International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-3, September 2019

Detection and Classification of Paddy Crop Disease using Deep Learning Techniques

Usha Kiruthika, Kanagasuba Raja S, Jaichandran R, Priyadharshini C

Abstract*: Agricultural production plays a vital role in Indian economy. The biggest menace for a farmer is the various diseases that infect the crop. Quality and high production of crops is involved with factors like efficient detection of diseases in the crop. The disease detection though Naked-eye observation of expert can be prohibitively expensive and requires meticulous and scrupulous analysis to detect the disease. The existing systems on disease detection is not efficient enough in terms on real time basis. This paper presents an effective method for identification of paddy leaf disease. The proposed approaches involves pre-processing of input image and the paddy plant disease type is recognized using Gray-Level Co-occurrence Matrix (GLCM) technique and classifiers namely Artificial Neural Networks is used for better accuracy of detection. This method will be very useful to farmers to detect paddy diseases beforehand and thus prevent over usage of pesticides which in turn affects the crop production.

***Keywords* : Paddy disease detection, Artificial Neural Networks, Grey Level co-occurrence matrix.**

I. INTRODUCTION

Agriculture provides the major contribution for the Indian economy. The rural households are dependent on agriculture for their livelihood. Detection of crop diseases in various stages of crop cultivation is challenging for farmers. The disease detection though Naked-eye observation of expert can be prohibitively expensive and requires meticulous and scrupulous analysis. Expert guidance to help farmers is either inaccessible or non-existent in most cases. Hence, there is a demand for a system to assist farmers in detecting the disease during each stage of crop cultivation. The biggest menace for a farmer is the various diseases that infect the crop. Quality and quantity of high production of crops is involved with factors like efficient detection of diseases in the crops. A farmer may be unaware of the disease that affects the crop during various stages of the crop cultivation. India is an agriculture based country and agriculture is considered as the key for human civilization. Paddy is considered as one of the major crops in the Indian agriculture. The cultivation of the paddy crop for optimum yield and quality production is becoming a major challenge due to plant diseases. Plant disease is the major cause for the reduction of quantity and quality of the agricultural crops. Fungal and bacterial diseases generally affect paddy crops. The major diseases occurring in paddy crop are Brown Spot Disease (BSD), Leaf Blast Disease (LBD), and Bacterial Blight Disease (BBD).

The proper detection of these diseases at early stage is very important for the effective growth of the crop. The proposed system focuses to effectively detect the diseases and classify them thereby preventing the damage of the crop due to diseases. The proposed methodology can be classified into four phases: the first phase is the capturing of the image. This capturing of the image is done using a digital camera or a web camera. Once the image is captured filtering techniques are applied to remove the noise and unwanted data. The second phase is the segmentation of the captured image which is done using SLIC(Simple Linear Interactive Clustering). SLIC is a superpixel based algorithm. A superpixel is a group of pixels with similar characteristics. It is a colour based segmentation method. The third phase is the feature extraction of the segmented image. The feature extraction is carried out using the histogram approach. The fourth phase is the classification of the diseases which is done by comparing the processed image values with the stored dataset values. The classification is carried out using the deep learning technique Artificial Neural Network(ANN). After all these phases an indication message is sent to the farmer through a GSM module. The message contains details about the percentage of the crop affected and the suitable pesticide to be used will also be recommended.

II. RELATED WORKS

Crop diseases create a major problem in cultivation of food or in food industry. Due to increase in population, there is parallel increase in food production. Xanthomonasoryzae is a bacterium that causes bacterial blight of rice and it leads to almost 50% of the worldwide annual yield loss [1] [2]. Prakash et al. (2017) listed and compared several techniques of image processing in order to identify diseases in plants [3]. The paper focused on identification of diseased part of crops using image processing techniques.

Manuscript published on 30 September 2019
* Correspondence Author
Usha Kiruthika*, SRM Institute of Science and Technology, Kattankulathur Campus, Chennai. Tamil Nadu, India.
Kanagasuba Raja S, Department of IT, SRM Eswari Engineering College, Chenaai, Tamil Nadu, India.
Jaichandran R*, Department of CSE, Aarupadai Veedu Institute of Technology, Vinayaka Mission's Research Foundation (Deemed to be University), Rajiv Gandhi Salai, Old Mamallapuram Road, Paiyanoor 603104, Kanchipuram (DT), Tamil Nadu, India,rjaichandran@gmail.com.
Priyadharshini C, Department of Information Technology, Easwari Engineering College, Ramapuram Campus, Chennai.

