Turkish Journal of Computer and Mathematics Education *Vol.12 No.10 (2021), 2585-2592*

Research Article

Canker Detection in Citrus Plants with an Efficient Finite Dissimilar Compatible Histogram Leveling Based Image Pre-processing and SVM Classifier

Shoby Sunny[1], M. P. Indra Gandhi[2]

[1]Research Scholar, Department of Computer Science, Mother Teresa Women's University, Kodaikanal, India.
[2]Assistant Professor, Department of Computer Science, Mother Teresa Women's University, Kodaikanal, India.

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

Abstract-The massive loss in agricultural yield mainly caused by leaf and fruit diseases. These diseases reduce both quality and quantity of agricultural products. As a major provenience of nutrients like vitamin C, citrus plants such as lemons, mandarins, oranges, tangerines, grapefruits, and limes are commonly grown fruits all over the world. As a result of various plant diseases, citrus producing companies make a huge amount of waste every year whereby 50% of citrus peel is destroyed every year. The disease canker is one of the mentionable leaf and fruit diseases. The main goal of this paper is to recognize and classify the canker disease precisely from the contrived leaf images by employing image processing techniques to detect plant leaf diseases from digital images. We offer a method consists of two phases to enhance the clarity of leaf images. The primary stage uses Finite Dissimilar Compatible Histogram Leveling (FDCHL) in preliminary step which advances the dissimilar level of disease influenced leaf image, segment the region of interest using fuzzy feature selection. The second phase by adopting the Support Vector Machine classifier to find out the canker leaf image and implements these methods in lemon citrus canker disease identification. Experimental results show effective accuracy detection and reduced execution time of canker disease detection.

Keywords: Canker Disease, Image Processing Techniques, Histogram Leveling, Gray-Level Co-Occurrence Matrix and Support Vector Machine.

1.Introduction

Agricultural production is directly turned down by plant diseases. The detection and classification of plants lesions are the main tasks to improvise the quality of plant production for economic growth. To increase the plant production for better economic growth, the process of recognition and classification of plants lesions are crucial. The occurrence of the disease on the plant result in substantial loss in both quality as well as the quantity of agricultural product. This can create the negative impact on the countries whose economies are primarily dependent on the agriculture. Hence the detection of the disease in plants is significant to avoid the financial loss.

In recent years, the leaf disease canker in citrus plants becomes one of the major acute diseases. Long period and traditional citrus medicinal plants such as lemon, orange are contrived by a canker disease which is a bacterial disease that affect the premature leaves and fruits of citrus plants. The proposed approach considers lemon leaves for classification of citrus canker disease because of high commercial cultivation crop. Lemon is an important source of vitamin C and contains flavonoid compounds that have distinct antioxidant and anti-cancer properties [2]. At the beginning stage of the disease, canker can be recognized by suddenly appeared some white spongy spots which is then turns into grey or brown later as shown in fig.1. The contrived location is characterized by oily margins or yellowish ring (lesions), which can be found on both sides of the leaves. This disease can be detected by the appearance of lesion on groves, stems and leaves. The symptoms appear as yellowish spots or halos on leaves that gradually enlarge to 2 – 4 mm dark brown pustules [7].

Figure 1 Citrus Canker Lesions in leaf and fruit

This citrus canker disease is caused by the bacterium Xanthomonas Axonopodis PV. Citric (XAC). The infection of citrus canker results in defoliation, dieback, tarnished fruit, reduced fruit quality, premature leaf and fruit and at last the trees will produce no fruits. Citrus canker is highly infectious and can be spread rapidly by wind, rain, landscaping equipment, people work in field, moving infected or exposed plants or plant parts and it is difficult to eradicate. Detecting citrus canker at the early stage is the key to control and spreading of this disease.

