Inconsistent Cluster Analysis With Disease Feature Enhancement (ICADFE) For American Cotton Leaf Disease Recognition

Kapil Prashar, Rajneesh Talwar, Chander Kant

ABSTRACT--- The broad leaves of cotton plant carry various visible disease symptoms. The ability of visual analysis by experts motivated the development of the plant disease recognition model. There are several visual feature descriptors, which can be primarily distinguished on the basis of pattern, texture or color. This system has been developed for the convenience of the farmers, who can avail the benefit by submitting the pictures of infected cotton leaves on the interface and the plant disease recognition system will return type of disease. In this paper, a dynamic feature descriptor is designed with inconsistent cluster analysis (ICA) and disease feature enhancement (DFE), which are combined as hybrid descriptor known as ICADFE for the recognition of the cotton plant disease. The ICADFE is found to improve the detection accuracy (approx. 80%), precision (approx 95%) and f1-measure (approx. 88%) on average in comparison with traditional shape and texture based feature descriptors such as scale invariant feature transform (SIFT), speeded up robust features (SURF) and fast retina keypoints (FREAK) with multi-category SVM (mSVM) for disease recognition.

Keywords: Leaf disease recognition, Inconsistent Cluster Analysis, Machine Vision

I. INTRODUCTION

According to the statistics of the Ministry of Public Affairs, nearly 50-60% people are completely or partially involved in the agricultural practices in India, whereas nearly 60-70% of total population is entirely dependable over agriculture for their daily needs[2]. The cotton crop falls in the broad leaf type crops, where the disease symptoms are clearly visible and can be spotted over the leaf with the image processing models, which inspired research in this field. Methods have been developed to determine the type and stage of the leaf-borne cotton crop disease after evaluating the shape and texture of the given disease sample in order to recognize the crop disease, which can be proved to be efficient and helpful for the novice or un-skilled farmer and laborers to understand the crop disease in order to help them apply the control measures in the earlier stages. The proposed model is based upon the difference of Gaussian (DoG) & difference of hessians (DoH) based shape analysis from the given cotton crop samples, whereas the texture features are extracted in the form of histogram of oriented gradients (HoG) and Statistical Analysis of Structural Information (SASI) features for the texture feature description. In the proposed model, the overall working methodology and results are based upon the accuracy of statistical measures for the assessment of the overall performance of this model in comparison with the existing models.

II. RELATED WORK

Ole Mathis Opstad Kruse et al. [1] have proposed the visual feature extraction model based upon the color-based pixel and pixel group analysis. S. Arivazhagan et al. [2] have worked upon the robust machine learning and flexible classification method for the recognition of the leaf-borne disease. A.H.Kulkarni et al. [3] have worked towards the extraction of the wide variety of the features from the given samples, which includes the shape, vein structure and color based features from the given objects in the images. Qinghai He et al. [4] team has worked on the utilization of the flexible and high contrast enhancement of the features as well as the de-noising for the filtering of the Gaussian, salt & peers and spike noise from the obtained samples in order to improve the overall accuracy for the cotton leaf images. Prof. Sanjay B. Dhaygude et al. [7] has designed and implemented the texture based feature description for the detection and classification of the leaf-borne diseases with the probabilistic classification based upon the support vector machine (SVM). Haiguang Wang et al. [8] studied the various models for the robust and flexible classification methods for the leaf-borne diseases, where the powerful feature extraction based upon the color, pattern and texture have been deeply studied to understand their impact in detail.

III. DESIGN & IMPLEMENTATION

The proposed model has been designed by combining the shape and texture based feature descriptors along with the probabilistic classification based upon the SVM classification algorithm. The proposed model utilizes the shape and texture based features for the localization and extraction of the disease regions, which is followed by the disease classification in the given samples. The shape and texture oriented feature are designed upon the basis of different of Gaussian (DoG), difference of hessians (DoH), inconsistent cluster analysis, computes the combination of directional derivatives on X and Y axes using Gaussian

Revised Manuscript Received on April 19, 2019.

Kapil Prashar, Research Scholar, I.K.Gujral Punjab Technical University, Kapurthala Road, Jalandhar, Punjab, India.

Rajneesh Talwar, Principal, CGC Group of Colleges, Jhanjeri, Dist.Mohali, Punjab, India

Chander Kant, Assistant Professor, Dept. of Computer Science & Applications, Kurukshetra University, Kurukshetra, Haryana, India.

