Journal of Artificial Intelligence and Capsule Networks (ISSN: 2582-2012)
www.irojournals.com/aicn/

Automatic Disease Detection in the Rice Leaves Employing a Support Vector Machine

D. Kiruthika[1], N. Somasundaram[2], R. Tamilselvi[3], M. Parisa Beham[4]

ECE Department, Sethu Institute of Technology, Pulloor, Kariapatti, Tamilnadu, India.

Email: [1]keerthisanju195@gmail.com, [2]somukrish0714@gmail.com, [3]tamilselvi@sethu.ac.in, [4]parisabeham@sethu.ac.in

Abstract

Rice is a major agricultural crop around the world. Crop diseases, on the other hand, have the potential to reduce yield and quality greatly, posing a major danger to global food supplies. As a result, disease control is essential for rice production. Accurate and prompt disease diagnosis is critical to disease control success, which allows pesticide control measures to be implemented. The most common method for diagnosing rice leaf diseases is a manual decision-making based on disease appearance. There aren't enough skilled workers in the area, for such tasks to be completed on time. As a result, a more effective and convenient way of identifying rice leaf diseases is required. Therefore, this research creates an automatic diagnosis approach for rice leaf disease detection using deep learning. The proposed solution is built with deep learning techniques and a huge dataset containing 2,000 images of three types of rice diseases such as leaf blast, sheath blight, and brown spot, and healthy leaf. The proposed model's robustness is improved by using its real-world rice leaf datasets as well as publicly available online datasets. With an accuracy of 96.0%, the proposed deep-learning-based strategy proved successful in automatically diagnosing the three discriminative diseases of rice leaves. Furthermore, 99.25% of the time, the algorithm accurately detected a healthy rice leaf. The results demonstrate that the suggested deep learning model gives a highly effective technique for identifying rice leaf infections, and is capable of quickly and reliably identifying the most common rice diseases.

Keywords: Rice diseases, deep learning, leaf blast, sheath blight, sesame leaf spot.

DOI: https://doi.org/10.36548/jaicn.2023.2.006
Received: 12.04.2023, received in revised form: 12.05.2023, accepted: 27.05.2023, published: 06.06.2023

1. Introduction

Rice is essential in many countries' diets and is one of the world's most popular meals. It is estimated that over three billion people choose it as their primary diet. [1]. The crop diseases, pose a significant threat to the world's food resources by significantly reducing crop output and quality. As a result, disease prevention is critical to rice cultivation. [2]. Crop diseases endanger the world's food supply by significantly reducing crop output and quality. As a result, disease prevention is critical to rice cultivation [3]. The most frequent and deadly illnesses at the time are the leaf brown spot and leaf blast diseases, and sesame leaf spot and bacterial blight, on the other hand, are currently the most prevalent and deadly diseases [4]. There are various patterns and geometries among these three ailments.

- Rice Leaf Blast [5]: It is one of the most affecting diseases to grain yield. The small fragments on the leaves grow into spindly spots with an ashy centre. Older lesions are ellipsoidal or in the shape of a spindle, with whitish dark grey cores, with a reddish brown-colour border.

- Bacterial Sheath Blight [6]: The blight disease is caused by the bacterium Xanthomonas oryzae. The primary symptoms are yellowish-white or golden-yellow marginal necrosis found in a leaf, curling and drying back from the tip, and the midrib rib remaining intact.

- Brown leaf Spots [7]: The fungus Bipolaris oryzae causes sesame leaf spot disease. Rice sesame leaf spot is a fungal plant disease that typically affects the leaves and glumes of adult host plants, as well as plants, the sheaths stems, and grains.

In order to guide the farmers in identifying the disease in the early stages, the proposed system aims to develop an autonomous device that helps in detecting, classifying and diagnosing the diseases in the rice leaves. Rice leaf disease prediction is an application that detects and provides users with remedial measures for diseases in the crop. Using the deep learning-based KNN-Multi SVM algorithm [8], the image is processed for effective remedial measures. Once the leaf disease is identified and matched with existing data, effective remedial measures, such as what action should be taken in response to the disease, is provided. The database is trained with the proposed deep learning model in the first module, and the images are tested with the proposed model. In the second module, the proposed model provides some corrective measures, such as which fertilizers to use and what actions to take, by comparing it to the datasets.