Digital image processing [9] and image analysis technology based on the advances in real time applications such as microelectronics, computers, medicine and biology and it able to circumvents the problems. In this paper a new model for enhancement of pre-processing image with efficient contrast and to predict the canker disease in citrus plant (lemon) by classifier is implement. This approach aims to use contrast enhancement techniques [12] to enhance the image quality and to classify the citrus canker affected leaf by Support Vector Machine classification. This system which can provide more accurate results related to the identification and classification of disease. From an innovation perspective, the research contributions are as follows,

I. To enhance the quality and contrast of citrus leaf image by employing a Finite Dissimilar Compatible Histogram Leveling (FDCHL) enhancement techniques.

II. Presenting a framework for citrus canker diseases detection in citrus lemon leaf classification by implementing Support Vector Machine.

The paper, is organized in five Section II describes the related works. Section III describes the proposed methodology. Section IV represents results and discussion and the paper is concluded in section V

2.Related works

Ali et al., (2017), This paper presents a technique to detect and classify major citrus diseases of economic importance. Kinnow mandarin being 80% of Pakistan citrus industry was the main focus of study. Due to a little variation in symptoms of different plant diseases, the diagnosis requires the expert's opinion in diseases detection. The inappropriate diagnosis may lead to tremendous amount of economical loss for farmers in terms of inputs like pesticides. For many decades, computers have been used to provide automatic solutions instead of a manual diagnosis of plant diseases which is costly and error prone. The proposed method applied DE color difference algorithm to separate the disease affected area, further, color histogram and textural features were used to classify diseases.

Badnakhe et al., (2011), presents indirect contribution for the Improvement of the Crop Quality. It is a Machine learning based recognition system which will going to help in the Indian Economy. The paper will propose the technique to classify and identify the different disease affected plant. Digital Analysis of crop color is the important. Now it's becoming popular day by day. It is also of the cost effective method. Because changed in the color are a valuable indicator of crop health and efficiency and survaibility. Then it can be measured with visual scales and inexpensive crop color. This proposed work is giving of the better technique to do the classification of crop disease. We can easily develop an application. In future the experimental results indicate that the proposed approach is a valuable approach, which can significantly support an accurate detection of leaf, steam, and root diseases in a little computational effort.

Piyush et al., (2012), proposed an algorithm for disease spot segmentation using image processing techniques in plant leaf. Disease spots are different in color but not in intensity, in comparison with plant leaf color. So we color transform of RGB image can be used for better segmentation of disease spots. In this paper a comparison of the effect of CIELAB, HSI and YCbCr color space in the process of disease spot detection is done. Median filter is used for image smoothing. Finally, threshold can be calculated by applying Otsu method on color component to detect the disease spot. Experimental result shows that noise which is introduced because of background, vein and camera flash; can be wiped out using CIELAB color model.

Shaikh et al., (2018), discussed about a brief idea to solve this problem by continuously monitoring crops using „Agri-Robo" and techniques called Image Processing. Image Processing give the good solution to above crisis. Image processing gives fast, automatic and accurate solution to user. We developed an agri-robo system to monitor crops and for identifications and monitoring of diseases & pesticides. This agri-robo not only detects disease but also spray pesticides to protect them from disease. The robot helps the farmer to take informed decision locally or allows connecting with other existing services. This agri-robo find diseases on various infected leafs. This system result in detection of cotton diseases and spray the pesticides of disease in proper amount when needed.

3. Proposed Methodology

Image processing has play a terribly important role in agriculture field because of widely accustomed observe the crop disease with high accuracy. This paper focused mainly to improve the image quality with greater clarity by FDCHL enhancement techniques in pre-processing stages and to detect the canker disease by classifiers. The following image processing techniques steps are used to detect the disease and Support Vector Machine are applied to get optimal solution of the canker disease are shown in fig. 2.