Kernel, non-maxima suppression to eliminate the outliers and dependence of state (known as hysteresis) along with the multi-category enabled support vector machine (SVM) for the disease feature analysis of the affected cotton leaves.
In this paper, the low level feature descriptors are employed for the automatic recognition of the disease in the cotton leaves to evaluate their performance. These texture-based feature descriptors are used to extract the low level features from the target images. The low-level features are known to extract the strongest features in the given image, hence, their performance is being evaluated in comparison with the proposed feature descriptors. The SIFT feature descriptor utilizes the different of Gaussians (DoG) based technique to describe the minima matrix of the input image. The Gaussian kernel, known as scale-space kernel, is used for the SIFT feature descriptor in this scenario. This Gaussian function is defined with L(x,y,σ), whereas the variable-scale Gaussian is computed with equation G(x,y,σ) over the image matrix I(x,y). Hence, the final equation can be described as following:

3.1 Shape based Features:

In this paper, the low level feature descriptors are employed for the automatic recognition of the disease in the cotton leaves to evaluate their performance. These texture-based feature descriptors are used to extract the low level features from the target images. The low-level features are known to extract the strongest features in the given image, hence, their performance is being evaluated in comparison with the proposed feature descriptors. The SIFT feature descriptor utilizes the different of Gaussians (DoG) based technique to describe the minima matrix of the input image. The Gaussian kernel, known as scale-space kernel, is used for the SIFT feature descriptor in this scenario. This Gaussian function is defined with L(x,y,σ), whereas the variable-scale Gaussian is computed with equation G(x,y,σ) over the image matrix I(x,y). Hence, the final equation can be described as following:

$$L(x,y,\sigma) = G(x,y,\sigma) * I(x,y) \quad \textbf{....Eq. 1}$$

Where x and y give the location of pixels in the image matrix, σ denotes the intensity of the Gaussian filter and * gives the convolution procedure in the pixel (x, y), and

$$G(x,y,\sigma)= e^{-(x*x + y*y)/2\sigma*\sigma} \div (2\pi\sigma*\sigma) \quad \textbf{....Eq. 2}$$

Where σ denotes the intensity of the Gaussian filter, x and y denotes the given pixel position and value in the image matrix. The scale space extrema is computed with symbol e, in the above equation.

The SURF descriptor is the feature descriptor with increased execution speech, and uses the difference of Hessian matrix instead of Gaussian. The fast computation under SURF is achieved with box-type convolution filter. The sum of the image I on the specific location $(x,y)^T$ represents the sum of all intensities to compute the integral image $\mathbf{I_s(x)}$. The equation to compute the integral image can be given as:

$$I_s(x)= \sum_{i=0}^{i <= x} \sum_{j=0}^{j <= y} I(i,j) \quad \textbf{....Eq. 3}$$

The integral image computation utilizes the sequence of the pixels, their intensities and summation procedure. The x and y give the number of rows and columns respectively, whereas i and j denote the current pixel position. I(i,j) denotes the value of current pixel, the sum of which is computed along with its neighbors. After the computation of the integral image, the hessian matrix is computed over the integral image matrix. The determinant of Hessian is observed by using the scale selection method. Each of the given point (x,y) in the integral image I is used to compute the Hessian matrix H(x,σ), where σ denotes the scale of the Hessian. The equation can be given as following:

$$H(x,\sigma)= \begin{bmatrix} L_{xx}(x,\sigma) & L_{xy}(x,\sigma) \\ L_{yx}(x,\sigma) & L_{yy}(x,\sigma) \end{bmatrix} \quad \textbf{.....Eq. 4}$$

Where $L_{xx}(x,\sigma)$, $L_{xy}(x,\sigma)$, $L_{yx}(x,\sigma)$ and $L_{yy}(x,\sigma)$ denote the different convolutions of second order derivative of Hessian, which are computed with $(\sigma^2/\sigma x^2)*g(\sigma)$. The FREAK feature descriptor is computed in the form of binary mask over the Hessians matrix computed under the SURF algorithm. The binary mask is prepared after computing the threshold value to describe the pixel with 0 or 1. The binary mask is further used as the feature descriptor rather than the Hessian matrix prepared by the SURF algorithm.

3.2 K-means analysis:

Image clustering is applied to the image data under the proposed disease recognition model. The k-means based image clustering is learnt to segment the images multiple clusters by fragmentation of pixel frequencies in the target image. The k-means algorithm is applied over the set of observations $(x_1,x_2,x_3.....x_n)$, which are pixels (Gray or RGB) in the target image, to partition the n no. of observations to k clusters, where (k! >n) and (k <=n) are satisfied.

The **figure 1** shows the results of different executions of k-means algorithm on the same image. The cluster sequence is noticeable in all of the rounds. The results of first and second rounds are similar, whereas second and third clusters shuffled in 3rd run. The 4th and 5th run also produced the results similar to 3rd run, which again got shuffled in 6th run. This problem is resolved by using the inconsistent cluster analysis method explained in the following section.

