Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 7, July 2021:8423 - 8440

Research Article

Multi-model Approach for Grading & Classification of Paddy Leaf Diseases

Anoop G L[1*], Nandini. C[2]

Abstract

Paddy is one of the major crops in world. The yield of the crop is directly related to type of disease affected and level of infection. Detecting and classifying diseases automatically or semi-automatically will help the farmers to take necessary steps and to improve the crop yield. In this paper, we are proposing a system to classify and grade the level of diseases affected to paddy leaves. The proposed system consists of four major steps. Initially in the first step, the system extracts leaf part from input image by eliminating background in L*a*b color space. In second step, proposed an algorithm to segment and binarize the input image in YCbCr color space, this algorithm extract the disease part and convert to binary image for further process. In third step, extracted features like mass, margin, shape, texture and color features based on the paddy leaf disease characteristics. These extracted features are fed to support Vector Machine (SVM) classifier to classify the paddy leaf diseases like, rice blast, brown spot, leaf blight, hispa and healthy leaves. In step four, Fuzzy Inference System (FIS) is used to grade the level of disease by considering infection percentage and maximum radius of disease part by setting 9×6 fuzzy rules. Finally, the performance of proposed segmentation & binarization algorithm is evaluated by extracting GLCM features and also by extracting proposed features and achieved 95.12% & 96.54% efficiency respectively. The similarity percentage and disagreements between manual prediction and Fuzzy Inference System to grade the level of disease affected is 90.66%, and 9.34% respectively.

Keywords: L*a*b, YCbCr, Paddy leaf disease, SVM, Fuzzy Inference System (FIS).

1. Introduction

Paddy is one of the important crops in world. India is the second largest producer of rice [19] & [20]. The weather conditions, malnutrition, diseases and pests are major impact on yield and also the planning of harvest. The infectious disease spread and environmental damages causes the crop failure indirectly. The loss of crop yield affects the financial stage of farmers and country's economy.

[1]Research Scholar, VTU Research Resource Centre, Belagavi, Karnataka 590018, India.
[2]Department of CSE, Dayananda Sagar Academy of Technology & Management, Bengaluru 560082, India.
E-mail: [1*]gl.anoop1@gmail.com (Corresponding author), [2]laasyanandini@gmail.com.

Symptoms of diseases are usually manifested in rice leaves. Major leaf diseases of rice are rice blast, brown spot, leaf blight, hispa, smut etc. The rice crop yield is affected by 70-80%, 50%, 20-25% and up to 20% from diseases like rice blast, brown spot, leaf blight and hispa respectively. These diseases are in different shape, size and discoloration [12]. The detection and classification of paddy leaf disease is an essential research topic. The general image processing steps to detect leaf disease are, pre-processing, segmentation, binarization, feature extraction and classification [3].

Plant diseases are one of the underlying causes for the decrease in the number of quantity and quality of the farming crops. Nalini et al., [1] proposed a novel deep neural network (DNN) classification model for paddy leaf disease detection using image data. In this model the region of disease infected area was extracted using k-means clustering, thresholding technique has been applied to eliminate the uninfected regions. Gray Level Co-occurrence Matrix (GLCM) features are extracted from preprocessed and segmented images and used DNN & crow search algorithm in training and fine-tuning process to classify paddy leaf disease images.

Matin et al., [2] applied AlexNet technique with 5-convolutional layers and 3-fully connected layers to classify bacterial blight, leaf smut and brown spot rice leaf diseases and used image augmentation to increase the size of dataset. Ramesh and Vydeki [3] proposed Optimized Deep Neural Network with Jaya Optimization Algorithm (DNN_JOA) to recognize and classify paddy leaf disease like bacterial blight, sheath rot, blast and brown spot disease along with normal leaves. In preprocessing, the captured leaf images from the farm field will be converted from RGB to HSV color space, the disease and non-diseased parts are extracted by converting hue and saturation part to binary image. The background clustering is used to segment the disease, normal portion and background. The color and texture features (GLCM) are extracted and used DNN_JOA to classify the paddy leaf diseases.

Satgunalingam and Thaneeshan [4] presented a methodology to develop an automatic system to detect brown spot, narrow brown spot and paddy blast diseases. The canny edge detection, region growing and multilevel thresholding techniques are used to segment the infected part in paddy leaf. Color, shape and GLCM features are extracted from segmented image and used multiclass support vector machine to classify the diseases. Satgunalingam and Thaneeshan [4] expressed the limitation of the proposed study regarding segmentation, the accuracy of segmentation in which overlapped and irregular boundary lesions.

Kumar et al., [5] presented a framework detect and classify rie leaf diseases like: leaf smut, brown spot and bacterial leaf blight using AlexNet model. Sreenivasulu and Pedda [6] proposed a system to give defense mechanism against diseases for plants using Convolution Neural Network (CNN) to predict the diseases. Image database is obtained from internet and segregated [6]. Swetha and Shravani [7] proposed a method to diagnose and classify rice diseases like: smut, tungro of rice, rice blast and blight. In this proposed method shape and color features are features from segmented images and classified diseases using Support vector machine (SVM) and k-Nearest Neighbor (k-NN) classifiers. kumar et al., proposed system accuracy for SVM and k-NN classifiers are 91.2% and 89.54% respectively.

Nidhis et al., [8] proposed image processing methods to detect leaf disease affected to paddy leaves and also calculate the severity level of disease infected to leaves, uses k-means clustering for segmentations, point feature matching features are used to classify diseases and mass of white pixels are used to estimate severity level of diseases. K-means clustering is used for segmentation and GLCM features are extracted to detect the plant leaf diagnosis [9]. Pinki et al., [10] proposed an automated system for diagnosis of leaf blast, brown spot and bacterial blight paddy leaf diseases and fertilizers or pesticides are advised based on the severity of leaf diseases. K-means clustering is used to segment the disease part in image, SVM classifier is used to classify disease by extracting visual content features from segmented image.

Sharada et al., [11] uses public dataset that consist of 54,306 diseased and healthy leaves images collected under controlled conditions. Deep Convolutional Neural Network (D-CNN) with AlexNet and GoogLeNet deep learning architecture used for classification process, trained and tested for different distribution sets of image data. Reinald and Vladimir [12] proposed an automated system to detect paddy leaf diseases using HSI color space image analysis. Hue-Saturation histogram intersection between healthy and test leaf images to obtain the outlier region, and k-means clustering to cluster the related regions from outlier.

Identifying leaf structure, background elimination, segmentation and binarization and feature extraction makes huge impact on classification and grading the level of leaf diseases, i.e., disease part in leaves, feature selection and extraction makes direct impact on performance of classification, leaf area detection directly related to estimate the infection ratio inter to grade the level of disease. From survey it is noticed that, segmentation and binarization is limited to clustering, features extracted are also limited to GLCM and color features to classify paddy leaf diseases. In this paper we are proposing a system to classify and grade the level of paddy leaf diseases, it consists of four sections.