2. Related Works

For decades, agricultural researchers have used computer vision technology to estimate crop yields, detect nutritional deficiencies in crops, estimate crop geometric sizes, and identify weeds. Crop diseases have been diagnosed using computer vision methods such as image processing, support vector machine, pattern recognition, object recognition and hyper spectral detection.

Bashir et al., [9] proposed classifying three rice crop diseases sesame leaf spot, bacterial leaf blight and false smuts using the SVM. The features were extracted using the BoW and the SIFT. For classification, clustering with K-means, a Brute-Force matcher, and Support Vector Machine (SVM) were used. The authors reported 94.16% average accuracy, 91.6% recall, and 90.9% precision. Joshi et al., [10] suggested that the rice disease control and monitoring could benefit from image processing. The shape and colour-based engineering elements were used. Also using standard classifiers for classification was recommended. The dataset of only 115 images was used to diagnose the disease, which was divided as 30% testing and 70% training. It was discovered that K-nearest Neighbour (k-NN) had an overall accuracy of 87.02%, while MDC had an accuracy of 89.23%.

Ahmed et al., [11] discussed machine-learning approaches for rice leaf disease. Classic classifiers Naive Bayes estimation (NB), Logistic Regression (LR), a Decision Tree (DT), J48 DT, and K-NN were among the methods developed. When using the J48 DT, an accuracy of 97.9% was reported. Research [12] suggested k-means for partitioning the affected leaf part and extract characteristics based on texture, colour and shape. SVM is the mean accuracy on the training set was 93.43%, and the mean accuracy on test data was 73.33%. Shrivastava et al., [13] used colour features to identify rice plant diseases. The method was then tested with seven different classifiers, Logistic regression model, Random Forest, K-NN, Navies Bayes, SVM, and discriminate classifiers. With SVM, the model achieved the highest level of accuracy, with a mean precision of 94.65%.

Bari et al., [14] proposed an improved RCNN- region-based CNN for real-time identification of rice leaf disease. The adoption of a localised proposal network improves the suggested Faster R-CNN. As a result, R-CNN can detect the object precisely and produce probable regions. The model achieved 99.17%, 98.85%, and 98.09% efficiency rates, respectively. Sowmyalakshmi and colleagues [15] proposed the CNNIR-OWEN-based method for classifying rice diseases, which was relate to as Residual Networks (ResNet) v2, and an ideal weight extreme learning machine (WELM) was used in a new CCNN-based inception. The advanced technology integrated the Internet of Things for image capture and histogram segmentation for disease detection. Then, using ResNet v2's deep learning inception [15], features were extracted. The Flower Pollination Algorithm (FPA) was used to tune the WELM for classification. On average, a 94.2% accuracy rate was achieved.

The preceding makes the case for developing a method that may help rice farmers in early identification of rice illness, thereby increasing production and quality while decreasing costs. The prime objectives of agricultural technology research are to increase output. A justification has been presented for what differentiates rice growers from other agricultural workers, necessitating researchers to pay close attention to rice illnesses and promote early detection and prevention. The three billion people worldwide who use rice, in addition to various other crops that are similarly or even more vital, require research in this subject. The main purpose is to create a device that incorporates innovative Deep Learning (DL) algorithms to recognize, classify, and diagnose rice disease. The goal is to suggest innovative system that could increase the accuracy of diagnosis over current techniques that use comparable databases or datasets.

3. Proposed Methodology

A novel deep learning model for categorising rice leaf disease is developed. Figure 1 presents how the proposed method can differentiate three classes: brown spot, leaf blast, and bacterial leaf blight and the healthy leaf. The proposed deep learning method strategy includes pre-processing stages that include noise removal, resizing of the image, and data augmentation process to increase the dataset's size. Although the potential over fitting issues are not addressed, the literature review reveals that the majority of research in the body of existing methods state that only small datasets are used in the literature, which results in over fitting. The GLCM technique is then used to extract the characteristics. CNN employs flatten, dense, and softmax layers for feature reduction. The algorithm used for classification is Multi-SVM. Figure 1 depicts the proposed flow diagram.

