of different image processing and machine learning operations applied in rice leaf disease detection.
- Section 3: Methodology. This section gives an overview of the working process.
- Section 4: Implementation. This section presents our proposed work for rice leaf disease detection. It describes the detailed methodology of our proposed work.
- Section 5: Results and Discussion. This section presents the results we have achieved.
- Section 6: Conclusion. The summary and scope of our research report has been discussed here.

II. Literature Reviews

In [13], the image is treated as a matrix of M rows and N columns. Their work extracted color texture using chromatography concepts of CIE XYZ color space. They have extracted color features using CIE L*a*b* color space. They used those features to form a simple metric indicating the roundness of the spot. They have achieved 70% accuracy with 50 sample images.

Authors in [14] divided their work into two phases. First, they prepare the healthy leaf image using HS histogram, and then they prepare diseased leaf image by extracting significant colors. After that, they have used image segmentation to apply the outlier detection method.

Q. Yao et al. [15] used Support Vector Machine (SVM) for detecting rice diseases. The achieved accuracy is 97.2%. After segmenting the diseased spot from the image, shape and texture features were extracted from the spot. Bacterial leaf blight, rice sheath blight, and rice blast disease were classified using SVM.

Orillo et al. [16] used 55 images for Bacterial leaf Blight, 37 images for Brown Spot, and 42 images for Rice Blast. After reducing noise and contrast adjustment, image segmentation was done for feature extracting. Then the back-propagation neural network is applied for the classification.

In this [17], 300 images were used for classification. After converting the image to HSV color format, K-Means clustering is used for image segmentation. Statistical features such as Mean Value, Standard Deviation, and GLCM are calculated for feeding to artificial neuron network (ANN). 90% and 86% accuracy achieved on test dataset for the infected and healthy images, respectively.

Research work in [18] used two datasets. One has 48 images, and another has 23 images. They have used the Simple Linear Iterative Clustering (SLIC) algorithm for over-segmentation of images. They extracted three features, namely color, shape, and texture. They have extracted texture from Gray Level Co-occurrence matrix (GLCM). They have used Random Forest classifier for classification.

The authors of this [19] paper used K-means clustering for the segmentation of the diseased area of the rice leaves. For classification, they used the Support Vector Machine (SVM). They achieved 93.33% accuracy on the training data set but 73.33% accuracy on the test dataset.

The author in [20] used fermi energy-based segmentation to extract the infected parts from the images. The shape of the infected part was extracted using the DRLSE method. To reduce computation time, they have used rough set theory.

III. Methodology

The methodology of the research work can be divided into five steps. They are image preprocessing, image segmentation, feature extraction, training model, and detecting healthy and infected leaves in Raspberry Pi. Block diagram of the system is shown in Fig. 2.

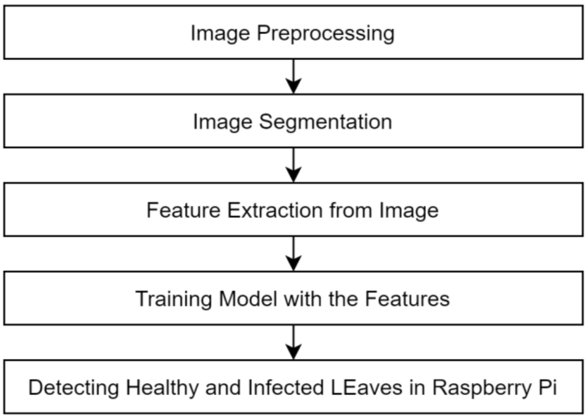

Fig. 2. Block Diagram

The system architecture of our proposed system is shown in Fig. 3.

