

DETECTING MILDEW DISEASES IN CUCUMBER USING IMAGE PROCESSING TECHNIQUE

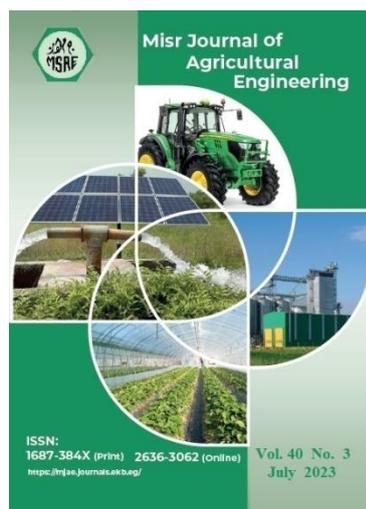
Alaa Aladdin I. Eldesoky¹, Heba M. Abd El-Nabi², and El-Sayed E. Omran^{3&*}

¹ Agricultural Specialist, General Administration of Plant Quarantine of Port Said and North Sinai, Port Said, Egypt.

² Prof. of Plant Pathology, Plant Pathology Dept., Fac. of Ag., Suez Canal U., Ismailia, Egypt.

³ Prof. of Soil and Water, Soil and Water Dept., Fac. of Ag., Suez Canal U., Ismailia, Egypt.

* E-mail: ee.omran@gmail.com



© Misr J. Ag. Eng. (MJAE)

Keywords:

Cucumber; Downy Mildew; Powdery Mildew; Machine Learning; Image Processing.

ABSTRACT

In Egypt, Cucumber is a crucial cash crop, and its farming could significantly benefit the country's agriculture-based economy. Meanwhile plant disease detection manually is costly and time consuming. This study aims to improve early identification of downy and powdery mildew diseases using machine vision by comparing the effectiveness of several detection methods and developing a real life application. This approach will involve five steps including: image acquisition, pre-processing, feature extraction, post-processing, and classification. In which the experiment was conducted on two greenhouses where 931 images were obtained and used with five key features to train and evaluate the proposed methods. The classification performance of three machine learning algorithms, named discriminant analysis (DA), support vector machine (SVM) and K-nearest neighbors (KNN), were compared. The results indicated that the fine gaussian SVM achieved the highest classification accuracy rate of 96%, where fine KNN got 95.8%, and quadratic DA obtained the lowest value 92.8%. Additionally, the suggested method has a practical application that enables automatic mildew disease detection via personal computers, eliminating the need for sample collection and laboratory analysis. This method could also be extended to identify other plant diseases and pests and track disease progression as the study moves forward.

1. INTRODUCTION

The Egyptian government has opted in modernizing the agricultural sector which will pave the way to industrialization and economic transformation. It has embraced protected agriculture technology by planning to build 100,000 feddans of greenhouses to achieve economic return thanks to rationalizing water consumption. According to estimations from FAO, about 613 thousand tons of cucumber and gherkins are being produced from approximately twenty-seven thousand hectares in Egypt of which eleven thousand tons were exported in 2020 (FAOSTAT, 2020). Cucumber is exposed to a variety of diseases such as powdery mildew, downy mildew which reduce yields (Yousuf and Dar,

2016; Ni and Punja, 2019 and Daunde et al., 2020). The deteriorating effects of climate change pose a challenge to agriculture transition. Hot climate helped fungal diseases to spread to new regions, which affected the agricultural industry and reduced our ability to export our crops. Plant diseases cost a huge amount of production expenses, which consider an obstacle in productivity in the global agricultural economy.

Powdery Mildew (PM) mainly affects cucumber in the middle and final stages resulting in significant production loss (**Lin et al., 2019**). *Podosphaera xanthii* (*P. fusca*) and *Erysiphe cichoracearum* are responsible for causing powdery mildew diseases. It affects lower regions of the plant earlier as the pathogens prefer shaded regions. So, it appears on older leaves first (**Wspanialy and Moussa, 2016**). Conidia produced profusely in the white powdery mycelium, and these spores spread quickly through wind to nearby leaves or plants as well as traveling over long distances. It can remain viable for 7-8 days (**Verhaar 1998 and Sharma et al., 2016**). The symptoms appear as a white circular powdery patch emerges on the upper and lower surfaces of infected leaves and stems (**Pawar et al., 2016; Bandamaravuri et al., 2020 and Daunde et al., 2020**). The disease develops at temperature of 27-35°C with a relative humidity of more than 70%. It would take 3-7 days after initial infection (**Ni and Punja, 2021 and Mostafa et al., 2021**).