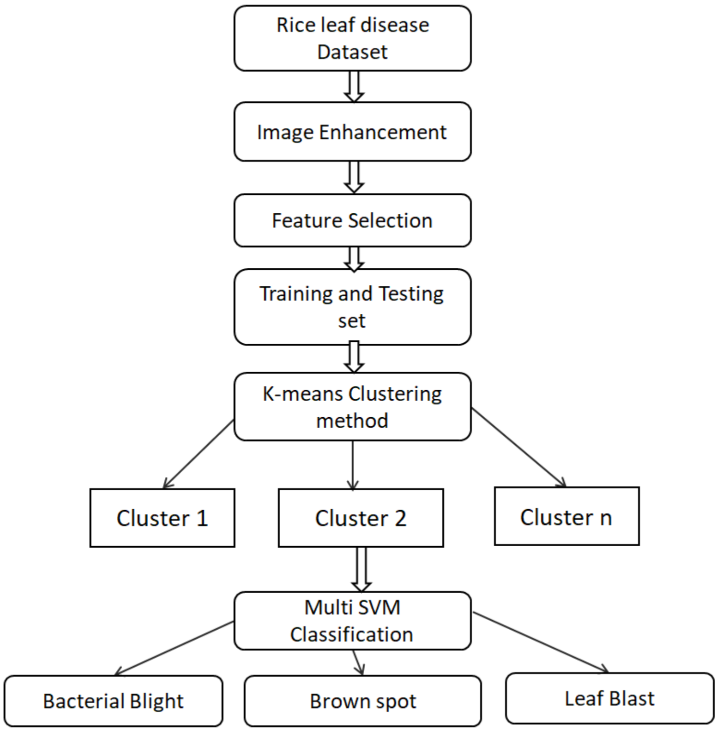


Figure 1. Block Diagram

3.1 Steps in Pre-processing

i) Image Resizing: The acquired image may be of varied sizes in practical uses. It is strongly advised that the input images must be resized in the range of 256*256 pixel values. The resizing method employs bilinear interpolation, which averages the adjacent four nearest pixel values to determine the new pixel value.

$$P = Q_1 + Q_2 + Q_3 + Q_4 \qquad [1]$$

Where "P" is denoted as the new pixel value, and the adjacent pixels are referred as Q_1, Q_2, Q_3, and Q_4.

ii) Contrast Enhancement: Contrast enhancement is used to distinguish between an image's object and background. The two kinds of "contrast enhancement" techniques are linear and nonlinear. Contrast stretching, a piecewise "linear transformation" function, is used in this work. The points (r1, s1), and (r2, s2) decide the transformation structure and produce a range of the level of intensity in output image.

iii) Conversion of Color Space: This conversion is highly required as the RGB color space is device dependent and used in computer monitors. It is also affected by the amount of light. This process is simple to perform. The steps to be followed are creating a transformation structure and converting the image by applying it. Here, the L*a*b color model is used as it performs better in separating the chrominance information.

iv) K-means Clustering: This is utilized to segment the color in a*b space. The number of clusters is fixed. When k = 3, the proposed model forms three clusters: the diseased part, the normal green leaf, and the color of the background. The centroid of each cluster is predetermined. Every data point is associated with the proper centroid. The initial stage of sorting is completed after all of those data points have been computed. In subsequent steps, new centroids are determined from the previous step. These new centroids are now used for data point matching. This is repeated in a loop until k-centroid's locations cannot be changed. This algorithm is intended to minimize the error functions.

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^{(j)} - c_j \right\|^2 \qquad [2]$$

Where, J is the error function, $x_i^{(j)}$ is a data point. The cluster centre is defined as the square of the distance between the cluster and datapoint. The algorithm's flow is explained in the below steps:

- The primary k centroids are positioned in the clustered objects' space.
- The centroid nearest to each data point is selected and the data points are connected with that k-centroid.
- A novel collection of k-centroid points is computed.
- The above steps must be repeated until new points are to be discovered. Following that, the input image is originated based on the clusters. For the feature extraction, the exact cluster is chosen in which the diseased segment region is presented.

v) Feature Extraction Process: This process helps to reduce the resources required to accurately represent a large set of data. The classification method performs poorly when faced with a high number of variables. This texture analysis can reveal certain distinct and underlying characteristics. To compute statistical texture features, the Grey Level Co-occurrence Matrix (GLCM) is used. Contrast, Correlation, Energy, and Homogeneity are the features extracted for image analysis. Based on the demand in the output image, the statistical parameters are used. The co-occurrence matrix element p (i,j) is considered. In other words, it indicates the probability of a pixel shifting from a grey level of i to a grey level of j. The equations below can then be used for describing the GLCM parameters.

