Malaya Journal of Matematik, Vol. 8, No. 4, 2173-2176, 2020
https://doi.org/10.26637/MJM0804/0144

Enhanced disease detection in plant leaves using segmentation and feature extraction techniques

Vijayabharathi Durairaj[1*] and Chellammal Surianarayanan[2]

Abstract
Plant diseases will decrease the yield, which will have a negative impact on the economy. Image processing techniques are essential for efficient detection of diseases in plants. Despite the existence of various research works for detection of plant diseases, obtaining enhanced accuracy is still an open challenge. In this work, kernel based Fuzzy C-Means clustering for segmentation along with Discrete Wavelet Transform based feature extraction and deep learning based autoencoder based classification are employed towards enhancing efficient detection of diseases in plant leaves. To test the proposed approach, experimentation has been carried out with different datasets in two different manners, namely, (i) with segmentation and (ii) without segmentation. Results are compared and recommendations are suggested.

Keywords
Kernel based Fuzzy C-Means Clustering, Discrete Wavelet Transform based feature extraction, deep learning based autoencoder, disease detection in plant leaves.

[1,2] *Department of Computer Science, Bharathidasan University Constituent Arts and Science College [Affiliated to Bharathidasan University], Navalurkuttapattu, Tiruchirappalli-620027, Tamil Nadu, India.*
***Corresponding author**: [1]vijayabharathiphd@gmail.com; [2]drschellammal@gmail.com
Article History: Received **14** September **2020**; Accepted **22** November **2020**

Contents

1. Introduction

Plants have become a significant energy source provided to solve the global warming issues [1]. Numerous samples of the collapse of the state economy due to the letdown of agriculture globally. Diseases in plants is the primary threat to the economy that directly impacts production and quality of crops. Disease in plants are caused by pathogens such as bacteria, virus and fungi. Also, each pathogen has its own specific infection and signs with particular lesions. Manual detection of diseases in plants is time consuming and also required expert knowledge [2]. With advent of modern technologies, image processing and machine learning have become viable tools for automatic and efficient detection of diseases in plants [3]. In this work, to enhance the accuracy of detection of diseases, kernel based Fuzzy C-Means clustering is used for image segmentation along with Grey Level Coherence Matrix (GLCM) features obtained with Discrete Wavelet Transform technique and with deep learning based autoencoder technique for detection of affected leaf images.

The organization of the paper is as follows. Section 2 describes the related literature. Section 3 presented the proposed approach. Section 4 describes the experimentation and results obtained using two different machine learning algorithms, namely, Support Vector Machine and deep learning based autoencoder technique. Section 5 concludes the work.

2. Related Work

The research work [4] illustrates different techniques for detection of diseases affecting plants. In [5] the use of kernel-based fuzzy C-Means clustering for segmentation is discussed. In [6] Artificial Neural Network (ANN) based method is proposed for pests' affected plant images. Others have illustrated various types of PD and techniques for inhibiting their growth. In [7], the authors demonstrated a framework for classifying images of using gradient-based features. In another work [8], Simple Iterative Linear Clustering Algorithm (SLIC) method is used group the color features (super pixels) which in turn