Downy Mildew (DM) is one of the most prevalent and destructive foliar diseases of cucumber in open fields and under the greenhouses. *Pseudoperonospora cubensis* is an obligate parasite which grows and reproduces on the living plant tissue. Sporangia are the main source of inoculum that spreads from infected plants to neighbor plants or far sites by wind currents or rain splashes. It only takes almost one hour for Sporangia to germinate, after reaching the host's plant and while free water is present on the leaves. The Sporangia production can occur at temperatures of 5-30°C, with at least 6 hours of high humidity (**Lebeda and Cohen, 2011; Savory et al., 2011 and Sharma et al., 2016**). The disease cannot develop at hot temperatures exceeding 35°C (**Sharma et al., 2016**). The symptoms appear as pale green patches on the upper surface of the leaves, which are gradually turning into angular or rectangular yellow dots. The patches are irregular or uneven in appearance and delimited by the leaf veins. The symptoms on young, newly developing leaves are rare. In severe infection, the lesions dilate and coalesce, the tissues turn necrotic and brown, resulting in wilting and death of large areas of the leaf surface, giving a burnt appearance within 4 to 10 days from the first symptoms (**Lebeda and Cohen, 2011 and Shoukry et al., 2021**). Unfortunately, this disease is spreading rapidly and is difficult to be controlled (**Shoukry et al., 2021**).

For maximizing productivity and reducing agricultural waste and production losses caused by diseases, the methods in the precision agriculture (PA) sector need to be adopted. Many researchers have begun to identify crop diseases based on visual images. Using integrated disease management strategies to reduce dependence on fungicides. Different computer visions have been adopted such as support vector machines (SVM), artificial neural networks (ANN), and convolutional neural network (CNN) for automated detection and classification of cucumber crop diseases (**Pujari et al., 2016; Jiang et al., 2020, Mia et al., 2021**).

This paper aims to establish an integrated disease detection system for cucumber mildew diseases. The objectives of the current study were to: 1) Integrated machine learning with RGB camera for automatic detecting cucumber disease. 2) Design a graphical user interface

(GUI) that is easy to use by farmers to help with the feedback of real life applications for further development for both the detection software and the GUI in use.

2. MATERIALS AND METHODS

2.1 Study Area

The experimental work conducted in the experimental greenhouse of faculty of agriculture, Suez Canal University (Latitude 30.62°N, Longitude 32.26°E), and another greenhouse in Ismailia – kilo eleven, Ismailia, Egypt (Latitude 30.66°N, Longitude 32.22°E). The images were captured during the 1st of January to the 30th of March in 2020, and during the 1st of January to the 30th of March in 2021 to have a wide range of images with different environments.

2.2 Samples Collection and Tools Used

At the start of the project, images of healthy leaves were captured using two different cameras - GoPro HERO8 Black with 12 megapixels, 4000×3000 pixels and iPhone XS with 12 megapixels 3024×4032 pixels. Images of the plant leaf were obtained at the beginning of the work as it was still not infected. In the middle of February, downy mildew infections were observed by Professor Heba Abdelnaby and the symptoms appeared as yellow dots with irregular shapes that were delimited by leaf veins on the upper surface of old leaves. Later, in the end of March, powdery mildew symptoms emerged as circular, white powdery patches on the upper and lower surfaces of the infected leaves. The graphical user interface (GUI) was developed using a Dell Inspiron 5520 i7-3632QM-16GB-Win10 device, and the code was written in MATLAB R2021b. To remove outliers and reduce data, the HIS-PP (Hyperspectral Imaging-Based Plant Phenotyping) standalone software platform developed by **ElManawy *et al.* (2022)** was used.

2.3 Steps of Image Processing

Crop disease recognition research with image processing technology began in the 1980s and 1990s (**Ying *et al.*, 2009**). Digital image processing is a process where the input is an image, and the output may be an image, or the details associated with that image. The images for healthy and infected leaf by downy and powdery mildew of cucumber captured using different cameras for use it in building the model. There are five main stages were used such as image acquisition, image pre-processing, feature extraction, post-processing, and classification. **Figure 1** show the overall steps for building the model using an algorithm developed with MATLAB program.

2.3.1 Image Acquisition

In this study, the target was to detect downy and powdery mildew in cucumber leaves. So, different cameras were used to capture different images. These cameras capture RGB images standing for three-color channels R: Red, G: Green and B: Blue. The images captured at a different time in the day from 10 am to 3 pm in a different illumination with no zoom or flash used. The images were acquired using high resolution camera with holding it horizontally with a suitable distance about 25 – 40 centimeters to have a wide range of vision towards the leaves. The healthy leaves carefully captured to be free from any diseases, bruise, blemishes, or any other visible damages. To build up a database of these diseases, each sample (healthy

or diseased) captured and saved in the computer to develop the detection and classification algorithms with MATLAB (MathWorks, MA, USA).