In another paper [4] study related to several diseases identified within the fruits is proposed. An automated approach is proposed to identify the diseases that affect the inner parts of fruits. The image is preprocessed to reduce noise and as per the analysis, the bacterial blight disease affects most of the fruits crops. To identify the disease at an early stage, singular valued analysis is proposed. An image segmentation using super pixel-based graph representation is introduced in this paper [5] and the techniques of SLIC super pixels are adopted. It is efficient than the pixel-based segmentation algorithms. In this paper [6], the paddy plant disease is detected and recognized. Haar-like features and AdaBoost classifier is used to detect the affected portion of the leaf [6]. This proposed system [7], preprocess the input image and GLCM is used for texture feature extraction. The region-based segmentation is done using k-means clustering. The disease is prediction using the KNN classifier [7]. This paper [8] proposes the advanced Neural Network (NN) to process hyperspectral data. A review on NN mechanism, types, models, and classifiers that use different algorithms is given. This paper highlights the current state of imaging and non-imaging of the hyperspectral data. A power tool for detection and diagnosis has emerged through the hybridization of NN hyperspectral approach. The ratio of different spectral bands of pure disease spectrais called as Spectral Disease Index (SDI). This paper introduces Neural Network techniques and they also explain the difficulties of hyperspectral data [8]. This paper explains about some of the plant diseases that affect the productive part of the crop and cause heavy loss in production [8].

III. METHODOLOGY

The proposed system aims to detect and classify the diseases in paddy crop and recommend the suitable amount of pesticides to be used for higher productivity. The most commonly affected diseases taken into consideration are leaf blast disease, brown spot disease and bacterial blight disease. Bacterial leaf blight commonly affects the leaves. The symptoms are elongated lesions on the leaf tip. The lesions usually appear in yellow or white color. Brown spot disease also affects the leaves. The symptoms are round to oval shape. The color is reddish brown to dark brown. Leaf blast disease affects the leaves. The symptoms are small spots on the leaf scattered in a non-linear fashion. The proposed system aims to detect and classify the diseases in paddy crop and recommend the suitable amount of pesticides to be used for higher productivity. The figure 2 provides the architecture diagram of the proposed system. The system consists of following modules.

A. Preprocessing

Preprocessing is used to suppress noise or any irregularities present in the captured image that is used for further processing. Noise reduction is a typical preprocessing technique and Median filter is used for this purpose. Median filter is a non-linear filter used to remove noise and enhance the quality of the image to produce better results at later stages. This technique removes the noise by preserving the edges of the image under certain conditions. In any method used for enhancing an image, the noise in the image has to be minimized, but intrinsic information of the image has to be kept intact. The median filter looks at the nearby pixels in the image to decide whether it is a part of the surrounding or not. It does not replace the pixels by just calculating the mean value instead, the median filter calculates the median values and replaces the values. In our proposed system 3x3 dimension (9 elements) is taken and the median value is calculated and the values are replaced which does not degrade the quality of the image. The median value is calculated by considering the surrounding values by arranging them in numerical order and replacing the other values with the centre value. The median filter is more advantageous than the mean filter since the median value is one of the values in the group of pixels and does not create any unrealistic new pixel as in the case of mean filter.