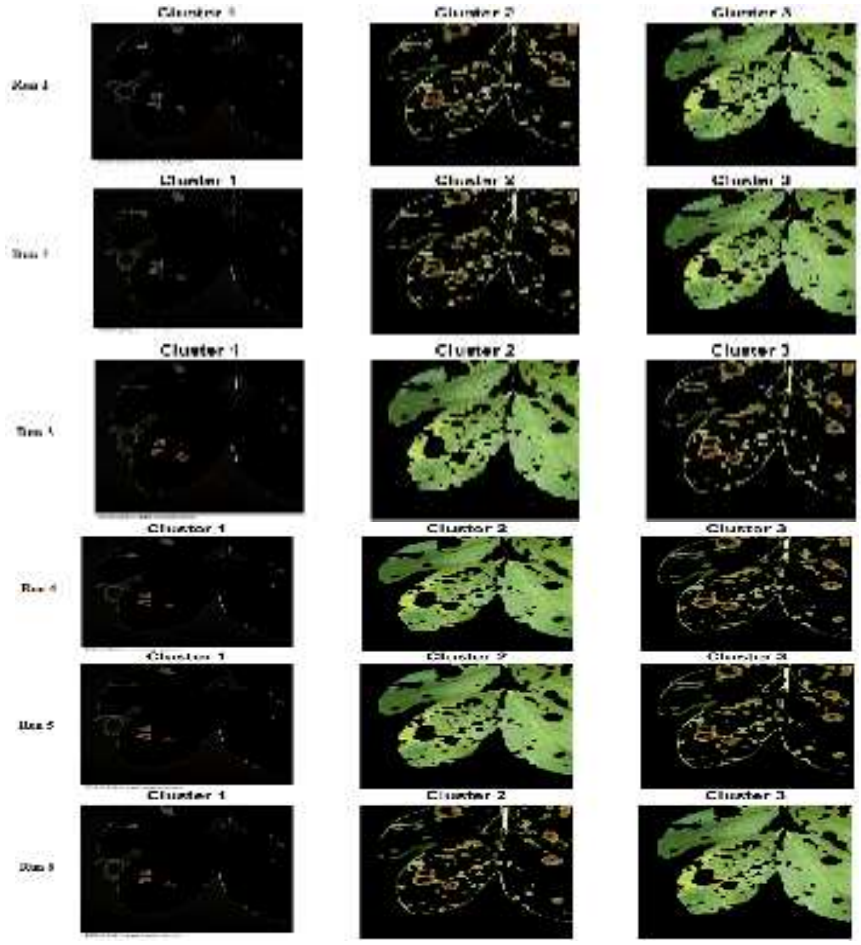

Figure 1: Proving the inconsistency of the K-means algorithm

The k-means clustering method aims to select the sets or clusters with minimum variance or difference in within-cluster sum of squares (WCSS). Technically, the objective of k-means algorithm is to discover the following (Eq. 5):

$$\sum_{i=1}^{k} \sum_{x \in S_i} \|x - \mu_i\| = \sum_{i=1}^{k} |S_i| (\text{Var}) S_i$$

....Eq. 5

Where μ_i denotes the mean value of data points (or pixels), denoted S_i, in image matrix. k gives the number of clusters, n represents the number of entities (or pixels) in cluster k. But there is problem of inconsistency associated with k-means clusters, which prevents the crop disease recognition model from detecting the correct cluster containing the disease symptoms. Hence, layer analysis model is used to select the required cluster.

3.3 Inconsistent Cluster Analysis:

The inconsistent cluster analysis is the method to determine the correct cluster containing the disease region in the target image. The adaptive inconsistent cluster analysis is applied to determine the target cluster containing the disease symptoms. The figure 2 explains the procedure used to determine the target cluster from the clusters produced by k-means algorithm. The inconsistent cluster analysis is performed with the supervision of training data prepared for selective clusters, which is used for classifier learning.

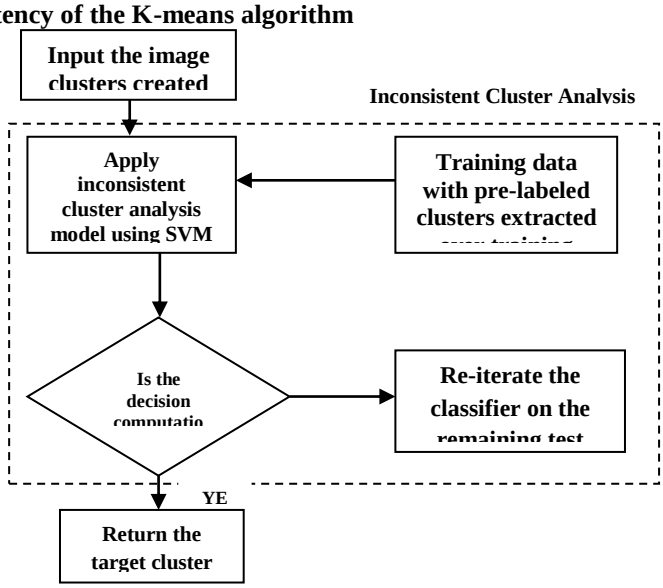

Figure 2: Inconsistent Cluster Analysis using SVM based supervised learning

In the above figure, the training data depicts the selective cluster data prepared to extract the desired clusters in the target image. The cluster selection is demonstrated in the figure 3, where the first three placeholders show the extracted clusters in the fixed sequence, whereas the 4th placeholder shows the selected cluster. This inconsistent cluster problem is evidently resolved, and its clearly shown

in figure 3. The inconsistency of the clusters can be noticeably verified according to the results shown by the inconsistent cluster selection approach in the proposed model. The result has been demonstrated on 6 step executions, where the k-means clustering approach is showing the inconsistent behavior, which is naturalized by inconsistent cluster selection as shown in figure 3.