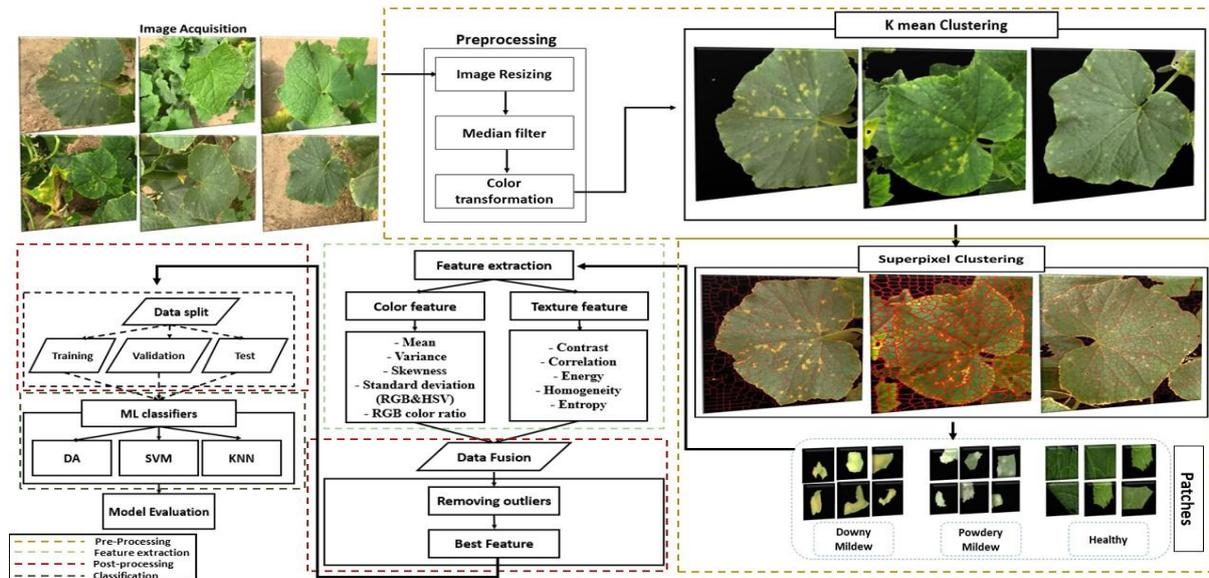


Fig. 1. Flowchart of overall proposed system architecture

2.3.2 Image Pre-Processing

The goal of pre-processing is to improve image data by suppressing unwanted distortions or enhancing visual properties that are important for subsequent processing and analysis. In this step, the captured image is at the lowest level of abstraction. So, the aim of pre-processing is to obtain the interested region and to improve image data from noise that added during the acquisition process or enhances image for increasing feature visibility for other processing phases.

2.3.2.1 Resizing

To simplify processing and shorten processing time, image resolution should be kept to a minimum. However, it can affect the system's overall performance and the accuracy of the disease severity estimation. To ensure a high accuracy of disease detection in a reasonable length of time, a good balance of computing time and accuracy should be achieved. Therefore, different resolution scales were tested, by running the background removal process for each image with multiple resolutions. The computational time and accuracy would be compared to find the appropriate resolution scale to use (**Table 1**). The resolution for iPhone XS and GoPro HERO8 has been reduced to 1000x750 pixels to unify the image resolutions as a first step in improving and speeding up subsequent processing.

Table 1: The image scale for each camera and the time consuming for removing the background for a single image.

Camera Type	Original Resolution (Width x Height)	Quarter Resolution (Width x Height)	Background subtraction before resizing	Background subtraction After resizing
GoPro HERO8	4000×3000	1000×750	40 – 55 second.	7- 9 second.
iPhone XS	4032×3024			

For image resizing, the Bicubic interpolation method was used because it produces clear images, which are smoother and have less interpolation distortion (Han, 2013 and Titus and Geroge, 2013).

2.3.2.2 Image Smoothing

Vector median filter is applied to remove noise and enhance the unclear image parts that caused by collective device, environment and so on. Blurry edge of disease spot and spots of disease leaf always occurs in color image of cucumber disease. So, a 3×3 median filter is used to enhance cucumber images, stress some useful information, and get rid of weakened information.

2.3.2.3 Image Segmentation (Background subtraction)

The component of cucumber disease's image is a bit complicated, as the image contains a leaf as an object with a background texture like soil, all in the same image, also the leaves have disease spots and normal parts. So, segmentation is a two-layer problem in essence. Therefore, there are two working processes. Namely removing the unwanted parts of the image, which is the background. Then splitting the leaf into parts, where disease spots with light or deep color are being arranged on leaf disorderly. The classification method used is called K- means clustering.

