SEMARAK ILMU PUBLISHING
202103268166(003316878-P)

Semarak International Journal of Machine Learning

Journal homepage:
https://semarakilmu.com.my/journals/index.php/sijml/index
ISSN: 3030-5241

Olive Leaf Disease Detection using Improvised Machine Learning Techniques

Qusay Bsoul [1,*], Malik Jawarneh [2]

[1] Information Security and Cybersecurity Department, Philadelphia University
[2] Faculty of Computing Sciences, Gulf College

ARTICLE INFO

Article history:
Received 12 January 2025
Received in revised form 12 February 2025
Accepted 12 March 2025
Available online 21 March 2025

Keywords:
Artificial Intelligence; classification; Olive Leaf Disease; Machine Learning Techniques

ABSTRACT

Plants are integral to human life, and so, plant health is important. Regularly monitoring of plant health and plant disease detections are important in property agriculture. In agriculture, the use of image processing techniques run by computers in solving agricultural problems is increasingly common, particularly in the classification and identification of crop disease. Such usage could preserve the technical and commercial well-being of agriculture. This study demonstrated the application of support vector machine and image processing-enabled approach to detect and classify Olive leaf disease. It comprises seven steps that begin with a presentation of a digital color picture of a sickly leaf, followed by the step of image denoising using mean function, image enhancing using CLAHE method, image segmentation using fuzzy C Means algorithm, image feature extraction using PCA, and disease detection and classification using PSO SVM, BPNN, and random forest algorithms. The results showed high accuracy of the proposed PSO SVM in Olive leaf disease classification and detection.

1. Introduction

Just like in other domains, the use of image processing in agricultural domain has been increasing [1,2]. Image processing is a rapidly progressing technological discipline, and it involves the use of various tools for image capturing. Among the common ones include cameras and satellites. These images are then processed using different analysis techniques run by a computer, to generate the sought after data. In the domain of agriculture, the use of image and data processing has helped in solving many problems related to agriculture, including in the classification and detection of plant related diseases. Using image processing, the diseased plant parts like stem, fruit or leaf can be singled out, while the diseased area can be measured. Equally, diagnosis of the disease can be formulated based on image color, shape, and size.

Image processing involves the application of various techniques to an image, so that the image could be improved and information can be obtained from it [3,4], and one image as an input can be

* *Corresponding author.*
E-mail address: qbsoul@philadelphia.edu.jo

https://doi.org/10.37934/sijml.5.1.6473

used in producing several images. Some images would require enhancement or modification in order that it could be utilized in a different setting, and image processing is among the common techniques used in image enhancement or modification. Not only that, through image processing techniques, images can be segmented, and some of its features can be extracted. Additionally, image enhancement may involve adjustment of some aspects of the image like its sharpness, color and brightness, and reduction of noise.

Some images are large in size, and in dealing with such images, segmentation is commonly done, whereby the image is broken down into several smaller images, especially in digital photograph recognition. Among the methods commonly used to segment image include texture-based methods, thresholding methods, and color-based methods. In feature extraction, the image dimensionality would be reduced leaving only the most discernible and crucial aspects of the image. This method could rapidly match large pictures while decreasing the amount of feature representations, simultaneously. Meanwhile, in image categorization, each picture is stationed in a discrete category in accordance with some predetermined criteria.

Considering the significant impact of plant diseases on mankind, the agriculture industry needs to be properly taken care of through (among others), correct classification and identification of crop diseases [5]. In property agriculture, the health of plant needs to be consistently monitored, while diseases need to be correctly identified. Utilizing some detection and classification methods, disease affecting plants could be identified and classed; some methods could only identify specific disease and symptoms, while others could identify certain disease and symptoms from a vast range of diseases and symptoms. The use of picture processing in identifying plant disease begins with the presentation of a digital color image of a sick plant parts like leaf, fruit or stem, to a computer system specifically developed for disease identification. The digital color image of the sickly plant part needs to have clear background so that the disease could be clearly identified without being distracted by other irrelevant elements or objects. Also, the capture settings of the image need to be controlled to ease disease identification [6,7].

In the identification and classification of plant related diseases, especially those related to Olive plant, an outstanding approach was proposed in this study, namely an approach involving the use of support vector machines and image processing. Accordingly, a framework comprising seven steps was presented (see Figure 1).