1) Contrast: It describes colours particularly the relationship that exists between image's darkest and brightest areas.

$$\text{Contrast} = \sum_{i,j} |i-j|^2 p(i,j) \qquad [3]$$

2) Correlation: It is the process of sliding a filter mask, also known as a kernel, over an image and measuring the sum of products at each position.

$$\text{Correlation} = \sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i \sigma_j} \qquad [4]$$

3) Energy: It returns the sum of the GLCM's squared elements. Range = [0 1] for a constant picture, the energy is 1. The possession Energy is sometimes referred to as uniformity, energy uniformity, and angular second moment.

$$\text{Energy} = \sum_{i,j} p(i,j)^2 \qquad [5]$$

4) Homogeneity: It eeturns a value indicating how near the distribution of items in the GLCM is to the GLCM diagonal.

$$\text{Homogeneity} = \sum_{i,j} \frac{p(i,j)}{1+|i-j|} \qquad [6]$$

5) Entropy: Entropy quantifies the randomness of intensity values in a neighbourhood.

$$\text{Entropy} = -\sum P_i \log_2 P_i \qquad [7]$$

6) Average intensity values: The area R is considered, where a [m, n] indicates the brightness of the pixels. Taking a sample mean of this over pixels, results in the average intensity m_a in that location.

$$m_a = \frac{1}{\Lambda} \sum_{(m,n)\in R} a[m,n] \qquad [8]$$

7) Standard Deviation: R is the pixel intensity in an image's region. The sample standard deviation (S_a) is an estimate of the standard deviation of this brightness and is given by,

$$s_a = \sqrt{\frac{1}{\wedge-1}\sum(a[m,n] - m_a)^2} \quad [9]$$

The statistical characteristics previously described are also collected from the diseased region of the leaf.

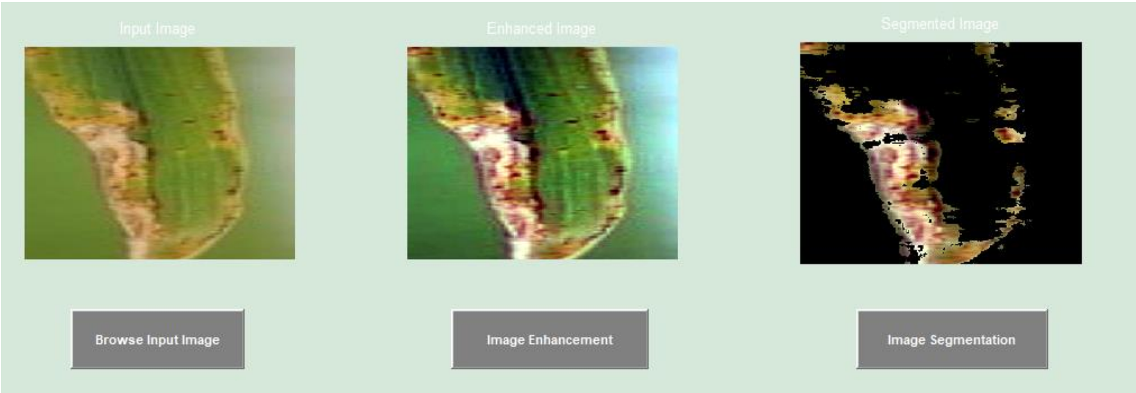
3.2 Multi class Support Vector Machine – Classification Process

"SVM" is a supervised classification algorithm. Based on the example image, the supervised classification algorithm provides training data with labels. In this work, each disease is regarded as a different class, and labels are allocated accordingly. This data for training is a set of examples that will be utilised to create a database for the purpose of learning. Each training set is inspected and an output is created in supervised classification. The labels for the class should be applied to the test data once it is provided. SVMs are efficient and precise even with limited training samples. SVM also beat Neural Networks in classification, according to previous study. As a result, SVM is used in this work to conduct supervised classification.