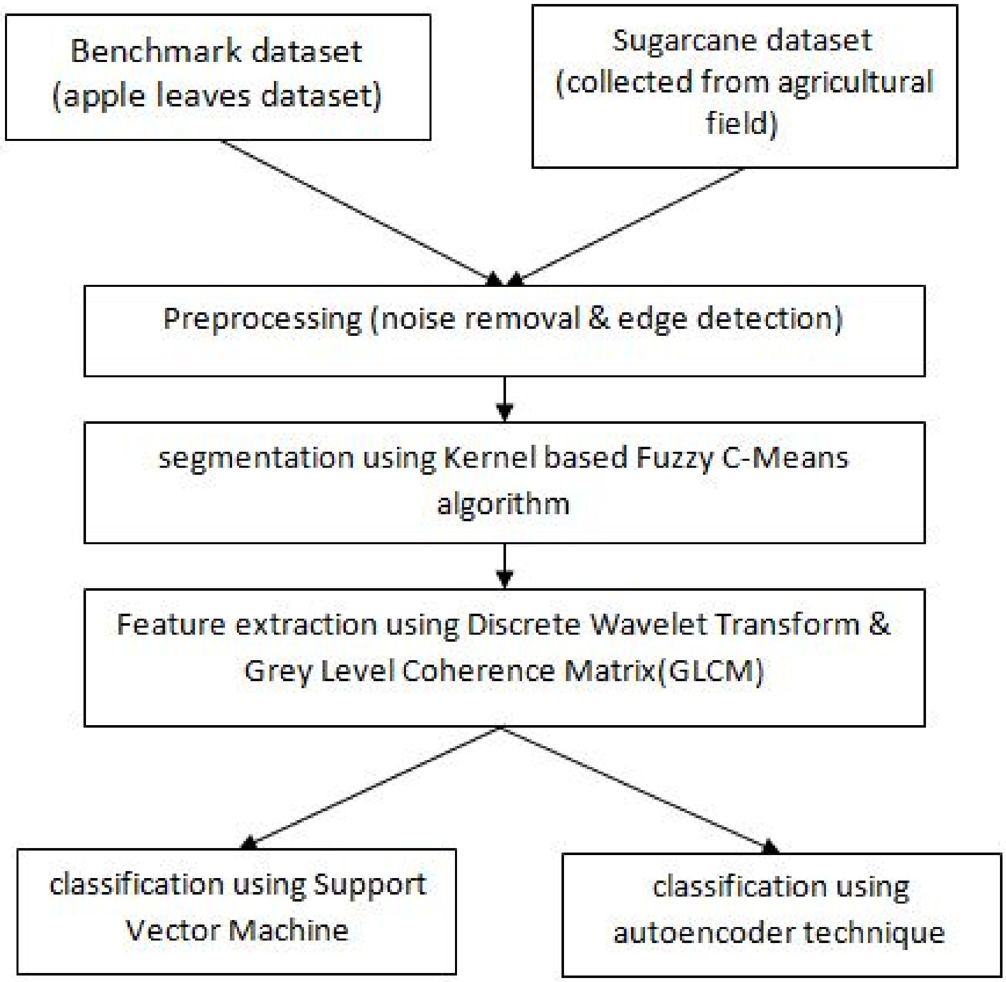

Figure 1. Proposed method of disease detection in plant leaves

are used to train the ANN to classify whether the super pixels that are neither healthy nor healthy. In [9], the authors recommended better method for diagnosing apple leaf diseases using Deep Neural Network (DNN).

In contrast to the above research works, the objective of the proposed work is to segment the preprocessed image using kernel based Fuzzy C-Means clustering and to extract the significant features using Discrete Wavelet Transformation. In addition, it is proposed to analyse the usefulness of the kernel based FCM for disease detection.

3. Proposed Approach

An approach is proposed for efficient detection of disease in plant leaves. The schematic of the proposed method is given in Fig. 1. As in Fig. 1, two datasets are constructed in this work. One is benchmark dataset is constructed using publicly available website (www.kaggle.com). This dataset is constructed using 100 apple leaf images. The images are converted into a size of 256*256 pixels. Another dataset is constructed by capturing sugarcane images from agricultural land in Thanjavur district. Initially, the data is preprocessed using Hessian of Laplacian of Gaussian (HLoG) filter. Preprocessed images are given for segmentation using kernel based Fuzzy C-Means clustering in order to facilitate accuracy feature extraction. Leaf lesions can be identified using segmentation techniques. In this work, we have used the Kernel-based FCM algorithm which is proposed in different research works [10,11] for segmentation. As mentioned in [10,11], the objective function of the above segmentation is defined through (3.1) and (3.2)

$$J_{\mathrm{KFCM}} = 2\left[\sum_{a=1}^{i}\sum_{b=1}^{c} u_{ab}^{m}\left(1-K\left(m_a, v_b\right)\right) + \sum_{a=1}^{i}\sum_{b=1}^{c} \varphi_a u_{ab}^{m}\left(1-K\left(\bar{m}_a, v_b\right)\right)\right] \tag{3.1}$$

$$u_{ij} = \frac{\left(\left(1-K\left(m_a, v_b\right)\right)+\varphi_a\left(1-K\left(\bar{m}_a, v_j\right)\right)\right)^{-1/(m-1)}}{\sum_{a=1}^{N} u_{a,b}^{m}\left(K\left(m_a, v_j\right)+\varphi_a\left(1-K\left(\bar{m}_a, v_j\right)\right)\right)^{-1/(m-1)}} \tag{3.2}$$