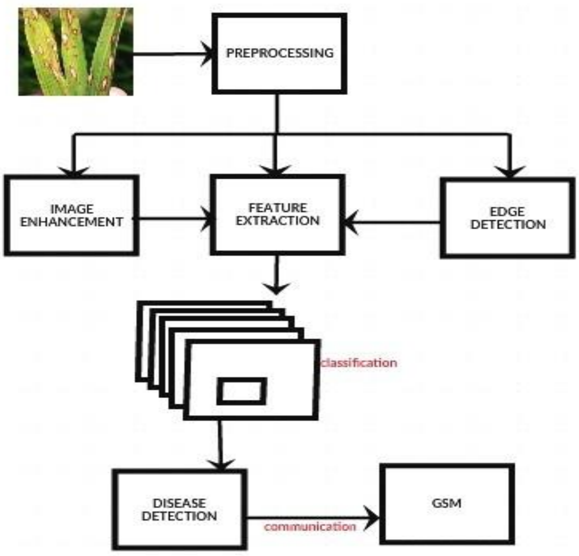

Fig.1. Flow chart of proposed system

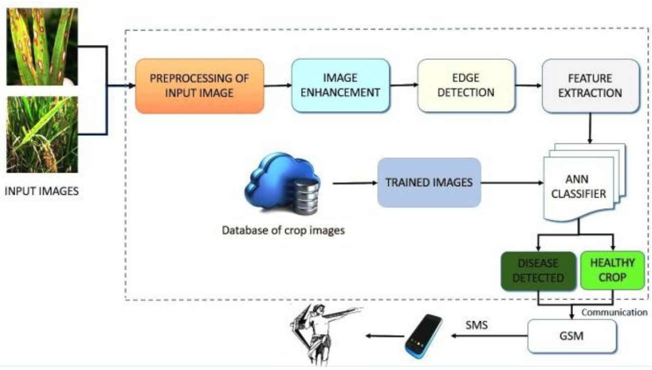

Fig.2. Architecture diagram for detection of disease in paddy crop

B. Edge Detection

Edge detection is a technique used for finding the boundaries of objects within the captured images. In our proposed system, the Sobel edge detector is used for detecting the edges of the leaf of the paddy crop and the edges of the diseased portion. Sobel edge detector is used to detect the object contour.

It performs a gradient measurement on the captured image and finds regions of high spatial frequency that match with edges. The approximate absolute gradient magnitude is calculated for the input grayscale image at each point of the image.

1. The gradient of the image is calculated for each pixel position in the image. The magnitude of

   $$\Delta f = mag(\Delta f) = [G_x^2 + G_y^2]^{1/2}$$

   the vector is calculated as, Where Gx is for x direction and Gy for y direction. the matrix value of the entire image is used
2. To find the x-direction derivative, subtract the first row value from the third row value using the mask

   $$Gx = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)$$
3. To find the y-direction derivative ,subtract the first column value from the third column value using the mask $Gy = (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)$
4. Using these values of Gx and Gy find the gradient. The advantage of using Sobel edge detector is that it provides better noise suppression and image smoothing.

C. Background Subtraction

The background of the captured image often varies according to the time and conditions. So the foreground image must be obtained by subtracting the background image. This subtraction of the background features helps in enhancing the foreground image and getting better results in further stages. To identify the object in the image morphological image difference operator is applied to subtract the background from the original image. This binary operation subtracts, image B from image A where A is the background image and B is the original image. This is defined as A-B=def{x|x ∈ A and x ∉ B}.

D. Segmentation

Segmentation is the technique that is used to extract the required portion from the image. In our system segmentation is used to find out the amount of diseased portion in the captured image. This is done by using the Simple Linear Interactive Clustering (SLIC) algorithm. The SLIC technique is a superpixel based algorithm that segments the image based on color intensities. After the edges of the image are obtained the small openings are connected by specifying the disc size i.e. for example disc size = 4 connects edges which are disconnected by a space of 4 pixels or less. The resulting image is called the dilated image. Then the diseased portions which are very small can be neglected. The larger portions are obtained by specifying the pixel size of the group and the obtained image is called the area open image. The image is represented in superpixels which reduces the processing complexity since each superpixel represents multiple pixels. The intensities of the different region are obtained and the regions with similar intensities are clustered together. The intensity of the diseased portion of the leaf is less compared to the healthy portions of the leaf.

We have found our method to perform better with respect to three aspects. First we use super pixels which consolidate multiple pixels and provide accurate information regarding boundaries. Instead of processing individual pixels, super pixels are processed which reduces complexity. The segmentation result of the algorithm matches object boundaries well and still is more efficient when compared to pixel based algorithms. Second, we use position information of pixels to improve spectral clustering technique in our work. Finally, boundary focused region merging improves our segmentation result. We have made the threshold to be adaptive according to the distribution of color information and the number of super pixels that form a cluster. We also calculate the difference of adjacent boundaries for determining whether boundaries have to be merged or not.