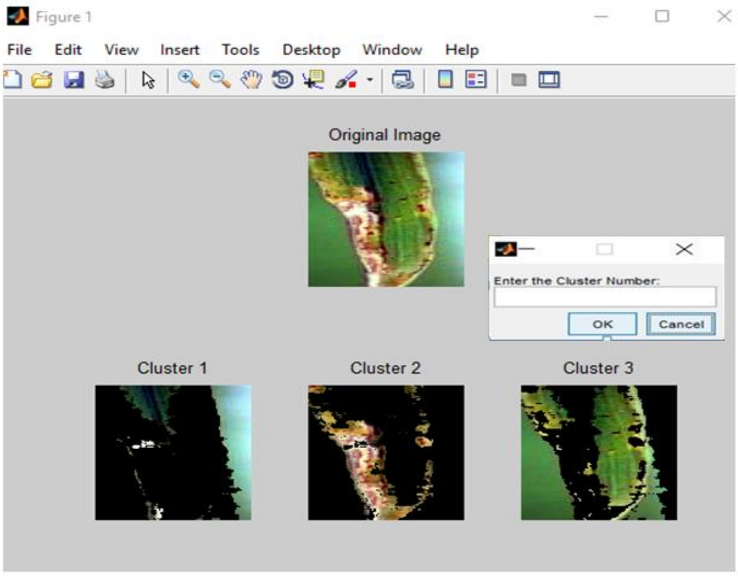
SVMs are binary classifiers used for multiple classifications. For this multiple categorizations, one-versus-one and one-versus-all approaches are applied. The division of "N classes into N two-class machines is included in 1AA. Classification would be conducted by comparing class 1 to non-class 1 (class 2,3), class 2 to non-class 2 (class 1,3), or class 3 to non-class 3 (class 1,2), as illustrated in Fig. 3. The other technique 1A1 requires constructing N(N 1)/2 classification combinations. Because of its simplicity, 1AA is utilised in disease classification. This work is divided into three categories: leaf blast, sheath blight, and brown spot. In addition, another category for healthy leaves is also chosen.

4. Results and Discussion

The initial step is to load the input image (the diseased leaf of rice). It is improved by carrying out pre-processing operations. The input image has been resized and the image enhancement process has been done, as shown in Fig. 2. The RGB image is converted into lab color space by color space conversion. K denotes clustering and k = 3 results in three clusters: backdrop, leaf part, and diseased leaf section, as shown in Fig.2. The diseased cluster is chosen, and the disease is classified using characteristics. This disease is identified by using multiclass SVM classification.



a) MATLAB GUI Results



b) Applying K-Means Clustering Algorithm to the Segmented Image

Figure 2. Obtained Threshold Output and Selective Cluster Output of K-means Clustering

Data is projected from the input area to feature space using the kernel function. This work employs the RBF kernel. The leaf disease has been identified and is shown in Fig. 3. Table 1 summarizes the four key features extracted from the GLCM.

Table 1. Features Extracted using GLCM Algorithm

Leaf disease name	Contrast	Correlation	Energy	Homogeneity
Rice blast	0.466189	0.764235	1.946578	0.965425
	0.823086	0.923497	0.600498	0.929365
	0.726493	0.831754	0.534931	1.915442
	1.459754	1.949783	0.513287	0.927489
	0.981728	0.319754	0.698345	0.942166
Sheath blight	0.566489	0.965489	0.549312	0.941247
	0.465871	0.945662	0.642183	0.942158
	0.471140	1.913547	0.654916	0.953448
	1.706945	0.924578	1.648798	0.943281
	0.913784	0.873452	0.788345	1.934612
Leaf Spot	0.542431	0.958874	0.521497	0.943221
	0.421578	0.943125	0.654257	0.943665
	0.746312	0.921537	0.653983	0.934162
	0.403983	0.943154	0.649871	0.943622
	0.925412	0.846218	0.786325	0.953478
Healthy leaf	1.462541	0.865821	0.433414	0.893001
	0.843254	0.831453	0.428656	0.879583
	0.496321	0.954399	1.380498	0.921547
	1.544872	1.948414	0.468722	1.913325
	1.487512	0.745488	0.632541	0.910025