First, the image converted to two types of color space such as CIE Lab color space and HSV color space. This conversion enables us to quantify the visual difference present in the RGB image.

Second, a k-means clustering method applied. K-means clustering attempts to divide a dataset into K clusters with each data point belonging to the cluster with the closest mean, which serves as the centroid of the cluster. Given a data matrix $X = [x_1, \dots, x_n]$ and k initial clustering centroids, k-means clustering targets to partition each x_i into k clusters, in which x_i belongs to its nearest cluster according to the squared Euclidian distance. For dataset $X = [x_1, \dots, x_n]$, the main steps are as follows:

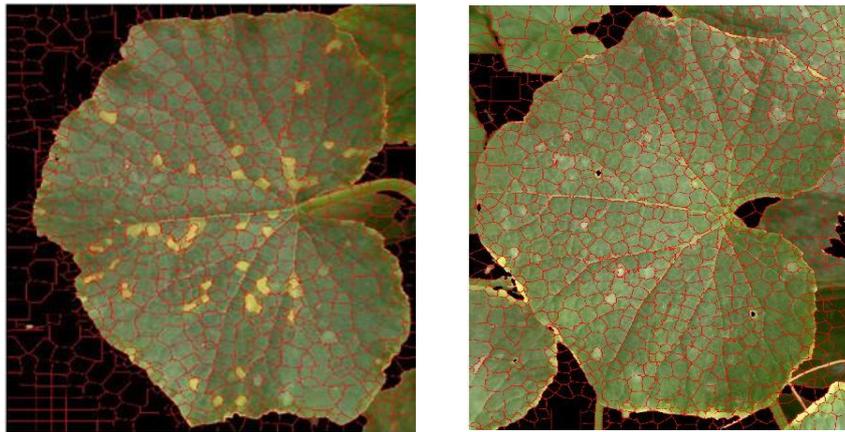
1. Choose k initial cluster centers.
2. Compute point-to-cluster-centroid distances of all observations to each centroid.
3. Assign each observation to the cluster with the closest centroid.
4. Compute means of the observations in each cluster to obtain k new centroid locations.
5. Repeat steps 2 through 4 for 10 times until there is no change in the cluster assignments.

The value k represents the number of clusters, in which each cluster has a defined centroid. The K-means clustering was used to partition the leaf image into two clusters in which the first one includes the entire leaf (foreground) and second cluster include the background. K-means uses squared Euclidean distances between the objects and the corresponding cluster depending on each channel of LAB color space and HSV color space to find out the best color channel in each color space give the best result. Then ten replications of the above steps were used until reaching the best clusters assignments.

2.3.2.4 Superpixel Clustering (Building the Dataset)

Superpixel is a pre-processing method to group the neighbor pixels with the same or similar features in an image into perceptually meaningful homogeneous regions. In diseased leaf

segmentation, a superpixel can be used to reduce the complexity of images from hundreds of thousands of pixels to just a few hundred superpixel. As shown in **Fig. 2** that the leaf image is clustered into several small irregular superpixel regions, and the original pixels in a single superpixel look the same. Therefore, after background removal, a simple linear iterative (SLIC) superpixel clustering is used to extract group the region of interest for healthy, downy, and powdery mildew to save them as a dataset for each class (healthy, downy mildew and powdery mildew). The number of images saved for healthy and infected regions of cucumber leaves, that was used in the experiment, were 345,345,345 images for normal, downy mildew and powdery mildew respectively. All superpixel images were re-sized to 224 x 224 pixels, then color and texture features were extracted.

**1****2****Fig. 2. Images clustered by Superpixel.****(1) For downy mildew. (2) For powdery mildew**

2.3.3 Features Extraction

The aim of this process is to find and extract features that can be used to determine crop diseases. In image processing, image features usually include color, shape, and texture features. The shape features were not used a lot for image analysis as it depended on the distance between CCD camera and the leaves (**Vakilian and Massah, 2013**). Therefore, the color and texture features of the healthy and diseased leaf images are extracted to obtain the unique features, which represent these images. The color features such as mean, variance, skewness, standard deviation, and color ratio are calculated for each channel of RGB color space and HSV color space for all the healthy and diseased parts. The gray level co-occurrence matrix (GLCM) method is adopted to extract the texture feature. The different textural features such as energy, entropy, contrast, homogeneity, and correlation are computed for all the healthy and diseased parts.