2. Literature Survey

Agricultural image processing eases both scholars and practitioners in resolving some agriculture related issues, such as crop diseases, by facilitating them in disease identification and classification. The ability in detecting and classifying plant diseases is important in preserving the well-being of the agricultural industry. As plant diseases can affect livelihood of man in general, plant health needs to be consistently monitored and their detected diseases need to be properly managed. Accordingly, studies on plant disease identification and classification have been carried out involving several varieties of plants and methods of detection and identification.

Sanyal [10] examined mineral insufficiencies and illnesses of rice leaf. On rice leaves, mineral deficiencies lead to brown spots called lesions that emerge in various shapes and sizes on the leaves. In this study, the author examined 400 rice leaf images in terms of color and texture – different colors and textures were extracted from the images, which were then fed to an ANN equipped with a single hidden layer, namely the Multilayer Perceptron (MLP). In this study, the RGB images were changed to HSI color space. The images were then segmented using entropy-based thresholding. Next, the segmented images were transformed into a grey scale image, which were then analysed utilizing an edge detection technique. Self-organizing maps were used to categorize the detected diseases.

Utilizing an intelligent system equipped with MLP-ANN and support vector machine (SVM), Meunkaewjinda *et al.,* [11] attempted to identify diseases on grape leaves. The authors used MLP-ANN to distinguish the object of the image namely the grape leaves, and its background. Then, the diseased segments of the leaves were identified using SVM. The identified diseases were then classed using a multiclass SVM. Relevantly, Pagola *et al.,* [12] employed RGB alterations, PCA and softmax regression in determining the levels of nitrogen in barley leaves. The methods used were compared in terms of accurateness, by comparing the results obtained from the proposed method with the results produced from a chlorophyll meter. Based on their examination, the examined barley leaves did not have sufficient nitrogen level. In another study by Carmargo and Smith [13], cotton plant diseases were diagnosed using picture pattern classification algorithm. The images of cotton plants were segmented, and the output was fed into an SVM with an on-one approach, to classify the diseases. The authors employed texture characteristics to their approach and were successful.

Utilizing SVM-based technique on images of cucumber leaves, Jian and Wei [14] detected diseases on Cucumber plant leaves. Image data features were extracted using basic thresholding method, and the authors trained an SVM via using these features. Radial basis function kernel, polynomial kernel, and sigmoid kernel function were the methods used on SVM in comparing the model's performance of the model. Among the three methods used on SVM, radial basis function kernel appeared to be the most effective, as shown by the outcomes. A spectrometer could detect deficiency of nutrient in palm plants. The use of the method begins with segmentation of images of palm plants based on color similarities. In this study, color and texture features were extracted from the segmented images using an algorithm. The obtained features were fed to fuzzy classifiers to class the obtained data.

Malnourished tomatoes were examined using a classifier, involving the conversion of the L * a * b * and RGB color spaces into each other, in the extraction of color and texture features from the tomato leaf images. The extraction of color and texture of tomato leaf images was performed using fourier transforms, wavelet packets, and percent intensity histograms. Then, the extracted features were classified using fuzzy K-nearest neighbor model. The approach used in this study had 82.5% accurateness level. Utilizing neural networks, Wang et al. [15] classified wheat and grapevine diseases on captured images. The images were segmented using K-means. Then, the features of the segmented images were extracted – specifically, the color, shape, and texture features were extracted. The extracted features were classified using MLP, Radial Basis Function (FBF), Generalized Regression, and Probabilistic ANNs, and among these methods used, the authors reported that RBF had the highest level of accurateness.

Gupta [17] demonstrated the application of image processing and a classifier called SVM-Cuckoo Search classifier in identifying plant diseases on images. Histogram equalization was used for contrast enhancement. K-means clustering was used in segmentation process. SVM-Cuckoo Search classifier used for data classification showed 95% accurateness level.

In their study, Mwebaze and Owomugisha [16] introduced a method that identifies five forms of disorders and five stages of disease development, in the attempt to identify diseases and the severity of these diseases on plant leaves. In extracting features from the leaf images, the authors utilized color and ORB feature transformations. Then, the extracted features were fed to an SVM classifier to classify the diseases. In addition to the said methods, the authors also created a mobile application hosted on a remote server.

2. Methodology

As mentioned, this study demonstrated an approach to detect and classify olive leaf disease. For the purpose, a new support vector machine and image processing were applied. There are seven steps included in the proposed approach. This approach, as displayed in Figure 1, begins with image acquisition, followed by image denoising using mean function. After denoising, the image is enhanced using CLAHE method, following by the process of image segmentation which involves the use of fuzzy C Means algorithm. Feature extraction as the fifth step, involves the use of PCA, and then, diseases were classed using PSO SVM, BPNN, and random forest algorithms. Figure 2 illustrates the process.