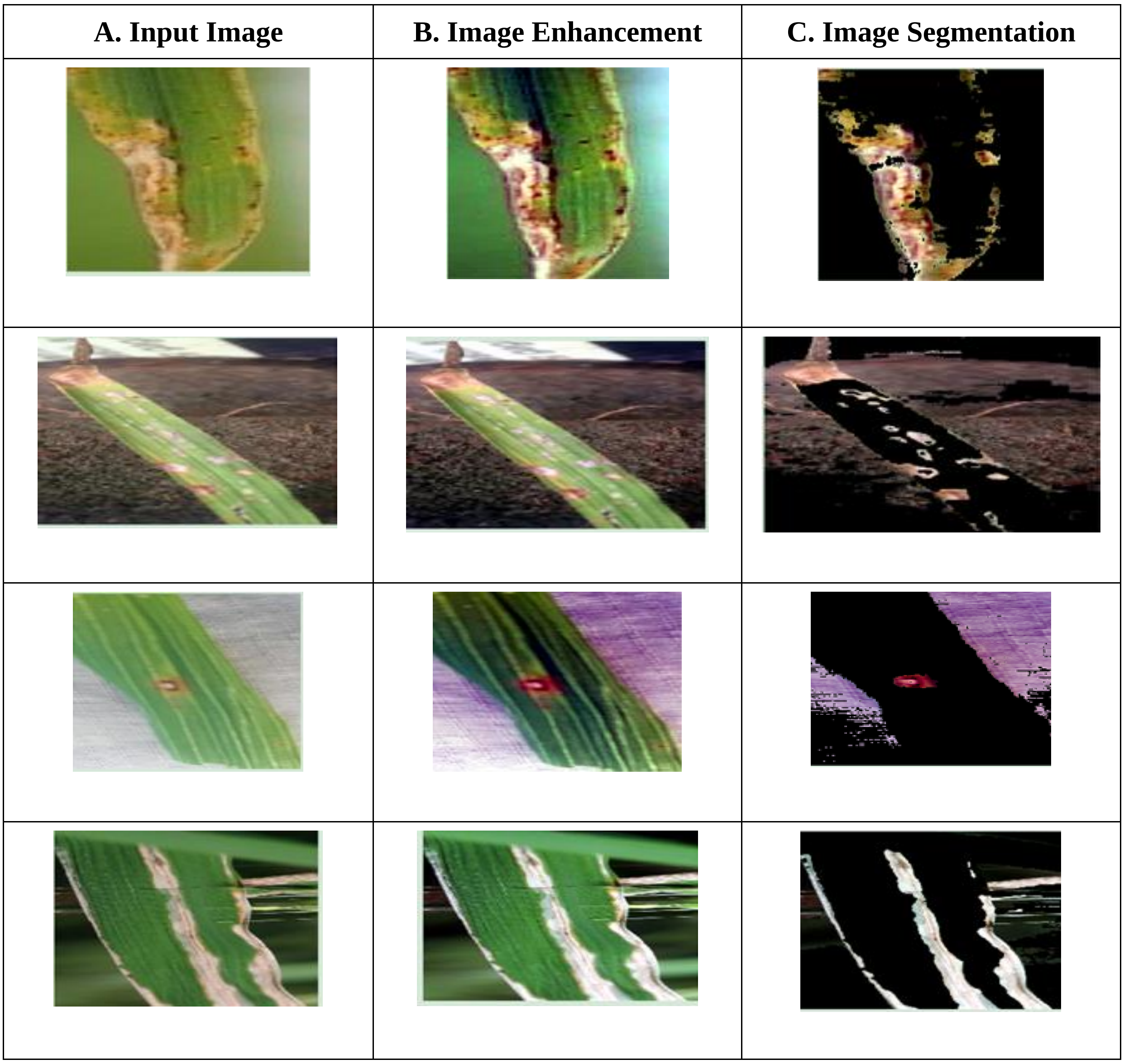


Figure 3. A) Input Image B) Pre-Processing Steps C) Segmentation Process

Figure 3 shows the results the leaf disease segmentation and classification methods. ROI and the L*a*b color model were extracted using KNN clustering algorithm. For pre-processing, the color median filter was used, and color texture characteristics were collected for extracting features. The multiclass SVM was utilized to diagnose leaf illnesses, and the outcomes will assist agriculturalists in disease classification. Additional real-time disease images will be collected in the future, as well as agricultural application software would be implemented.

5. Conclusion

In the proposed research, a novel deep learning approach is presented for accurately diagnosing three classes of diseased rice leaves and the healthy rice leaf. From the leaf data, three rice diseases are appropriately recognized. In addition to healthy leaves, the database for rice leaves contains three diseases: bacterial leaf blight, leaf blasts, and brown spots. The rice leaf disease is segmented using the Multiclass SVM classifier and employing Grey Level Co-occurrence Matrix (GLCM) texture features for classification. The quantitative performance of the proposed system is determined by measuring GLCM parameters. The suggested multiclass SVM on rice leaf disease detection shows 96.08% accuracy rate, and outperforms some of the prevailing methods. Further in the future, the study aims to classify many more rice leaf diseases in different plants and crops, and a neural network will be used to increase the classification accuracy.