2.3.4 Post Processing

After completing the previous processing, we have data that includes all the color and texture features of the extracted superpixel images, so post-processing is used to remove the outliers and obtain the best features to use in the classification process. First, the data is reduced as outliers are removed using HIS-PP. Then the fifty-seven features that are calculated for all the super pixel images are analyzed with HIS-PP by using SFS-forward Technique “Sequential

forward Selection”. SFS-Forward is a wrapper method for finding the optimal feature subset using a greedy search algorithm, which does not assume the full search and evaluating the relevance of every possible feature subset. SFS-Forward starts with an empty subset and adds features through sequential iteration to the subset to check the effect of the features based on the performance of a predefined classifier (Moghimi *et al.*, 2018).

2.3.5 Classification

It is the final stage in disease detection. It is an identifying rule according to the selected features and assigning each disease to any one of the predetermined classes. Data will run with the MATLAB® Classification Toolbox version 2021b (MathWorks, Natick, MA, USA) for training, validating, and testing data with discriminant analysis, support vector machine, and K-nearest neighbors classifiers.

The steps applied in our study are as follows:

- (i) Dataset will be formed based on the parameters included.
- (ii) Use three main models, such as DA, SVM Kernels and KNN with different hyperparameters applied.
- (iii) The cross validations will be done with hold out validation with 20 % and test with 10% of the dataset.
- (iv) Accuracy rates will be calculated and compared with one another through these classification models.
- (v) The confusion matrix of the resulting best models for each classifier has been calculated.
- (vi) The best model will save in .mat file to use in building the Graphical User Interface.

3. RESULTS AND DISCUSSION

This section represents the development of our software, showcasing the performance and efficiency that we have achieved through our approach to disease detection.

3.1 Comparing Color Features

To show the best color space to have a more accurate outcome and toned disease detection. The image converted into two types of color space. CIELAB color space and HSV color space. The K-means clustering used to partition the leaf image into two clusters in which the first one includes the entire leaf (foreground) and the second one includes the background. In **Figure 3**, we show the result of using color spaces such as HSV and LAB by analyzing each channel of these color space separately, or by merging different channels within each color space. We have determined that the A channel of the LAB color space is particularly effective in subtracting objects of interest based on theoretical observation. After running forty images and take its average to choose the right centroids for background subtraction. The initial centroid was set to 105 for foreground part and 135 for background part, enabling us to automatically segment the other images.

As shown in **Fig. 4** some errors may occur during the background subtraction process due to the spread of infection and the death of plant tissues, which can cause the color of the affected leaf spots to become indistinguishable from the color of the soil. This can lead to a decrease in accuracy, but the main purpose of the program is to detect disease in their early stages and prevent them from reaching this stage of infection.

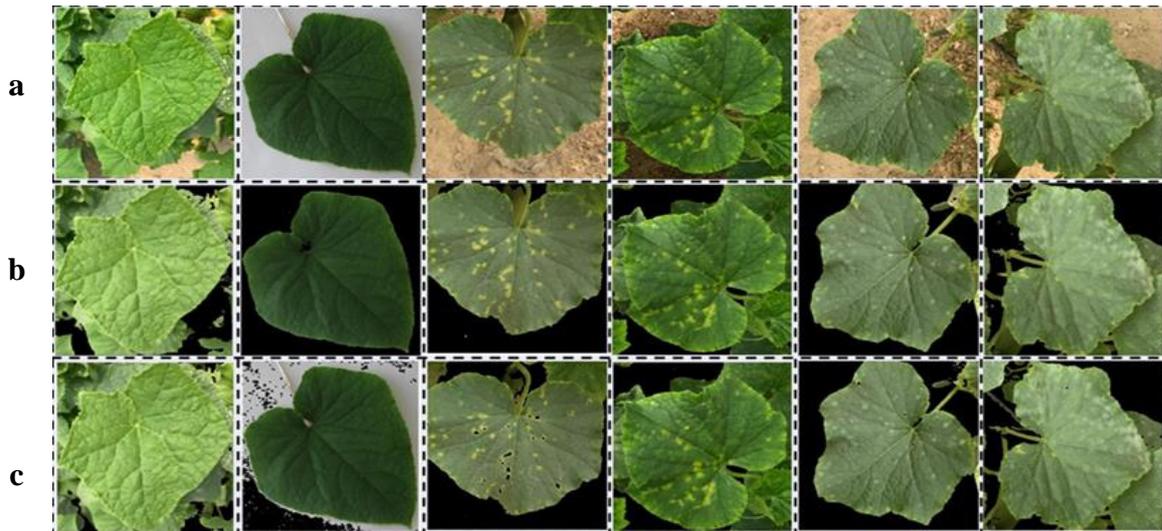


Fig. 3. Some representative images of K-mean Segmentation. (a) raw image. (b) LAB color space. (c) HSV color space.