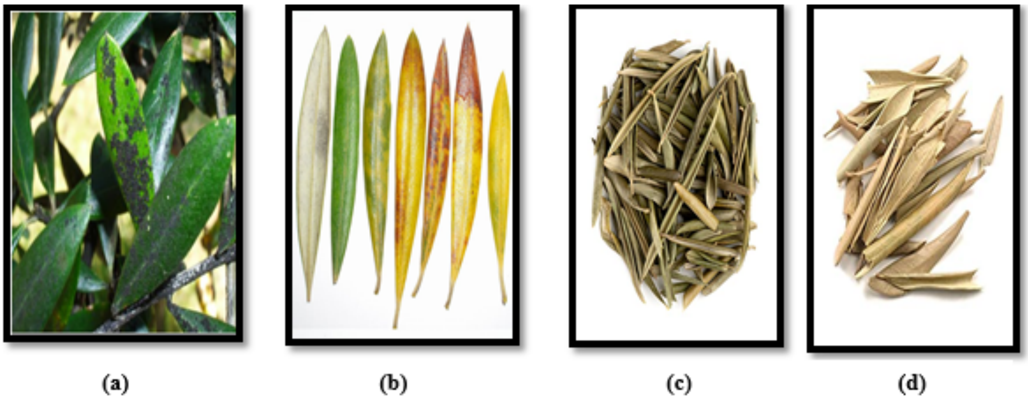

Fig. 1. Images of olive leaf disease [8]

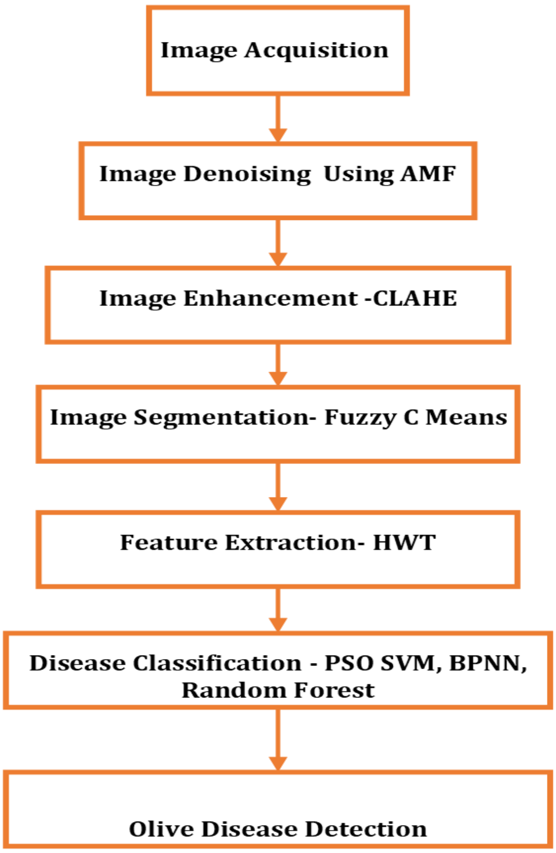

Fig.2. Image processing-enabled methodology for detection and classification of olive leaf disease

Past studies have shown that Adaptive median filtering (AMF) could effectively remove unwanted noise from photos, and also determine the image pixels that are impacted by impulse noise. Hence, the right action could be determined. Relevantly, unprompted noise results from substantial fraction of misaligned pixels within an image. Notably, the noise-free pixels within the surrounding pixels are obscured by the median value of the noise-free pixels within the nearby pixels [18].

In order that an image could be identified, it is important that the background extraction process does not encroach the image quality. Accordingly, pixel value histograms and the nearby region's value histograms may be produced with CLAHE. It should be noted that CLAHE will constrain the highest contrast modification to the pinnacle of local histogram which is the highest contrast enhancement factor, by specifying the clip level representing the maximum, and so, the image will appear cleaner. For this reason, the use of CLAHE in mammograms has been appropriate because it results in clearer mammograms, allowing more minute details to be discernible [19]. The signal and the noise can also be easily distinguished with CLAHE. However, CLAHE could result in grainy images.