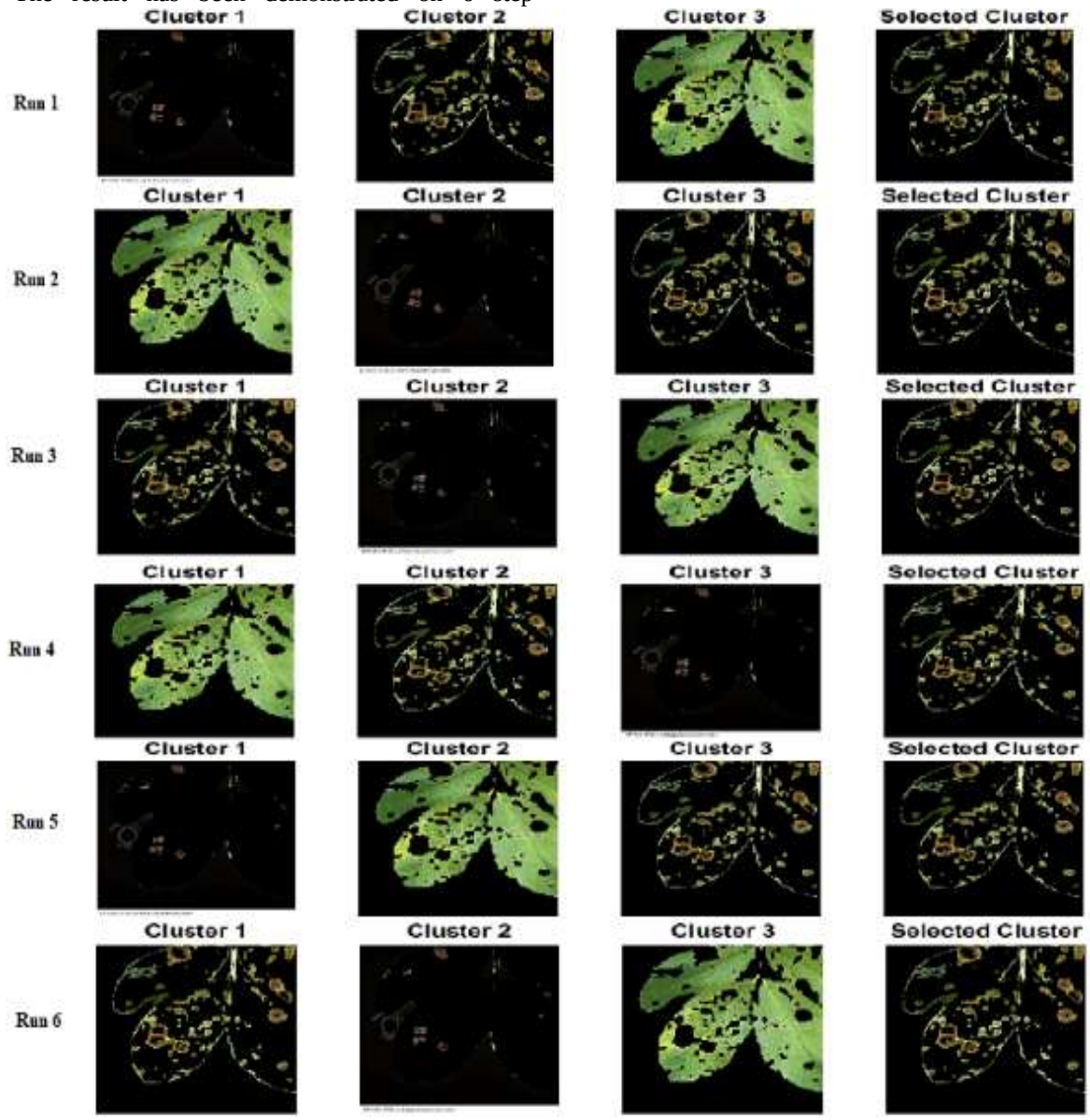

Figure 3: Selection of clusters using inconsistent cluster selection approach

3.4 Disease Feature Enhancement:

The image features learnt using inconsistent feature analysis method, are sometimes insignificant due to the use of unsupervised clustering, which must be transformed to produce consistent features. A reliable edge detection model has created to determine the regions containing the features in target cluster, which computes the combination of directional derivatives on X and Y axes using Gaussian Kernel, non-maxima suppression to eliminate the outliers and dependence of state (known as hysteresis). The X and Y derivatives of the target image cluster are computed by using the following kernel function (Eq. 1) and its derivative (Eq. 6).

$$K=(6 * \sigma) + 1 \qquad \text{....Eq. 6}$$

Where K is the kernel function and σ denotes sigma (or frequency) to transform the input value. The following equations compute the first bi-directional derivative on X (Eq. 8) and Y (Eq. 7) axes respectively.

$$Y1(i,j)=-((i-\frac{k-1}{2}-1)/2\pi\sigma^2)e^{(-(i-\frac{k-1}{2}-1)^* \quad (i-\frac{k-1}{2}-1)+(j-\frac{k-1}{2}-1)^* (j-\frac{k-1}{2}-1)/2\sigma^*\sigma} \qquad \text{.....Eq. 7}$$

$$X1(i,j)=-((j-\frac{k-1}{2}-1)/2\pi\sigma^2)e^{(-(i-\frac{k-1}{2}-1)^*(i-\frac{k-1}{2}-1)+(j-\frac{k-1}{2}-1)^* (j-\frac{k-1}{2}-1)/2\sigma^*\sigma} \qquad \text{.....Eq. 8}$$

