The Journal of Supercomputing
https://doi.org/10.1007/s11227-021-04245-x

A novel method to improve computational and classification performance of rice plant disease identification

K. S. Archana[1] · S. Srinivasan[2] · S. Prasanna Bharathi[3] · R. Balamurugan[4] · T. N. Prabakar[5] · A. Sagai Francis Britto[6]

Accepted: 30 November 2021

Abstract

Rice is a major food crop that plays an important role in the Indian economy. It is the most consumed staple food, greatly in demand in the market to meet the requirements of a growing population, which is only possible with increased production. To meet this demand, rice production should be increased. To maximize crop productivity, measures must be taken to eradicate rice plant diseases, namely, brown spot, bacterial leaf blight, and rice blast. In the proposed method, the modified K-means segmentation algorithm is used to separate the targeted region from the background of the rice plant image. Following segmentation, features are extracted through the three parameters of color, shape and texture. A novel intensity-based color feature extraction (NIBCFE) proposed method is used to extract color features, while the texture features are identified from the gray-level cooccurrence matrix (GLCM) and bit pattern features (BPF), and the shape features are extracted by finding the area and diameter of the infected portions. Thereafter, unique feature values are identified through the novel support vector machine-based probabilistic neural network (NSVMBPNN) to classify the images. A comparison in terms of performance is made using three classifiers, namely naïve Bayes, support vector machine and probabilistic neural network. This proposed method achieved better accuracy than the other three methods based on different performance measures. Finally, the result was validated under the fivefold cross-validation method with final accuracies of 95.20%, 97.60%, 99.20% and 98.40% for bacterial leaf blight, brown spot, healthy leaves and rice blast, respectively.

Keywords Rice plant disease classification · Modified K-means segmentation · Novel intensity-based color feature extraction · Gray-level cooccurrence matrix · Novel support vector machine-based probabilistic neural network

K. S. Archana
archana.se@velsuniv.ac.in

Extended author information available on the last page of the article

1 Introduction

The economy of developing South Asian countries is largely dependent on agriculture [18, 25]. Every year, there is a reduction in overall crop cultivation owing to erratic climatic conditions, an issue that can no longer be neglected. Detecting disease growth in its initial stages is critical for farmers who can put appropriate measures in place to arrest the spread of the disease before it devastates an entire crop [31]. The majority of plant diseases are bacterial, fungal and viral. Viral diseases are brought on by environmental changes, while fungal and bacterial diseases are brought on by fungi and germs on the leaves, respectively. From the symptoms appearing on the leaves, plant disease can easily be detected, owing to the thorough and relentless research being carried out today [24]. Across a wide range of crops, farmers choose the most appropriate vegetable and fruit crops. Such crop cultivation yields excellent returns, with production quality to match. Rice is considered the nation's staple crop because it is India's primary food source and cereal crop [13]. Consequently, there is a need to take adequate precautions to keep crops disease-free. Fungal and bacterial diseases affect paddies [27]. The most common rice disease infections are bacterial leaf blight, brown spots and rice blast. Because of these infections, the quality of the grains is highly reduced which places a high burden on farmers.

1.1 Rice diseases

A large number of pathogenic microorganisms attack rice plants and cause heavy losses in the field [32]. There are more than 30 different kinds of diseases that attack rice plants and reduce the quality and quantity of the product. According to current statistics, the typical rice diseases are rice blast, bacterial leaf blight and brown spot [4].

1.1.1 Bacterial leaf blight

Bacterial leaf blight is an origin of a bacterial pathogen by Xanthomonas oryzae pv. Oryzae. The disease mainly occurs on weeds, at favorable temperatures of relative humidity and by infected plants. The major symptoms of the disease are yellowish strips on leaf blades with a wavy margin, which later wilt and die.

1.1.2 Rice blast

Rice blast is one of the major fungal infections named Pyricularia grisea. This pathogen affects all portions of the plant, such as the leaf, neck, and collar. The disease mainly occurs because of low soil moisture, large day-night temperature, and cool temperature. The initial disease symptoms appear as gray color, green lesion spots with dark green borders and spindle shape, which causes severe infection later and affects severe yield loss to the entire farm.

1.1.3 Brown spot

Brown spot is the origin of a fungal pathogen named Helminthosporiosis. It occurs in every stage of the crop. During its development, the infection causes death to young plants and reduces grain quality. The initial symptoms of the disease are a minute brown dot, which later turns into a small, circular, yellow–brown dot. These diseases are responsible for severe yield loss to humans in several ways. Therefore, the goal of plant disease management is to predict disease in earlier stages to avoid economic loss.

To overcome the challenges of rice plant diseases, computer vision has been used to detect and classify rice plant diseases from early symptoms [25, 26]. In the past few decades, interest in research on plant disease detection using imaging has grown substantially. In the agricultural industry, classification and crop disease identification receive serious consideration in technical and financial terms [33, 17]

Image processing in agricultural applications offers the following benefits:

- Recognizing the shape and size of fruits
- Identifying the stem, fruit and infected leaf
- Recognizing infected portions through color

The system brings together experts and farmers on a common platform. Since the existing techniques demonstrate poor accuracy in classifying rice plant images, a novel identification and classification approach is proposed in this work. The objectives of this research are.

- To segment rice plant disease images using the modified K-means segmentation technique
- To extract color-based features using a new intensity-based color feature extraction technique
- To classify rice plant disease images using a novel support vector machine-based probabilistic neural network classifier.

The remaining sections of this paper are organized as follows. Related work on detecting and classifying rice plant diseases is discussed in Sect. 2. The working procedure of the proposed methodology is explained in Sect. 3. The performance of the proposed methodology is estimated and compared in Sect. 4. Finally, the research is concluded in Sect. 5.

2 Related work

Garcia et al. [9] proposed an approach for discovering plant diseases. The resultant color histograms were used to detect rice plant disease, and a pairwise classification was carried out. Plant disease testing was carried out on a dataset comprising 82 biotic and abiotic components, and nearly 12 kinds of plant species were found to have blighted leaves. An extended examination was carried out on a wide range of images to assess the benefits and deficiencies of the proposed algorithm and compared with existing algorithms for the performance of the proposed algorithm. The limitation of this work was in the challenges confronted in terms of its real-time adaptation.