After segmentation, Haralick's texture features [12] which are obtained using DWT based GLCM as given in equation

(3.3) has been used

$$G = \begin{bmatrix} p(1,1) & p(1,2) & \cdots & p(1,N_g) \\ p(2,1) & p(2,2) & \cdots & p(2,N_g) \\ \vdots & \vdots & \ddots & \vdots \\ p(N_g,1) & p(N_g,2) & \cdots & p(N_g,N_g) \end{bmatrix} \quad (3.3)$$

The proximity of a square 2D pixel image (horizontal, vertical, left and right diagonal can be defined in any four directions and four such metrics can be calculated. The extracted features are classified using two different algorithms, namely deep learning based autoencoder and Support Vector Machine.

4. Experimentation

The proposed approach is tested using MATLAB (version 2018a) with a 3.0 GHz Intel i3 processor, 1TB hard disk and 8GB RAM. To analyze the performance of the proposed approach, two experiments are carried out, one with clustering and one without clustering. Experiments have been carried out for the above mentioned two different datasets, one apple dataset and sugarcane dataset with split ration of 70:30 for training and testing. The results are analyzed using different evaluation measures, sensitivity (SE), specificity (SP), accuracy (AC), recall (R) and precision (P). The measures are defined as follows.

Sensitivity

Sensitivity is expressed as the sum of true positive divided by true positive and false negative. This is given in (4.1)

$$SE = \frac{tp}{tp+fn}. \quad (4.1)$$

Accuracy

In term accuracy the term accuracy. It is the ratio of number of correct predictions to the total number of input samples. It is expressed as in (4.2)

$$AC = \frac{tp+tn}{tp+fp+tn+fn} \quad (4.2)$$

Precision

It is the number of correct positive results divided by the number of positive results predicted by the classifier. It is expressed as in (4.3)

$$P = \frac{tp}{tp+fp}. \quad (4.3)$$

Specificity

Sensitivity is the proportion of actual positives which are correctly identified as positives by the classifier. It is defined as in (4.4)

$$SP = \frac{tn}{tn+fp}, \quad (4.4)$$

where tp, tn, fp and fn represent the sum of a true positive, false positive, true negative and false negative. As for the classification, there are 4 assessment criteria, with sensitivity (SE), specificity (SP), accuracy (AC) Recall (R) and Precision (P).

The results obtained for apple dataset and sugarcane dataset using the two classifiers with and without segmentation are given as in Table 1 and Table 2 respectively.

The performance of the classifiers are shown in graphical representation in Fig. 2 and Fig. 3