Clustering occurs based on the intensity value of the pixels. Here, the preprocessed pixel values of the image are broken down into various classes, allowing pixels in alike class to be compared, but the same cannot be said between pixels in different classes. Clustering algorithms are many in number, and clusters can also be subsets of bigger dataset. The subsets can determine the classification of clustering methods, and the subsets can be either fuzzy or crisp. In general, fuzzy clustering algorithms are the better ones for clustering, and this type of algorithm, especially fuzzy c-means (FCM), could break the image into many clusters that overlap with one another at certain degree. In this study, fuzzy c-means was employed in image processing, to find object clusters inside a given image. A spatial element was added to enhance FCM algorithm, making it more accurate in clustering images with noise [20].

Haar wavelet transformation is considered as the most straightforward type of wavelet transform [21]. All wavelet transformations use the Haar transform as a sampling procedure. Utilizing the Haar transform, a signal could be decreased by half. Additionally, PSO SVM simplifies and speeds up the binary linear classification, in determining the target groups. A dot or a point represents each data, and culture diversity in the data will extend the data. The additional instances are used in determining the allocation location of target class. Support vector machine algorithms, which are non-linear classification method, may be used for unlabeled input datasets [22]. However, this study employed an unsupervised learning approach because no objective classes were to be apportioned to the instances. Also, more instances may be added in creating function-based clusters.

Haykin and Anderson were the creators of the popularly used back propagation technique in learning algorithms. This method entails a learning process rather than a network. BPN is suitable when dealing with simple pattern recognition and mapping tasks. A training pair is made up of an input and a target [23], and utilizing algorithm examples, the network is trained, specifically to generate fitting output for each pattern of input. The weights of the network are modified as needed.

Random Forest is a common method for classifying. It is a decision-tree-based classifier. The trees in the model are created with bootstrap sample of the data and random sampling of features. Bagging and random selection can be used in tree creation. Accordingly, during the development of the forest, the class prediction accurateness of trees is significantly affected by the relationship between tree pairs. In this situation, error rates may be expected. Issues from regression and classification can be ranked using this strategy [24-26].

4. Results and Discussion

A total of 440 images of olive leaves were used in the study experiments. From the 440 images, 260 were images of sickly leaf, and 180 images were of healthy leaf. A total of 260 images were used in training the machine learning classifiers. Mean function was used for denoising, while CLAHE was used for improving, the images. Fuzzy C Means algorithm was the method used in image segmentation, and features from these segmented images were extracted with PCA. Then, PSO SVM, BPNN, and random forest algorithm were applied in disease classification.

These parameters below were employed in this study to compare results; Figs 3–7 can be referred.

$$Accuracy=\frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (1)$$

$$Sensitivity=\frac{(TP)}{(TP+FN)}` \quad (2)$$

$$Specificity=\frac{(TN)}{(TN+FP)}` \quad (3)$$

$$Precision=\frac{(TP)}{(TP+FP)}` \quad (4)$$

$$Recall=\frac{(TP)}{(TP+FN)}` \quad (5)$$

Referring to the above: TP = True Positive, TN = True Negative, FP = False Positive, and FN = False Negative.