References

[1] Mishra, R.; Joshi, R.K.; Zhao, K., “Genome Editing in Rice: Recent Advances, Challenges, and Future Implications”, Front. Plant Sci.9,(2018):1359-1361.

[2] Sethy, P. K., Barpanda, N. K., Rath, A. K., and Behera, S. K., “Image processing techniques for diagnosing rice plant disease: a survey”, Proc. Comput. Sci. 167, (2020): 516–530.

[3] Zhang, J.; Huang, Y.; Pu, R.; Gonzalez-Moreno, P.; Yuan, L.; Wu, K.; Huang, W. “Monitoring Plant Diseases and Pests through Remote Sensing Technology: A Review”, Comput. Electron. Agric.165, no. 104943(2019):250-257.

[4] S. Miah, A. Shahjahan, M. Hossain, and N. Sharma, “A survey of rice diseases in bangladesh,” International Journal of Pest Management.31, no. 3, (1985):208–213.

[5] Kapil Simkhada, “Rice Blast, A Major Threat to the Rice Production and its Various Management Techniques”, Turkish journal of Agriculture-Food science and technology.10, no.2(2022):12-26.

[6] Mohsen Mohamed Elsharkawy, Raghda M. Sakran, “Induction of Systemic Resistance against Sheath Blight in Rice by Different Pseudomonas Isolates”, National center for biotechnology information.12, no.3(2022):349.

[7] S. Sunder, Ram Singh, "Brown spot of rice: an overview", Indian Phytopathology.67, no.3(2014):201-215.

[8] J. S. Raikwal, Kanak Saxena, "Performance Evaluation of SVM and K-Nearest Neighbor Algorithm over Medical Data set", International Journal of Computer Applications.50, no.14(2012):35-39.

[9] Bashir, K.; Rehman, M.; Bari, M. "Detection and classification of rice diseases: An automated approach using textural features", Mehran Univ. Res. J. Eng. Technol. 38(2019):239–250.

[10] Joshi, A.A.; Jadhav, B.D. "Monitoring and Controlling Rice Diseases Using Image Processing Techniques", In Proceedings of the 2016 International Conference on Computing, Analytics and Security Trends (CAST).10(2016):19–21.

[11] Ahmed, K.; Shahidi, T.R.; Irfanul Alam, S.M.; Momen, S. "Rice Leaf Disease Detection Using Machine Learning Techniques", In Proceedings of the 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), Dhaka, Bangladesh,12(2019):24–25.

[12] Shrivastava, V.K.; Pradhan, M.K. "Rice Plant Disease Classification Using Color Features: A Machine Learning Paradigm", J. Plant Pathol. 103(2020):17–26.

[13] Bari, B.S.; Islam, M.N.; Rashid, M.; Hasan, M.J.; Razman, M.A.M.; Musa, R.M.; AbNasir, A.F.; Abdul Majeed, A.P.P. "A Real-Time Approach of Diagnosing Rice Leaf Disease Using Deep Learning-Based Faster R-CNN Framework", Peer J Comput. Sci. 7(2021):30-36.

[14] Sowmyalakshmi, R.; Jayasankar, T.; AyyemPiIllai, V.; Subramaniyan, K.; Pustokhina, I.V.; A. Pustokhin, D.; Shankar, K. "An Optimal Classification Model for Rice Plant Disease Detection" Comput. Mater. Contin. 68(2021):1751–1767.

[15] Xiayoue xie, Yuan ma, "A deep learning based real time detector for grape leaf disease using improved CNN", frontiers in plant science.751, no.11,(2020):5-14.

Author's biography

D.Kiruthika is a Junior Research Fellow, in the Department of Electronics and Communication Engineering at Sethu Institute of Technology. She has B.E degree in ECE and M.E degree in Communication systems. Her research involves the challenges and solutions for latest agricultural problems. Her major research domain is Bio-medical signal and Image processing.

N. Somasundaram is a Project assistant, in Sethu Institute of Technology. He has B.E degree in ECE and M.E degree in Embedded System Technologies. His research interest includes Agro-based and Bio-medical related projects. His major research domain is Signal and Image processing.