Figure 5 reveals the accumulation of water droplets on the leaf surface, which can make false disease symptoms and lead to misdiagnosis.

Images were converted to LAB, HSV, YCbCr color space. The K-means clustering was used to partition the leaf image into five clusters for extracting the disease part. K-means used squared Euclidean distances between the objects and the corresponding cluster. **Figure 6** shows that the YCb channels of the YCbCr color space gives the best result in subtracting the objects of interest which will be used in the program for user interface.



Fig. 4. Some images with high disease severity. (a) raw image. (b) leaf after segmentation.



Fig. 5. The raw image after spraying (a), and applying the k-mean clustering (b)

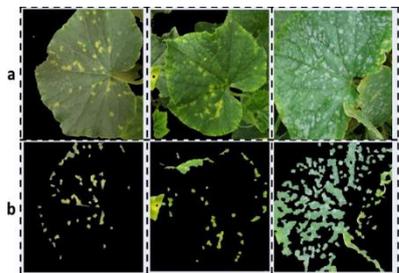


Fig. 6. Images of diseased leaf (a), and segmented images (b) for testing process.

3.2 Superpixel Clustering

Figure 7 shows different colors and shapes of different superpixel images of infected cucumber's leaves which were used in building the model as the color, size and distribution of an infected leaf symptom are not always the same. In addition, the symptoms of an infected leaf change as the plant disease develops.

3.3 Post Processing

In order to remove outliers and feature data reduction, **Table 2** shows the data obtained from superpixel was reduced as the outliers were removed using HIS-PP software.

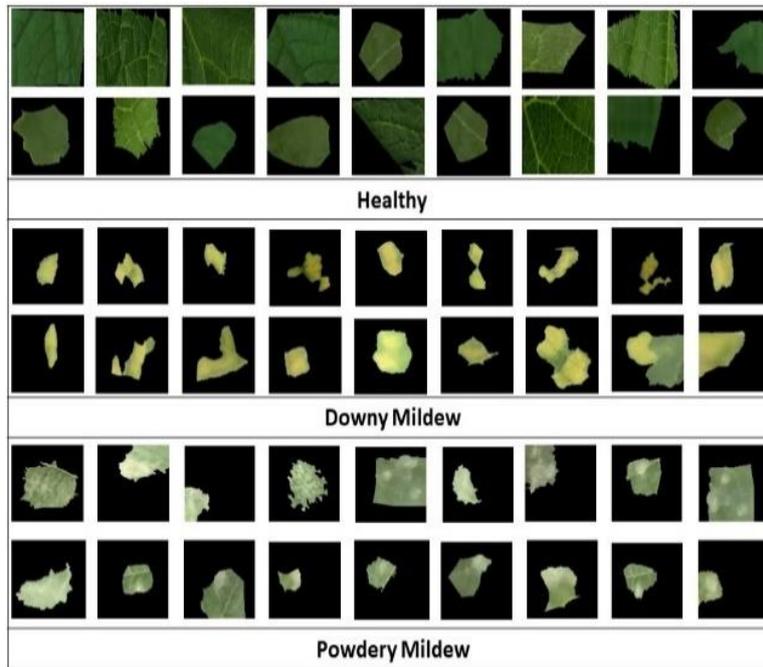


Fig. 7. Examples of data set used in building the model.

Table 2: Number of images extracted from superpixel clustering after removing outliers using HIS-PP.

	Type of image	No. of images
Superpixel of Cucumber leaves	Normal	306
	Downy Mildew	313
	Powdery Mildew	312

The color features (Mean, Variance, Skewness, Standard deviation, and the color ratio) for each color space (R, G, B/H, S, V) were extracted, along with the color ratio for (R, G, B) and texture features (texture, correlation, energy, homogeneity, and entropy) for each color space (R, G, B/H, S, V). In total, there were around fifty-seven features. The analysis was performed using the HIS-PP, and a technique called SFS was used to select the most appropriate features for use in the classification. The results of the analysis showed that five features gave the best performance: skewness S, V entropy, R standard deviation, Skewness G, G standard deviation.

3.4 Cucumber Diseases Classification

3.4.1 Data Coding

The extracted data were run with the MATLAB® Classification Toolbox version 2021b (MathWorks, Natick, MA, USA) for train models, validate and test data with Discriminant analysis, Support Vector Machine and K-Nearest Neighbors. The validation process was done based on 20% of the data and 10% of the data was used for testing as shown in Table 3.

Table 3: Number of final data images used in the training and validation process.