The proposed technique is effective and the results are consistent. We have compared our technique with to k-means method. The number of distance calculations to be made is reduced since super pixels are used. It results in optimized search space which is proportional to the number of super pixels. The complexity is reduced and it is linear with the number of pixels and does not depend on the number of super pixels. Another improvement in the proposed method compared to other methods is the fact that we apply weighted distance measurement. Because of this, the size and compactness of super pixel is controllable. The SLIC algorithm divides the input image into k blocks for a fixed parameter k. Cluster center is determined for each block. Clustering is done based on the distances of cluster centres repeatedly until convergence. Since the search space is limited, complexity is less and the super pixels are evenly sized.

E. Feature Extraction

Feature extraction represents a large set of data within a reduced amount of resources. In our paper, we use the Gray-Level Co-occurrence Matrix (GLCM) for feature extraction. In our algorithm we use statistical texture analysis for calculating texture features based on intensities. The GLCM method extracts second order statistical features. When creating GLCMs, an array of offsets describing direction and distance based pixel relationships are specified. Contrast, Energy, Correlation, Homogeneity are calculated using the GLCM feature. These features are represented using the variables E,H,I,S. These values are found for the input image and stored which will be used during the classification stage for the identification of the diseases.

1. Contrast

Contrast is a metric to describe the color intensity of that pixel and the neighboring pixel in the image. If they are equal and the image is constant then the contrast is 0. The equation is:

$$\text{Contrast} = \sum_{i,j=0}^{N-1} (P_{i,j})(i-j)^2$$

2. **Energy**
   Energy values can vary between 0 and 1. For a constant image, the energy would be 1. The energy is described by the following equation

$$\text{Energy}= \sum_{i,j=0}^{N-1} (P_{i,j})^2$$

3. **Homogeneity**
   Homogeneity is another feature metric that measures similarity among the pixels. It also ranges between 0 and 1. For a diagonal GLCM, the value of homogeneity metric would be 1. The equation of the Homogeneity is described by:

$$\text{Homogeneity}=\sum_{i,j=0}^{N-1} \frac{(P_{i,j})^2}{[1+(i-j)^2]}$$

4. **Correlation**
   Correlation is a measurement of how correlated a pixel is with its neighbourhood. Its range is in between -1 and 1.

$$\text{Correlation}=\sum_{i,j=0}^{N-1} P_{i,j}\left(\frac{(i-\mu)(j-\mu)}{\sigma^2}\right)$$

F. Classification

Classification is for classifying each image to the appropriate disease. It has to be precise for automatic recognition of diseases in leaves. Classification takes the output of the feature extraction as a input and the extracted feature values are compared with the training data set. In our system, ANN (Artificial Neural Network) is used to classify affected and healthy region in the image. ANN based on two modes of operation 1) Training mode 2) Using mode. ANN takes the signal between the images as real number and the output is computed as a non-linear function by weighting and adding the inputs, where weights can adjust themselves. Weight can either increase or decrease the strength of the signal of the image. ANN has threshold value that the signal of the image is sent only if it crossed by aggregated threshold. Back propagation based neural network with triangle function(Step Function, Sigmoid function, Sign Function) using three hidden layers perform different level of transformation. The process continues until the output matches any of the images in the data set.

IV. RECOMENNDATION OF PESTICIDES

On finding the percentage of disease affected in the plant, GSM (Global System for mobile communication) is used to intimate farmer about the disease through message. GSM is a Digital cellular technology used to send voice and data services operates at 850MHz, 900MHz,1800MHz,1900MHz frequency bands. Message sent to farmers contains the percentage of the diseases affected in the paddy crop along with the required amount of pesticides to be used in the affected regions. The three most commonly affecting diseases in paddy crop are Bacterial Blight, Brown Spot, leaf Blast.