Barbedo [5] explored an innovative approach to the segmentation of plant leaves with symptoms of disease ranging across a broad spectrum. This work employed color channels and the Boolean approach to binary masks and is therefore compact and stronger than state-of-the-art approaches that are designated automatically. The performance of the proposed method was tested against a large database comprising 77 diseases in 11 plant species. The efficacy of this method was compared with manual segmentation, and its advantages were reinforced. [6] described leaf diseases using high-resolution multispectral stereo images and three automatic classification methods. Two cameras captured perfectly illuminated single sugar beet leaves in a laboratory. The information from the two sensors was fused to generate 3D leaf models. The potential for pixelwise contextual classification, with its ability to eliminate the errors associated with the process, was investigated.

Gayathri Devi and Neelamegam [10] utilized image processing to detect leaf diseases automatically, focusing particularly on paddy leaves. A hybrid approach comprising the discrete wavelet transform, grayscale cooccurrence matrix, and scale-invariant feature transform was used to extract the relevant features. The back-propagation and K-nearest neighborhood neural networks, alongside the naïve Bayes and multiclass SVM, were used to differentiate between diseased and healthy plants and classify them accordingly thereafter. The performance of several classification approaches in disease categorization was studied. MATLAB software was implemented to arrive at the experimental outcomes. The work concluded that the multiclass SVM delivers superior precision when compared to other classifiers.

Phadikar, Sil, and Das [22] classified rice plant diseases, extracting features from the affected portions in rice plant images. The affected regions were isolated from the background by deploying a Fermi energy-based segmentation approach. The symptoms of the disease were classified in line with field specialists' opinions. Classifier issues were reduced by utilizing rough set theory (RST); there was, consequently, a corresponding reduction in information loss. Finally, a rule-based classifier was constructed using the features considered, with the ability to discover different kinds of diseased rice plant images. The classifier provided better outcomes than state-of-the-art classifiers.

(Kaur and Bhardwaj [15] developed an image processing system to detect rice plant diseases such as bacterial blight, leaf blast, and brown spot. The affected portion of the rice plant was detected during the initial stages of the infection using the KNN and clustering classifiers. Their research can help the farming community take adequate steps to stem rot and make a profit.

Abu Bakar et al. [1] advocated a combined approach to detect rice leaf blast (RLB) disease. This work analyzed images based on the hue saturation value (HSV) color space and carried out pattern recognition using a multilevel thresholding method. The outcomes were categorized into three stages of infection: initial stage, spreading stage, and final stage.

Chouhan, Kaul, Singh, and Jain [8] introduced a bacterial foraging optimization-based radial basis function neural network (BRBFNN) to automatically identify and classify plant leaf disease. This work utilized bacterial foraging optimization to assign an optimum weight to the radial base function of the neural network and improved the speed and accuracy of the network. Fungal diseases such as early blight, leaf spot, late blight, common rust, cedar apple rust, and leaf curl were identified. The proposed approach demonstrated superior effectiveness in terms of classifying and identifying plant leaf disease.

Nurhikmah and Sany [20] introduced a rice plant disease recognition approach based on pattern recognition and image processing. Nearly 38 (shape, texture and color) classifying features were extracted from every portion of the leaf image. Genetic algorithms (GA) and correlation-based features (CFS) were combined to decrease the dimensions of the feature space and maximize the accuracy of rice disease identification. Mobile-based applications were used to determine plant quality.

Hussein et al. [12] identified disease from the collected images of crops such as tomato, cotton and rice. The images were preprocessed to resize them to a constant size. Fuzzy histogram equalization (FHE), followed by image segmentation, was carried out using color-based K-means. Finally, the performance of the four feature extraction approaches was compared. The results showed the superior outcomes produced by the four approaches considered, including color moments, texture-based features, shape-based features, and the percentage of leaf area infected (PI).

Latha et al. [17] proposed a disease identification method, especially for apple leaves, using deep conventional neural networks. This work generated an adequate number of pathological images and designed an innovative AlexNet-based architecture. Their research indicated that the proposed framework provides superior outcomes in apple leaf disease detection with improved accuracy and a quick convergence rate. Furthermore, their image generation method has been enhanced with the use of a strong conventional neural network framework.

(Orillo, Dela Cruz, Agapito, Satimbre, and Valenzuela, 2014) used digital image processing to address the problem of negligence in the manual inspection of rice plants and studied at length the three generic diseases affecting Philippine farmlands. The proposed work used the backpropagation neural network for accuracy and enhanced image processing performance and obtained 100% accuracy. Majumdar et al. [18] discussed several image processing approaches used for the early detection of disease in plants and described the strengths and limitations of each. Their study offers scope for future research in plant disease segmentation. The superior performance classification methods were analyzed, and the issues therein were optimized to determine the best-performing method.

Patil [21] categorized cotton leaf disease using images captured on a mobile phone. The cotton farming community attempted to combat an early manifestation of the disease in the groves with the application of an elective fungicide. The captured images were processed using segmentation methods. Color and texture feature extraction methods were used to extract texture, shape, color and boundary features. The features extracted were used to identify disease spots.

3 Proposed methodology

*Image analysis procedur*e: This section briefly discusses the proposed work. Figure 1 shows the architecture diagram for disease identification in rice plants; it is fully based on an automatic approach. The fundamental idea of this research is to identify the three common rice plant leaf diseases, including healthy leaves, that cause devastating losses in rice cultivation. The dataset comprises images of diseased plant leaves. This architecture reports the identification and classification of disease in rice plants through image processing techniques such as (1) Image acquisition: input image through digital camera, (2) Image preprocessing: filter the unwanted noise in the picture, (3) Image segmentation: target diseased region, (4) Feature extraction: retrieve useful information from the unique feature to classify the result, followed by (5) Classification: detection and classification of different diseases, namely, Class 1 showing bacterial leaf blight, Class 2 brown spot, Class 3 healthy leaves, and Class 4 rice blast.

3.1 Rice disease images acquisition

Rice leaf images were obtained from Neduvasal in the Pudukkottai District of Tamil Nadu, India. Using controlled lighting, 109 images were collected in all, adding to the 16 from the Indian Rice Research Institute (IRRI) website. The images were captured using a Canon EOS 3000D camera.