In section-1, we have proposed an algorithm to extract leaf by eliminating background in L*a*b color space, this process suppress the error or true negative in outer region of leaves while extracting the disease part from leaves. In section-2, proposed an algorithm to segment and binarize the paddy leaves to extract the disease part in leaves in YcbCr color space along with morphological operation, k-means clustering and logical operation. In section-3, selected mass, mass-margin, shape, texture and color features based on the paddy leaf disease characteristics, this helps to classify the diseases by using SVM classifier. In section-4, used Fuzzy inference system to grade the level of disease by considering infection percentage and maximum radius of disease part. The experimental result section consists of two parts, they are: dataset description and result analysis. In dataset description, the types of datasets, number of images belonging to different classes of diseases is described. In result analysis, evaluated performance of proposed segmentation & binarization with GLCM features and proposed features, tabulated the accuracy of proposed system and compared with existing approaches. Further the performance of fuzzy inference system and manual prediction of grading the level of paddy leaf diseases is evaluated, and described by using confusion matrix.

2. Proposed System

In proposed system, we have extracted leaf by elimination background, extracted disease part using proposed segmentation and binarization technique, extracted mass, margin, shape, texture and color features to classify paddy leaf diseases and grade the level of diseases infected by using fuzzy inference system. The complete architecture of proposed system is shown in Figure 1.

2.1. Leaf Extraction / Background elimination

Automatic plant recognition and disease analysis may be streamlined by an image of a complete, isolated leaf as an initial input. The proposed background elimination or leaf extraction is described in this section. The input images are in RGB color space, and it is resized to 256×256 pixels.

Let I (x, y, 3) be the input RGB image with size 256×256×3. Initially, convert input image I from RGB color space to L*a*b color space by applying color transformation structure shows in Figure 2(b), L*a*b color space is a 3-axis color system. Where L is lightness, a & b are the color dimensions shows in Figure 2(c), (d) and (e) respectively. After converting to L*a*b color space applied image subtraction between 'L' and 'b' components and it is represented as following equation, respective resultant image for sample input is showed in Figure 2(f).

$$L\_sub\_b = L - b$$

Further, eliminate the background pixels by filling value zero (Black color) in input image I (x, y, 3) as defined in Algorithm-1 and showed resultant image for sample input is showed in Figure 2(h). Figure 2(g) shows the boundary detection of leaf, i.e., left, right, top & bottom margin/boundary.

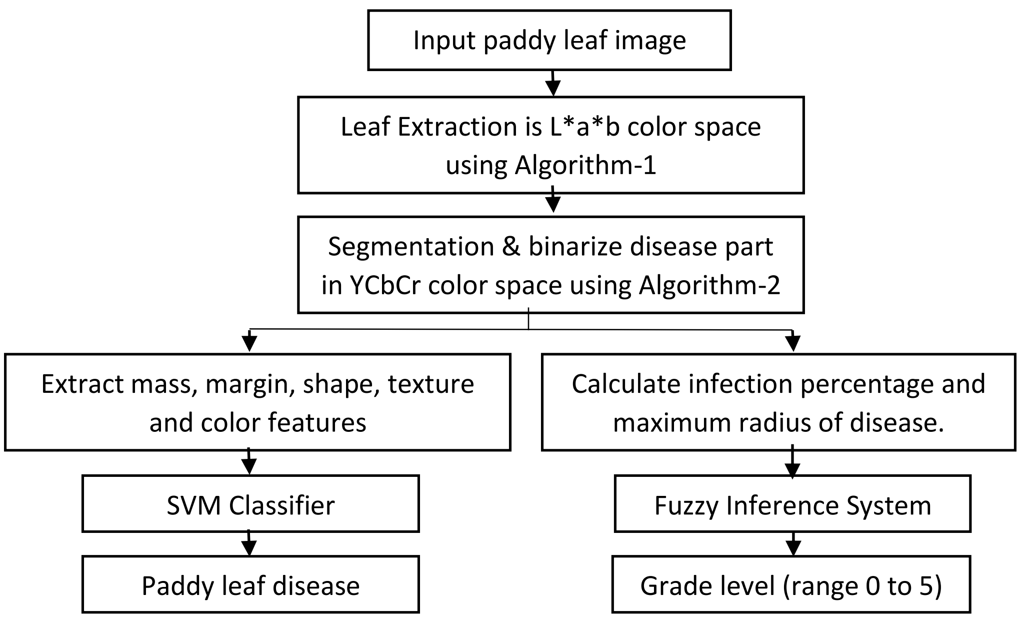


Figure 1: Proposed System Architecture.

Algorithm-1: Leaf Background Elimination(I)
//Input: RGB leaf image I (x, y, 3) with background of size 256×256×3, where x & y are co-ordinate values, i.e., number of rows and columns respectively.
//output: RGB leaf image I with background pixel values as '0'. (Background eliminated leaf image)

Lab= convert image I from RGB to L*a*b color space by applying color transformation structure.

Let L ← Lightness, 'a' & 'b' are color dimensions.

L_sub_b= L – b;

for each row r←1 to x

  for each column c← 1 to y

$$I(r,c,:) = \begin{cases} 0 & if(L\_sub\_b\,! = 0) \\ break & Otherwise \end{cases}$$

  for each column c← y down to 1

$$I(r,c,:) = \begin{cases} 0 & if(L\_sub\_b\,! = 0) \\ break & Otherwise \end{cases}$$

for each row r ← x down to 1

  for each column c← 1 to y

$$I(r,c,:) = \begin{cases} 0 & if(L\_sub\_b\,! = 0) \\ break & Otherwise \end{cases}$$

  for each column c← y down to 1

$$I(r,c,:) = \begin{cases} 0 & if(L\_sub\_b\,! = 0) \\ break & Otherwise \end{cases}$$