Where X and Y denote the derivatives, sigma is derivation frequency, i & j give the data point position in 2-D matrix, and denote rows & columns respectively. Then, the following method is applied to compute second derivative w.r.t. original pixel value.

$$R= i - \frac{k}{2} \qquad \text{.....Eq. 9}$$

$$C = i - \frac{k}{2} \quad \text{.....Eq. 10}$$

Where R and C represent the rows and columns respectively, K denotes the kernel, and i and j gives the current row and column respectively.

$$X2(i,j)=X1(i,j)+I(R+K1-1,C+K2-1)*X1(K,1) \quad \text{....Eq. 11}$$

$$Y2(i,j)=Y1(i,j)+I(R+K1-1,C+K2-1)*Y1(K,1) \quad \text{.....Eq. 12}$$

Where X2 and Y2 denote the second derivatives of the image pixels, K1 and K2 denote the steps in range of kernel from 0 to K (Kernel value), X1 & Y1 represent the first derivatives and the i & j denote the current location of pixels in the terms of rows and columns respectively. Now, compute the magnitude of the second derivatives using following equations. The magnitude is computed between the Kernel ranges of $1 + fx(\frac{K}{2})$ to $H - fx(\frac{K}{2})$ vertically and $1 + fx(\frac{K}{2})$ to $W - fx(\frac{K}{2})$ horizontally, where H and W are number of rows and columns in the image matrix.

$$G(i,j)= \sqrt{X2(i,j)2 + Y2(i,j)2} \quad \text{....Eq. 13}$$

Where G denotes magnitude, X2 and Y2 are second derivatives computed using Eq. 11 and 12 respectively. To compute the non-maxima suppression, the kernel ranges similar to magnitude applied over following method:

$$T(i,j)= \left\{ \begin{matrix} 5, \\ \frac{Y2(i,j)}{X2(i,j)} \end{matrix} \right\} \begin{matrix} ifX2(i,j) == 0 \\ otherwise \end{matrix} \quad \text{.Eq. 14}$$

Where T is term frequency, and computed over X2 and Y2 derivatives of target image. The non-maxima suppression is created over the magnitude matrix, where the values falling in certain range are transformed to 0 on the basis of decision taken term frequency as shown in following equation:

$$NS=G \quad \text{....Eq. 15}$$

$$NS(i,j)= \left\{ \begin{matrix} 0, if\,(-0.4142 < T(i,j) <= 0.4142) and (G(i,j) < G(i,j+1)) or (G(i,j) < G(i,j-1)) \\ 0, if\,(0.4142 < T(i,j) <= 2.4142) and (G(i,j) < G(i-1,j+1)) or (G(i,j) < G(i+1,j-1)) \\ 0, if\,(abs(T(i,j) > 2.4142) and (G(i,j) < G(i-1,j)) or (G(i,j) < G(i+1,j)) \\ 0, if\,(-2.4142 < T <= 0.4142) and (G(i,j) < G(i-1,j-1)) or (G(i,j) < G(i+1,j+1)) \\ NS(i,j), otherwise \end{matrix} \right\}$$

.....Eq.16

Where NS denotes non-maxima suppression matrix, and certain set of ranges (T) are applied to determine the outliers, which are converted to absolute zero in order to compute the non-maxima suppression matrix. Further, the hysteresis method is applied to determine the significance of each pixel in the feature matrix selected through the inconsistent cluster analysis method. The following method is used to determine the feature significance:

$$H=NS \quad \text{.....Eq. 17}$$

$$\mathbf{H(i,j)=} \left\{ \begin{matrix} 1, if H(i,j) >= MaxThresh \\ 2, If\,(H(i,j) < MaxThresh) and (H(i,j) >= MinThresh) \\ 0, If H(i,j) < MaxThresh \end{matrix} \right\}$$

.....Eq. 18

Where H denotes the hysteresis component, and MaxThresh & MinThresh are the maximum and minimum thresholds respectively, which are fixed as**1.6 (max) and 0.03 (min)**. At last, the hysteresis matrix is normalized on the basis of few situation based rules, which are deployed on run-time according to data pre-analysis.