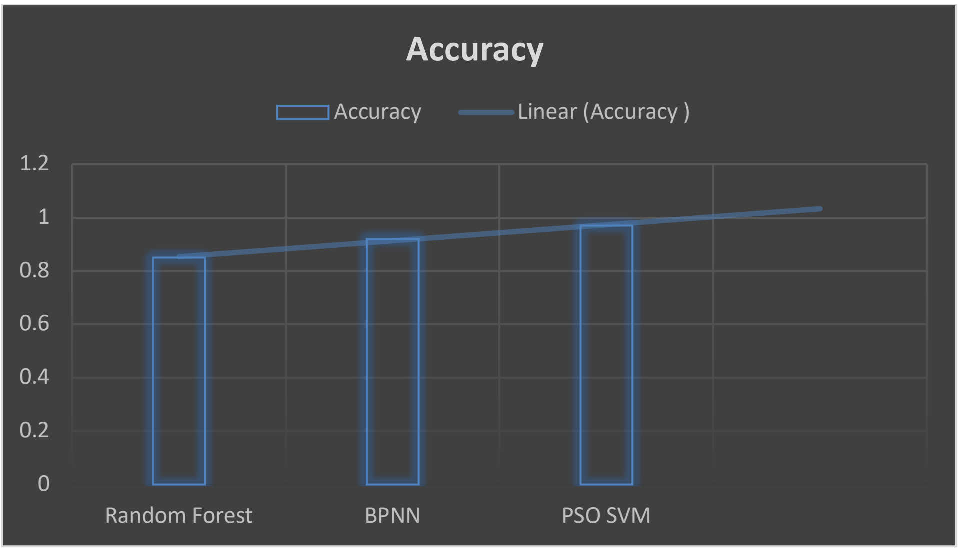

Fig. 3. Accuracy of classifiers for olive leaf disease classification

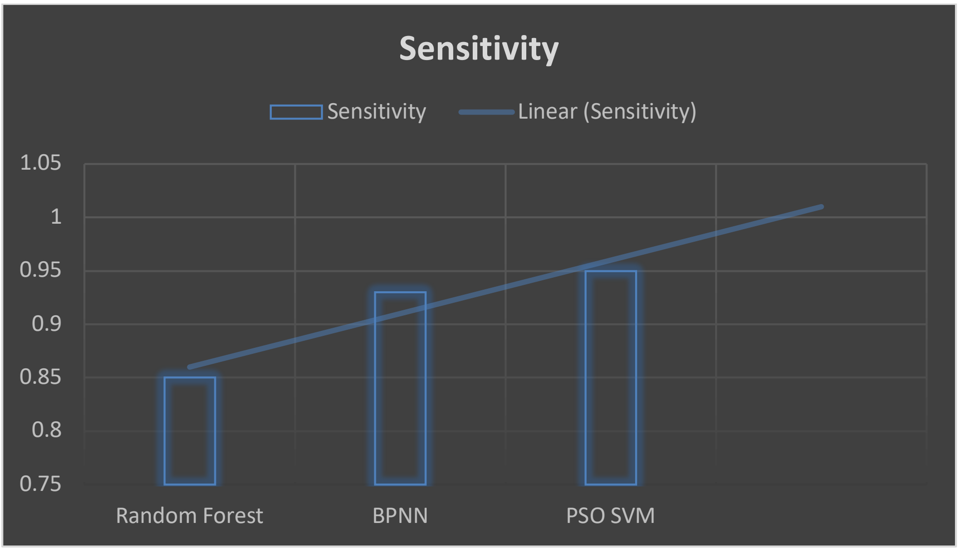

Fig. 4. Sensitivity of classifiers for olive leaf disease classification

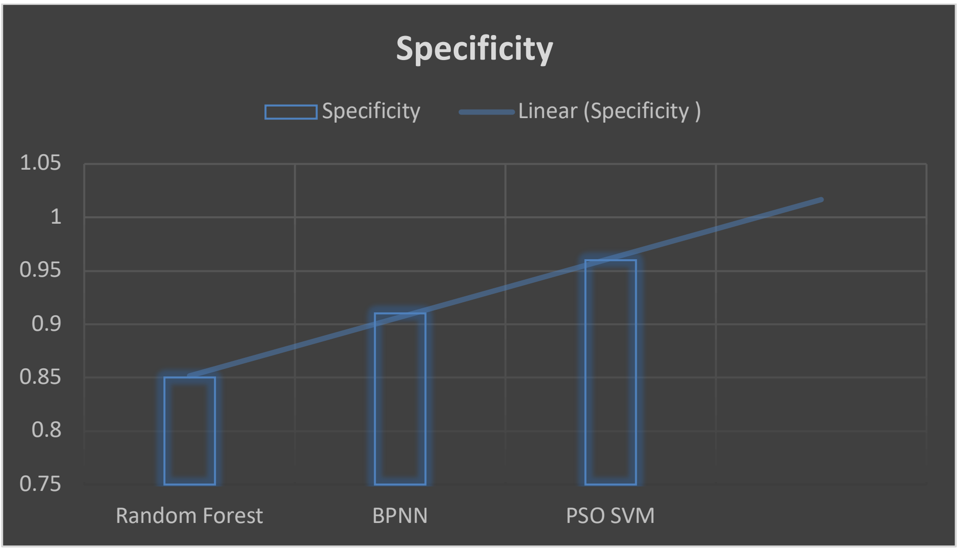

Fig. 5. Specificity of classifiers for olive leaf disease classification

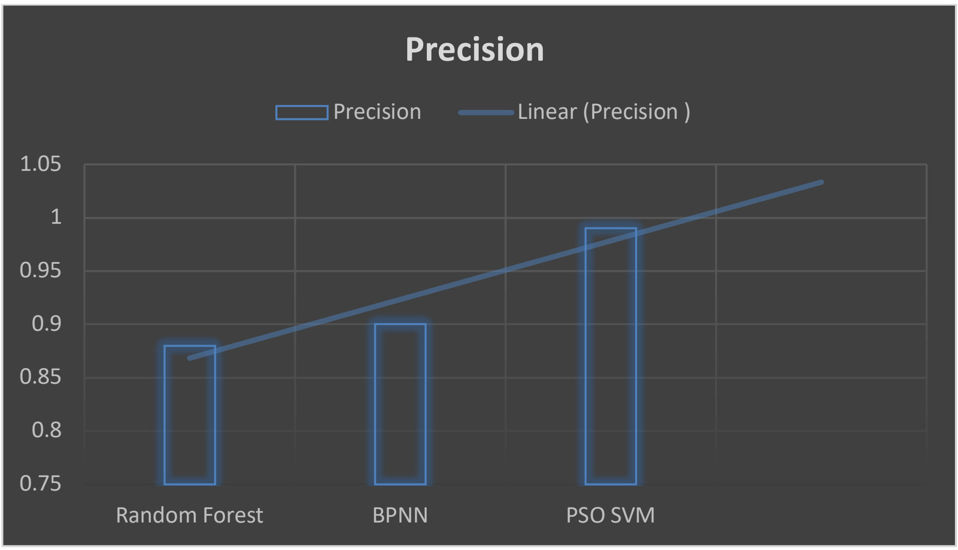

Fig. 6. Precision of classifiers for olive leaf disease classification

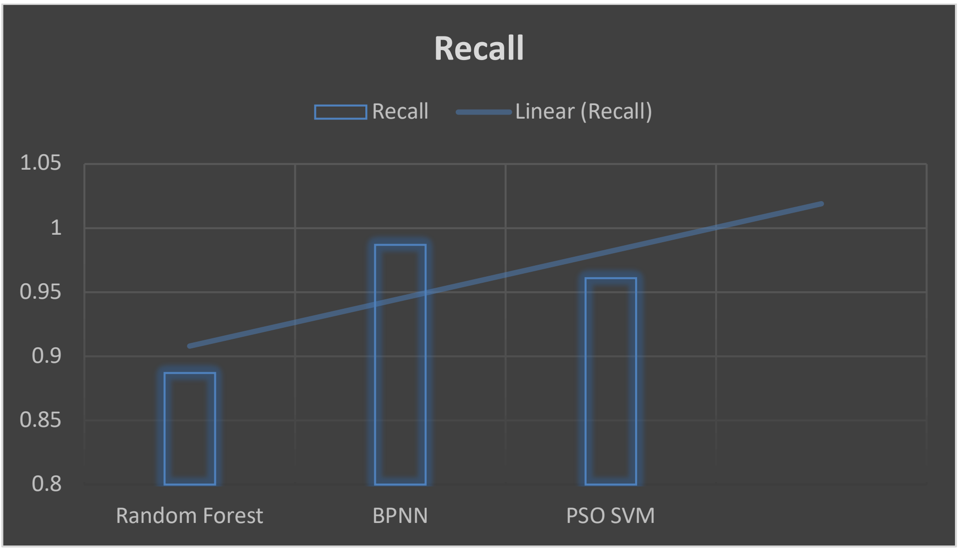

Fig. 7. Recall of classifiers for olive leaf disease classification

4. Conclusions

Agricultural image processing technology is increasingly progressing. As the technologies progress, the management of many issues associated with agricultural becomes simpler and faster, especially with the use of acquired images and data processing. Plant health should be consistently monitored, and the diseases should be detected promptly so that more serious problems could be prevented. In addition, effective crop diseases classification and diagnosis could improve the agricultural business. Accordingly, this study demonstrated the application of a classification approach with the capability of disease classification. The approach proposed includes a new support vector machine and image processing technique, for the classification and detection of olive leaf disease. Computers are the main tool used in image processing to analyse the acquired images. The approach proposed in this study has seven stages, which begin with image acquisition, followed by image denoising, image enhancement, image segmentation, image feature extraction, and lastly, disease classification from the extracted features, and then disease detection. The progression of the stages involves the use of methods such as mean function in image denoising stage, CLAHE in image enhancement stage, fuzzy C Means algorithm in image segmentation stage, PCA in feature extraction stage, and PSO SVM, BPNN, and random forest algorithms in disease classification from the extracted features, as the final stage before disease detection. The results proved good accurateness of PSO SVM in the classification and detection of olive leaf diseases.

Acknowledgement
This project has been conducted at the Information Security and Cyber security Department, Faculty of Information Technology, Philadelphia University under the funding provided by the Deanship of Scientific Research and Graduate Studies. We want to express our appreciation to the sponsor and everyone who has contributed to the project.