Type	Training Set	Validation Set	Testing Set	Total
Healthy	214	61	31	306
Downy Mildew	220	61	32	313
Powdery Mildew	219	61	32	312
Total	653	183	95	931

3.4.2 Performance Evaluation of Different Classifiers

Table 4 shows the confusion matrix for the best hyperparameter of the resulting discriminant analysis model, SVM model and KNN model. It also presents the performance of the three detection methods in detecting cucumber leaves, which presents the number of images to be classified (input) and the number of images that have been classified as healthy, downy, and powdery mildew disease and the percentage of accuracy, precision, sensitivity, and f-measure. It is clear to see that the accuracy of quadratic discriminant, fine gaussian SVM and fine KNN method in detecting infected leaves are 92.8, 96 and 95.8 %, respectively, which is suitable and robust enough for practical applications. Based on the accuracy and f-measure, the detection by fine gaussian SVM is better than the other two methods of quadratic discriminant, and fine KNN in distinguishing the healthy leaf resulting in 100% accuracy. The fine KNN is better than the other two methods of quadratic discriminant, and fine gaussian SVM in distinguishing the downy and powdery diseased leaf resulting in 95.22 and 96.9 accuracy respectively.

Table 4: The Confusion Matrix for best hyperparameter in each DA, SVM and KNN classifier.

Different Hyperparameters	Predicted Class / True Class		Input images (%)	Downy (%)	Healthy (%)	Powdery (%)	Accuracy (%)	Precision (%)	Sensitivity (%)	F-Measure (%)
	Quadratic Discriminant	Downy	32	28		4	90.4	93.3	87.5	90.3
Healthy		31		30	1	98.4	100	96.8	98.4	
Powdery		32	2		30	89.7	85.7	93.8	89.6	
Sum		95				92.8	93	92.7	92.8	
Fine Gaussian SVM	Downy	32	28		4	93.75	100	87.5	93.3	
	Healthy	31		31		100	100	100	100	
	Powdery	32			32	94.45	90	100	94.7	
	Sum	95				96	96.3	95.8	96	
Fine KNN	Downy	32	30	2		95.3	96.7	93.8	95.2	
	Healthy	31		30	1	95.3	93.9	96.8	95.3	
	Powdery	32	1		31	96.9	96.9	96.9	96.9	
	Sum	95				95.8	95.8	95.8	95.8	

3.4.3 Comparison

Results show that the three classification methods have different efficacy in detecting the disease with an accuracy of 92.8%, 96% and 95.8%, respectively. Results from this study were more significant than those from (Zhou *et al.*, 2015), which proposed an SVM classifier to detect downy mildew cucumber disease which only showed 90% accuracy. Our results were also higher than the results from (Zhang *et al.*, 2017), which proposed an SVM classifier to recognize five kinds of cucumber diseases: Scab Angular, Powdery Mildew, Downy Mildew, Anthracnose and Scab. While the obtained total accuracy was 91.48%, the accuracy of powdery mildew and downy mildew was equal to 90.64, and 92.85% separately. As for the result, the detection of downy mildew was close to the SVM that we proposed. But it was less in the detection of powdery mildew than the proposed model. The researcher

employed a novel technique that combines super pixels with expectation maximization (EM). Utilizing this technique in our experiment led to a classification result that was superior when superpixel were combined with k-means.

Due the little amount of data utilized in this study, the overall performance of SVM and kNN classifiers in detecting powdery mildew disease offered about similar accuracy with 96%, which was higher than the results from (Mahmud *et al.*, 2018) where the accuracy was 91.86% and 83.20%. Ma *et al.*, (2018) proposed SVM, RF and DCNN classifiers to recognition four cucumber diseases, anthracnose, downy mildew, powdery mildew, and target leaf spot. The accuracy of the SVM and RF, was 81.9% and 84.8% respectively, while DCNN achieved good recognition results, with an accuracy of 93.4%. The accuracy of the SVM classifier is less than what we got in this experiment. The accuracy of powdery mildew and downy mildew for the SVM classifier were different, as it gave a close result to the proposed SVM in detecting downy mildew disease at 94.6%, while the F-measure of detecting powdery mildew disease was at 80% which is less than the proposed SVM.

Khan *et al.*, (2022) in his study provided an automated framework using deep learning and best feature selection for the recognition of cucumber leaf diseases. His framework's accuracy was higher at 98.4%.

3.5 Graphical User Interface (GUI)

A Graphical User Interface (GUI) is designed and implemented using MATLAB for computer vision programming.