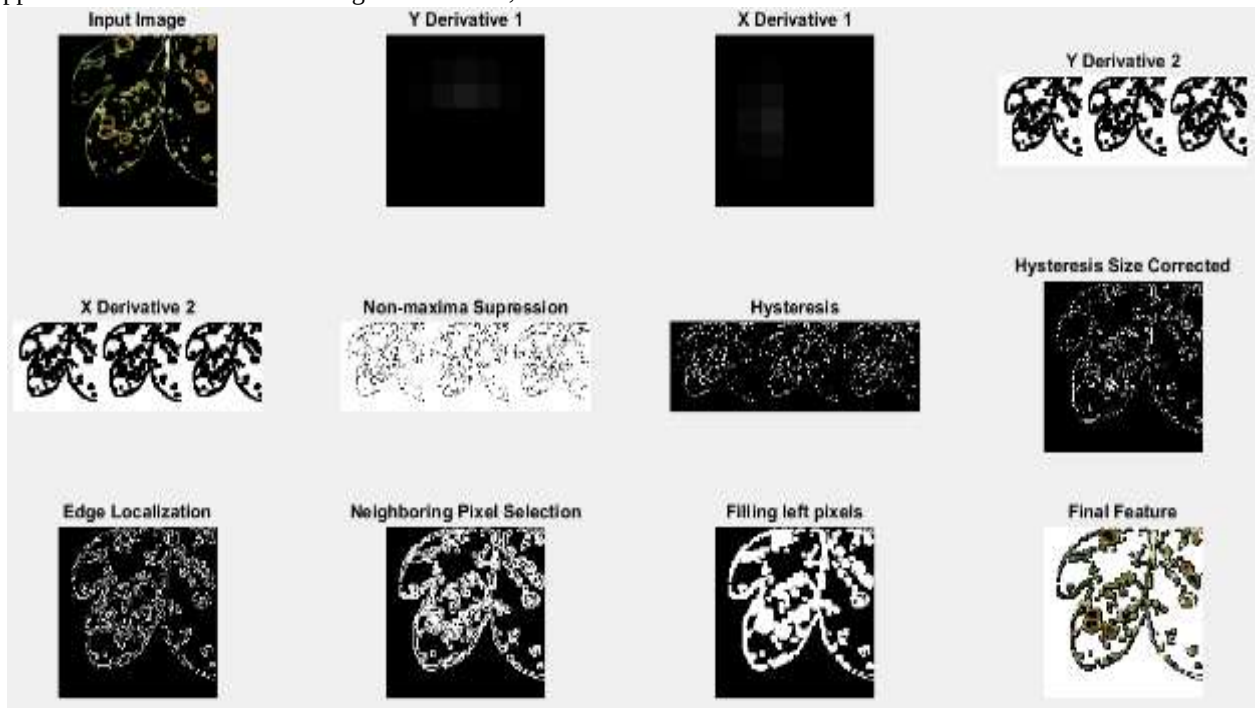

Figure 4: Step by step representation of disease feature enhancement

In figure 4, the disease feature enhancement is explained step by step. The visualization begins with original feature image before enhancement in the first placeholder. The multi layered derivatives are described in the first row (columns 2-4) and 2nd row (column 1). In 2nd row (column 2) the non maxima derivative is extracted on the basis of X and Y derivatives. Afterwards, the hysteresis size is corrected from the hysteresis matrix, which is computed for multiple Kernel values.

Then, the edge localization is performed on the hysteresis image after size correction. After edge localization, the neighboring pixel selection is performed, which is followed by gap filling method using morphological operations. Finally, the enhanced disease feature is extracted from the target image w.r.t. the binary mask.In the proposed model, we have combined the SIFT, SURF and FREAK to eliminate the false recognition cases. These feature descriptors work to determine the shape and texture of the maxima points in the given image matrix. The shape and texture features describe the strong visual pattern in the given image in the form of shape and texture in order to recognize the cotton disease. The proposed model uses the hybrid feature descriptor for combining the shape and texture based features has been developed as the improved classification model for plant disease detection from the infected leaves. The shape and texture features are considered to be the best features in the case of estimation and localization of the disease infected area in the given set of images. The SIFT feature descriptor model utilizes the difference of Gaussians for the creation of minima matrix, whereas the SURF descriptor utilizes the difference of Hessians for the similar purpose. The support vector machine (SVM) model has been incorporated for the purpose of classification of the plant disease in the given sample over the extracted features with SIFT, SURF and FREAK extractors. The following algorithm describes the workflow of the proposed model in detail:

Proposed Algorithm

1. Load the testing image matrix into runtime memory
2. Compute the clusters over the testing image using K-means algorithm
3. Return the clusters in similar sized image matrices
4. Load the training data for inconsistent cluster analysis (ICA)
5. Configure the mSVM (multi support vector machine) classification
6. Train mSVM classification model with ICA training data matrix and label vector
7. Classify the clusters using the mSVM trained with ICA data
8. Select cluster according to the classification decision, Set sigma, maximum & minimum threshold values
9. Compute the kernel function (Eq. 6), Compute the first derivatives horizontally and vertically (Eq. 7 and 8)
10. Compute the second derivatives horizontally and vertically (Eq. 11 and 12)
11. Compute the magnitude matrix over second derivative (Eq. 13)
12. Store the magnitude value matrix as term frequency matrix (Eq. 14)
13. Normalize the term frequency for each magnitude value to determine the outliers to create the non-maxima suppression by filtering out the outliers (Eq. 15 and 16)
14. Compute the hysteresis matrix over the non-maxima suppression matrix (Eq. 18)
15. Load training data for disease classification with disease labels, Train mSVM classifier with training data
16. Evaluate the testing feature computed on step 16 to determine the plant disease
17. Return the detected type of disease

The multilayered decision logic has been incorporated in this algorithm, where all the features including Hysteresis, inconsistent cluster analysis, etc are analyzed under the layered design in order to finalize the observation. The voting based classification has been used over the observations provided by SVM with all visual features collectively.