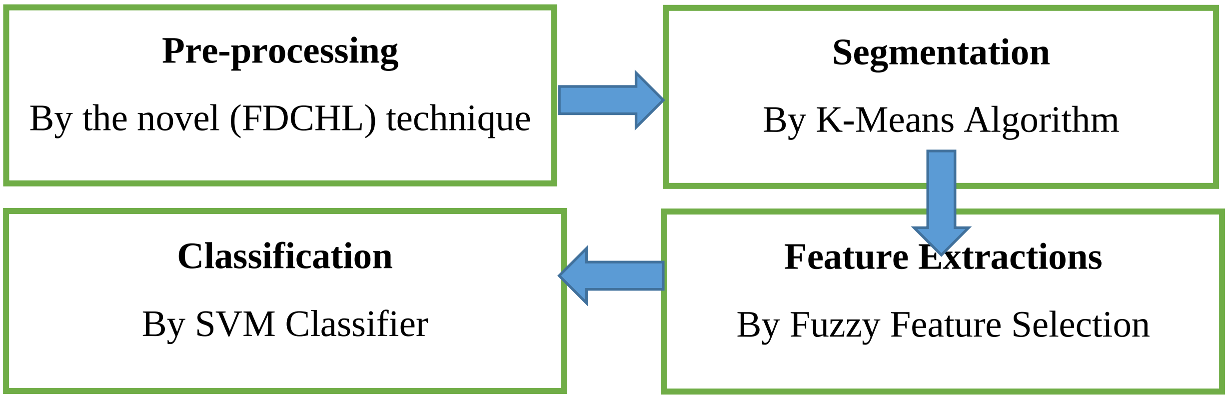

Fig. 2 System architecture for detection and classification of infected plant leaves

3.1. Collection of data

The data set were gathered using an advanced DSLR (Canon EOS 1300D) having a sensor with CMOS system and resolution of 5202-3465 (Mpix). The sensor size for Canon EOS 1300D is 14.9-22.3 (mm). RGB color range is chosen for each of the images in the JPG combination, including 256 shades for each RGB layer and 8 pixels for each shading layer.

3.2. FDCHL Pre-processing

The FDCHL to enhance the color images consists of the following steps:

Step 1: Dividing the intensity image into non overlapping contextual regions. Intensity components are partitioned into 8 x 8 non-overlapping contextual regions.

Step 2: Calculating the histogram of each region.

Step 3: Clipping the histogram of each region by the clip-limit value. The clipping rule is given by the following statements:

If H_{region} (i) > N_{clip} **then**

H_{region_clip} (i)=N_{clip}

Else if (H_{region} (i)+N_{avgbin}) > N_{clip} **then**

H_{region_clip} (i)=N_{clip}

Else H_{region_clip} (i) = (H_{region} (i)+N_{avgbin})

where H_{region} (i) is a local histogram of each region at i-th gray level. H_{region_clip} (i) represents clipped histogram of the region, N_{clip} denotes the actual clip-limit which is defined by

$N_{clip} = Min_{clip} + \text{roundn}(V_{clip} * (N_{pix} - Min_{clip}))$

where V_{clip} is clip-limit value in the range [0, 1] defined by the user. N_{pix} denotes the total number of pixels in the region. Min_{clip} is the minimum average of total pixels, N_{pix}, per total bins, N_{bin}, in the local histogram. Min_{clip} is defined by:

$$Min_{clip} = round(\frac{N_{pix}}{N_{bin}})$$

Thus, the total number of pixels, $N_{\Sigma clip}$, denotes the remain pixels from the clipped histogram. The average of the remain pixels to redistribute to each bin is calculated by

$$N_{avgbin} = \frac{N_{\Sigma clip}}{N_{bin}}$$

Step 4: Enhancing intensity values in each region. The clipped histogram, H_{region_clip}, is transformed to cumulative probability, P_{input} (x), which is provided to create transfer function. Rayleigh forward transform is given by

$$y = y_{min} + \sqrt{2\alpha^2 ln\left(\frac{1}{1 - P_{input}\ (x)}\right)}$$

where Y_{min} is the lower bound of the intensity value. α is a scaling parameter of Rayleigh distribution that is defined depending on each input image. The output probability density of each intensity value, Y, can be derived as

$$p(y) = \frac{y - y_{min}}{\alpha^2} exp\left(\frac{(y - y_{\min})^2}{2\alpha^2}\right) \; for \; y \geq \; y_{min}$$