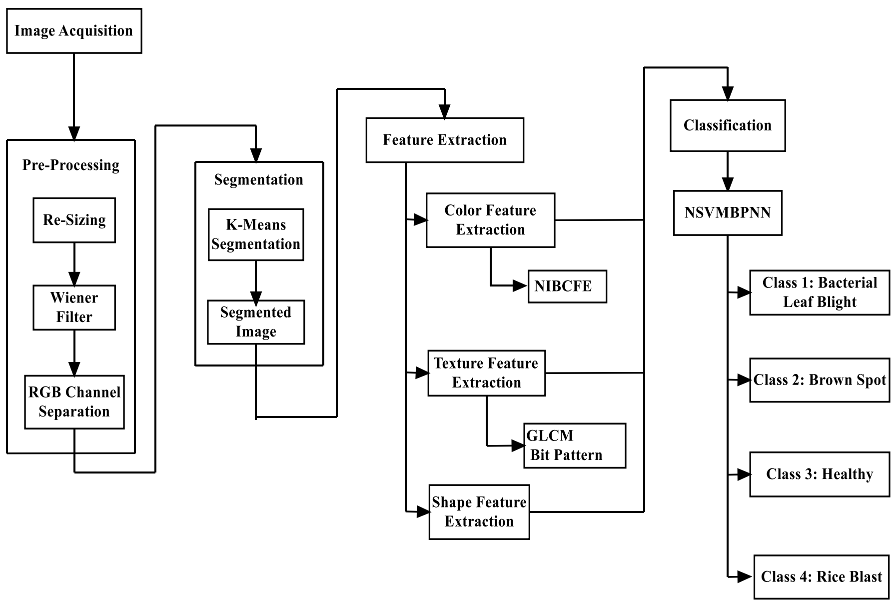


Fig. 1 Rice disease identification and classification procedure

Camera specifications: The image-capturing device, a Canon EOS 3000D digital camera, has a calibration gray card with additional features that include

(a) Resolution ranging from 1-18 megapixels, which is 25 times larger than that of mobile camera phones,
(b) An RGB 24-bit color frame grabber with 1584 x 3456 pixels and a resolution of 18–55 mm focal lengths, and
(c) The image-capturing software of MV Tools.

3.1.1 Image-capturing times/seasons

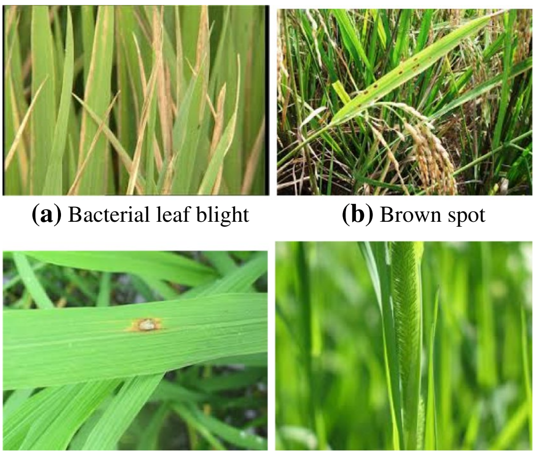
Diseased leaf samples of different rice plants, as well as healthy leaves, were collected in the (late) spring of 2017.

Location: The samples were collected on agricultural land constituting 2.471 acres.

Image collection: The dataset used in this research comprises different rice plants and their disorders, such as *brown spot, bacterial leaf blight,* and *rice blast,* in addition to images of healthy leaves, all of which are depicted as follows in Fig. 2: (a) Bacterial leaf blight, (b) Brown spot, (c) Rice blast, and (d) Healthy leaf.

3.2 Rice diseases preprocessing

During image capture, noise will appear. In the presence of noise, the true image details may be lost. Preprocessing not only removes unwanted noise but also enhances some image features for further processing [14]. Imaging sensors may also be affected by various factors, such as the quality of the sensing element and the environmental conditions [27]. The resulting image may be affected by the sensor temperature, charge-coupled device [CCD] and illumination level. Another fact is that the image is corrupted due to atmospheric conditions or lighting defects. For this reason, the noise should be



(a) Bacterial leaf blight **(b)** Brown spot **(c)** Rice blast **(d)** Healthy leaf

Fig. 2 Images of the four rice diseases filtered through various techniques. Here, the filtering process is carried out using the Wiener filter.

3.2.1 Wiener filter

The acquired images are preprocessed, wherein they are resized and filtered to eliminate superfluous pixels. Corrupting noises are filtered using the Wiener filter, which is based on a statistical approach and the desired frequency response. This Wiener filter minimizes the mean square error as much as possible. As part of it, the frequency domain and discrete Fourier transform (DFT) are used in the Weiner filter to remove unwanted noise in an image. It is estimated by

$$W(f1,f2) = \frac{H*(f1,f2)S_{rr}(f1,f2)}{|H(f1,f2)|^2 S_{rr}(f1,f2) + S_{mn}(f1,f2)} \tag{1}$$

Images can be resized for the transformation of images, displays and storage purposes, as shown in Fig. 3.

The performance metrics of the Wiener filter give the best results, outperforming other filters such as the salt-and-pepper, Gaussian noise, and median noise filters [2, 3].

3.2.2 RGB solor transformation

The symptoms of a damaged pathogen can be differentiated by the color model. To analyze the color image, there are several models to be represented [29]. The model represented may use RGB, which is transformed to the HSV color space model. This RGB image comes from three initials of the channel, red, green and blue, with an additive color model in different methods to combine the large array of colors [33]. In this research work, the HSV space model can be applied to discriminate pathogens and healthy regions from rice plants. The following equations are used to transfer the image from the RGB to the HSI color model.

$$r1 = \frac{R}{G} + \in \tag{2}$$

(a) Input image **(b)** Resized image **(c)** Wiener filter image

Fig. 3 Weiner filter algorithm in rice blast image

$$r2 = \frac{B}{G} + \in \tag{3}$$

Where R, G and B represent the red, green and blue channels of the RGB color representation, respectively, and are arbitrarily small values that help avoid division by zero. That is, the smaller the value, the greener the pixel and, in theory, the healthier that particular area of the leaf, as shown in Fig. 4.

3.3 Segmentation

Once preprocessed, the image is segmented for classification into different groups or clusters for final analysis [30]. The modified K-means technique is used to segment the region from the background image [2, 3]. Figure 5 shows that Cluster 2 gives the affected portion from the infected area.

3.3.1 Modified K-means segmentation

In this proposed method, a collection of data is grouped into a distinct number of clusters by dividing it. According to the algorithm, first, the RGB value is calculated using a different lab conversion model by the following equations. Then, the image of every pixel has the color compounds of red, green and blue. The pixels are identified based on the center point of unlabeled N-dimensional K-cluster points.