IV. EXPERIMENTAL CLASSIFICATION RESULTS AND ANALYSIS

For the purpose of result analysis, a number of samples are tested under this research, which involves the different grades of disease with variable disease covered regions. The image data for crop diseases also includes the following image (Figure 5), which carries the minimum disease covered regions, which can be analyzed with difficulty. This classification model is the proposed case is capable to detect the infected leaves even with the minimum disease covered regions.

Figure 5: Dataset image after clustering (a) Infected leaves (clusters) (b) Healthy leaves (clusters)

The results have been evaluated in the form of the statistical parameters and the performance measures in order to understand the effectiveness of the proposed model. The proposed model has been tested with 10% ratio for cross validation and remaining 90% data is used for the classifier learning. The results of the Inconsistent Cluster Analysis based Disease Feature Enhancement (ICADFE) model are collected and evaluated against the existing models for disease recognitions, as per shown in following table. The statistical parameter of accuracy, precision, recall and f1-measure are computed for each of the feature descriptor, as shown in the following table:

Table 1: Performance evaluation with cross validation ratio at 10%

	SIFT	SURF	FREAK	Proposed Model
Accuracy	65.00	45.00	40.00	79.55
Precision	62.50	45.00	41.67	96.67
Recall	90.91	100.00	50.00	78.38
F1-Measure	74.07	62.07	45.45	86.57

The table 1 contains the performance measures of the cross validation testing with selected samples from the given dataset. The proposed ICADFE descriptor has been recorded with 79.55% of accuracy, which outperformed SIFT (65%), SURF (45%) and FREAK (40%). Similarly, the ICADFE descriptor has outperformed using precision (96.67%) in comparison with SIFT (62.50%), SURF (45%) and FREAK (41.67%).

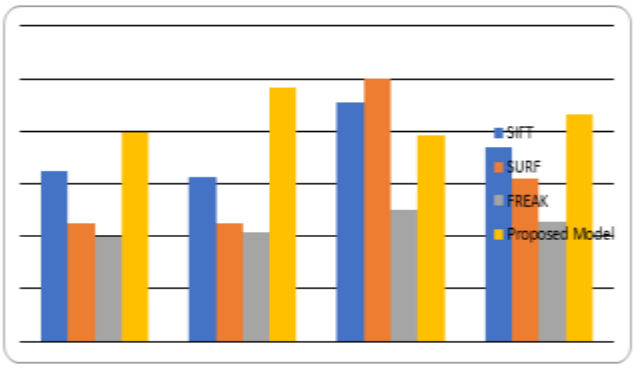

Figure 6: Comparison of Accuracy, Precision, Recall & F1-Measure of SIFT, SURF, FREAK & our Proposed Model

Table 2: Performance evaluation with cross validation ratio at 20%

	SIFT	SURF	FREAK	Proposed Model
Accuracy	70.00	68.00	60.00	81.81
Precision	73.91	79.49	57.14	94.44
Recall	91.89	79.49	66.67	85
F1-Measure	81.93	79.49	61.54	89.47

The table 2 contains the performance measures of the cross validation ratio as 20% testing with selected samples from the given dataset. The proposed ICADFE descriptor has been recorded with 81.81% of accuracy, which outperformed SIFT (70%), SURF (68%) and FREAK (60%). Similarly, the ICADFE descriptor has outperformed using precision (94.44%) in comparison with SIFT (73.91%), SURF (79.49%) and FREAK (57.14%). Most of the errors are based upon classification of plant leaves infected with Alternaria and Bacterial Blight. In the earlier

stages, both of the diseases are nearly not distinguishable, which becomes the primary reason behind the false positive cases during the classification. The proposed model has been recorded with the variable leveled values as per the other feature descriptors with support vector machine classification. The only exception remains with recall, where SIFT descriptor (91.89%) is observed with higher recall than ICADFE (85%).

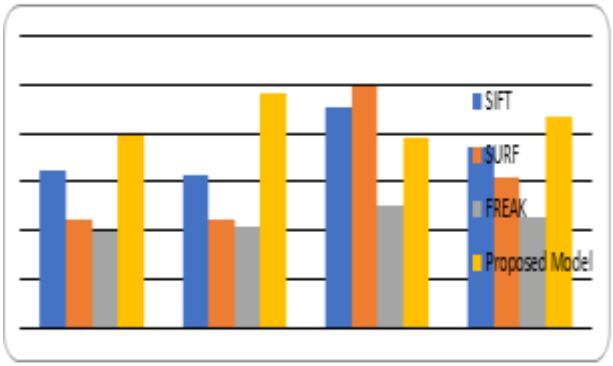

Figure 7: Comparison of Performance Evaluation of SIFT, SURF, FREAK & Proposed Model feature descriptors on basis of Accuracy, Precision, Recall & F1-Measure

V. CONCLUSION

The proposed model is made capable of working with the variations in the variety of the diseases defined among the given image dataset of disease affected leaves of cotton, which are determined from the local or global image features such as texture, pattern or color. Also the visual orientations along the size of disease covered regions, occlusion and several color illumination variations are used for the purpose of image classification. The proposed model has been designed to produce the rigid and improved results than the existing models, which has been analyzed from the obtained results. The proposed model has been found improved in the terms of the all of the performance parameters, which includes the precision, recall, f1-measure and accuracy. The proposed model has been found highly accurate at approx 80%, against the 67.50%, 56.50 and 50% percent readings obtained from the other disease feature descriptors.