Bacterial Blight commonly affects the leaves. The symptoms are elongated lesions on the leaf. The lesions usually appear in yellow or white colour, can be treated by using Terramycin 17, Brestanol, Agrimycin 500 and a combination of Agrimycin 100 + Fytolan will give an effective control of the blight phase of the disease. The combination of Agrimycin 100 + Fytolan spraying can give good control of the disease and are economically effective. Brown spot disease also affects the leaves. The symptoms are round to oval shape. The colour is reddish brown to dark brown, can be treated by spraying 1g of ediphenphos or 2g of mancozeb or 2.25g zineb in 1litre of water. Leaf Blast affects the leaves of the crop. The symptoms are small spots on the leaf scattered in a non-linear fashion .Seed can treated by pseudomonas fluorescence of 10g in 1 litre of water for 30 minutes. Seed root tip can be treated by pseudomonas fluorescence of 4g in 1 litre of water for 20 minutes. Foliar spray of pseudomonas fluorescence of 4g in 1 litre of water for 20 minutes after transplanting.

V. EXPERIMENTAL RESULTS

The experimental results of the proposed system are given as follows. The image of the paddy leaf is captured and it is given as the input to the system to recognize the kind of disease and the percentage of the leaf that is affected. This figure explains to select an input image that is to be processed to detect the disease present in it.

Figure 4 is the resized image of the source image.As every image would be in different sizes, it is commonly resized to a fixed frame size to enable consistent processing.

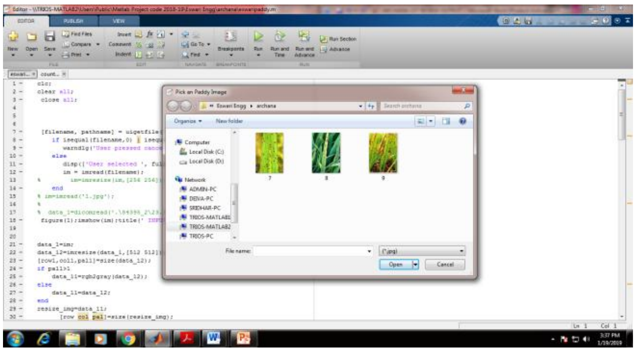

Fig. 3. Select an input image

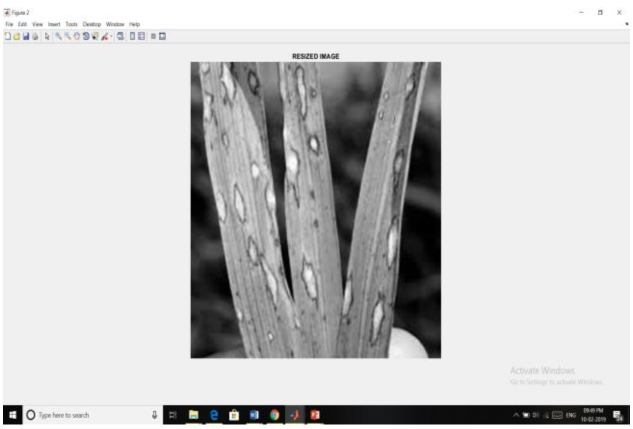

Fig. 4. Resized image

Retrieval Number C5506098319/2019©BEIESP
DOI: 10.35940/ijrte.C5506.098319
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