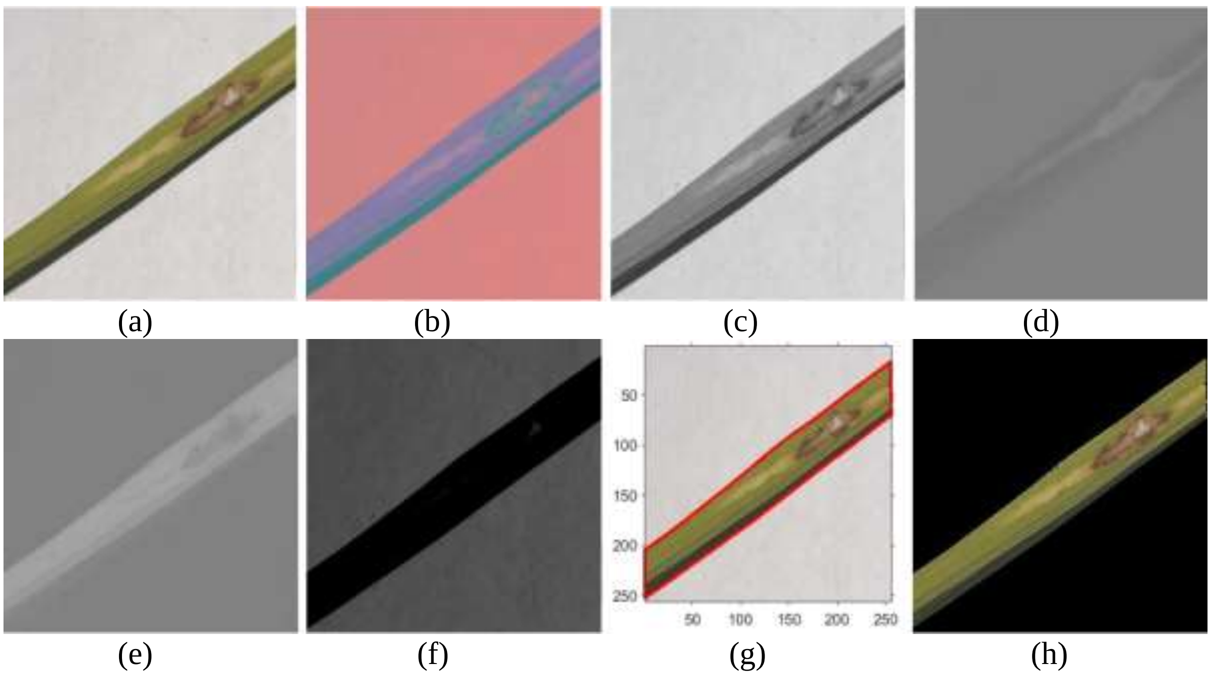


Figure 2: Leaf background elimination: (a) Resize RGB input image. (b) L*a*b color space image. (c) L-component. (d) a-component. (e) a-component. (f) L subtract b (L-b) image. (g) Left, right, top and bottom leaf boundary detection. (h) Leaf extracted by eliminating background by filling black color.

2.2. Segmentation and Binarization

For segmentation and binarization used morphological operations, k-means clustering, averaging and binary AND & OR operations on YCbCr color space as described in Algorithm-2, i.e., Segmentation and Binarization Algorithm. In proposed segmentation and binarization algorithm, initially converts the input image to YCbCr color space showed in Figure 3(b). ‘Y’ in YCbCr defined the luma component, Cb & Cr are the blue and red difference chroma components are showed in Figure 3(c), (d) and (e) respectively. The input image is in unit-8, so the range of Y is [16, 235], range of Cb & Cr is [16,240]. In proposed algorithm there are two ways of segmentation and binarization. One is by using morphological operations, another by using k-means clustering.

Finally, by performing AND operation both technique results are combined to achieve the final binarized image with Region Of interest (ROI).

In morphological processing, applied closing & opening for each component of YCbCr. i.e., Y, Cb and Cr. Further, subtract resultant of closing and opening. Suppress the background and segment the disease part in leaf by averaging result of subtracted components, then improve the intensity level by adjusting image intensity values. Further, binarized the images by applying OR & AND operation as showed in Algorithm-2. The morphological binarization and segmentation process is showed in Figures 3(f)(g)(h)(o) & (p). In k-means clustering technique, selected k-value as 5 and applied on YCbCr image is showed in Figure 3(i)-(m), by this, it segments the background and disease infected region. Further convert to binary image by selecting infected region cluster as showed in Figure 3(n).

Algorithm-2: Segmentation and Binarization (I(x, y))
//Input: Paddy leaf image I(x, y), where x & y are the co-ordinates that indicate row and column number of image.
//Output: Segmented and Binarized Image with Region of Interest (ROI) $bin_{ROI}(x, y)$ & bin_{canny} ← edge detected image in binary (mass margin image).
Step 1: Convert input RGB image I(x,y) to YCbCr color space.
Let Y ← luma component.
Cb ← blue difference chroma components and
Cr ← red difference chroma components.

Step 2: Create structure element as follows:

se1 = (square, 3)

se2 = (octagon, 9)

Step 3: For each component in YCbCr, apply morphological operations like closing and opening, perform arithmetic subtraction between closing & opening operations as shown below.

$$cl_Y = Y \bullet se1 \qquad cl_{Cb} = Cb \bullet se1 \qquad cl_{Cr} = Cr \bullet se1$$

$$op_y = Y \circ se2 \qquad op_{Cb} = Cb \circ se2 \qquad op_{Cr} = Cr \circ se2$$

$$MP_Y = cl_Y - op_Y \qquad MP_{Cb} = cl_{Cb} - op_{Cb} \qquad MP_{Cr} = cl_{Cr} - op_{Cr}$$

Step 4: Apply k-means clustering with k-value as 5 on YCbCr image, and convert selected cluster image to binary image. Let $Bin_{k\text{-}means}$ is the binary image evaluated by applying k-means clustering to YCbCr image.

Step 5: Calculate Average of three morphological images as shown below.

$$Avg_{YCbCr} = (MP_Y + MP_{Cb} + MP_{Cr})/3$$

Step 6: Binary conversion using Otsu's method. Let $bin_Y,\ bin_{Cb},\ bin_{Cr}$ & bin_{Avg} are the binary images converted from $MP_Y,\ MP_{Cb},\ MP_{Cr}$ & Avg_{YCbCr} images to binary image using Otsu's method respectively.

Step 7: The final Binary image bin_{ROI} is evaluated by using following equations:

$$bin_{YCbCr}(x,y) = \begin{cases} 1 & bin_Y(x,y) = 1 \text{ and } bin_{Cb}(x,y) = 1 \text{ and } bin_{Cr}(x,y) = 1 \\ 0 & Otherwise \end{cases}$$