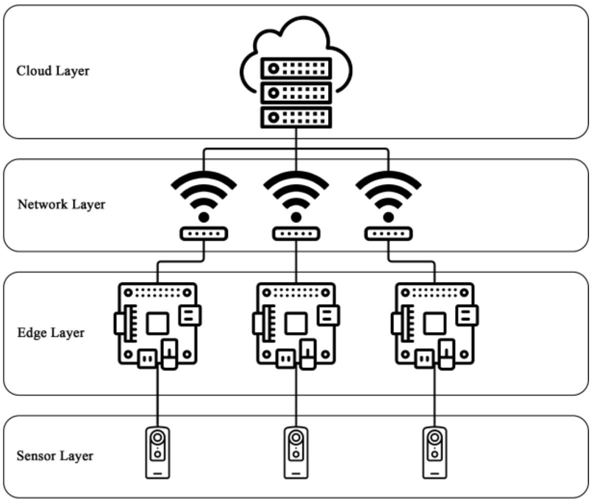

Fig. 3. System Architecture

IV. Implementation

We are working with an image dataset [8] from Kaggle. The dataset is not well sorted. In the healthy folder, there are some infected leaf images, and in the infected folder, there are some healthy leaf images. After separating them, we discovered a lower number of healthy leaf images than infected leaf images. Furthermore, the picture sizes are not consistent.

As we have to make an image classification model which can be run on Raspberry Pi, we are not classifying individual diseases. We will just find out if the leaf is infected or not. That is why we have labeled our dataset in two classes called infected and healthy. For the training dataset, we have taken 310 infected leaf images and 190 healthy leaf images. Furthermore, we have taken 50 infected leaf images and 30 healthy leaf images for the testing dataset. All of the images were selected at random.

A. Components

1) Raspberry Pi 3 Model B v1.2: We are using Raspberry Pi as our edge device. It is packed with Quad-Core 1.2GHz 64bit CPU. It comes with 1GB of RAM. The classification of the leaves is done on Raspberry Pi.

2) Raspberry Pi Camera Board v1.3: We will use a camera module to capture rice leaves images for detecting purposes. It is a 5MP resolution camera. The camera module can be connected to Raspberry Pi through the CSI Camera Port of Raspberry Pi.

B. Creating Model

We are creating an image classification model which can be run on Raspberry Pi. For that reason, we are not going to use complicating filters and image processing methods. Our goal is to get higher performance with a lesser computational load.

1) Image Preprocessing: For reducing computation time, we have resized the images to 256 × 256 pixels. Then we have changed the color format of images from RGB color format to HSV color format. Because it is easier to separate colors in HSV color format. HSV represents Hue, Saturation, and Value part of an image. The output is shown in Fig. 4.

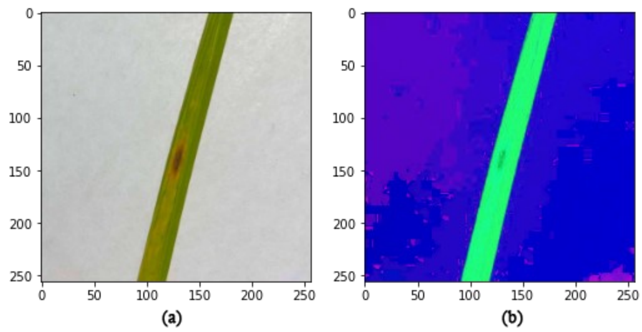

Fig. 4. (a) Image in RGB Format, (b) Image in HSV Format

2) Image Segmentation: After that, we have removed the background from the images. For that, we have performed image segmentation. In segmentation, we have first made a mask that will extract green color from the images, shown in Fig. 5.

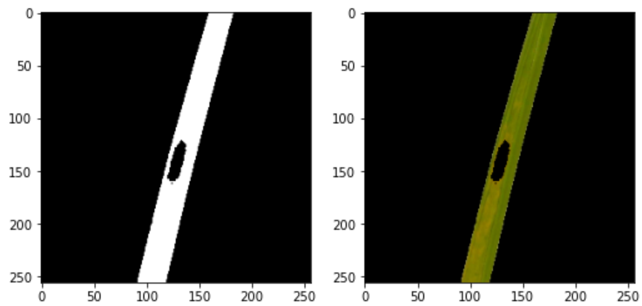

Fig. 5. The output of the green mask

Then we have made another mask that will extract brown color from the images, shown in Fig. 6.