In the future, the proposed model can be improved by using the Self-learning AI techniques deep learning with the help of the neural network along with the shape and curve based features, which can further improve the performance of the script recognition models.

VI. ACKNOWLEDGEMENT

We are highly thankful and grateful to the Department of Research, Innovation & Consultancy of IK Gujral Punjab Technical University, Kapurthala Road, Jalandhar for giving us an opportunity and providing us all the required assistance to write this Research Paper. The constant guidance and encouragement received from Dr. A.P.Singh, Dean RIC, I.K Gujral Punjab Technical University has been of great help in carrying out the research work and is acknowledged with reverential thanks. The authors would like to express a deep sense of gratitude and thanks profusely to Dr.Pankaj Rathore, Director, Regional Research Station, Punjab Agricultural University, Faridkot, Punjab and Dr.Paramjit Singh, Director, Regional Research Station, Punjab Agricultural University, Bathinda, Punjab. Without his wise counsel and able guidance, it would not have been possible to complete the research work in this manner.

REFERENCES

1. Ole Mathis Opstad Kruse, José Manuel Prats-Montalbán, Ulf Geir Indahl, Knut Kvaal, Alberto Ferrer, Cecilia Marie Futsaether "Pixel classification methods for identifying and quantifying leaf surface injury from digital images", Computers and Electronics in Agriculture, Elsevier 108 (2014) 155–165.
2. S. Arivazhagan, R. Newlin Shebiah, S. Ananthi, S. Vishnu Varthini, "Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture feature," CIGR Journal, vol. 15, no. 1, March 2013.
3. A.H.Kulkarni, Dr.H.M.Rai, Dr.K.A.Jahagirdar & P.S.Upparamani in the paper titled "A Leaf recognition technique for plant classification using RBPNN and Zernike Moments", IJARCCE Journal, Vol. 2, Issue 1, January 2013
4. Qinghai He, Benxue Ma, Duanyang Qu," Cotton pests and diseases detection based on image processing," TELKOMANIA, Vol 11, no 6, June 2013
5. P. Revathi, M. Hemalatha "Advance Computing Enrichment Evaluation of Cotton Leaf Spot Disease Detection Using Image Edge detection", IEEE-20180, 26-28 July 2012.
6. Yinmao Song, Zhihua Diao, Yunpeng Wang, Huan Wang "Image Feature Extraction of Crop Disease," 2012 IEEE Symposium on Electrical & Electronics Engineering (EEESYM)
7. Prof. Sanjay B. Dhaygude, Nitin P. Kumbhar, "Agricultural plant leaf disease detection using image processing" International Journal of Advanced Research in Electrical, Electronics & Instrumentation, Vol 2, Issue 1, Jan 2013
8. Haiguang Wang, Guanlin Li, Zhanhong Ma, Xiaolong Li, "Image Recognition of Plant Diseases Based on Principal Component Analysis and Neural Networks" 2012 8th International Conference on Natural Computation (ICNC 2012)
9. Anand Kulkarni & Ashavin Patil R.K. "Applying image processing technique to detect plant diseases," International Journal of Modern Engineering Research(IJMER), ISSN 2249-6645, Vol 2, Issue 5, pp 3661-6264, Sep-Oct 2012
10. Tushar J. Haware, Ravindra D. Badgujar and Prashant G. Patil, "Crop disease detection using image segmentation," World Journal of Science & Technology, ISSN 2231-2587, Vol 2(4):190-194,2012
11. Kamljot Singh Kailey and Gurjinder Singh Sahdra "Content based image retrieval (CBIR) for identifying image based plant disease," IJCTA, Vol 3(3), 1099-1104, May-June 2012
12. Viraj Gulhane and Dr.A.A.Gurjar "Detection of diseases on cotton leaves and its possible diagnosis," IJIP, Vol(5), Issue(5),2011
13. H. Al-Hiary, Bani-Ahmad, M-Reyalat and M.Braik, "Fast and accurate detection and classification of plant diseases", IJCA, Vol 17, No. 1, March 2011

14. Piyush Chaudhari, Anand K. Chaudhari, Dr.A.N. Cheeran and Sharda Godara,"Color transform based approach for disease spot detection on plant leaf", IJCST, Vol 3, Issue 6, June 2012
15. Dheeb Al Bashish, Malik Braik and Sulieman Bani Ahmad "Detection and classification of leaf disease using K-Mean based segmentation and neural network based classification," International Technology Journal, ISSN 1812-5638, pp 267-275, 2011
16. Sanjay B.Patil and Dr. Shrikant K.Bodhe "Leaf disease severity measurement using image processing," International Journal of Engineering Technology (IJET), Vol 3(5), 297-301, 2011