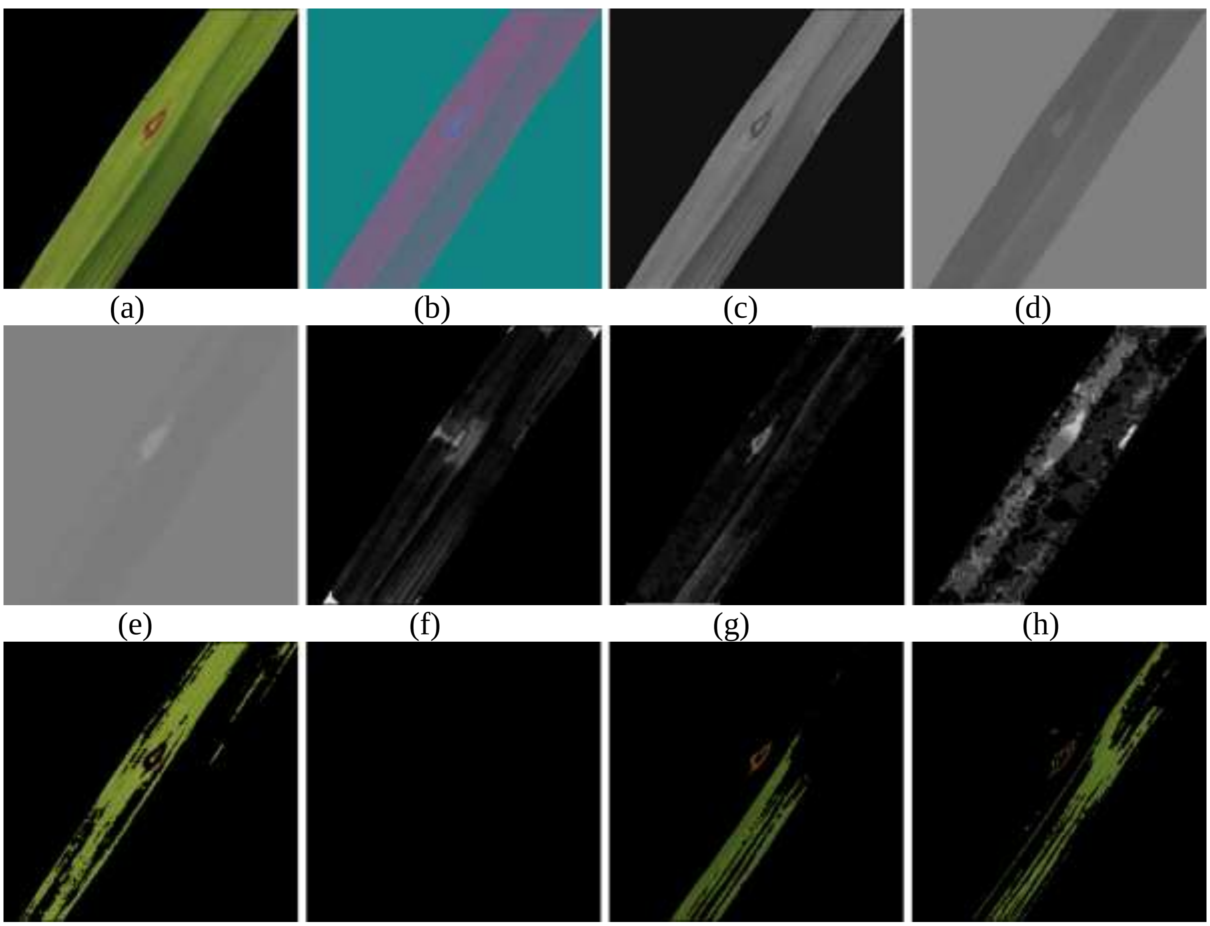
$$bin_{YCbCr_{Avg}}(x,y) = \begin{cases} 1 & bin_{YCbCr}(x,y) = 1 \text{ or } bin_{Avg}(x,y) = 1 \\ 0 & Otherwise \end{cases}$$

$$bin_{ROI}(x,y) = \begin{cases} 1 & bin_{YCbCr_{Avg}}(x,y) \text{ and } bin_{k-means}(x,y) \\ 0 & Otherwise \end{cases}$$

Apply canny edge detection technique to extract the margin of disease part.

$$bin_{canny} = canny(bin_{ROI})$$

Finally, AND operation is applied on binary image for morphological processing and k-means clustering technique to obtain the final binary image with ROI showed in Figure 3(q). Canny edge detection is used to extract the margin of disease mass is shown in Figure 3(r), the leaf disease part in RGB color space is shown in Figure 3(s). Connected component analysis is used to extract the disease part in input image, by creating bounding box as shown in Figure 3(t).



(a) (b) (c) (d)

(e) (f) (g) (h)

(i) (j) (k) (l)

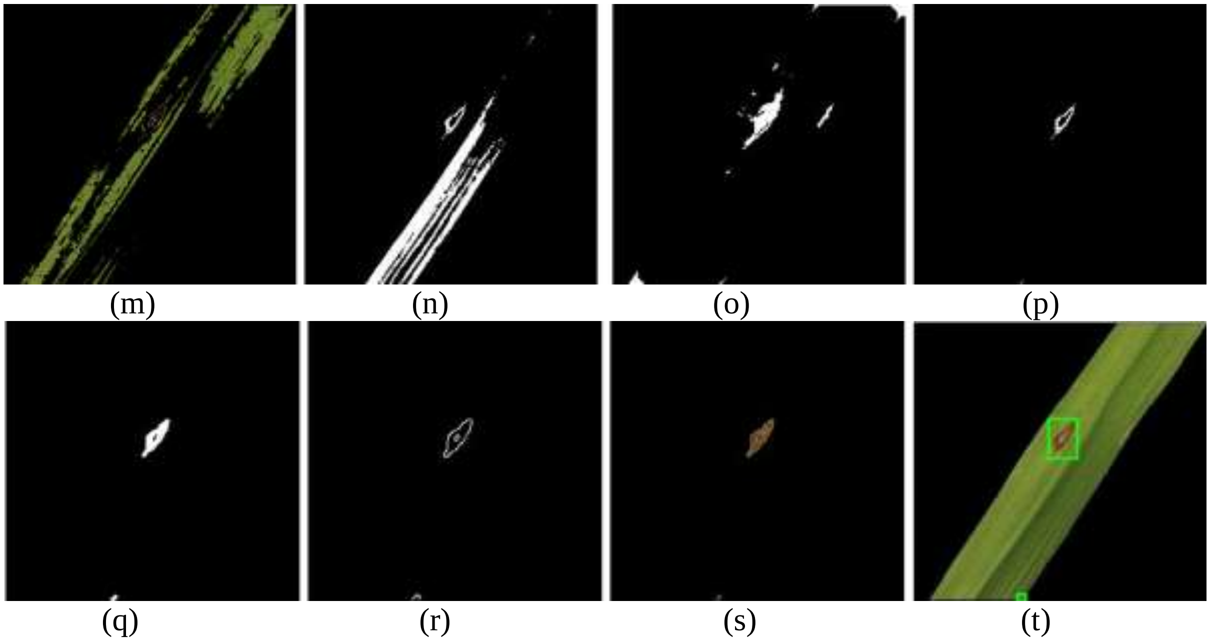


Figure 3: Segmentation and binarization process. (a) Input image. (b) YCbCr. (c) Y-component. (d) Cb-component. (e) Cr-component. (f) MP_Y (g) MP_{Cb} (h) MP_{Cr} (i) to (m) k-means clustering images (5-clusters). (n) Binary of k-means selected cluster. (o) Avg_{YCbCr} (p) bin_{YCbCr} (q) Finale dilated binary image. (r) Canny edge detection. (s) ROI in RGB color space. (t) Detected disease part in leaf.