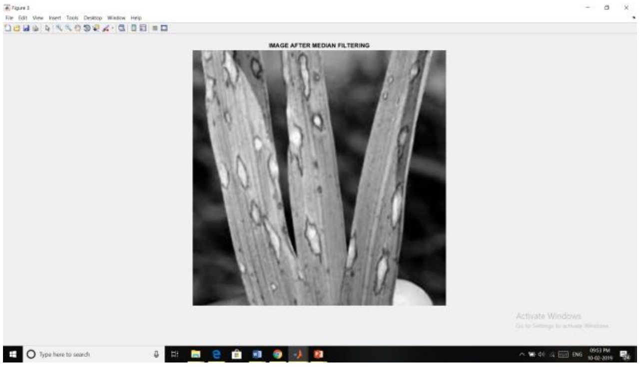

Fig. 5. Image after applying median filter

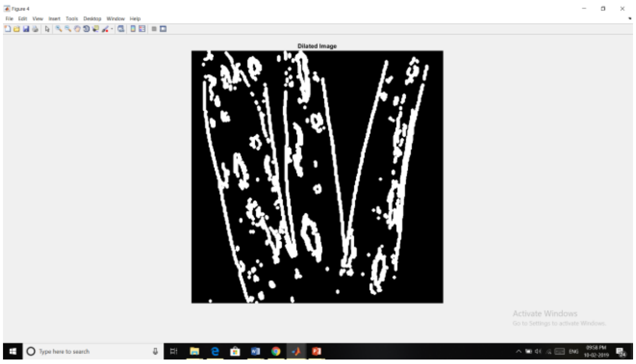

Fig. 6. Dilated image

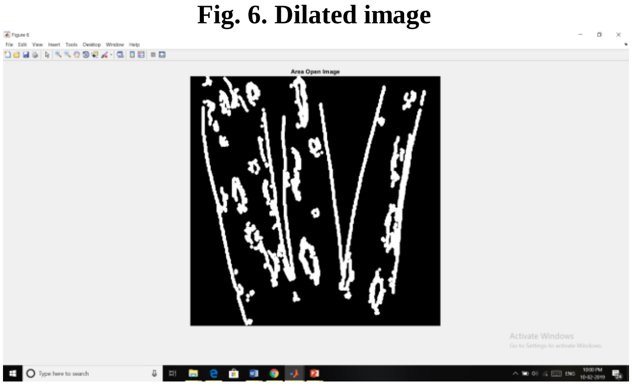

Fig. 7. Area open image

This figure 5 shows the resulting image after applying the median filter. The median filter is applied to remove the noise or any irregularities that are present in the captured image. The image is preprocessed at this stage and this is carried out to get better results in further stages. After the preprocessing stage, the edges of the paddy leaf and diseased portions are detected using the sobel edge detector. The image we get after preprocessing is called the dilated image. This image is converted to binary image by taking the values in 0's and 1's.It is represented in figure 6.