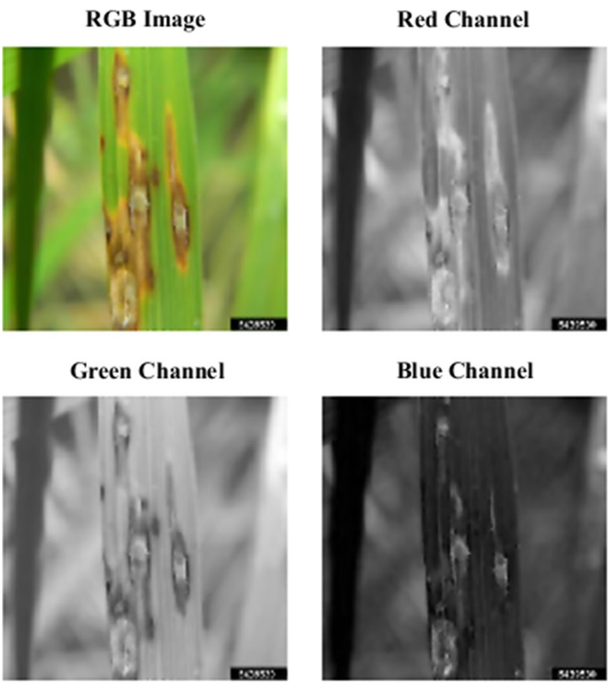


Fig. 4 RGB components in bacterial leaf blight image

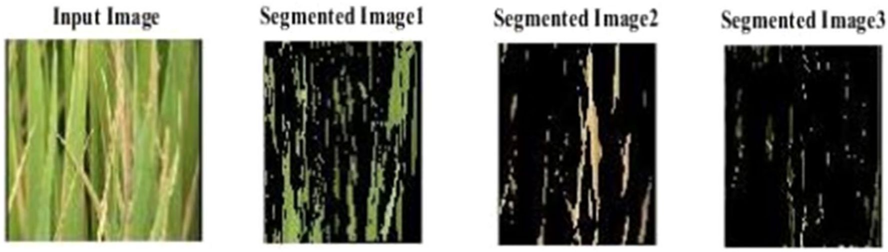


Fig. 5 Different Clusters in Modified K-Means Segmentation

$$\mathrm{xi}\ (\boldsymbol{r},\boldsymbol{g},\boldsymbol{b}) = 1 - \boldsymbol{ni} \sum (\boldsymbol{yi}(\boldsymbol{r},\boldsymbol{g},\boldsymbol{b})) \quad \boldsymbol{i} = 1, 2, 3, \ldots \boldsymbol{k} \tag{4}$$

By calculating the new centroids of xi for each pixel, the new values are assigned to find the new clusters by using the Euclidean distance between the respective clusters yi. Finally, an automated implication method is developed and applied to identify the diseased region of rice plant leaves through computer vision techniques.

3.4 Feature extraction

After the preprocessed image is segmented, feature extraction is carried out. This method aims to extract meaningful information from the segmented object and to represent the attributes for further processing [28]. Therefore, the independent features are identified through color, shape and texture. In general, feature extraction means extracting the information from the existing data so that the classifier can be trained. Hence, the new information is extracted to find the unique features by using the algorithm of color, texture and shape [7]. The unique features are identified through feature extraction using the proposed algorithm based on color, shape and texture. Finally, the selected unique features are extracted for the final classification to differentiate the disease.

3.4.1 Color-based feature extraction

In the feature extraction method, one of the most important visual feature extractions is a color-based method [23]. The fundamental method for representing the color contents is a color histogram using numerous color models. Hence, the color features are extracted using the novel intensity-based color feature extraction (NIB-CFE) algorithm. Here, the color distribution in an image involves an enormous mass of image data. Attributes are obtained from the image color distribution using the color cooccurrence matrix. The matrix helps to estimate the probability of a pixel and that of its neighbors for constructing information about a particular color. In addition, the matrix displays the spatial information of the image.

Table 1 Features Extracted from Shape

Features	Equation	Features	Equation
Area	$\iint I(x,y)dxdy$	Diameter	$\frac{C}{d}\pi$

Table 2 Features Extracted from Gray level co-occurrence matrix

Features	Equations	Features	Equations
Entropy	$\sum_{i,j,k=0}^{G-1} P(i,j,k) * \log(P(i,j,k))$	Contrast	$\sum_i \sum_j P(i-j)^2 * (i,j,k)$
Energy	$\sum_{i,j,k=0}^{M-1} \left(P_i,j,k\right)^2$	Prominence	$\sum_{i,j=1}^{N} \left(i+j-\mu_i-\mu_j\right)^2 P(i,j)$
Homogeneity	$\sum_{x,y=0}^{N-1} \frac{P_{xy}}{1+(x-y)^2}$	Shade	$\mathrm{sgn}(A)\lvert A\rvert^{1/3}$
Correlation	$\sum_i \sum_j \frac{(ix_j)(\mu_i-\mu_j)}{\sigma_x\sigma_y}$	Cluster shade	$\sum_{i,j=1}^{N} \left(i+j-\mu_i-\mu_j\right)^4 P(i,j)$

3.4.2 Shape-based feature extraction

Here, shape features are extracted using the area and diameter of the segmented image. Shape-based feature extraction is the process of extracting features based on the shape of the affected region. The area is represented by the actual number of white pixels in the affected region. The diameter can be represented as the diameter of a boundary with the same area as the region, as tabulated in Table 1.

Here, the unique features are extracted from the formulas mentioned earlier through the image pixel. Hence, from all the images, those features are identified for further analysis from each category.

3.4.3 Texture-based feature extraction

The texture feature helps to calculate the information about the selected region of an image [34]. The texture feature is the most useful and unique feature that is used for the final classification method. In this work, the most relevant metrics of GLCM and the bit pattern features are used to identify the features.

3.4.3.1 GLCM In texture features, a large amount of information is retrieved from the image content through GLCM [23]. Hence, to estimate the image properties, the most well-known texture analysis methods, namely, the gray-level cooccurrence matrix (GLCM), are used. The properties of the texture are identified by the intensity value at the pixel of interest. Finally, the texture features are extracted in various analyses, such as entropy, energy, homogeneity, correlation, contrast, prominence, shade, and cluster shade, from GLCM, as tabulated in Table 2.

3.4.3.2 Bit pattern feature Bit pattern features (BPFs) [11] are also used to analyze the texture features. This bit pattern feature helps to characterize the shape, edge and image contents. From a set of training bitmap images, binary vector quantization generates a representative bit pattern book code that is utilized to index the images [19]. All the features of the image are used to describe the location and character of the object, as shown in Fig. 6.