2.3. Feature Extraction & Classification using SVM Classifier

Feature extraction is one of the important parts in classification of paddy leaf diseases. The general shape of masses and color in paddy leaf diseases are visually distinguishable. These characteristics can be used as a feature to classify paddy leaf diseases using SVM classifier. In Figure 4, showed some common shape and margin features of masses taken.

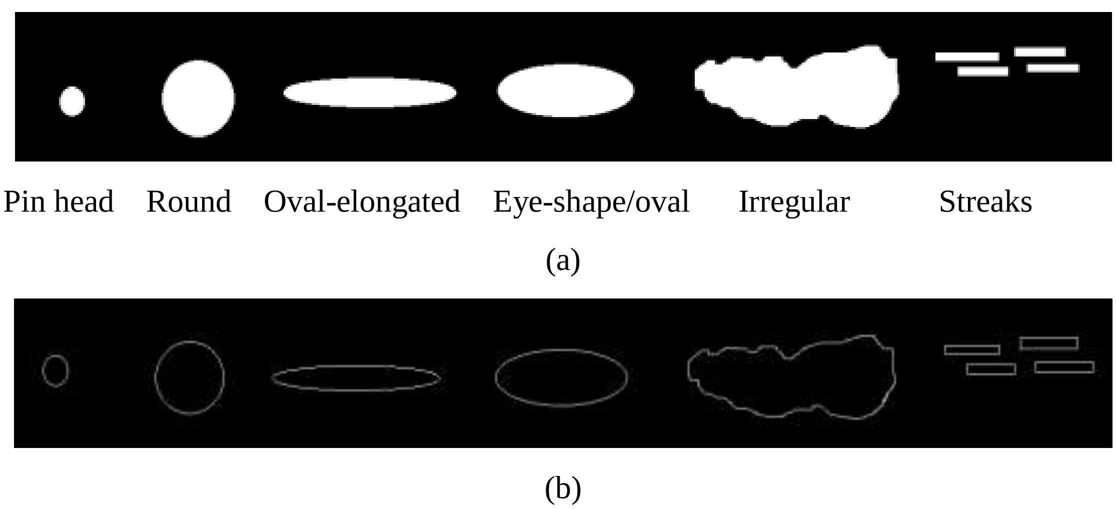


Figure 4: Common shape and margin of paddy leaf disease (Rice blast, brown spot, leaf blight and hispa). (a) mass shapes (b) mass margin

The paddy leaf diseases can be discriminated by observing the size, shape and color characteristics. The paddy leaf blast is pin head size, circular and oval elongated in light brown color. The brown spot disease in paddy leaf is circular and eye shaped in dark brown color. Leaf bight is irregular in shape with linear yellow color. Hispa disease in paddy leaf is white streaks parallel to the midrib.

In this section, we are proposing global and local feature extraction to classify the diseases. The global features extracted from entire image, where as local features are extracted on ROI detected by bounding box. Color features, eccentricity and solidity are extracted as a local feature, whereas remaining features listed in Table 1 are extracted as a global feature.

Area, circularity, dispersion, eccentricity, elongatedness, Equivalent diameter, solidity, maximum & minimum radius, orientation, standard deviation and total number of connected components (bounding boxes) are the features extracted as shape & mass characteristics from segmented disease mass. Similarly, from mass margin, extracted features like: contrast, correlation, entropy, perimeter, skewness, shape index and standard deviation. The color features are mean and standard deviation in RGB and Lab color space. Segmented color image, bin_{ROI} & $bin_{canny}$are the input images to extract features. The various features used to classify paddy leaf disease, their expressions and the impact of features on diseases are listed in Table 1.

In this proposed feature work, 21 features are extracted based on mass, margin shape, texture and color, these features are effective in classification of paddy leaf disease. The extracted features for sample images are present in Table 2.

Table 1. Mass, margin, shape, texture and color features

Sl. No.	***Features***	***Expression***	***Impact***
1	Area (A) / Total Mass	Total number of pixels in mass	Hispa masses are smaller in size, leaf blast is bigger in size
2	Number of Connected Components (TBB)	Number of Bounding Boxes in mass image	The number of brown spots in a leaf are more in general compared to other disease
3	Perimeter (P)	Total pixels in edge border of mass	To differentiate the regular and irregular polygons. Perimeter, maximum and minimum radius will vary for irregular polygon compared with regular polygon.
4	Maximum radius (Max_R)	Maximum distance from center to boundary of mass	
5	Minimum radius (Min_R)	Minimum distance from center to boundary of mass	
6	Circularity as circle (C1)	$\sqrt{\frac{A}{\pi * {Max_R}^2}}$	Useful to differentiate circular or oval blast and brown spot masses.
7	Circularity as ellipse (C2)	$\sqrt{\frac{Min_R}{Max_R}}$	

8	Dispersion (Dp)	$\left(\frac{Max_R}{A}\right)$	This determines leaf blight irregular shapes
9	Eccentricity (Ecc)	$\sqrt{1-\frac{b^2}{a^2}}$ Where, a← Major axis (longest diameter of ellipse) b← Minor axis (shortest diameter of ellipse)	Uses to identify elliptical mass
10	Solidity (Sol)	$\left(\frac{Area}{Convex\ Area}\right)$	Determines the image object's roundness.
11	Equivalent Diameter (EqD)	$\sqrt{4*\frac{A}{\pi}}$	Differentiate round, oval from irregular shape or mass
12	Elongatedness (Eln)	$\left(\frac{A}{(2*d)^2}\right)$ where d← the maximum thickness of a holeless region.	Differentiate the regular oval and irregular masses
13	Entropy (Ent)	$\sum(p*log_2(p))$ Where p ← Normalized histogram counts	Identifies the amount of disorder and speculations present in mass.
14	Standard deviation of mass (σ_{mass})	$\left(\sqrt{\frac{1}{N}\sum_1^N \lvert x_i-\mu\rvert^2}\right)$ Mean $\mu=\frac{1}{N}\sum_1^N x_i$ Where, x← input image, N← Total number of pixels.	The dispersion of set of data from mean, higher the deviation, more spread apart the data
15	Standard deviation of edge (σ_{edge})		
16	Contrast (cont)	$\sum_{i,j}\lvert i-j\rvert^2\,x(i,j)$	Contrast differences distinguish the diseases
17	Correlation (corr)	$\sum_{i,j}\frac{(i-\mu i)(j-\mu j)\,x(i,j)}{\sigma_i\sigma_j}$	1 or -1 for a perfectly positively or negatively correlated image. Correlation is NaN for a constant image.
18	Shape Index (SI)	$\left(\frac{P}{2*\ Max_R}\right)$	Differentiate thin polygons from regular polygon
19	Skewness (Sk)	$\left(\frac{E\ (x-\mu)^3}{\sigma^3}\right)$ Where, E(x) represents the expected value of quantity x.	If skewness is negative, the data spreads out more to the left of the mean than to the