Step 5: Reducing abruptly changing effect, the output from the transfer function is re-scaled using linear contrast stretch. The re-scale function still keeps the original shape of the transfer function to compress noise background and to design the color of output continuously. The linear contrast stretch is calculated by

$$y = \frac{x - x_{min}}{x_{max} - x_{min}}$$

where x is the input value from the transfer function, X_{min} and X_{max} denotes the minimum and maximum value of the transfer function.

Step 6: Interpolating by using bilinear of the neighboring sample points from the center pixel of contextual regions to form the enhanced in each region for the whole image as the new intensity, 1'. [7].

3.3. K-Means Clustering Procedure for Segmentation

Imagine a picture of resolution of x × y and the picture has to be congregate into k number of group.

Say p (x, y) be an input pixel to be cluster and C_k be the cluster centers. The algorithm is as following:

1. Initialize number of cluster k and centre.
2. For each pixel of an image, calculate the Euclidean distance d, between the center and each pixel of an image using the relation given below.

$d = \|p(x, y) - C_k\|$

3. Assign all the pixels to the nearest centre based on distance d.
4. After all pixels have been assigned, recalculate new position of the centre using the relation given below.

$$C_k == \frac{1}{k} \sum_{y \in C_k} \sum_{x \in C_k} p(x, y)$$

5. Repeat the process until it satisfies the tolerance or error value.
6. Reshape the cluster pixels into image.

3.4. Feature Extraction

3.4.1. Fuzzy curves

It is a nonlinear continues curve, which establishes a connection between a specific input and the output, performing a projection of the multidimensional input and output space on the probed input-output space [15]. For each date point, we can create a fuzzy membership function $\mu_{I,k}$ (x_i) using $\mu_{i,k}(x_i)$ from the fuzzy rule: IF x_i is $\mu_{I,k}$ (x_i) THEN y is y_k , that can be thought of as fuzzy rule for the output y with respect to each feature variable x_i ,

$$\mu_{i,k}(x_i) = \exp\left(-(\left(\frac{x_{i,k} - x_i}{\sigma_i}\right)^2\right)$$

each Gauss function is located at point ($x_{i,k}$, y_k) the parameter σ_i has a fixed value per feature variable, x_i which equals 5%～20% of the x_i variable range. Using the centroid defuzzification technique, we defuzzify these fuzzy membership functions to produce a fuzzy curve c_i for each feature variable x_i by the following formula:

$$c_{i,k}(x_i) = \frac{\sum_{k=m}^{m} y_k \times \mu_{i,k}(x_i)}{\sum_{k=1}^{m} \mu_{i,k}(x_i)}$$

The above equation provides a continuous curve, which approximates the input output data, and behaves as a fuzzy model. The mean square error used to estimate the quality of the approximation,

$$MSE_{c_i} = \frac{1}{m}\sum_{k=1}^{m}\left(c_{i,k}(x_i) - y_k\right)^2$$

Using fuzzy curves, we can mechanically and rapidly identify the important features from the set of candidate features.