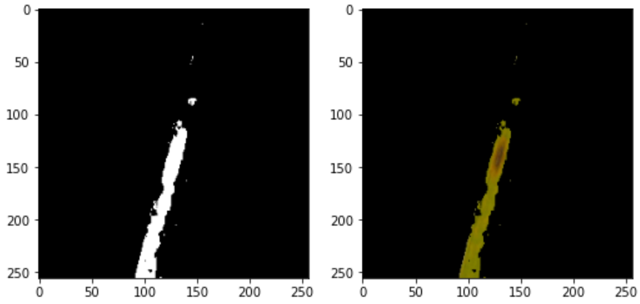

Fig. 6. The output of the brown mask

After that, we have combined these two masks, which will extract leaf from the images. After applying segmentation, the output image becomes like in Fig. 7.

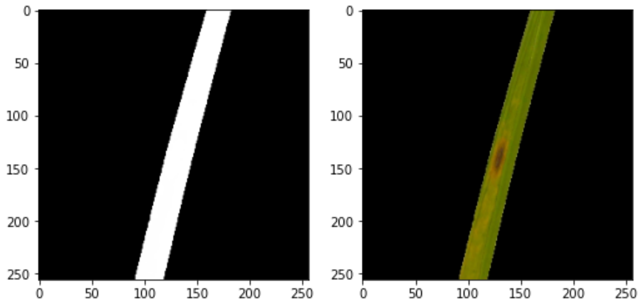

Fig. 7. The output of the final mask

3) Feature Extraction: We have used three feature descriptors to extract features from the images. For extracting color features, we have used Color Histogram. The color distribution in an image is represented by a histogram. Calculating an image's histogram is extremely useful because it offers insight into some of its properties, such as total range, contrast, and brightness. We have used Hu Moments for extracting shape features. A weighted average of the image pixel intensities is used to measure the image moment. Hu moments are moments that are not affected by translation, scaling, or rotation. Texture features were extracted using Haralick. A Gray Level Co-occurrence Matrix, or GLCM, is used to compute Haralick texture features. Haralick Texture is a texture-based approach for quantifying images.

4) Training Model: We have trained our model with six machine learning algorithms. All the models are validated using the 10-Fold Cross-Validation technique. The algorithms are,

a) Random Forest [RF]

b) Naïve Bayes [NB].

c) Decision Trees [DT].

d) Logistic Regression [LR]

e) K Nearest Neighbors [KNN].

f) Support Vector Machine [SVM].

A comparison of the machine learning algorithms is shown in Fig. 8. We can see that Random Forest has performed the best. That is why we have trained the whole dataset with the Random Forest algorithm.

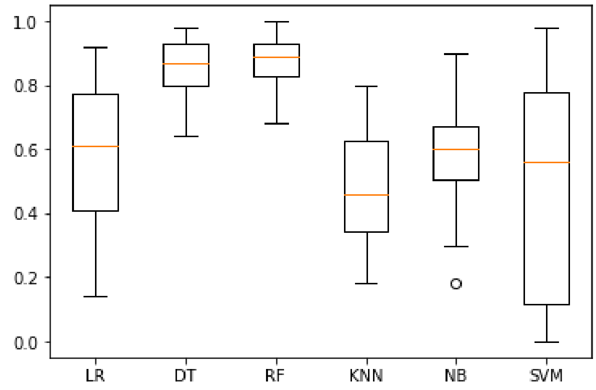

Fig. 8. Comparison of machine learning algorithms.

5) Detecting Healthy and Infected Leaves: Our main focus of this research is to implement edge computing concept in rice leaf disease detection. That is why after training our image classification model, we have exported the image classification model. Then we have transferred the image classification model to our Raspberry Pi. Then we have executed our program. We could detect healthy and infected leaves with our system without any hassle. We have shown in Fig. 9 and Fig. 10.

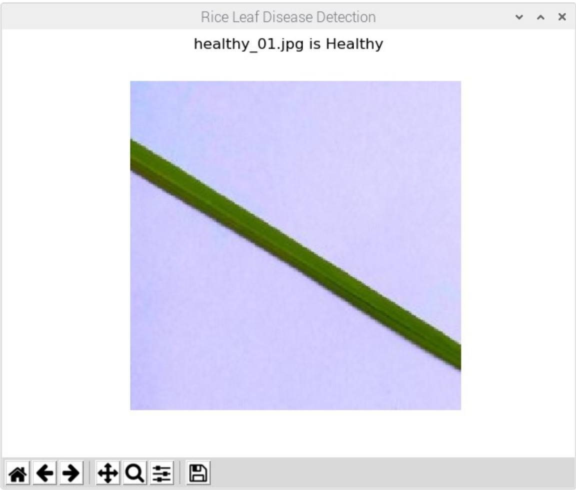

Fig. 9. Classified as Healthy

Fig. 10. Classified as Infected

V. Results and Discussion

A. Model Accuracy

After performing prediction on our test dataset, we achieved an impressive 97.50% accuracy. The classification report of our image classification model is given in Table I.