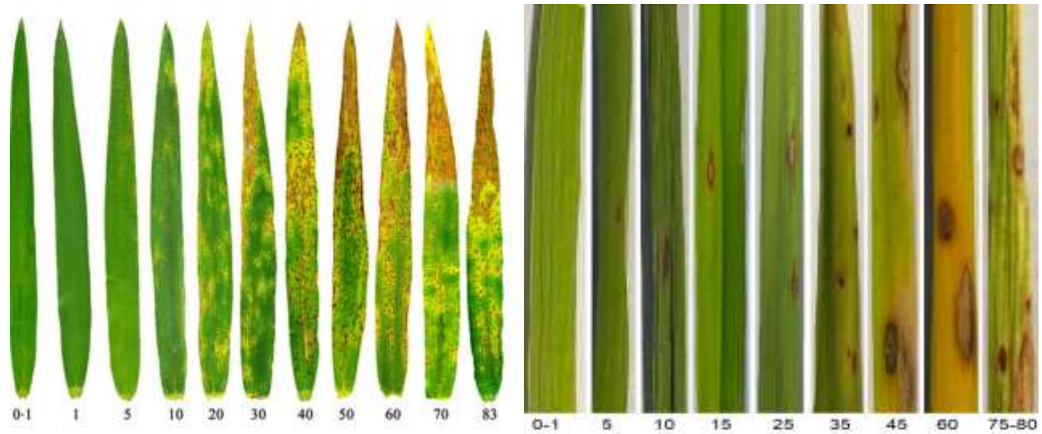
			right. If it's positive, the data spreads out more to the right.
20	Mean ($\mu - RGB \& LAB$)	$\left(\frac{1}{N}\sum_{1}^{N} x_i\right)$	Useful to classify disease, diseases have different color property
21	Standard deviation (σ) for RGB & LAB color space	$\left(\sqrt{\frac{1}{N}\sum_{1}^{N} \lvert x_i - \mu \rvert^2}\right)$	Diseases have different color property

Table 2. Feature measures for sample images

Features	Sample image1 (Rice Blast)	Sample image2 (Brown spot)	Sample image3 (Leaf blight)	Sample image4 (Hispa)
Area	861	458	1982	681
TBB	5	1	7	2
Perimeter	282	126	468	177
Max_R	148.3671	72.0101	154.7417	291
Min_R	28.5107	18.1833	28	62.6070
C1	0.062953	0.09459	0.091579	0.0285
C2	0.438364	0.50250	0.425378	0.4638
Dp	0.17232	0.157227	0.078074	0.4273
Ecc	0.857905	0.96447	0.881988	0.9701
Sol	0.352381	0.611	0.363636	0.2497
EqD	6.863663	12.964	6.770275	16.4682
Eln	0.009778	0.02208	0.020693	0.0020
Ent	0.186396	0.10310	0.382497	0.1366
σ_{mass}	0.076584	0.03467	0.042896	0.0491
σ_{edge}	0.049285	0.02188	0.029829	0.0327
cont	2317.232	1.3490e+03	4212.038	8.0499e+03
corr	0.614104	-0.90097	0.005797	-0.9947
SI	0.950345	0.8748	1.512197	0.3041
Sk	13.7959	19.2189	10.35151	15.2007
	162.4494774	152.72	211.0041515	141.9466
$\mu - RGB$	122.3414634	118.1	183.2501297	143.3171
	48.86527294	32.84	142.3788272	53.9356
$\mu - LAB$	35.73325141	36.8	35.0252843	38.5462
	24.61744425	18.58	26.07428345	17.4590
σ for RGB	24.68990983	21.11	23.98391577	16.4837
	16.908859	13.73	27.06759251	18.8729
σ for LAB	7.428834532	2.5660e+04	10.08539616	6.4526

2.4. Leaf disease severity level estimation

The paddy leaf diseases are initially in pin head size, in later stage, it becomes large circular, oval, elongated or irregular based on the types of disease. The quantity of disease knowledge is a fundamental to decision making, crop loss and to evaluate the effectiveness of treatment. Standard area diagram (SAD) examples to aid in severity estimation of a spot blotch severity on wheat leaves is shown in Figure 5(a). SAD example for paddy leaf blast disease, the severity estimation image set/levels are defined in Figure 5(b).



(a) (b)

Figure 5: Standard Area Diagram (SAD) example images to estimate the levels if disease (a) For spot blotch severity on wheat leaves [13]. (b) For leaf blast severity on paddy leaves.

The intensity, color and disease part of leaves are extracted using proposed Segmentation and Binarization algorithm as defined in Algorithm-2. The binary-mass image (bin_{ROI}) extracted by proposed Algorithm-2 is used to estimate the infection percentage. The infection of paddy leaf diseases is estimated by using following equation.

$$\text{Infection Percentage } (IP) = \frac{Area_{disease}}{Area_{leaf}} \times 100$$

Where, $Area_{disease}$ is the total number of white pixels of disease part detected by Algorithm-2. $Area_{leaf}$ is the total number of pixels in leaf area.

The bacterial leaf blight is occurring as a stripe with waxy margins on edges of leaf and hispa disease is a white patch that is parallel to leaf veins. These diseases grow along vertically more than width/horizontal. Due to these characteristics, Maximum radius (Max_R) along with IP is considered to grade the level of diseases.

In this study, Mamdani Fuzzy Inference System (MFIS) is used to grade the level of leaf infection. In this system 9×6 rules based on the membership functions are considered as input to MFIS. 9×6 rules defined as IP×Max_R, i.e., [0-1, 5, 10, 15, 25, 35, 45, 60, 75] are the classes for IP & [0-60 pixels, 61-120 pixels, 121-180 pixels, 181-240 pixels, 241-360 pixels] are the classes for Max_R. The result of grading will be from '0' to '5'. Zero (0) being healthy/not infected, '5' being the highest infected. Example of rule definition is: if IP is '75' and Max_R is '275' then the leaf infection grade is '5'.

3. Experimental Results

3.1. Dataset Description

The proposed system is evaluated using two datasets. First dataset is collected from Kaggle [13]. In this dataset there are three classes of paddy leaf disease images and healthy leaves, three disease classes are brown spot, leaf blast and hispa. Kaggle paddy leaf disease dataset contain 523 brown spot images, 779 leaf blast images, 565 hispa disease images and 1488 healthy leaf images. Another dataset is created by capturing diseased paddy leaves using mobile phone camera. Images are captured in sunlight and phone flash light, with different resolution, different distance between camera and leaves different and with different angles. Totally 578 images are collected with three classes, it includes 200 brown spots, 160 leaf blast and 218 leaf blight paddy leaf disease images.

3.2. Result Analysis

In leaf background elimination technique, the performance of proposed system works better for simple background images, if the background is complex i.e., background and leaf colors are similar, the efficiency of eliminating background by filling black color is less.