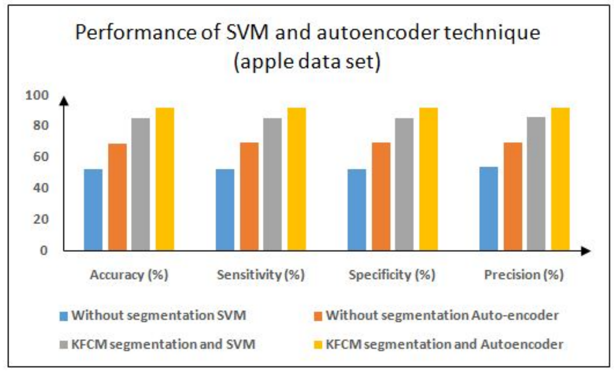

Figure 2. Graphical representation of results obtained for apple dataset

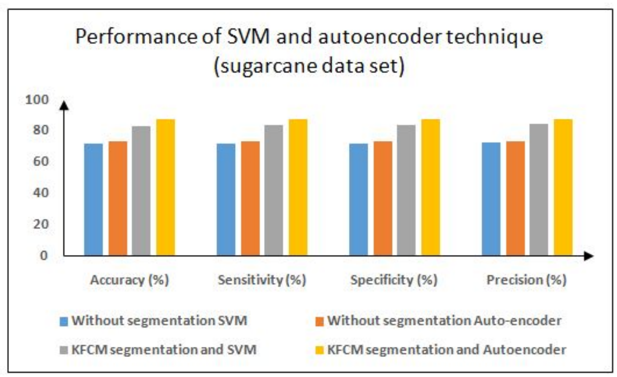

Figure 3. Graphical representation of results obtained for sugarcane dataset

5. Conclusion

In this work, an approach is proposed to enhance the accuracy of classification of leaf diseases in plants using Kernel based FCM and DWT based feature extraction. After feature extraction, the classification is performed using two different classifiers, namely, SVM and deep learning based autoencoder. The proposed approach has been tested using two different datasets. The performance obtained using the two algorithms are compared using different evaluation measures. From Table 1 and Table 2 it is found that deep learning based autoencoder performs better than SVM.

Table 1. Results obtained for apple leaf dataset (number of samples = 100)

Technique	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)
Without segmentation SVM	52	52.5	52.5	54
Without segmentation Auto-encoder	69	69.5	69.5	69.5
KFCM segmentation and SVM	85	85	85	85.5
KFCM segmentation and Autoencoder	92	92	92	92

Table 2. 2 Results obtained for sugarcane leaf dataset (number of samples = 100)

Technique	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)
Without segmentation SVM	72	72	72	72.5
Without segmentation Auto-encoder	73	73	73	73.5
KFCM segmentation and SVM	83	83.5	83.5	84
KFCM segmentation and Autoencoder	87	87	87	87.5

References

[1] J. G. A. Barbedo, Digital image processing techniques for detecting, quantifying and classifying plant diseases, *Springer Plus,* 2(660)(2013).

[2] S. Rout, A. Fulsunge, Plant disease detection in image processing using MATLAB, *IJIRSET*, 6(6)(2017), 10373-10381.

[3] S. Tadmore, B. Mahalakshmi, A survey on plant disease detection and classification using different machine learning algoriths, *IJCSE,* 7(2)2019, 338-341.

[4] G. Tripathi and J. Save, An image processing and neural network based approach for detection and classification of plant leaf diseases, *J. Impact Factor*, 6(4)(2015), 14-20.

[5] N. Vinushree, B. Hemalatha, and V. K. Kaliappan, *Efficient kernel-based fuzzy cmeans clustering for pest detection and classification,* in 2014 World Congress on Computing and Communication Technologies (WCCCT), (2014), 179-181.

[6] M. H. Javed, M. H. Noor, B. Y. Khan, N. Noor, and T. Arshad, *K*-means Based Automatic Pests Detection and Classification for Pesticides Spraying, *Int. J. Adv. Comput. Sci. Appl.*, 8(11)(2017), 236-240.

[7] K. Venugoban and A. Ramanan, Image classification of paddy field insect pests using gradient-based features, *Int. J. Mach. Learn. Comput.,* 4(1)(2004), 1-5.

[8] I. F. Salazar-Reque, S. G. Huamán, G. Kemper, J. Telles and D. Diaz, An algorithm for plant disease visual symptom detection in digital images based on superpixels, *Int. J. Adv. Sci. Eng. Inf. Technol.,* 9(1)(2019), 194-203.

[9] M. Turkoglu, D. Hanbay and A. Sengur, Multi-model LSTM-based convolutional neural networks for detection of apple diseases and pests, *Journal of Ambient Intelligence and Humanized Computing,* (2019), 1-11.

[10] NehaTomar, Vyom Kulshreshtha, Pankaj Sharma, Segmentation of Medical Images using Adaptively Regularized Kernel-based Fuzzy C-Means Clustering, *International Journal of Computer Applications,* 178(13)(2019).

[11] Rehna Kalam, Ciza Thomas and M Abdul Rahiman, Gaussian Kernel Based Fuzzy CMeans Clustering Algorithm For Image Segmentation, *Comput. Sci. Inf. Technol,* (2016), 47-56.

[12] Mohanty, P. Sharada, David P. Hughes, and Marcel Salathé. Using deep learning for image-based plant disease detection, *Frontiers in Plant Science*, 7(2016), 14-19.

⋆⋆⋆⋆⋆⋆⋆⋆⋆
ISSN(P):2319 − 3786
Malaya Journal of Matematik
ISSN(O):2321 − 5666
⋆⋆⋆⋆⋆⋆⋆⋆⋆