3.5 Classification in rice plants

Feature extraction from the segmented image is followed by classification. Image classification is an important task to extract and classify information in various applications [15]. In this paper, the most common techniques, namely, support vector machines, probability neural networks, and naïve Bayes classifiers, are analyzed for the classification of rice plant leaves.

3.5.1 Novel support vector machine-based probabilistic neural network (NSVMBPNN)

The query images are classified based on the feature extraction values using the novel support vector machine-based probabilistic neural network (NSVMBPNN). To determine whether the query image is affected, the SVM classifier involves two stages: training and testing. In this approach, input features are trained by giving them to the learning algorithm. SVM is a binary-class classifier that establishes appropriate margins between the two classes by constructing a hyperplane in a high-dimensional feature space. The PNN comprises four layers: input, pattern, summation and output. The input layer has n values to be classified. The dot product between the input and weight is executed in the pattern layer and divided by a certain threshold, which is then inserted into the radial basis function. Each pattern in

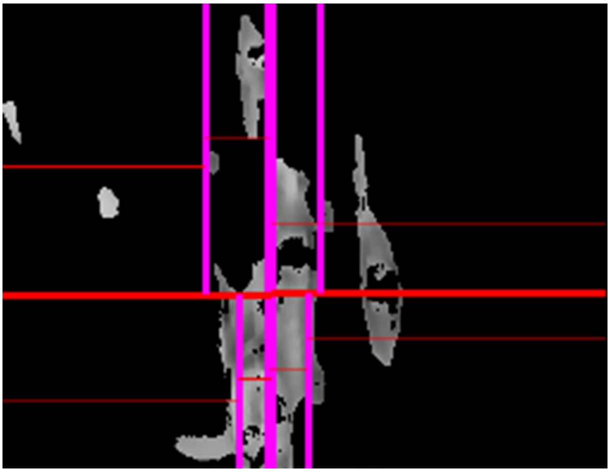


Fig. 6 Bit Pattern feature (BPF) in bacterial leaf blight image

every class is added to generate a population density function for every class in the summation layer. Finally, the input is classified at the output decision layer based on the different performance measures used. The advantage of the PNN is that it offers fast training speed and is robust to noise. By combining the SVM and PNN, the proposed approach offers better classification results based on the algorithm. Figure 7 shows the rice plant disease detection and identification samples.

3.5.2 Performance measures

The overall performance has drawn the severity level of rice plant disease from their early symptoms to detect and classify the disease. The performance evaluations are calculated under the confusion matrix of precision and accuracy for result classification. The confusion matrix holds. Here, the infected areas are assumed to hold the

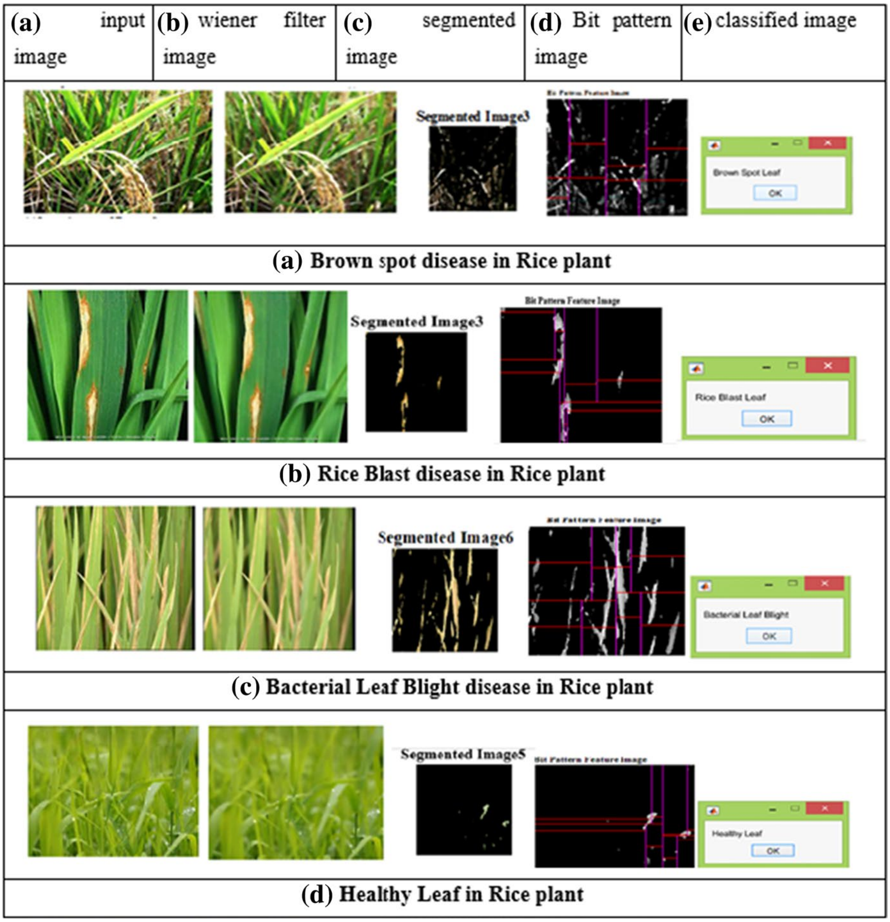


Fig. 7 Different disease identification in rice plants using the NSVMBPNN method

Table 3 Confusion matrix for between two classes

True class	Predicted class
A1 (Infected Region) TN	TP
A2 (normal Region) TN	FP

TP –True Positive, *TN*-True Negative, FP –False Positive, TN- True Negative

Table 4 Various Performance Metrics used for comparison

Metrics	Equations	Metrics	Equations
Accuracy	$\frac{TP+TN}{TP+FN+TN+FP} \times 100\%$	Precision	$\frac{TP}{TP+FP} * 100\%$
Sensitivity	$\frac{TP}{TP+FN} * 100\%$	Recall	$\frac{TP}{TP+FN} * 100\%$
Specificity	$\frac{TN}{TN+FP} * 100\%$		

true value, and normal regions are assumed to hold the false value. Hence, Table 3 shows a confusion matrix for binary classification, where TP, TN, FP, and FN.

To evaluate the model, the classification problem of the proposed method has to be measured. Parameters such as TP, TN, FP and FN are found, and further, the quality of the model is measured in terms of accuracy, sensitivity, specificity, precision and recall, as shown in Table 4.

Accuracy is an important metric for comparing techniques. where sensitivity, specificity, precision and recall describe how well the diagnosis predicts the true presence or absence of the disease. This estimates the accuracy and inaccuracy rate of the recognition method for detection and classification.