References
[1] Ji, Miaomiao, Lei Zhang, and Qiufeng Wu. "Automatic grape leaf diseases identification via UnitedModel based on multiple convolutional neural networks." *Information Processing in Agriculture* 7, no. 3 (2020): 418-426. https://doi.org/10.1016/j.inpa.2019.10.003

[2] Kaur, Parminder, Husanbir Singh Pannu, and Avleen Kaur Malhi. "Plant disease recognition using fractional-order Zernike moments and SVM classifier." *Neural Computing and Applications* 31 (2019): 8749-8768. https://doi.org/10.1007/s00521-018-3939-6

[3] Raghuvanshi, Abhishek, Umesh Kumar Singh, Guna Sekhar Sajja, Harikumar Pallathadka, Evans Asenso, Mustafa Kamal, Abha Singh, and Khongdet Phasinam. "Intrusion detection using machine learning for risk mitigation in IoT-enabled smart irrigation in smart farming." *Journal of Food Quality* 2022, no. 1 (2022): 3955514. https://doi.org/10.1155/2022/3955514

[4] Liang, Wan-jie, Hong Zhang, Gu-feng Zhang, and Hong-xin Cao. "Rice blast disease recognition using a deep convolutional neural network." *Scientific reports* 9, no. 1 (2019): 1-10. https://doi.org/10.1038/s41598-019-38966-0

[5] Hemamalini, V., S. Rajarajeswari, S. Nachiyappan, M. Sambath, T. Devi, Bhupesh Kumar Singh, and Abhishek Raghuvanshi. "Food quality inspection and grading using efficient image segmentation and machine learning-based system." *Journal of Food Quality* 2022, no. 1 (2022): 5262294. https://doi.org/10.1155/2022/5262294

[6] Jiang, Peng, Yuehan Chen, Bin Liu, Dongjian He, and Chunquan Liang. "Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks." *Ieee Access* 7 (2019): 59069-59080. https://doi.org/10.1109/ACCESS.2019.2914929

[7] Jasti, V. Durga Prasad, Abu Sarwar Zamani, K. Arumugam, Mohd Naved, Harikumar Pallathadka, F. Sammy, Abhishek Raghuvanshi, and Karthikeyan Kaliyaperumal. "Computational technique based on machine learning and image processing for medical image analysis of breast cancer diagnosis." *Security and communication networks* 2022, no. 1 (2022): 1918379. https://doi.org/10.1155/2022/1918379

[8] Fenu, Gianni, and Francesca Maridina Malloci. "DiaMOS plant: A dataset for diagnosis and monitoring plant disease." *Agronomy* 11, no. 11 (2021): 2107. https://doi.org/10.3390/agronomy11112107

[9] Ansari, Arshiya S., Malik Jawarneh, Mahyudin Ritonga, Pragti Jamwal, Mohammad Sajid Mohammadi, Ravi Kishore Veluri, Virendra Kumar, and Mohd Asif Shah. "[Retracted] Improved Support Vector Machine and Image Processing Enabled Methodology for Detection and Classification of Grape Leaf Disease." *Journal of Food Quality* 2022, no. 1 (2022): 9502475. https://doi.org/10.1155/2022/9502475

[10] Sanyal, Pritimoy, and Suresh C. Patel. "Pattern recognition method to detect two diseases in rice plants." *The Imaging Science Journal* 56, no. 6 (2008): 319-325. https://doi.org/10.1179/174313108X319397

[11] Meunkaewjinda, A., P. Kumsawat, K. Attakitmongcol, and A. Srikaew. "Grape leaf disease detection from color imagery using hybrid intelligent system." In *2008 5th international conference on electrical engineering/electronics, computer, telecommunications and information technology*, vol. 1, pp. 513-516. IEEE, 2008. https://doi.org/10.1109/ECTICON.2008.4600483

[12] Pagola, Miguel, Rubén Ortiz, Ignacio Irigoyen, Humberto Bustince, Edurne Barrenechea, Pedro Aparicio-Tejo, Carmen Lamsfus, and Berta Lasa. "New method to assess barley nitrogen nutrition status based on image colour analysis: Comparison with SPAD-502." *Computers and electronics in agriculture* 65, no. 2 (2009): 213-218. https://doi.org/10.1016/j.compag.2008.10.003