3.4.2. Fuzzy surfaces

The fuzzy surface is an extension of the fuzzy curve; it is based on the simple idea: Independent feature do a better job of approximating the output than dependent inputs [15]. Fuzzy surface can be thought of as a "two-dimensional" fuzzy curve. Fuzzy surface is defined as [16]:

$$s_{i,j}(x_i, x_j) = \frac{\sum_{k=1}^{m} y_k \times \mu_{i,k}(x_i) + \sum_{k=1}^{m} y_k \times \mu_{j,k}(x_j)}{2\sum_{k=1}^{m} \mu_{i,k}(x_i) \times \mu_{j,k}(x_j)}$$

Where x_i and x_j are feature variables, $s_{i,j}(x_i, x_j)$ is a two-dimensional surface in feature space corresponding to a fuzzy rule:

IF x_i is $\mu_{i,k}(x_i)$ and x_j is $\mu_{j,k}(x_j)$, THEN y is y_k

As with fuzzy curves, the mean square error, MSE_{c_i} for the fuzzy surfaces is be used to evaluate the approximation of the input output data.

$$MSE_{S_{I,J}} = \frac{1}{m}\sum_{k=1}^{m}\left(s_{i,j}(x_i, x_j) - y_k\right)^2$$

3.4.3. Feature selection procedure of image

Using all spot features presented in the previous section as candidate features set. We choose 50 spots for each class disease leaf image, there are three classes disease leaf, the sample data are 150 (m=150). Output variable y_k is class label. We use the procedure to isolate the important and independent features [16]

1) Use fuzzy curves to rank all candidate features in order of significance- with in ascending MSE_{c_i} order.
2) Use the most important feature and each of other remaining features, chosen by the fuzzy curve, to create fuzzy surfaces, rank the features in order of dependence with the most important feature (in ascending MSE_s).
3) Use the second most important feature and each of remained features in previous step to create new fuzzy surfaces, once again.
4) Repeat fuzzy surfaces process until enough important inputs are obtained, or no remaining inputs are left.

3.5. Classification using Support Vector Machine (SVM)

The Support Vector Machine (SVM) classifiers are adopted differentiate citrus leaf disease. SVM is used to classify disease on their texture feature. The most attractive feature of the SVM is the maximum-margin hyper plane, the soft margin and the kernel function. For classifying any two linearly separable classes there may exist many separating lines that correctly classify the data. Among these lines the SVM select the line, which maximizes the distance between the separating hyper-planes. To explain it clearly we label the training data

$\{xi, yi\}$, $i = 1, \ldots\ldots, l$, $yi \in \{-1, 1\}$, $\mathbf{x}i \in \mathbf{R}d$

where l denotes the total no of training sample and d denotes the dimension of the feature vector. Suppose we have some hyper-plane, which separates the positive from the negative examples (a separating hyper-plane). The points $\boldsymbol{x}$ which lie on the hyper-plane satisfy

$\boldsymbol{w} \cdot \boldsymbol{x} + b = 0$,

where $\boldsymbol{w}$ is normal to the hyper-plane, $|b|/\|\boldsymbol{w}\|$ is the perpendicular distance from the hyper-plane to the origin, and $\|\boldsymbol{w}\|$ is the Euclidean norm of $\boldsymbol{w}$. Let $d+$ ($d-$) be the shortest distance from the separating hyper- lane to the closest positive (negative) instance. Define the "margin" of a separating hyper-plane to be $d+ + d-$. For the linearly separable case, the support vector algorithm simply looks for the separating hyper-plane with largest margin. This can be formulated as follows: suppose that all the training data satisfy the following constraints:

$\boldsymbol{x}i * \boldsymbol{w} + b = +1$ for $yi = +1$ (1)

$\boldsymbol{x}i * \boldsymbol{w} + b = -1$ for $yi = -1$ (2)

These can be combined into one set of inequalities:

$yi(\boldsymbol{x}i * \boldsymbol{w} + b) - 1 \geq 0\ \forall i$ (3)

So, $d+ = d- = 1/\|\boldsymbol{w}\|$ and the margin is simply $2/\|\boldsymbol{w}\|$. Thus we can find the pair of hyper-planes, which gives the maximum margin by minimizing $\|\boldsymbol{w}\|2$, subject to constraint (3). Other important characteristic of the SVM is the soft margin that gives the user flexibility to choose the parameter to roughly control the number of examples. Another feature is the kernel function that projects the non-linearly separable data from low-dimensional space to a space of higher dimension so that they may become separable in the higher dimensional space too.