Paddy leaf disease classification results are produced by Multiclass SVM classifier, used 75% of image data to train the SVM classifier and 25% to test the performance of proposed systems. Training and testing data distribution for each class are shown in Table 3. Tested our proposed segmentation and binarization system by extracting GLCM features and also by extracting proposed mass, margin, shape, texture and color features.

Accuracy of proposed system is evaluated as follows:

$$Accuracy = \frac{No.of\ truly\ detected\ Disease\ Images}{No.of\ Images\ used\ for\ testing} \times 100$$

The performance analysis for brown spot, leaf blight, hispa, leaf blast and healthy leaves with proposed system is shown in table 4. From the Table 4, it is noticed the proposed feature extraction technique increase the efficiency of proposed segmentation and binarization technique by 1.42% when compared to GLCM features. The performance analysis and comparison between existing system and proposed system is showed in Table 5.

Table 3. Training and testing data selection.

	Brown Spot	Leaf blast	Hispa	Leaf blight	Healthy	Total
Training Dataset	542	704	424	163	1116	2949
Testing Dataset	181	235	141	55	372	984

Table 4. The performance analysis of proposed segmentation & binarization system with different features.

Classes	No. of truly detected disease images (Proposed segmentation & binarization + GLCM features)	Accuracy (%)	No. of truly detected disease images (Proposed segmentation & binarization + proposed features)	Accuracy (%)

Brown spot	170	93.92	173	95.58
Leaf blast	222	94.46	224	95.31
Hispa	130	92.19	136	96.45
Leaf blight	50	90.90	52	94.54
Healthy	364	97.84	365	98.11
Total	936	95.12	950	96.54

Table 5. Performance analysis in percentage (%)

Authors	**Type of Diseases classified**	**Features**	**Classifier**	**Accuracy (%)**
Suresha et al., [18]	Blast and Brown Spot	Area, minor and major axis length and Perimeter	KNN	76.59
Satgunalingam & Thaneeshan [4]	Paddy blast, brown spot, narrow brown spot diseases.	GLCM & Color	Multi class SVM	87.5
Ramesh and vydeki [15]	Healthy and blast	GLCM	ANN	90 & 86
Swetha and Shravani [7]	Bacterial Leaf Blight, Brown Spot and Leaf Smut	color and spots geometry	SVM and k-NN	91.23
Phadikar and Sil [17]	Leaf blast and brown spot	Zooming algorithm	Self-Organizing Map (SOM) neural network	92
Ghosal and Sarkar [14]	Rice blast, leaf blight, brown spot and healthy	-	CNN With Transfer Learning	92.46
Ramesh and vydeki [3]	Bacterial blight, Blast, Brown spot, Sheath rot and normal leaves	Color and GLCM	Optimized Deep Neural Network with Jaya Optimization Algorithm (DNN_JOA)	94.53
Pothen and Pai [16]	Leaf blight, smut and brown spot	HOG	SVM	94.6
Proposed segmentation & binarization	Rice Blast, Brown Spot, leaf Blight, Hispa and Healthy leaves	GLCM	SVM	**95.12**
Proposed segmentation & binarization	Rice Blast, Brown Spot, leaf Blight, Hispa and Healthy leaves	Proposed Mass, mass-margin, shape, texture and color	SVM	**96.54**

Table 6. Comparison of manual and fuzzy inference system to grade the level of diseases.

	Grade-0	Grade-1	Grade-2	Grade-3	Grade-4	Grade-5	Manual Predict	Percentage (%)
Garde-0	24	1	0	0	0	0	25	96
Grade-1	1	28	1	0	0	0	30	93.33
Grade-2	0	2	30	3	0	0	35	85.71
Grade-3	0	0	2	21	1	0	24	83.33
Grade-4	0	0	0	2	19	0	21	90.47
Grade-5	0	0	0	0	1	14	15	93.33
Observed	25	32	33	25	21	14		

The confusion matrix to grade the level of paddy leaf diseases by using MFIS is shown in Table 6. Manual prediction column in Table 6, shows the number of images predicted manually. diagonal elements [24, 28, 30, 21, 19, 14] in Table 6 indicate the number of True Positive (TP) prediction of grading paddy leaf disease by using fuzzy inference system. The percentage of similarity between manual prediction system and prediction by using fuzzy inference system is calculated as shown below.

$$Similarity\ Percentage\ (SP) = \frac{Total\ number\ TP\ prediction\ by\ MFIS}{Total\ number\ of\ images} \times 100$$

150 images are considered to evaluate performance of grading the level of diseases, and predicted manually. The total number of TP prediction by MFIS is 136. The similarity percentage and disagreements between manual prediction and Fuzzy Inference System is 90.66%, and 9.34% respectively.

Conclusion

In this paper we have addressed the image processing techniques suitable for agricultural sector by proposing segmentation and binarization algorithm, feature extraction. The feature extraction is based on the disease characteristic to classify the paddy leaf diseases. The proposed system is evaluated using Rice Diseases Image Dataset downloaded from Kaggle and our own datasets. Initially, evaluated proposed segmentation and binarization algorithm by extracting GLCM features, and has achieved 93.92%, 94.46%, 92.19%, 90.90% and 97.84% of efficiency for classification of brown spot, leaf blast, hispa, leaf blight and healthy leaves respectively. The evaluation of proposed segmentation and binarization by extracting proposed features has achieved 95.58%, 95.31%, 96.45%, 94.54% and 98.11% efficiency for brown spot, leaf blast, hispa, leaf blight and healthy leaves respectively. This experimental result indicates the proposed feature extraction gives better classification. Finally, the proposed method is evaluated to grade the level of disease infection by using fuzzy rule-based prediction and achieved 90.66% of similarity between manual prediction and MFIS prediction. In future, additional features can be added to create fuzzy rules to achieve more similarities. In leaf extraction or background elimination, different approaches can be introduced to achieve better efficiency to extract leaf part in complex background images. The proposed methodology can be applied for various crop diseases.