4 Results and discussion

The selected features were sent to the classifier, and the results obtained from the different classifiers are promising. To verify the efficiency of the algorithm, the recognized experiments are carried out on the rice plant disease leaf dataset of bacterial leaf blight, rice blast, brown spot and including healthy leaves. For each kind of disease, 125 leaf images are collected in which 75% of images are randomly selected as the training dataset to train the classifier and the remaining are used to test the algorithm's performance. Finally, the results of the proposed novel support vector machine-based probabilistic neural network (NSVMBPNN) method are compared with the other three existing algorithms, such as the support vector machine, probabilistic neural network and naïve Bayes algorithm.

4.1 Performance comparison of NSVMBPNN

To test the effectiveness of our proposed technique, a comparison was performed with different classifiers: SVM, PNN and naïve Bayesian. From the results shown in Table 5, it seems that our proposed novel support vector machine-based probabilistic

Table 5 Classification performance of all comparison methods

Performance metrics	Support vector machine			
	Bacterial leaf blight	Brown spot	Healthy leaf	Rice blast
Accuracy	88.00	96.00	97.60	92.80
Sensitivity	93.33	91.42	86.66	73.33
Specificity	85.00	97.77	99.09	98.80
Precision	77.77	94.11	92.85	95.65
Recall	93.33	91.42	86.66	73.33
Performance metrics	Probabilistic Neural Network (PNN)			
	Bacterial leaf blight	Brown spot	Healthy leaf	Rice blast
Accuracy	95.10	93.63	98.20	95.23
Sensitivity	91.15	90.40	93.35	93.30
Specificity	92.10	94.45	100	95.85
Precision	97.65	86.53	100	87.57
Recall	91.13	91.40	92.30	92.33
Performance metrics	Naïve Bayes Algorithm			
	Bacterial leaf blight	Brown spot	Healthy leaf	Rice blast
Accuracy	90.40	92.00	93.60	92.00
Sensitivity	93.33	77.14	66.67	86.67
Specificity	88.75	97.78	97.27	93.68
Precision	82.35	93.10	76.92	81.25
Recall	93.33	77.14	66.67	86.67
Performance metrics	NSVMBPNN (Proposed Method)			
	Bacterial leaf blight	Brown spot	Healthy leaf	Rice blast
Accuracy	95.20	97.60	99.20	98.40
Sensitivity	100	91.43	93.33	93.33
Specificity	92.50	100	100	100
Precision	88.24	100	100	100
Recall	100	91.43	93.33	93.33

neural network (NSVMBPNN) performs better than other methods and effectively includes healthy leaves. Healthy leaves and rice blasts achieved the highest accuracy compared to all other diseases, as shown in Fig. 8. It seems that (1) For classifying bacterial leaf blight, our method achieves a classification accuracy of 95.20%, with a sensitivity of 100%, a specificity of 92.50%, a precision of 88.24% and a recall value of 100%. (2) For classifying brown spots, our method achieves a classification accuracy of 97.60%, with a sensitivity of 91.43%, a specificity of 100%, a precision

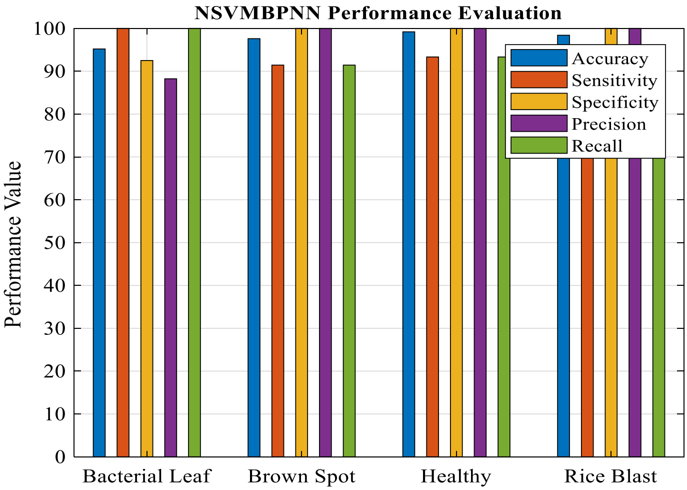


Fig. 8 A comparative analysis of the NSVMPNN algorithm

of 100% and a recall value of 91.43%. (3) For classifying healthy leaves, our method achieves a classification accuracy of 99.20%, with a sensitivity of 93.33%, a specificity of 100%, a precision of 100% and a recall value of 93.33%. (4) For classifying rice blast, our method achieves a classification accuracy of 98.40%, with a sensitivity of 93.33%, a specificity of 100%, a precision of 100% and a recall value of 93.33%. The overall detection accuracy of different classifiers is shown in Fig. 9.

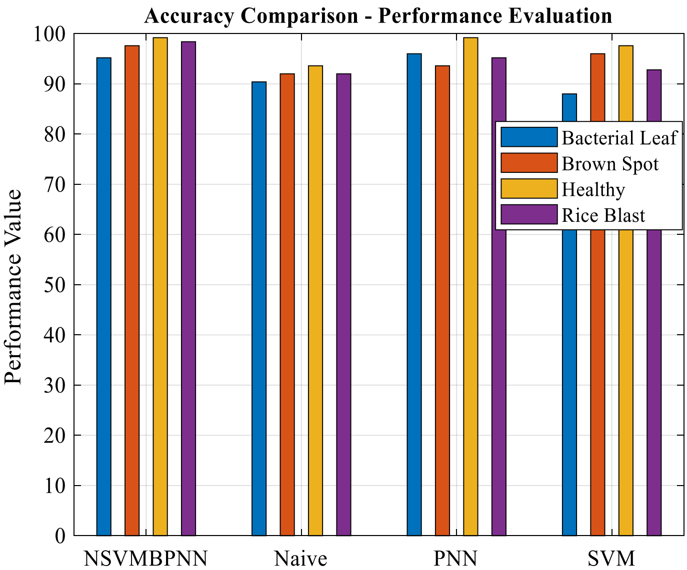


Fig. 9 Comparative results of the different classifiers

From the results, it can be seen that the proposed NSVMBPNN produces higher accuracy than the other classifiers, such as SVM, PNN and naïve Bayes, as shown in Table 6. Finally, to achieve a reliable measurement, all the performance results are obtained using k-fold cross validation. The main limitation of the proposed system is that bacterial leaf blight has quite 95.20% due to various factors, such as illumination, variable lighting, blurring, and fading. This is why the recognition process is also affected in terms of the overall accuracy rate. This can be improved and proposed as future work.