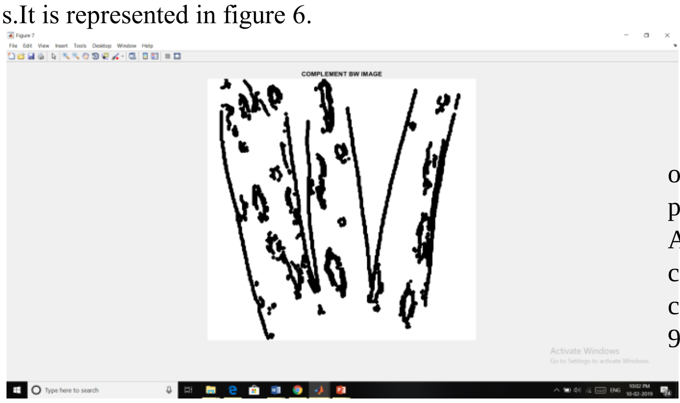

Fig.8. complement image

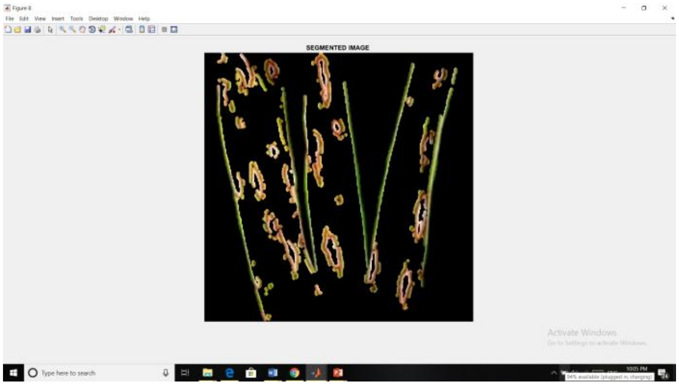

Fig. 9. Segmented image

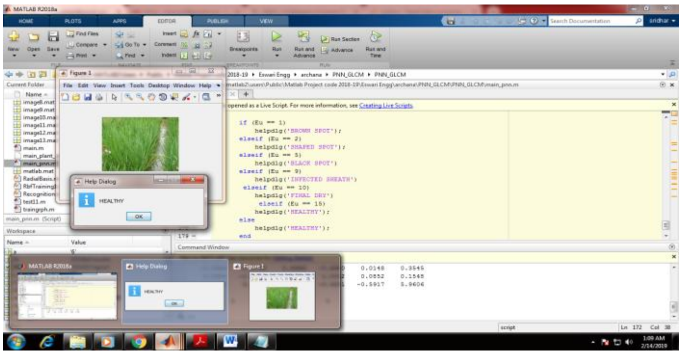

Fig. 10. Healthy crop detected

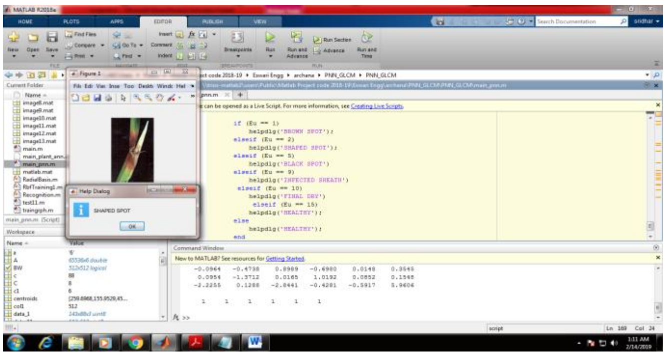

Fig. 11. Shaped spot disease detected

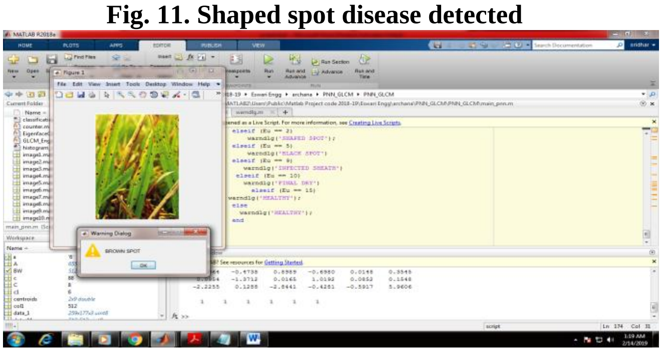

Fig. 12. Brown spot disease detected

Once the edges are detected, the areas with lesser number of pixels are eliminated. This is done by grouping neighbour pixels over a certain range. The resulting image is called as Area open image as shown in figure 7. Figure 8 represents the complement image where the white and black colors are complemented. This can be used for further processing.Figure 9 shows the segmented image.

The diseased portin and the edges of the paddy leaf are identified. The background portion of the input image is subtracted. After the process of segmentation the features of the segmented image are extracted. The features that are extracted are contrast, correlation, energy, homogeneity and they are represented in values which will be used to compare with the trained dataset and classify. The above figures 10, 11 and 12 show the results of classification.

Table- I: Performance table for paddy plant disease recognition with GLCM using ANN

Feature / Classifier	Precision (%)	Accuracy (%)	Recall (%)	Fscore (%)
SVM	86.66	91.10	86.66	86.66
ANN	90.60	93.33	90.00	90.14

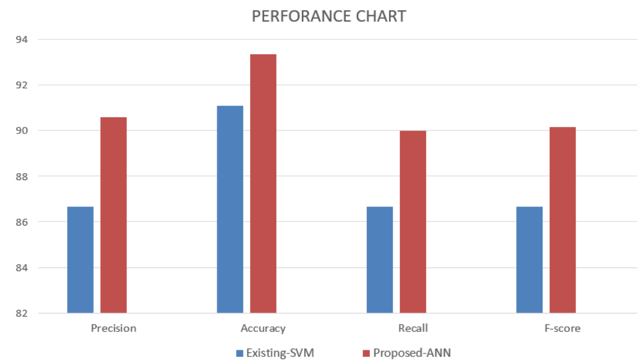

Fig. 13. Performance graph for paddy plant disease recognition with GLCM using ANN

The figure 10 shows the detection of a healthy paddy crop. Figure 11 and figure 12 show the images of disease detected. The diseases are classified by comparing the processed image with the trained dataset. This is done by using the deep learning technique Artificial Neural Network (ANN). Once the diseased portions are detected and the disease is identified it is notified to the farmer using a GSM module. The message contains details about the percentage of the crop that is affected, the identified disease and also recommends the suitable amount of pesticides that should be used. Figure 13 shows the comparison between the existing system that uses SVM for classification and the proposed system which uses Artificial Neural Network (ANN) for classification. It is observed that the proposed system provides an efficient way of detecting and identifying the diseases.