4. Experimental Results

We employ MATLAB in windows system to assess the performance of proposed methodology. Experiments are carried out on citrus (lemon) leaves to detect the canker disease. The experiment on the proposed methodology involves two phases firstly, the betterment of the quality of image is handled by applying the novel FDCHL pre-processing and then secondly, the process of detecting the canker disease in citrus is executed by adopting SVM classifier.

The performance of the proposed methodology is assessed through some specific metrics including Equal Error Rate, False Rejection Rate, False Acceptance Rate and Genuine Acceptance Rate by comparing it with different kind of classifier strategies such K-NN and Navies Bayes.

During the process of verification of the proposed system, the metric Equal Error Rate (EER) is considered to compute error rate in the operation of proposed system. On another hand, we measure the proposed system's both incorrect acceptance and incorrect rejection of canker leaves images through the metrics False Rejection Rate (FRR) and False Acceptance Rate respectively (FAR).

TABLE 1: Performance Comparison of Canker Detection

	Classifiers	FAR (%)	FRR (%)	GAR (%)	EER	Execution Time (ms)
Without FDCHL Pre-processing	KNN	3	2	97	0.72	0.64
	NB	4	4.5	96	0.62	0.57
	SVM	5	6	95	0.54	0.41
With FDCHL Pre-processing	KNN	4.7	4	95	0.39	0.30
	NB	5.3	6.7	94	0.33	0.21
	SVM	6.3	8.5	93	0.20	0.14

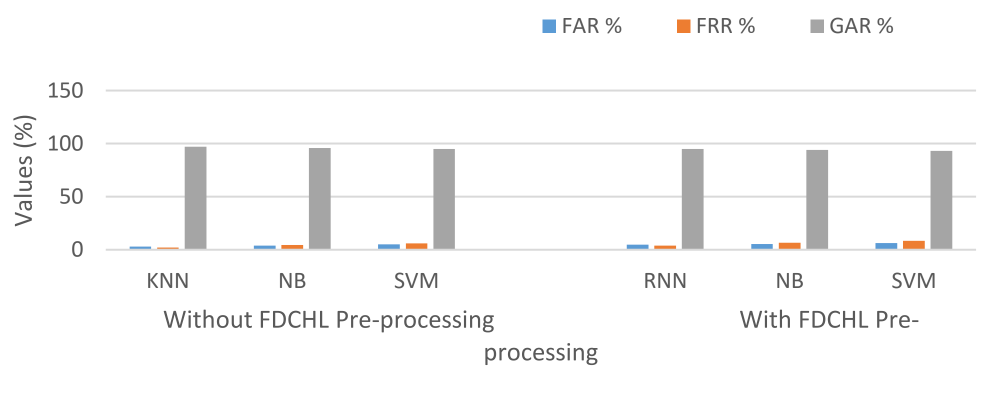

Figure 3 Accuracy Comparison for FDCHL Pre-processing

The above figure 3 shows Accuracy prediction of canker detection disease in citrus leaves in terms of FAR, FRR and GAR. It is clearly noted that the pre-processing step with FDCHL achieves efficient result integrated with SVM classifiers. The SVM classifiers achieves the good prediction rate for differentiating the canker disease.