TABLE I. Classification Report

	precision	recall	f1-score	support
healthy	0.94	1.00	0.97	30
infected	1.00	0.96	0.98	50
micro avg	0.97	0.97	0.97	80
macro avg	0.97	0.98	0.97	80
weighted avg	0.98	0.97	0.98	80

The confusion matrix of our image classification model is shown in Fig. 11.

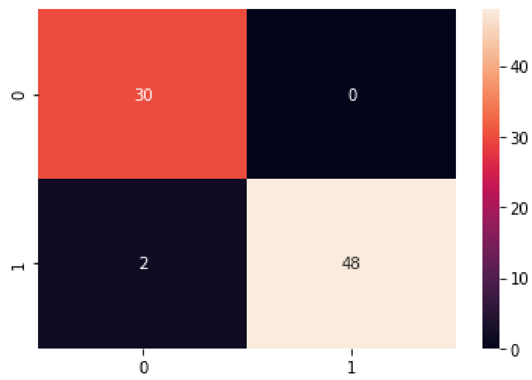

Fig. 11. Confusion Matrix

B. Raspberry Pi Performance

As a microcomputer, the performance we are getting from Raspberry Pi is pretty good. A total of 80 images were used to test the performance. We compared the performance of Raspberry Pi with a computer having a Hexa-Core 3.6GHz 64bit CPU and 8GB of RAM. Comparison of performance between Raspberry Pi and Computer is shown in Table II.

TABLE II. Performance Comparison

	Average time to classify one image
Raspberry Pi	0.20 sec
Computer	0.02 sec

C. Discussion

This paper has introduced the use of edge computing in agriculture. To the best of our knowledge, we are not aware of any other work in this field relating to edge computing. We are able to successfully identify healthy and infected leaves with the help of our system. We have got a good amount of accuracy from our image classification model. As an edge device, the performance of our system is promising. As we are classifying the images in our edge device, it will reduce the problems related to latency and connectivity.

We have randomly taken 30 images. Out of 30 images, only 4 images were the image of infected leaves. So, it can be said that our system can reduce approximately 86.66% of data transmission cost which will improve the latency.

VI. Conclusion and Future Scope

A. Conclusion

Rice leaf diseases can cause a significant amount of loss to the agriculture sector. This paper presents a system that will detect healthy and infected leaves automatically.

Many types of research have been done on detecting rice leaf diseases. But most of them use server-side data processing. The cost of a server is not cheap. That is why we have proposed a system that will classify the images on an edge device. We are using Raspberry Pi as our edge device, which is low in cost and power consumption. Our system can classify a leaf within 0.18 sec to 0.25 sec, and the accuracy of our image classification model is 97.50%. We are pleased with the performance of our system.

We hope our system will contribute a little to the advancement of the agriculture sector. The work has significant economic importance for Bangladesh.

B. Future Scope

Our current image classification model is built with machine learning. In the future, we are thinking of implementing deep learning for detecting leaf diseases. Currently, we are classifying all diseases as infected. Nevertheless, for future work, we plan to classify each disease by its name.

References

[1] "Rice in Bangladesh" Bangladesh Rice Knowledge Bank, http://www.knowledgebank-brri.org/riceinban.php. Accessed 02-01-2021.

[2] "Explaining Edge Computing" Youtube, uploaded by ExplainingComputers, 20-10-2019, https://www.youtube.com/watch?v=0idvaOCnF9E.

[3] "What is a Raspberry Pi?" https://www.raspberrypi.org/help/what-%20is-a-raspberry-pi. Accessed 04-01-2021.