4.2 Cross validation

Here, our model requires cross validation to estimate the result of our proposed method to find the quality of the model [16]. The dataset of 125 images is classified into 4 kinds of rice plant diseases. Then, the fivefold cross-validation method is used to train and test and compare the accuracy metrics using support vector machines, probabilistic neural networks, naïve Bayesian and NSVMBPNN. It gives better results between the challenges of four classes. Then, the dataset is divided into 5 mutual subsets of approximately equal size. Here, four subsets are used as the training set and the last subset for the validation set. The abovementioned procedure is repeated five times, so each subset is used once for validation.

The proposed NSVMPNN approach offers more effective classification results than the other existing techniques. K-fold cross validation is also analyzed, and the results are also compared with other existing algorithms. Finally, this proposed NSVMPNN approach is evaluated in relation to accuracy, and the results displayed in Fig. 10 offer better classification accuracy.

5 Conclusion

The detection of plant disease has, in recent times, become an interesting area of research in agriculture. Image processing plays a significant role in the identification and classification of rice plant disease, to which, currently, the concept of machine learning has been effectively applied. Several traditional studies have focused on detecting disease in rice plant images, but with a major drawback in terms of inaccurate classification results. Hence, a novel feature extraction-based integrated classifier was developed in this research. The input image was taken from the dataset and

Table 6 Classification accuracies of various classifiers

Performance metrics	Bacterial leaf blight	Brown spot	Healthy leaf	Rice blast
NSVMBPNN	95.20	97.60	99.20	98.40
Naïve Bayes	90.40	92.00	93.60	92.00
PNN	95.10	93.63	98.20	95.23
SVM	88.00	96.00	97.60	92.80

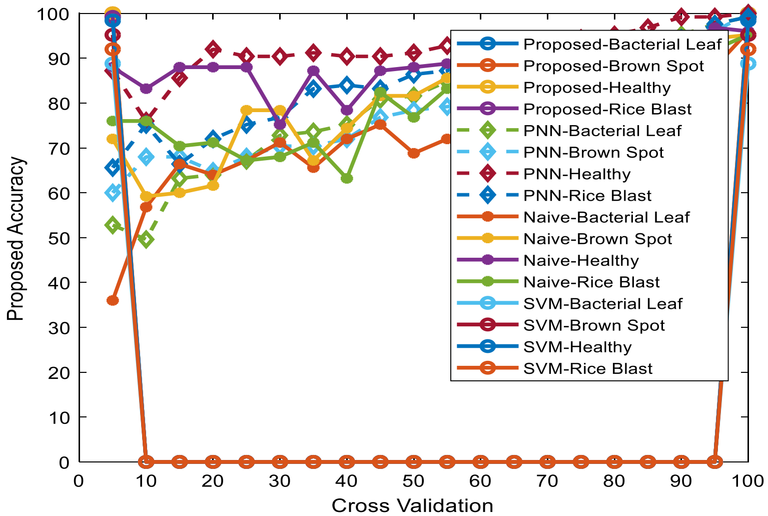


Fig. 10 Cross validation at various iterations

preprocessed. It was then resized, and the Wiener filter was applied to remove noise and blurred pixels. The preprocessed images were segmented using the modified K-means segmentation technique to identify affected regions in the image. Thereafter, features were extracted from the preprocessed images. Color-based features were extracted using the novel intensity-based color feature extraction (NIBCFE) technique, shape-based features with the area and diameter of the segmented image, and texture features using the gray-level cooccurrence matrix (GLCM) and bit pattern features (BPF). Finally, the novel support vector machine-based probabilistic neural network (NSVMPNN) was used to classify the images from the extracted feature values. The performance of the proposed methodology was evaluated and compared with existing techniques. From the evaluation results, it was concluded that the proposed NSVMPNN approach offers more effective classification results than the others (refer to Table).

Funding The authors received no specific funding for this study.

Declarations

Conflicts of Interest The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Abu Bakar MN, Abdullah AH, Abdul Rahim N, Yazid H, Misman SN, Masnan MJ (2018) Rice leaf blast disease detection using multi-level color image thresholding. J Telecomm, Electr Comp Eng 10(1–15):1–6