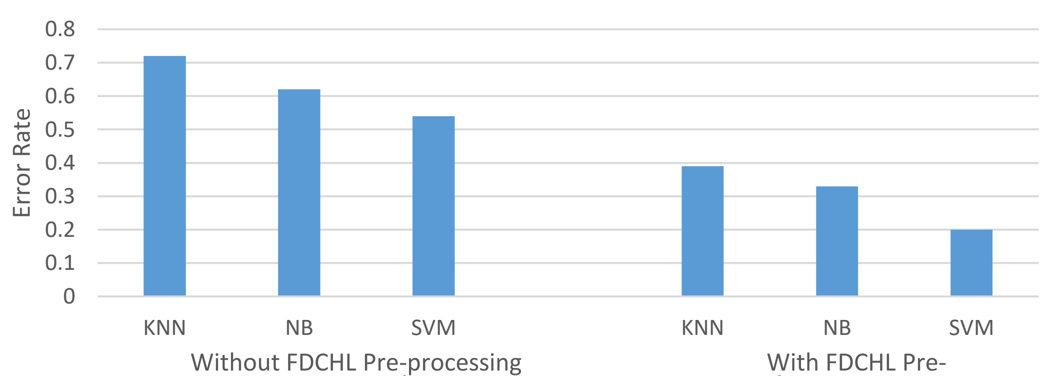

Figure 4 Error Rate Comparison

Figure 4 shows the Error rate comparison of various classifiers in terms of FDCHL Pre-processing. The SVM classifiers shows less error rate when compared to other classifiers.

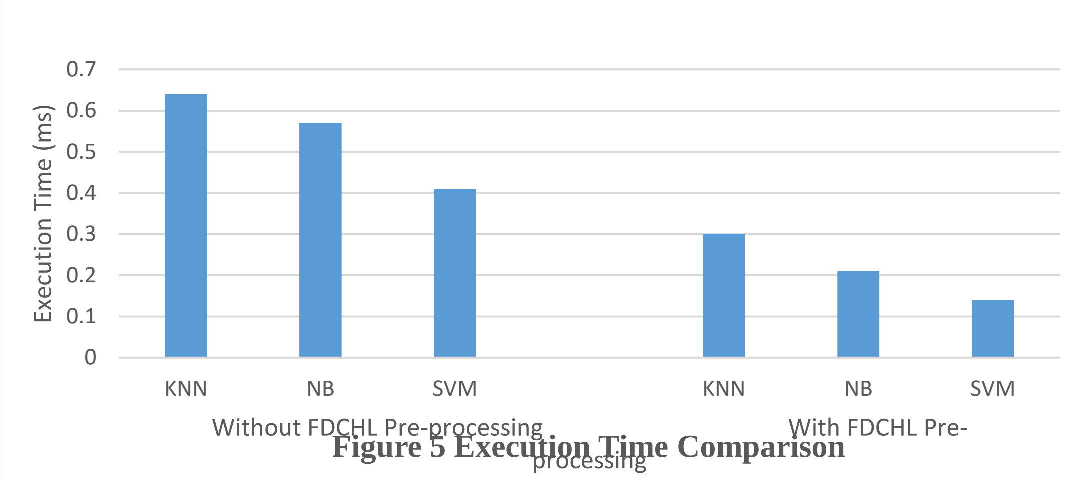

Figure 5 Execution Time Comparison

Figure 5 shows the Execution time comparison of various classifiers in terms of FDCHL Pre-processing. The SVM classifiers shows less execution time for detection of canker disease in citrus plant. Less execution and high accuracy will be the effective result for canker prediction.

5. Conclusion

Citrus is a vital plant grown mainly in the tropical areas of the world due to its richness in vitamin C and other important nutrients. Hence, this paper proposes an efficient methodology for detection of canker in citrus by applying FDCHL pre-processing to improve quality of image. First the sample leaves images are acquired and segmented into multiple parts. Then color and textures features are extracted and SVM classifiers are applied to detect the disease of the leaves. Experimental results express that our proposed enhancement methodology outperforms well in terms of image enhancement and the canker detection based classifiers achieves the efficient results of accurately detecting and differentiate the canker leaf disease.