VI. CONCLUSION

The proposed system offers detection of diseases that majorly affect the paddy crops using recent technologies such as deep learning and image processing. The system provides the detection of diseases at early stages and notifies the farmer about the necessary action to be taken. The system enables production of quality crops largely and minimizes the losses incurred by the farmer due to infected crops. It minimizes the excessive use of pesticides and ensures healthy crop is produced in each cycle. This proposed system considers the three major diseases that affects the paddy crop and the future work may take all the diseases that affects the crop with better dataset for more accuracy.

REFERENCES

1. X. Yin, X. Liu, J. Chen and D. M. Kramer, "Joint Multi-Leaf Segmentation, Alignment, and Tracking for Fluorescence Plant Videos," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 6, pp. 1411-1423, 1 June 2018.
2. Sujan Sarkar, Sanket Biswas, Avinaba Tapadar, and Pritam Saha, "AI Based Fault Detection on Leaf and Disease Prediction Using K-means Clustering," in *International Research Journal of Engineering and Technology (IRJET)*, Vol. 05, Issue 03, pp.926-931, Mar-2018.
3. R. M. Prakash, G. P. Saraswathy, G. Ramalakshmi, K. H. Mangaleswari and T. Kaviya, "Detection of leaf diseases and classification using digital image processing," *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, Coimbatore, 2017, pp. 1-4.
4. K. Kaur and C. Marwaha, "Analaysis of diseases in fruits using image proccessing techniques," *2017 International Conference on Trends in Electronics and Informatics (ICEI),* Tirunelveli, 2017, pp. 183-189.
5. Chi-Yu Hsu and Jian-Jiun Ding, "Efficient image segmentation algorithm using SLIC superpixels and boundary-focused region merging," *2013 9th International Conference on Information, Communications & Signal Processing*, Tainan, 2013, pp. 1-5.
6. K.Jagan mohan, M.Balasubramanian, and S.Palanivel, "Detection and recognition of diseases from paddy plant leaf images," in *International Journal of Computer Applications*, Vol. 144 – No.12, pp. 34-41, June 2016.
7. Jaskaran Singh, Harpreet Kaur, "Plant Disease Detection Based on Region-Based Segmentation and KNN Classifier," *Proceedings of the International Conference on ISMAC in Computational Vision and Bio-Engineering 2018 (ISMAC-CVB).*
8. Kamlesh Golhani, Siva K.Balasundram, Ganesan Vadamalai, and Biswajeet Pradhan,, "A review of neural networks in plant disease detection using hyperspectral data," in *Information Processing in Agriculture*, Vol. 5, Issue 3, pp. 354-371, September 2018.

AUTHORS PROFILE

Dr. Usha Kiruthika is currently working as an Assistant Professor in the Department of Computer Science and Engineering in SRM Institute of Science and Technology, Kattankulathur Campus. She has completed her Ph.D at Anna University and she has many publications to her credit. Her areas of interest include artificial intelligence and deep learning.

Dr. S.Kanaga Suba Raja, is working as an Associate Professor in Easwari Engineering College. He obtained his Ph.D degree in Computer Science in 2013 from Manonmaniam Sundaranar University, Tirunelveli. He has about 14 years of teaching and research experience. He has successfully guided many graduate and under-graduate students for their research projects. He has several papers published in International journals and conferences to his credit. His current research interests are Wireless Body Area Networks, Ad hoc and Sensor Networks and Mobile and Ubiquitous Computing.

Dr.R.Jaichandran is currently working as Head of the department CSE, Aarupadai Veedu Institute of Technology an ambit institution of Vinayaka Missions Research Foundation (Deemed to be University), Tamil Nadu, India. He has 13 years of experience in academics, industry, research, anddevelopment activities. Published 33 research papers in referred Journals and Conferences.

His area of Interest includes Wireless Sensor Networks, Internet of Things (IoT), Ethical Hacking, Big data Analytics, and Embedded systems.

C.Priyadharshini is currently working as an Assistant Professor in the Department of Inormation Technology in Easwari Engineering College, Ramapuram Campus. She has completed her M.E at Anna University and she has many publications to her credit. Her areas of interest include Wireless Networks.

Retrieval Number C5506098319/2019©BEIESP
DOI: 10.35940/ijrte.C5506.098319
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication