

Detection and Recognition of Diseases from Paddy Plant Leaf Images

K. Jagan Mohan
Assistant Professor
Dept. of CSE
Annamalai University
Annamalainagar – 608002

M. Balasubramanian
Assistant Professor
Dept. of CSE
Annamalai University
Annamalainagar – 608002

S. Palanivel
Professor
Dept. of CSE
Annamalai University
Annamalainagar – 608002

ABSTRACT

In agricultural field, paddy cultivation plays a vital role. But their growths are affected by various diseases. There will be decrease in the production, if the diseases are not identified at an early stage. The main goal of this work is to develop an image processing system that can identify and classify the various paddy plant diseases affecting the cultivation of paddy namely brown spot disease, leaf blast disease and bacterial blight disease. This work can be divided into two parts namely, paddy plant disease detection and recognition of paddy plant diseases. In disease detection, the disease affected portion of the paddy plant is first identified using Haar-like features and AdaBoost classifier. The detection accuracy rate is found to be 83.33%. In disease recognition, the paddy plant disease type is recognized using Scale Invariant Feature Transform (SIFT) feature and classifiers namely k-Nearest Neighbour (k-NN) and Support Vector Machine (SVM). By this approach one can detect the disease at an early stage and thus can take necessary steps in time to minimize the loss of production. The disease recognition accuracy rate is 91.10% using SVM and 93.33% using k-NN.

Keywords

Pre-processing, Disease detection, Disease recognition, Haar-Like features, AdaBoost classifier, SIFT features, k-NN classifier.

1. INTRODUCTION

India is agriculture based country that has many people working in the agriculture industry. The agricultural sector plays an important role in economic development by providing rural employment. Paddy is one of the nation's most important products as it is considered to be one of India's staple food and cereal crops and because of that, many efforts have taken to ensure its safety, one of them is crop management of paddy plants. Paddy plants are affected by various fungal and bacterial diseases. This work focuses on recognizing three paddy plant diseases namely Brown Spot Disease (BSD), Leaf Blast Disease (LBD), Bacterial Blight Disease (BBD). The proper detection and recognition of disease is very important in applying required fertilizer. BSD is one of the most common fungal disease and most damaging paddy plant disease. They are at first little, roundabout and dim chestnut to purple cocoa. Completely created sores are round to oval with a light cocoa to dark focus, encompassed by a rosy chestnut edge brought on by the poison delivered by the organisms. BBD is caused by bacteria *Xanthomonas oryzae*. It causes wilting of seedlings and yellowing and drying of leaves. LBD is caused by fungus *Magnaporthe oryzae*. Starting indications show up as white to dim green injuries or spots, with dull green outskirts. More seasoned sores on the leaves are circular or shaft formed and whitish to dark focuses with red to caramel fringes. Some take after jewel shape, wide in the middle and indicated either end.

The main objective of this work is to develop a system for classifying the paddy plant diseases using image processing technique.

2. RELATED WORK

P. R. Rothe [1], has proposed a work in light of example acknowledgment framework for distinguishing proof and order of three cotton leaf maladies in particular Bacterial Blight, *Myrothecium*, *Alternaria*. Image Segmentation is done utilizing dynamic shape model. Versatile neuro-fuzzy derivation framework utilized Hu's moments as components for the preparation technique. The arrangement precision is 85 percent.

Viraj A. Gulhane [2], has tended to the issue of finding of sicknesses on cotton leaf utilizing Principal Component Analysis and Nearest Neighborhood Classifier. It includes sicknesses like Blight, Leaf Necrosis, Gray Mildew, *Alternaria* and Magnesium Deficiency. The characterization precision is 95 percent.

Rong Zhou et.al., [3], has introduced a novel strategy for vigorous and early *Cercospora* leaf spot identification in sugar beet utilizing hybrid algorithm of template matching support vector machine. This technique is strong and practical for ahead of schedule identification and persistent quantization under normal light conditions. Sometimes the recognition gets to be mind boggling because of the unpredictable changes of outer environment.

John William ORILLO et.al., [4], has joined the computerized picture handling systems to dispose of the manual examination of maladies in rice plant that more often than not happen in Philippine's farmlands in particular Bacterial leaf blight, Brown spot and Rice blight. The sickness recognizable proof precision is 100 percent around 134 example images.

Auzi Asfarian et.al., [5], has endeavored to distinguish the four noteworthy paddy sicknesses in Indonesia to be specific leaf blast, brown spot, bacterial leaf blight and tungro. Fractal descriptors are utilized to dissect the composition of the sores. The infection distinguishing proof precision is 83 percent.

Kholis Majid et.al., [6], has added to a portable application for paddy plant malady identification framework utilizing fuzzy entropy and Probabilistic neural system classifier that keeps running on Android versatile's working framework. It includes for sorts of maladies in particular brown spot, leaf blast, tungro and bacterial leaf blight. The exactness of paddy sicknesses distinguishing proof is 91.46 percent.

3. OUTLINE OF THE WORK

This recognition system involves feature extraction and classification of Diseases using K-NN and SVM proposed work is described in Section 4. Disease detection is described in section 5. Feature extraction is described in Section 6 Image Classification

is described in Section 7. Experimental results are described in Section 8. Conclusion and future scope are described in Section 9.

4. PROPOSED WORK

The block diagram for proposed system is shown in the Fig. 1. It is an image recognition system for identifying the paddy plant diseases that first involves disease detection and then disease recognition. This work does not involve any colour and shape feature extraction techniques instead it uses SIFT (Scale Invariant Feature Transform) for extracting the local features in images that have sudden changes in intensities by filtering images at various scales and patches of interest.

Disease detection part uses Haar-like feature and AdaBoost (Adaptive Boosting) classifier to locate the disease affected portion of the paddy plant. Disease recognition part uses SIFT (Scale Invariant Feature Transform) feature extraction and two classifiers namely k-NN (k-Nearest Neighbours) and SVM (Support Vector Machine) to recognize the various categories of diseases like brown spot, leaf blast and bacterial blight.

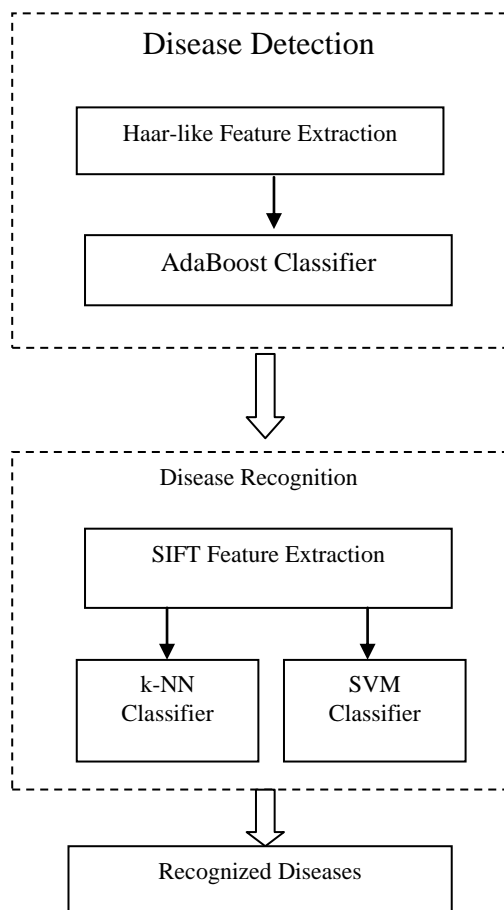


Fig 1. Block diagram for the proposed system

5. DISEASE DETECTION USING HAAR-LIKE FEATURE AND ADABOOST CLASSIFIER

5.1 Haar- Like Feature Extraction

Haar-like elements are advanced picture highlights utilized for identification purposes. They claim their name because of natural closeness with Haar wavelets. Verifiably working with just picture intensities (i.e., the RGB pixel values at every last pixel of picture) made the undertaking of highlight count computationally costly. Viola and Jones adjusted the thought of utilizing Haar wavelets and grew purported Haar-like elements.

A Haar like element considers adjoining rectangular region at a particular area in a detection window, whole up the pixel intensities in every locale and figures the distinction between these totals. This distinction is then contrasted with a scholarly limit that isolates infected locale from non-infected area. The principle point of preference of Haar such as highlight is its high computation speed. Haar like component of any size can be ascertained in consistent time.

A basic rectangular Haar like component can be characterized as the distinctions of the aggregate of pixels of zones inside the rectangle, which can be at any position and scale inside of the first picture. This altered list of capabilities is called 2-rectangle highlight. The value indicates certain characteristics of a particular area of the image. Rectangular Haar like feature type is shown in fig 2. which indicates (a) horizontal variations, (b) vertical variations, (c) horizontal changes and (d) diagonal variations of pixel in images respectively.

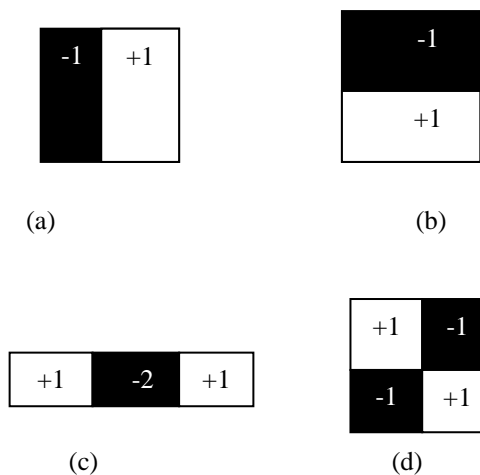


Fig. 2 Haar like rectangular feature type

The feature vector is given as $F(x)=w1.Sum(r1)+w2.sum(r2)$ where r1 and r2 represents the darker rectangular region and lighter rectangular region respectively. And w1 and w2 refers to the weights which can be negative(-1) and positive(+1) respectively.

5.2 Adaboost Classifier

Adaboost classifier another way to say "Versatile Boosting" is a machine learning meta calculation. It is regularly alluded to as the best-out-of-box classifier. Adaboost preparing process chooses just those elements known to improve correctness power of the model, lessening dimensionality and conceivably enhancing execution time as unessential components should not be registered. Adaboost classifier is shown in the fig.3

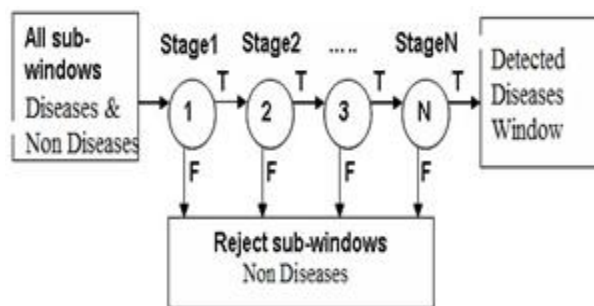


Fig. 3 Adaboost Classifier

Input image is subdivided into the number of sub-windows. Rectangle prototypes are measured independently in vertical and horizontal directions. Rectangle prototypes are applied to all sub-

windows of the image for getting complete set of rectangle features. Rectangle features are clustered into N-sets (N is normally 20). A classifier is taught using AdaBoost learning algorithm for each set of rectangular features. N-stage classifier is created from the N-sets of rectangle features.

5.3 Testing and Training of AdaBoost classifier

For paddy plant disease detection, totally 60 disease affected sample images are taken. Each image undergoes training and testing using AdaBoost classifier algorithm. Based on number of iterations, thus the training stages and False alarm rate is getting varied for every input image given to the classifier. The AdaBoost classifier is mainly used to convert the weak classifier to strong classifier. The various paddy plant disease images are taken such as brown spot, leaf blast and bacterial blight. Among the 60 diseased images only 50 images are detected accurately. The detection rate for diseased affected image is 83.33%.

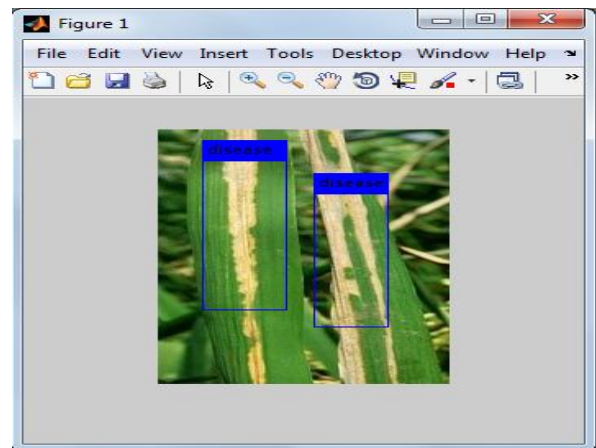


Fig. 6 Screen shot for disease detection of bacterial blight

The cropped view of diseased image after detection are shown in the figures below

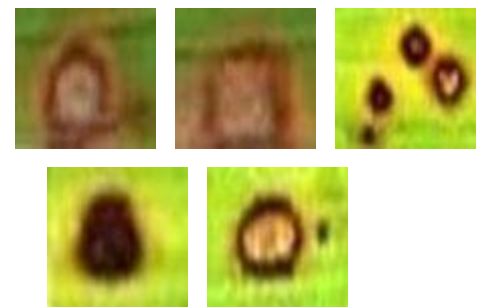


Fig.7 Cropped view of brown spot disease



Fig.8 Cropped view of leaf blast disease



Fig.9 Cropped view of bacterial blight disease

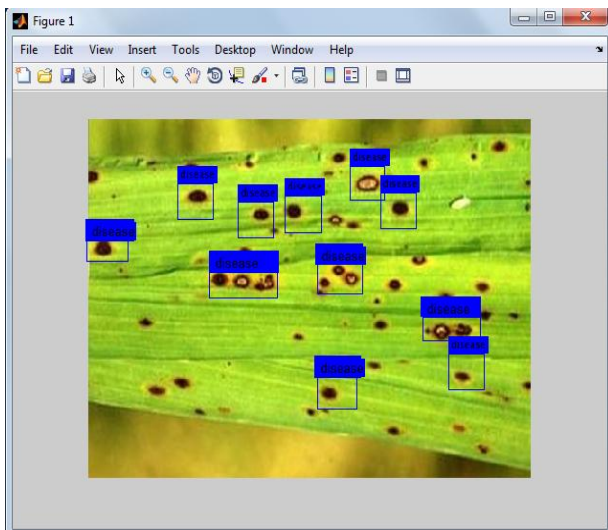


Fig.4 Screen shot for disease detection of brown spot

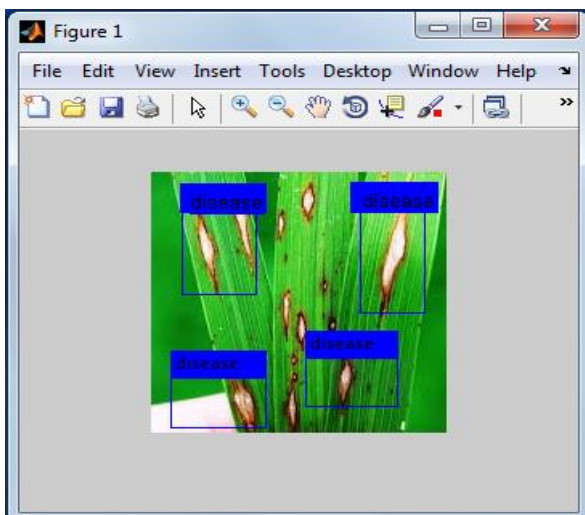


Fig.5 Screen shot for disease detection of leaf blast

6. FEATURE EXTRACTION IN DISEASE RECOGNITION

6.1 Scale Invariant Feature Transform(SIFT)

Filter is a strategy to extricate unmistakable elements from dark level pictures, by sifting pictures at different scales and fixes of fad that have sharp changes in neighborhood picture intensities. The SIFT calculation comprises of four noteworthy stages: Scale-

space extrema detection, Key point localization, Orientation Assignment and Representation of a key point descriptor.

The elements are situated at maxima and minima of a Difference of Gaussian (DoG) capacities connected in scale space. Next, the descriptors are processed as an arrangement of orientation histograms on 4x4 pixel neighborhoods, and every histogram contains 8 bins. This prompts a SIFT feature vector with 4x4x8 (i.e., 128) measurements on every patch.

6.1.1 Detection of Scale-Space Extrema

The principal stage is to develop a Gaussian "scale space" function from the input image. This is shaped by convolution of the original image with Gaussian elements of shifting widths. The scale space of a image is characterized as a capacity $L(x,y,\sigma)$ that is delivered from the convolution of a variable-scale Gaussian, $G(x,y,\sigma)$ with an input image, $I(x,y)$:

$$L(x,y,\sigma) = G(x,y,\sigma) * I(x,y)$$

Where * is the convolution operation in x and y, and

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$$

To effectively distinguish stable keypoint areas in scale space, David Lowe proposed utilizing scale-space extrema as a part of the difference of Gaussian capacity convolved with the image, $D(x, y, \sigma)$, which can be processed from the distinction of two adjacent scales isolated by a consistent multiplicative element k :

$$\text{DoG}(x, y, \sigma) = (G(x,y,k\sigma) - G(x, y, \sigma)) * I(x,y)$$

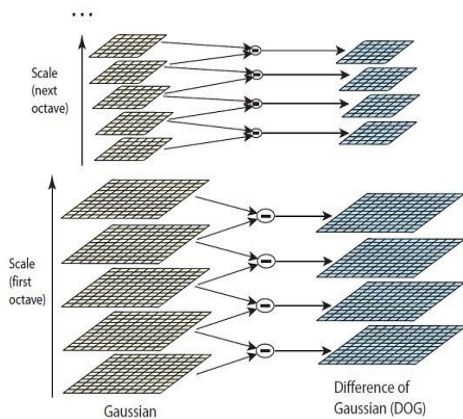


Fig.10 Difference of Gaussian

There are various explanations behind picking this function. To begin with, it is an especially effective capacity to register, as the smoothed images, L , should be processed regardless for scale space feature depiction, and DoG can be figured by simple image subtraction as appeared in the fig.10. To identify the nearby maxima and minima of $\text{DoG}(x, y, \sigma)$ every point is contrasted with its 8 neighbors at the same scale, and its 9 neighbors all over one scale as appeared in the fig.11. In the event that this value is the base or greatest of every one of these focuses then this point is an extrema.

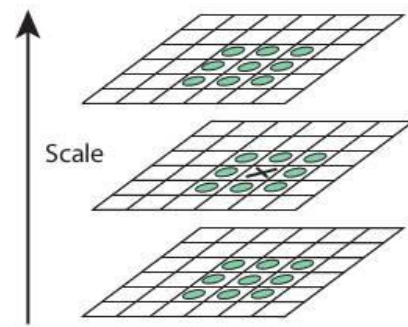


Figure 11. Detection of scale space extrema

6.1.2 Key Point Localization

This stage endeavours to dispense with a few points from the competitor rundown of key points by finding those that have low difference or are inadequately limited on an edge. The estimation of the key points in the DoG pyramid at the extrema is given by:

$$D(z) = D + \frac{1}{2} \frac{\partial D^{-1}}{\partial x} z$$

On the off chance that the threshold esteem at z is underneath limit esteem, then this point is barred.

To eliminate inadequately limited extrema we utilize the way that in these cases there is a huge principle curvature over the edge yet a little recede and flow in the opposite course in the distinction of Gaussian function. A 2x2 Hessian network, H , registered at the location and scale of the keypoints is utilized to discover the curvature. With these equations, the proportion of principle curvature can be checked proficiently.

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$$

6.1.3 Orientation Assignment

This step expects to allocate a steady orientation to the key points in light of nearby image properties. An orientation histogram is framed from the gradient orientation of test points inside of an area around the key points. A 16x16 square is picked in this usage. The orientation histogram has 36 bins covering the 360 degree scope of orientations. The gradient magnitude, $m(x, y)$, and orientation, $\theta(x, y)$, are pre-processed utilizing pixel contrasts:

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2}$$

$$\theta(x,y) = \tan^{-1} \left(\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)} \right)$$

Every sample is weighted by its slope extent and by a Gaussian-weighted roundabout window with a σ that is 1.5 times that of the size of the key focuses. Peaks in the orientation histogram compare to predominant headings of nearby inclinations. We find the most elevated peak in the histogram and utilize this peak and some other neighbourhood peak inside 80% of the height of this peak to make key points with that orientation. A few points will be provided with various orientations if there are numerous peaks of comparative greatness. A Gaussian circulation is fit to the 3 histogram values nearest to every peak to insert the peaks position for better precision. This registers the location, orientation and scale of SIFT features that have been found in the image. These features react emphatically to the corners and force angles. The length of the arrow demonstrates the greatness of the difference at the key focuses, and the arrow indicates from the dark to brighter side.

6.1.4 Keypoint Descriptor

In this stage, a descriptor is registered for the nearby image region that is as particular as could be expected under the circumstances at every hopeful key point. The image gradient magnitudes and orientations are tested around the key point area. A Gaussian weighting function with σ identified with the size of the key point is utilized to relegate a weight to the magnitude. We utilize an σ equivalent to one a large portion of the width of the descriptor window in this execution. To accomplish orientation invariance, the directions of the descriptor and the gradient orientations are pivoted in respect to the key point orientation. This procedure is demonstrated in fig.12 a 16x16 sample cluster is figured and a histogram with 8 bins is utilized. So a descriptor contains 16x16x8 components altogether.

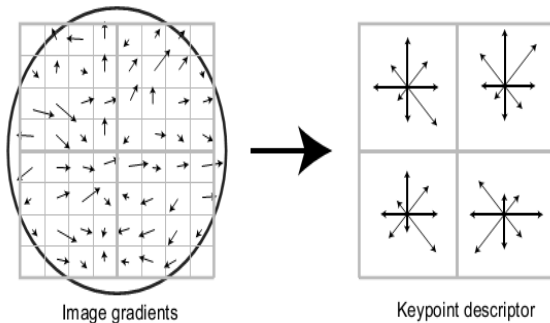


Fig.12 Building keypoint

These are weighted by a Gaussian window, demonstrated by the Overlaid circle. The image gradients are added to an orientation histogram. Every Histogram incorporate 8 directions showed by the arrows and is processed from 4x4 Sub locales. The length of every arrow compares to the whole of the gradient magnitude close to that course inside of the region.

7. CLASSIFICATION IN DISEASE RECOGNITION

7.1 SVM Classifier

Support vector machine (SVM) depends on the Structural Risk Minimization (SRM). Like RBFNN, Support vector machines can be utilized for pattern classification and nonlinear regression. SVM builds a linear model to appraise the decision function utilizing non-linear class limits taking into account support vectors. On the off chance that the information are straightly isolated, SVM trains linear machines for an ideal hyper plane that isolates the information without mistake and into the most extreme separation between the hyper plane and the nearest training points. The training points indicating the nearest the ideal isolating hyper plane are called support vectors. Fig. 13 demonstrates the architecture of the SVM. SVM maps the input image into a higher dimensional element space through some nonlinear mapping picked from the earlier. A direct choice surface is then developed in this high dimensional component space. Along these lines, SVM is a straight classifier in the parameter space, however it turns into a nonlinear classifier as an after effect of the nonlinear mapping of the space of the data designs into the highdimensional element.space.

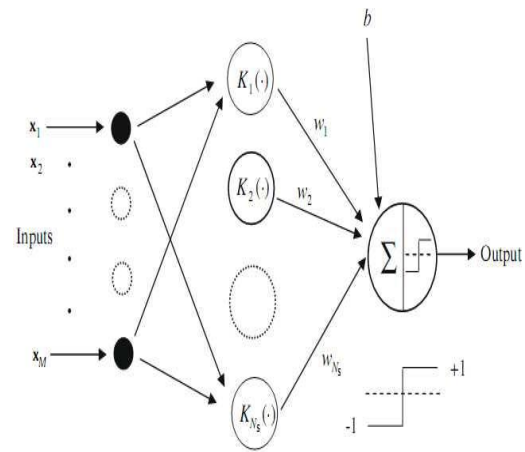


Fig.13 Architecture of SVM

7.1.1 SVM Principle

Support vector machine (SVM) boycott be utilized for ordering the got information (Burges, 1998). SVM are an arrangement of related directed learning strategies utilized for classification and regression. They fit in with a group of summed up straight classifiers. Give us a chance to denote a input vector (termed as example) by $x=(x_1, x_2, \dots, x_n)$ and its class label by y such that $y = \{+1, -1\}$. In this manner, consider the issue of isolating the arrangement of n-training patterns having a place with two classes.

7.1.2 SVM for Linearly Separable Data

A linear SVM is used to isolate data sets which are linearly separable. The SVM linear classifier extends the margin between the separating hyperplane. The features lying on the maximal margins are called support vectors. Such a hyper plane with maximum margin is called maximum margin hyperplane. In case of linear SVM, the discriminate function is of the form:

$$g(x) = \omega^t x + b$$

Such that $g(x_i) \geq 0$ for $y_i = +1$ and $g(x_i) < 0$ for $y_i = -1$. In other words, training samples from the two different classes are detached by the hyperplane $g(x) = \omega^t x + b = 0$. SVM finds the hyperplane those results in the largest separation between the decision function values from the two classes. Theoretically, this hyperplane can be found by minimizing the following cost function:

$$j(\omega) = \omega^t \omega$$

For the linearly separable case, the decision rules is scripted by an optimal hyperplane separating the binary decision classes are given in the following equation in terms of the support vectors:

$$Y = \text{sign} \left(\sum_{i=1}^N y_i (xx_i) + b \right)$$

Where Y is the outcome, y_i is the class value of the training sample x_i and represents the inner product. The vector tends to an input and the vectors $x_i, i = 1 \dots N_s$, are the support vectors.

7.1.3 SVM for Linearly Non- Separable Data

For non-linearly separable data, it depicts the data in the input space into a high dimension space with kernel function $\Phi(x)$, to find the separating hyperplane.

$$Y = \text{sign} \left(\sum_{i=1}^N y_i k(xx_i) + b \right)$$

7.1.4 Determining Support Vectors

The support vectors are the (transformed) training patterns. The support vectors are (equally) close to hyperplane. The support vectors are training samples that gives the optimal separating hyperplane and are the most difficult patterns to classify.

Informally speaking, they are the patterns most informative for the classification task.

7.2 k-NN Classifier

In example acknowledgment, the k-Nearest Neighbours algorithm (or k-NN for short) is a non-parametric strategy utilized for classification and regression. In both cases, the data comprises of the k nearest training samples in the component space. The yield relies on upon whether k-NN is utilized for classification or regression.

In k-NN classification, the yield is class participation. An object is well-organized by a larger part vote of its neighbours, with the object being assigned to the class most regular among its k closest neighbours (k is a positive whole number, ordinarily little). On the off chance that $k = 1$, then the object is basically allotted to the class of that single closest neighbour.

In k-NN regression, the yield is the property estimation for the object. This worth is the normal of the estimations of its k closest neighbors.

k-NN is a kind of example based learning, or sluggish realizing, where the capacity is just approximated locally and all calculation is conceded until characterization. The k-NN calculation is among the most straightforward of all machine learning calculations. The preparation illustrations are vectors in a multidimensional element space, each with a class mark. The preparation period of the calculation comprises just of putting away the component vectors and class names of the preparation tests.

In the order stage, k is a user characterized steady, and an unlabeled vector (an inquiry or test point) is arranged by assigning the label which is most incessant among the k preparing tests closest to that question point. A usually utilized separation metric for consistent variables is Euclidean separation. For discrete variables, for example, for content characterization, another metric can be utilized, for example, the cover metric (or Hamming separation). In the setting of quality expression microarray information, for instance, k-NN has likewise been utilized with connection coefficients, for example, Pearson and Spearman. Frequently, the grouping exactness of k-NN can be enhanced essentially if the separation metric is found out with specific calculations, for example, Large Margin Nearest Neighbor or Neighborhood parts examination.

A disadvantage of the essential "larger part voting" arrangement happens when the class conveyance is skewed. That is, illustrations of a more incessant class have a tendency to command the forecast of the new sample, since they have a tendency to be basic among the k closest neighbours because of their extensive number. One approach to defeat this issue is to weight the order, considering the separation from the test point to each of its k closest neighbors. Another approach to overcome skew is by reflection in information representation. For instance in a Self Organizing Map (SOM), every core is a assign (an inside) of a group of comparative focuses, paying little notice to their thickness in the first preparing information. k-NN can then be connected to the SOM.

A k-NN calculation commonly includes selecting a predefined number of days comparative in qualities to the day of fad. One of nowadays is arbitrarily re-examined to speak to the climate of the following day in the reproduction period. In spite of their intrinsic straightforwardness, closest neighbor calculations are viewed as adaptable and powerful. These techniques have been seriously examined in the field of insights and in example acknowledgment systems that go for recognizing distinctive examples. The closest neighbor approach includes synchronous inspecting of the climate variables, for example, precipitation and temperature. The testing is completed from the watched information, with substitution.

8. EXPERIMENTAL RESULT

8.1 SIFT Features with SVM

Support Vector Machine is used to construct the optimal separating hyper plane for various paddy plant disease features. For identifying a disease, paddy plant diseases features are extracted using SIFT from the input images for the three disease categories. In the training phase, seven dimensional feature vector is extracted from each diseased image and is given as input to the SVM model. The seven features are x position, y position, scale(sub-level), size of feature on image, edge flag, edge orientation, curvature of response through scale space. For training, seven features per diseases are extracted, but number of keypoint varies and it depends on the image complexity.

For recognition, seven disease features fed into the SVM model and the distance between each of the feature vectors and the SVM hyperplane is derived. The average distance is calculated for each model. The average distance gives better result than using distance for each feature vector. The recognition of the disease is decided based on the maximum distance.

8.2 SIFT Features with k-NN

In training phase, SIFT is applied to all paddy plant disease categories. In our work seven SIFT features are extracted from each key point. Number of pixels extracted from an input image is differ from image to image, as well as depends on the complexity of an image. The k-Nearest Neighbours algorithm is a non-parametric method used for classification and regression. In both cases, the input consists of the 3 closest training examples in the feature space. The training phase of the algorithm consists only of storing the feature vectors and class labels of the training samples.

In the classification phase, k is a user-defined constant (In this work $k=3$), and an unlabeled vector (a query or test point) is classified by assigning the label which is most frequent among the k training samples nearest to that query point. The k-NN recognizes the disease by using the test features compared with train features of different disease affected images. The recognition of the disease is based on the minimum distance value.

The paddy plant disease affected images for 3 classes namely Brown Spot, Leaf Blast and Bacterial Blight have been taken. Using SIFT, seven features namely x position, y position, scale(sub-level), size of feature on image, edge flag, edge orientation, curvature of response through scale space are extracted from each point. In training phase, SIFT is applied to all train image categories.

The images are then arranged through the framework that includes recognizable proof of nearby elements and representation of those elements as Scale Invariant Feature Transform (SIFT) descriptors, development of codebooks which gives an approach to outline descriptors into an altered length vector in histogram space and the multi-class grouping of the element histograms utilizing bolster vector machines (SVM) and k-Nearest Neighbors.

A linear SVM is utilized to characterize information sets which are straightly distinct. The SVM straight classifier tries to expand the edge between the isolating hyperplane. The examples lying on the maximal edges are called bolster vectors. In machine learning, support vector machines are managed learning models with related learning calculations that investigate information and perceive designs, utilized for order. The k-Nearest Neighbours (k-NN)

classification divides data into a test set and training set. For each row of the test set, the k nearest training set objects are found based on Euclidean distance and the classification is determined by majority vote with ties broken at random.

Total number of disease affected images (3 classes) = 120 images

Total number of training samples = 90 images (each class contains 30 images)

Total number of testing samples = 30 images (each class contains 10 images)

Table 1 Confusion Matrix for SIFT using SVM

Paddy Plant Diseases	TP	FN	FP	TN
Brown Spot	9	1	1	19
Leaf Blast	8	2	2	18
Bacterial Blight	9	1	1	19

Table 2 Confusion Matrix for SIFT using k-NN

Paddy Plant Diseases	TP	FN	FP	TN
Brown Spot	9	1	0	20
Leaf Blast	9	1	2	18
Bacterial Blight	9	1	1	19

Table 3 Performance table for paddy plant disease recognition with SIFT using SVM and k-NN

Feature \ Classifier	Precision (%)	Recall (%)	Accuracy (%)	F-score (%)
SIFT+SVM	86.66	86.66	91.10	86.66
SIFT+k-NN	90.60	90.00	93.33	90.14

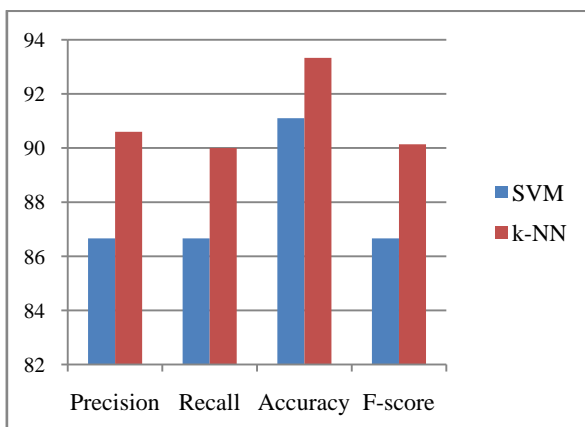


Fig.14 Performance chart for diseases recognition using SIFT with SVM and k-NN

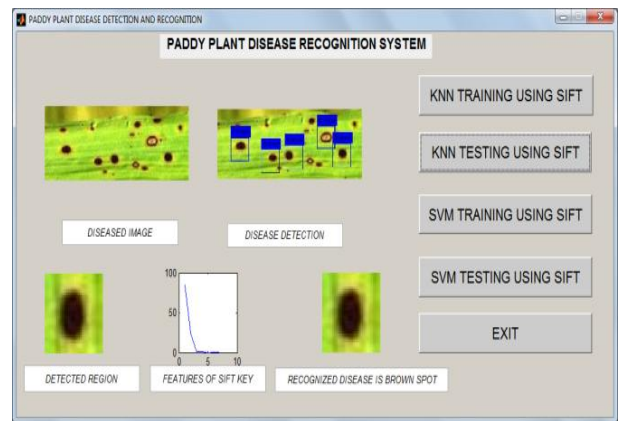


Fig.15 Screen shot for detection and Recognition of spot disease

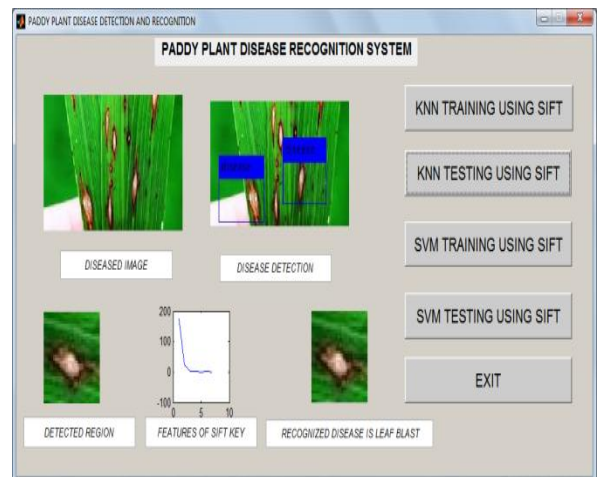


Fig.16 Screen shot for detection and Recognition of Leaf blast disease

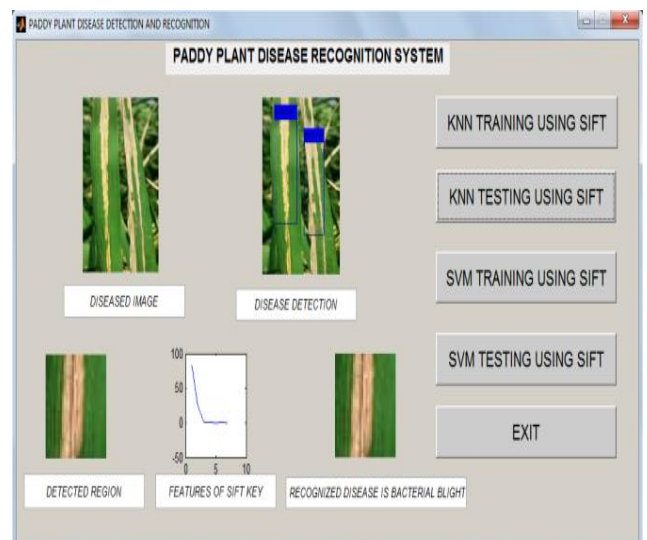


Fig.17 Screen shot for detection and Recognition of Bacterial Blight disease

9. CONCLUSION AND FUTURE SCOPE

The image processing techniques were used to deploy the classification system. In this work Scale Invariant Feature Transform (SIFT) is used to get features from the disease affected images. Then these features are taken to recognise the image using Support Vector Machine (SVM) and k-Nearest Neighbours

(KNN). This work mainly concentrates on three main diseases of paddy plant namely Brown spot, Leaf blast and Bacterial blight. It is useful to farmers and agriculture related researches. Experimental Result showed that the model is capable to predict the disease with accuracy of 91.10% using SVM and 93.33% using k-NN.

For future work, some alternative methods can be used to extract features and some other classifiers can be used to improve the result accuracy.

10. REFERENCES

- [1] P. R. Rothe, "Cotton Leaf Disease Identification Using Pattern Recognition Techniques", International Conference On Pervasive Computing, 2015.
- [2] Viraj A. Gulhane, Maheshkumar H. Kolekar, "Diagnosis Of Diseases On Cotton Leaves Using Principal Component Analysis Classifier", Annual IEEE India Conference, 2014.
- [3] Rong Zhou, Shun'ichi Kaneko, Fumio Tanaka, Miyuki Kayamori, Motoshige Shimizu, "Early Detection And Continuous Quantization Of Plant Disease Using Template Matching And Support Vector Machine Algorithms", First International Symposium On Computing And Networking, 2013.
- [4] John William Orillo, Jennifer Dela Cruz, Leobelle Agapito, Paul Jensen Satimbre Ira Valenzuela, "Identification Of Diseases In Rice Plant (*Oryza Sativa*) Using Back Propagation Artificial Neural Network", 7th IEEE International Conference, 2013.
- [5] Auzi Asfarian, Yeni Herdiyeni, Aunu Rauf, Kikin Hamzah Mutaqin, "Paddy Diseases Identification With Texture Analysis Using Fractal Descriptors Based On Fourier Spectrum", International Conference On Computer, Control, Informatics And Its Applications, 2013.
- [6] Kholis Majid, Yeni Herdiyeni, Annu Rauf, "I-Pedia: Mobile Application For Paddy Disease Identification Using Fuzzy Entropy And Probabilistic Neural Network", ICACIS, 2013.
- [7] Nunik Noviana Kurniawati, Siti Norul Huda Sheikh Abdullah, Salwani Abdullah, Saad Abdullah, "Investigation On Image Processing Techniques For Diagnosing Paddy Diseases", International Conference Of Soft Computing And Pattern Recognition, 2009.
- [8] Nunik Noviana Kurniawati, Siti Norul Huda Sheikh Abdullah, Salwani Abdullah, Saad Abdullah, "Texture Analysis For Diagnosing Paddy Disease", International Conference On Electrical Engineering And Informatics, 2009.
- [9] G.Anthonys, N. Wickramarachchi, "An Image Recognition System For Crop Disease Identification Of Paddy Fields In Sri Lanka", Fourth International Conference On Industrial And Information Systems, 2009.
- [10] Santanu Phadikar And Jaya Sil, "Rice Disease Identification Using Pattern Recognition Techniques", Proceedings Of 11th International Conference On Computer And Information Technology, 2008.
- [11] G.Anthonys and N. Wickramarachchi, 'An Image Recognition System for Crop Disease Identification of Paddy Fields In Sri Lanka', Fourth International Conference on Industrial and Information Systems (ICIIS), 28-31 December 2009.
- [12] Santanu Phadikar and Jaya Sil, 'Rice Disease Identification Using Pattern Recognition Techniques', Proceedings of 11th International Conference on Computer and Information Technology (ICCIT), 25-27 December 2008.
- [13] Qin Z and Zhang, 'Detection of rice sheath blight for in-season disease management using multispectral remote sensing', International Journal of Applied Earth Observation and Geoinformation, 2005.
- [14] J.B. Cunha, 'Application of Image Processing Techniques in the Characterization of Plant Leafs', Proc. IEEE Intl' Symposium on Industrial Electronics, 2003.
- [15] L. Lucchese and S.K. Mitra, 'Color Image Segmentation: A State-of-the-Art Survey', Proceeding of the Indian National Science Academy, Vol. 67A, No. 2, 2001, pp. 207-221.