References

1. Lee, H. W., Ho, W. W., Thangavel, R., Baskarathevan, J., & Alexander, B. J. (2019). Validation of qPCR assays for the detection of citrus canker. *New Zealand Plant Protection*, *72*, 282-282.
2. Sunny, S., & Gandhi, M. I. (2018). An efficient citrus canker detection method based on contrast limited adaptive histogram equalization enhancement. *Int. J. Appl. Eng. Res.*, *13*(1), 809-815.
3. Wetterich, C. B., de Oliveira Neves, R. F., Belasque, J., & Marcassa, L. G. (2016). Detection of citrus canker and Huanglongbing using fluorescence imaging spectroscopy and support vector machine technique. *Applied Optics*, *55*(2), 400-407.
4. Wang, L., Chen, S., Peng, A., Xie, Z., He, Y., & Zou, X. (2019). CRISPR/Cas9-mediated editing of CsWRKY22 reduces susceptibility to Xanthomonas citri subsp. citri in Wanjincheng orange (Citrus sinensis (L.) Osbeck). *Plant Biotechnology Reports*, *13*(5), 501-510.
5. Sabrol, H., & Kumar, S. (2015). Recent studies of image and soft computing techniques for plant disease recognition and classification. *International Journal of Computer Applications*, *126*(1).
6. Gavhale, K. R., Gawande, U., & Hajari, K. O. (2014, April). Unhealthy region of citrus leaf detection using image processing techniques. In *International Conference for Convergence for Technology-2014* (pp. 1-6). IEEE.
7. Murata, M. M., Omar, A. A., Mou, Z., Chase, C. D., Grosser, J. W., & Graham, J. H. (2019). Novel plastid-nuclear genome combinations enhance resistance to citrus canker in cybrid grapefruit. *Frontiers in plant science*, *9*, 1858.
8. Golmohammadi, M., Cubero, J., Peñalver, J., Quesada, J. M., López, M. M., & Llop, P. (2007). Diagnosis of Xanthomonas axonopodis pv. citri, causal agent of citrus canker, in commercial fruits by isolation and PCR-based methods. *Journal of Applied Microbiology*, *103*(6), 2309-2315.
9. Li, H., Lee, W. S., Wang, K., Ehsani, R., & Yang, C. (2014). 'Extended spectral angle mapping (ESAM)'for citrus greening disease detection using airborne hyperspectral imaging. *Precision Agriculture*, *15*(2), 162-183.
10. Madhu Yedla, Srinivasa Rao Pathakota and T. M. Srinivasa, Enhanced K-means Clustering Algorithm with Improved Initial Center, In International Journal of Science and Information Technologies, vol. 1(2), pp. 121–125, (2010).

11. Jia, H., Zhang, Y., Orbović, V., Xu, J., White, F. F., Jones, J. B., & Wang, N. (2017). Genome editing of the disease susceptibility gene Cs LOB 1 in citrus confers resistance to citrus canker. *Plant biotechnology journal*, *15*(7), 817-823.
12. Pu, Y. Y., Feng, Y. Z., & Sun, D. W. (2015). Recent progress of hyperspectral imaging on quality and safety inspection of fruits and vegetables: a review. *Comprehensive Reviews in Food Science and Food Safety*, *14*(2), 176-188.
13. Pitino, M., Armstrong, C. M., & Duan, Y. (2015). Rapid screening for citrus canker resistance employing pathogen-associated molecular pattern-triggered immunity responses. *Horticulture research*, *2*(1), 1-8.
14. Li, L., Li, J., Zhang, Y., & Wang, N. (2019). Diffusible signal factor (DSF)-mediated quorum sensing modulates expression of diverse traits in Xanthomonas citri and responses of citrus plants to promote disease. *BMC genomics*, *20*(1), 55.
15. Atta, B. M., Saleem, M., Ali, H., Ali, Z., & Zakria, M. (2019). Synchronous fluorescence spectroscopy for early diagnosis of citrus canker in citrus species. *Laser Physics*, *29*(8), 085604.
16. K. A. Abdul Nazeer and M. P. Sebastian, Improving the Accuracy and Efficiency of the k-means Clustering Algorithm, In Proceedings of the World Congress on Engineering, London, WCE, vol. 1, July (2001).