References

1. S. Nalini, N. Krishnaraj, T. Jayasankar, K. Vinothkumar, A. Sagai et al., "Paddy leaf disease detection using an optimized deep neural network," Computers, Materials & Continua, vol. 68, no.1, pp. 1117–1128, 2021.
2. Matin, M. , Khatun, A. , Moazzam, M. and Uddin, M. (2020) An Efficient Disease Detection Technique of Rice Leaf Using AlexNet. Journal of Computer and Communications, 8, 49-57. doi: 10.4236/jcc.2020.812005.
3. S. Ramesh, D. Vydeki, “Recognition and classification of paddy leaf diseases using Optimized Deep Neural network with Jaya algorithm”, Information Processing in Agriculture, Volume 7, Issue 2, 2020, Pages 249-260, ISSN 2214-3173.
4. Satgunalingam, V & Thaneeshan, R. (2020). Automatic Paddy Leaf Disease Detection Based on GLCM Using Multiclass Support Vector Machine. International Journal of Computer (IJC), 39(1), 97-106.
5. N.Pranay Kumar, D. R. N. S. K. L. V. (2020). Detection and Classification of Rice Leaf Diseases Using Deep Learning. International Journal of Advanced Science and Technology, 29(3), 5868- 5874.
6. Sreenivasulu S, P.Pedda Sadhu Naik, “PADDY LEAF DISEASE DETECTION USING CNN”, International Journal of Psychosocial Rehabilitation, Volume 24 - Issue 8, August 2020, Pages: 16126-16138, ISSN:1475-7192.
7. Naga Swetha R, V Shravani, “Monitoring of Rice Plant for Disease Detection using Machine Learning”, International Journal of Engineering and Advanced Technology (IJEAT), ISSN: 2249 – 8958, Volume-9 Issue-3, February, 2020.
8. Nidhis A.D., Pardhu C.N.V., Reddy K.C., Deepa K. (2019) Cluster Based Paddy Leaf Disease Detection, Classification and Diagnosis in Crop Health Monitoring Unit. In: Peter J., Fernandes S., Eduardo Thomaz C., Viriri S. (eds) Computer Aided Intervention and Diagnostics in Clinical and Medical Images. Lecture Notes in Computational Vision and Biomechanics, vol 31. Springer, Cham. https://doi.org/10.1007/978-3-030-04061-1_29
9. L. Devichandana, Dr. Kunjam Nageswara Rao, Dr. G. Sita Ratnam, “Paddy Leaf Disease Detection and Quantification using Computational Techniques”, International Journal of Management, Technology And Engineering, Volume 8, Issue XI, Page No:2373, NOVEMBER/2018, ISSN NO : 2249-7455.
10. Farhana Tazmim Pinki, Nipa Khatun, S.M. Mohidul Islam, “Content based Paddy Leaf Disease Recognition and Remedy Prediction using Support Vector Machine”, 20th International Conference of Computer and Information Technology (ICCIT), 22-24 December, 2017, doi: 10.1109/ICCITECHN.2017.8281764.
11. Mohanty Sharada P., Hughes David P., Salathé Marcel, “Using Deep Learning for Image-Based Plant Disease Detection”, Frontiers in Plant Science, VOLUME-7, 2016, PAGES=1419, DOI=10.3389/fpls.2016.01419, ISSN:1664-462X.
12. Reinald Adrian DL. Pugoy, Vladimir Y. Mariano, “Automated rice leaf disease detection using color image analysis”, Third International Conference on Digital Image Processing (ICDIP 2011), edited by Ting Zhang, Proc. of SPIE Vol. 8009, 80090F · © 2011 SPIE. doi: 10.1117/12.896494
13. Rice Diseases Image Dataset: An image dataset for rice and its diseases, by Huy Minh Do. API command: kaggle datasets download -d minhhuy2810/rice-diseases-image-dataset.
14. Retrieved from: https://www.kaggle.com/minhhuy2810/rice-diseases-image-dataset

15. S. Ghosal and K. Sarkar, "Rice Leaf Diseases Classification Using CNN With Transfer Learning," 2020 IEEE Calcutta Conference (CALCON), Kolkata, India, 2020, pp. 230-236, doi: 10.1109/CALCON49167.2020.9106423.
16. 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.
17. M. E. Pothen and M. L. Pai, "Detection of Rice Leaf Diseases Using Image Processing," 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2020, pp. 424-430, doi: 10.1109/ICCMC48092.2020.ICCMC-00080.
18. S. Phadikar and J. Sil, "Rice disease identification using pattern recognition techniques," 2008 11th International Conference on Computer and Information Technology, Khulna, 2008, pp. 420-423, doi: 10.1109/ICCITECHN.2008.4803079.
19. M. Suresha, K. N. Shreekanth and B. V. Thirumalesh, "Recognition of diseases in paddy leaves using knn classifier," 2017 2nd International Conference for Convergence in Technology (I2CT), Mumbai, 2017, pp. 663-666, doi: 10.1109/I2CT.2017.8226213.
20. Prajwalgowda B.S, Nisarga M A, Rachana M, Shashank S, Sahana Raj B.S, 2020, Paddy Crop Disease Detection using Machine Learning, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) NCCDS – 2020 (Volume 8 – Issue 13).
21. Rice production in India.
22. Retrieved from: https://en.wikipedia.org/wiki/Rice_production_in_India

Author’s Details

Anoop G L was born in Ganadalu village, Mandya district, Karnataka, India in 1989. He received the B.E. and M.Tech. degrees in computer science engineering from Visvesvaraya Technological University, Belagavi, Karnataka, India, in 2012.

Since August-2012, he has been an Assistant Professor with the Computer Science and Engineering Department, Dayananda Sagar Academy of Technology & Management, Bangalore. He is the author of 6 articles. His research interests include image processing, pattren recognition and datamining.

Mr. G L was a recipient of best paper award in International Conference on Recent Innovations in Engineering and Technology (ICRIET), Kuala Lumpur, Malaysia, 1st August 2016.

Dr. C. Nandini was born in Mysore, Karnataka, India in 1972. She received the B.E. and M.Tech. degrees in Electrical and Electronics engineering, from N.I.E, Mysore University, Karnataka, India in 1996. The Ph.D degree in computer science engineering from Visvesvaraya Technological University, Belagavi, Karnataka, India, in 2009.

From 1993 to 1998, she was a Lecturer in A.I.T chikkamagaluru, from 1998 to 1999 she was a senior Lecturer in Vidyavardhaka College of Engineering Mysore, from 1999 to 2010 she was an Assistant Professor & head CSE and MCA Department in Vidya Vikas Institute of Engineering Mysore, from 2010 to 2011 she was a professor in Dayananda Sagara College of Engineering, Bangalore, from 2011 she is vice principal, Professor, Head Department of CSE, Dayananda Sagar Academy of Technology & Management, Bangalore. She is the author of one book chapter and more than 80 articles. Her research interests includes computer vision, machine learning, IoT, image processing, pattren recognition, crypto biometrics and datamining. She is a reviewer for journal IEAE and IJAE. She is a life member of ISTE. She is a member of CSI and holds five patents.

Dr. Nandini was recognized as an Indian researcher by Washington DC News.net. And Air BBC. She won Rajiv Ghandhi Education Excellence award for outstanding achievements in the field of education in 2013, Research excellence award in Indo-American education summit in 2016, VGST award in the year 2016-17, she was a recipient of best paper award in International Conference on Recent Innovations in Engineering and Technology (ICRIET), Kuala Lumpur, Malaysia, 1st August 2016. HOD of the year by Institute of scholars in the year 2019, she was awarded with Women Research Award by VDGOOD International Scientist awards 2021.