[13] Camargo, Anyela, and J. S. Smith. "An image-processing based algorithm to automatically identify plant disease visual symptoms." *Biosystems engineering* 102, no. 1 (2009): 9-21. https://doi.org/10.1016/j.biosystemseng.2008.09.030

[14] Jian, Zhang, and Zhang Wei. "Support vector machine for recognition of cucumber leaf diseases." In *2010 2nd international conference on advanced computer control*, vol. 5, pp. 264-266. IEEE, 2010. https://doi.org/10.1109/ICACC.2010.5487242

[15] Wang, Haiguang, Guanlin Li, Zhanhong Ma, and Xiaolong Li. "Application of neural networks to image recognition of plant diseases." In *2012 International conference on systems and informatics (ICSAI2012)*, pp. 2159-2164. IEEE, 2012. https://doi.org/10.1109/ICSAI.2012.6223479

[16] Owomugisha, Godliver, and Ernest Mwebaze. "Machine learning for plant disease incidence and severity measurements from leaf images." In *2016 15th IEEE international conference on machine learning and applications (ICMLA)*, pp. 158-163. IEEE, 2016. https://doi.org/10.1109/ICMLA.2016.0034

[17] Gupta, T. "Plant leaf disease analysis using image processing technique with modified SVM-CS classifier." *Int. J. Eng. Manag. Technol* 5 (2017): 11-17.

[18] Yao, Qi, Mohammad Shabaz, Tarun Kumar Lohani, Mohammed Wasim Bhatt, Gurpreet Singh Panesar, and Raj Karan Singh. "3D modelling and visualization for vision-based vibration signal processing and measurement." *Journal of intelligent systems* 30, no. 1 (2021): 541-553. https://doi.org/10.1515/jisys-2020-0123

[19] Gupta, Ankur, and Nimrita Koul. "SWAN: a swarm intelligence based framework for network management of IP networks." In *International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007)*, vol. 1, pp. 114-118. IEEE, 2007. https://doi.org/10.1109/ICCIMA.2007.63

[20] Yang, Ming, Pawan Kumar, Jyoti Bhola, and Mohammad Shabaz. "Development of image recognition software based on artificial intelligence algorithm for the efficient sorting of apple fruit." *International Journal of System Assurance Engineering and Management* 13, no. Suppl 1 (2022): 322-330. https://doi.org/10.1007/s13198-021-01415-1

[21] Struzik, Zbigniew R., and Arno Siebes. "The Haar wavelet transform in the time series similarity paradigm." In *European Conference on Principles of Data Mining and Knowledge Discovery*, pp. 12-22. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999. https://doi.org/10.1007/978-3-540-48247-5_2

[22] Lokhande, Meghana P., Dipti Durgesh Patil, Lalit V. Patil, and Mohammad Shabaz. "[Retracted] Machine-to-Machine Communication for Device Identification and Classification in Secure Telerobotics Surgery." *Security and communication networks* 2021, no. 1 (2021): 5287514. https://doi.org/10.1155/2021/5287514

[23] Parab, Jivan, Marlon Sequeira, Madhusudan Lanjewar, Caje Pinto, and Gourish Naik. "Backpropagation neural network-based machine learning model for prediction of blood urea and glucose in CKD patients." *IEEE journal of translational engineering in health and medicine* 9 (2021): 1-8. https://doi.org/10.1109/JTEHM.2021.3079714

[24] Sandika, Biswas, Saunshi Avil, Sarangi Sanat, and Pappula Srinivasu. "Random forest based classification of diseases in grapes from images captured in uncontrolled environments." In *2016 IEEE 13th international conference on signal processing (ICSP)*, pp. 1775-1780. IEEE, 2016. https://doi.org/10.1109/ICSP.2016.7878133

[25] Ansari, Arshiya S., Malik Jawarneh, Mahyudin Ritonga, Pragti Jamwal, Mohammad Sajid Mohammadi, Ravi Kishore Veluri, Virendra Kumar, and Mohd Asif Shah. "[Retracted] Improved Support Vector Machine and Image Processing Enabled Methodology for Detection and Classification of Grape Leaf Disease." *Journal of Food Quality* 2022, no. 1 (2022): 9502475. https://doi.org/10.1155/2022/9502475

[26] Pallathadka, Harikumar, Malik Jawarneh, F. Sammy, Vipul Garchar, Domenic T. Sanchez, and Mohd Naved. "A review of using artificial intelligence and machine learning in food and agriculture industry." In *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, pp. 2215-2218. IEEE, 2022. https://doi.org/10.1109/ICACITE53722.2022.9823427