[4] "Rice Disease and Its Managment" Bangladesh Rice Knowledge Bank, http://knowledgebank-brri.org/rice-disease-and-its-managment. Accessed 02-01-2021.

[5] "Brown spot" Bangladesh Rice Knowledge Bank, http://www.knowledgebank.irri.org/training/fact-sheets/pest-management/diseases/item/brown-spot. Accessed 02-01-2021.

[6] "Rice hispa" Bangladesh Rice Knowledge Bank, http://www.knowledgebank.irri.org/training/fact-sheets/pest-management/diseases/item/blast-leaf-collar. Accessed 02-01-2021.

[7] "Blast (leaf and collar)" Bangladesh Rice Knowledge Bank, http://www.knowledgebank.irri.org/training/fact-sheets/pest-management/insects/item/rice-hispa. Accessed 02-01-2021.

[8] Huy Minh Do. "Rice Diseases Image Dataset" Kaggle, 16-11-2019, https://www.kaggle.com/minhhuy2810/rice-diseases-image-dataset.

[9] "Random forest" Wikipedia, https://en.wikipedia.org/wiki/Random_forest. Accessed 15-01-2021.

[10] Alexander S. Gillis. "internet of things (IoT)" Tech Target, https://internetofthingsagenda.techtarget.com/definition/Internet-of-Things-IoT. Accessed 19-01-2021.

[11] By Eric Hamilton. "What is Edge Computing: The Network Edge Explained", https://www.cloudwards.net/what-is-edge-computing. Accessed 15-01-2021.

[12] 2 IBM Cloud Education. "Machine Learning" IBM, 15-07-2020, https://www.ibm.com/cloud/learn/machine-learning.

[13] G. Anthonys and N. Wickramarachchi, "An image recognition system for crop disease identification of paddy fields in Sri Lanka," 2009 International Conference on Industrial and Information Systems (ICIIS), Sri Lanka, 2009, pp. 403-407, doi: 10.1109/ICIINFS.2009.5429828.

[14] Reinald Adrian D. L. Pugoy and Vladimir Y. Mariano "Automated rice leaf disease detection using color image analysis", Proc. SPIE 8009, Third International Conference on Digital Image Processing (ICDIP 2011), 80090F (8 July 2011); doi: doi.org/10.1117/12.896494

[15] Q. Yao, Z. Guan, Y. Zhou, J. Tang, Y. Hu and B. Yang, "Application of Support Vector Machine for Detecting Rice Diseases Using Shape and Color Texture Features," 2009 International Conference on Engineering Computation, Hong Kong, 2009, pp. 79-83, doi: 10.1109/ICEC.2009.73.

[16] J. W. Orillo, J. Dela Cruz, L. Agapito, P. J. Satimbre and I. Valenzuela, "Identification of diseases in rice plant (oryza sativa) using back propagation Artificial Neural Network," 2014 International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), Palawan, 2014, pp. 1-6, doi: 10.1109/HNICEM.2014.7016248.

[17] S. Ramesh and D. Vydeki, "Rice Blast Disease Detection and Classification Using Machine Learning Algorithm," 2018 2nd International Conference on Micro-Electronics and Telecommunication Engineering (ICMETE), Ghaziabad, India, 2018, pp. 255-259, doi: 10.1109/ICMETE.2018.00063.

[18] X. Mai and M. Q. -. Meng, "Automatic lesion segmentation from rice leaf blast field images based on random forest," 2016 IEEE International Conference on Real-time Computing and Robotics (RCAR), Angkor Wat, 2016, pp. 255-259, doi: 10.1109/RCAR.2016.7784035.

[19] F. T. Pinki, N. Khatun and S. M. M. Islam, "Content based paddy leaf disease recognition and remedy prediction using support vector machine," 2017 20th International Conference of Computer and Information Technology (ICCIT), Dhaka, 2017, pp. 1-5, doi: 10.1109/ICCITECHN.2017.8281764.

[20] H. B. Prajapati, J. P. Shah, and V. K. Dabhi, “Detection and classification of rice plant diseases,” Intelligent Decision Technologies, vol. 11, no. 3, pp. 357–373, 2017, doi: 10.3233/IDT-170301.