2. Archana KS, Sahayadhas A (2018) Comparison of various filters for noise removal in paddy leaf images. Int J Eng Technol 7(2):372–374
3. Archana KS, Sahayadhas A (2018) Automatic rice leaf disease segmentation using image processing techniques. Int J Eng Technol (UAE) 7(3.27):182–185. https://doi.org/10.14419/ijet.v7i3.27.17756
4. Archana KS, Sahayadhas A (2019) Computer vision for predicting unhealthy region of rice leaves - a review. Ind J Environm Prot 39(7):609–613
5. Barbedo JGA (2017) A new automatic method for disease symptom segmentation in digital photographs of plant leaves. Eur J Plant Pathol. https://doi.org/10.1007/s10658-016-1007-6
6. Bauer SD, Korč F, Förstner W (2011) The potential of automatic methods of classification to identify leaf diseases from multispectral images. Precision Agric 12(3):361–377. https://doi.org/10.1007/s11119-011-9217-6
7. Caglayan, A., Guclu, O., & Can, A. B. (2013). A plant recognition approach using shape and color features in leaf images. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *8157 LNCS*(PART 2), 161–170. https://doi.org/10.1007/978-3-642-41184-7_17
8. Chouhan SS, Kaul A, Singh UP, Jain S (2018) Bacterial foraging optimization based radial basis function neural network (BRBFNN) for identification and classification of plant leaf diseases: An automatic approach towards plant pathology. IEEE Access 6:8852–8863. https://doi.org/10.1109/ACCESS.2018.2800685
9. Garcia J, Barbedo A, Koenigkan LV (2016) ScienceDirect Identifying multiple plant diseases using digital image processing. Biosys Eng 147:104–116. https://doi.org/10.1016/j.biosystemseng.2016.03.012
10. Gayathri Devi T, Neelamegam P (2018) Image processing based rice plant leaves diseases in Thanjavur. Cluster Computing, Tamilnadu. https://doi.org/10.1007/s10586-018-1949-x
11. Hamouchene I, Aouat S, Lacheheb H (2014) Intelligent Systems for Science and Information 542:389–407. https://doi.org/10.1007/978-3-319-04702-7
12. Hussein MA, Abbas AH (2018) Comparison of features extraction algorithms used in the diagnosis of plant diseases. Ibn AL- Haitham J Pure Appl Sci. https://doi.org/10.30526/2017.ihsciconf.1785
13. Jagan K, Balasubramanian M, Palanivel S (2016) Detection and recognition of diseases from paddy plant leaf images. Int J Comput Appl 144(12):34–41. https://doi.org/10.5120/ijca2016910505
14. Kanagalakshmi, K., & Chandra, E. (2011). Performance evaluation of filters in noise removal of the fingerprint image. *ICECT 2011 - 2011 In:* 3rd International conference on electronics computer technology, *1*, 117–121. https://doi.org/10.1109/ICECTECH.2011.5941572
15. Kaur A, Bhardwaj V (2018) Rice Plant Disease Detection Based on Clustering and Binarization 5(4):245–249
16. Krstajic D, Buturovic LJ, Leahy DE, Thomas S (2014) Cross-validation pitfalls when selecting and assessing regression and classification models. J Cheminform 6(1):1–15. https://doi.org/10.1186/1758-2946-6-10
17. Latha A, Prasanna S, Hemalatha S, Sivakumar B (2019) A harmonized trust assisted energy efficient data aggregation scheme for distributed sensor networks. Cogn Syst Res 56(March):14–22. https://doi.org/10.1016/j.cogsys.2018.11.006
18. Majumdar, D., Kole, D. K., Chakraborty, A., Dutta Majumder, D., & Majumder, D. D. (2014). Review: detection & diagnosis of plant leaf disease using integrated image processing approach. *Int J Comput Eng Appl*, *VI*
19. Munisami T, Ramsurn M, Kishnah S, Pudaruth S (2015) Plant Leaf Recognition using shape features and colour histogram with k-nearest neighbour classifiers. Procedia Comput Sci 58:740–747. https://doi.org/10.1016/j.procs.2015.08.095
20. Nalini S, Krishnaraj N, Jayasankar T, Vinothkumar K, Sagai A et al (2021) Paddy leaf disease detection using an optimized deep neural network. Comput Mater Continua 68(1):1117–1128
21. Patil PSP, Zambre MRS (2014) Classification of cotton leaf spot disease using support vector machine 4(5):92–97
22. Phadikar S, Sil J, Das AK (2013) Rice disease classification using feature selection and rule generation techniques. Comput Electron Agric 90:76–85. https://doi.org/10.1016/j.compag.2012.11.001
23. Rawat P, Singh KD, Chaouchi H et al (2014) Wireless sensor networks: a survey on recent developments and potential synergies. J Supercomput 68:1–48
24. Rishi, N., & Gill, J. S. (2015). Detection and Classification of Plant Diseases by Image ProcessingRishi, N., & Gill, J. S. (2015). Detection and Classification of Plant Diseases by Image Processing.

An Overview on Detection and Classification of Plant Diseases in Image Processing, 3(5),. *An Overview on Detection and Classification of Plant Diseases in Image Processing*, *3*(5), 114–117.
25. Sanjeevi P, Prasanna S, Siva Kumar B, Gunasekaran G, Alagiri I, Vijay Anand R (2020) Precision agriculture and farming using Internet of Things based on wireless sensor network. Trans Emerg Telecommun Technol. 31(2):1–14. https://doi.org/10.1002/ett.3978
26. Sanjeevi P, Siva Kumar B, Prasanna S, Maruthupandi J, Manikandan R, Baseera A (2020) An ontology enabled internet of things framework in intelligent agriculture for preventing post-harvest losses. Complex Intell Syst. https://doi.org/10.1007/s40747-020-00183-y
27. Sarangi S, Umadikar J, Kar S (2016) Automation of agriculture support systems using Wisekar: a case study of a crop-disease advisory service. Comput Electron Agric 122:200–210. https://doi.org/10.1016/j.compag.2016.01.009
28. Sharma, A., & Dey, S. (2012). A comparative study of feature selection and machine learning techniques for sentiment analysis, 1. https://doi.org/10.1145/2401603.2401605
29. Shrivastava S, Singh SK, Hooda DS (2015) Color sensing and image processing-based automatic soybean plant foliar disease severity detection and estimation. Multimed Tools Appl 74(24):11467–11484. https://doi.org/10.1007/s11042-014-2239-0
30. Singh V, Misra AK (2017) Detection of plant leaf diseases using image segmentation and soft computing techniques. Inform Processi Agricul. https://doi.org/10.1016/j.inpa.2016.10.005
31. Sivakumar B, Sowmya B (2016) An energy efficient clustering with delay reduction in data gathering (EE-CDRDG) using mobile sensor node. Wireless Pers Commun 90(2):793–806. https://doi.org/10.1007/s11277-016-3214-z
32. Varma, P. (2017). Rice productivity and food security in India: A study of the system of rice intensification. *Rice Productivity and Food Security in India: A Study of the System of Rice Intensification*, https://doi.org/10.1007/978-981-10-3692-7
33. Varshney S (2016) Plant disease prediction using image processing techniques-a review. Int J Comput Sci Mob Comput 55(5):394–398
34. William, J.,O., Cruz, J. Dela, L. A., Jensen, P., S., & Valenzuela, I. (2013). Information technology communication and control, environment and management (HNICEM) The Institute of Electrical and Electronics Engineers Inc. (IEEE)-Philippine Section 12–16. Nanotechnology.

Authors and Affiliations

K. S. Archana[1] · S. Srinivasan[2] · S. Prasanna Bharathi[3] · R. Balamurugan[4] · T. N. Prabakar[5] · A. Sagai Francis Britto[6]

S. Srinivasan
srinivasan.me.03@gmail.com

S. Prasanna Bharathi
prasanns@srmist.edu.in

R. Balamurugan
balawow@gmail.com

T. N. Prabakar
prabakar.ece@sairamce.edu.in

A. Sagai Francis Britto
sf_britto@yahoo.co.in

[1] Department of Computer Science and Engineering, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Chennai, India

[2] Institute of Biomedical Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Saveetha Nagar, Thandalam, Chennai, Tamilnadu 602105, India

[3] Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, Vadapalani Campus, No 1, Jawaharlal Nehru Salai, Vadapalani, Chennai, Tamilnadu, India

[4] Department of Electronics and Communication Engineering, K. Ramakrishnan College of Engineering, Tiruchirapalli, Tamilnadu, India

[5] Department of ECE, Sri Sairam College of Engineering, Bengaluru 562106, India

[6] Department of Mechanical Engineering, Rohini College of Engineering and Technology, Palkulam, Kanyakumari District, Tamilnadu, India