Figure 8 shows the GUI elements, as it consists of the following steps which explains how the program work: -

- 1) Insert Image: to insert the image as the image show in the axis 1 (No 5), then crop the region of interest.
- 2) Segmentation: to classify the image into 5 clusters where another box containing 5 images is shown. From which you choose the part desired, and the chosen image shows in the axis 2 (No 6)
- 3) Feature Extraction: to extract the characteristic of the region of interest which are (skewness S, V entropy, R standard deviation, Skewness G, G standard deviation) and then the program compares this data with the previously extracted model which is fine gaussian SVM.
- 4) Results: to show the classification type as a text. No (7) Whether the region of interest is healthy, downy, or powdery mildew. Also, a small information about the disease and the treatment for it based on the recommendation of the agriculture ministry.

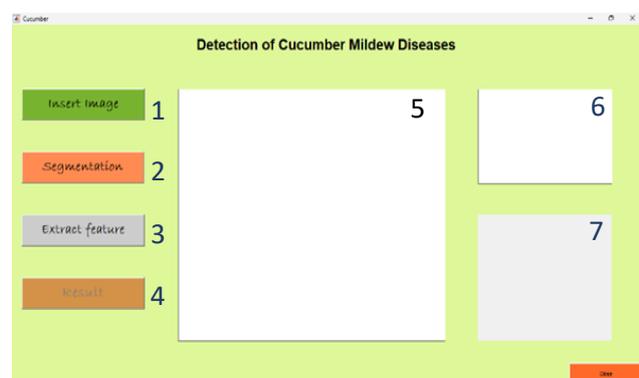


Fig. 8. The Graphical User Interface.

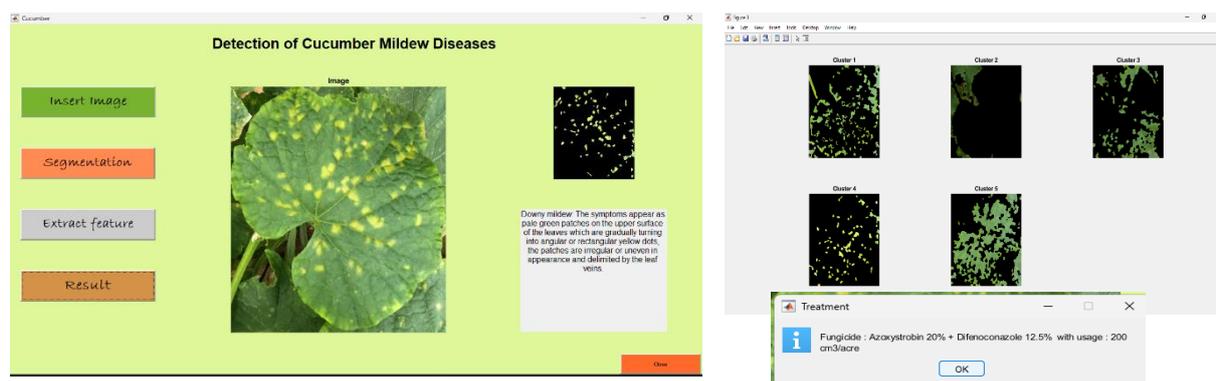


Fig. 9. The GUI show downy mildew detection. (a) The GUI. (b) The 5 clusters. (c) Treatment.

4. CONCLUSION AND RECOMMENDATIONS

The study's first objective was to identify the different tools and algorithms to develop an integrated image processing technique. To achieve this, the images of cucumber leaves were taken using GoPro HERO8 and iPhone XS and used four distinct image processing techniques to detect the disease. The image processing techniques included scaling down the image to 1000x750 pixels by using bicubic interpolation to unify the image resolutions, applying a vector median filter to remove noise, using the K-means clustering algorithm based on A channel in LAB color space to remove the background, and using the K-means clustering algorithm based on Ycb channels in Ycbr color space to extract the infected patches. Superpixel improved the accuracy of detecting the region of interest. Then, HIS-PP software was used to remove outliers and select the appropriate features for classification during the training and testing process. The study's results found that five features (skewness *S* in (HSV channel), *V* entropy in (HSV channel), *R* standard deviation in (RGB channel), Skewness *G* in (RGB channel), *G* standard deviation in (RGB channel), gave the best performance.

The study's second objective was to evaluate the performance of defect detection algorithms in detecting cucumber mildew disease. A main file was created, containing 931 rows of each sample: healthy, downy, and powdery mildew, with five chosen features shown by six columns for each row. The best accuracy results for the ML technique are fine Gaussian SVM at 96%, followed by fine KNN at 95.8%, while QDA's accuracy is lowered at 92.8%. The study's results showed that the proposed method provides superior accuracy when compared to other recent methods. Then the GUI was developed using MATLAB 2021b.

This study has some limitations, such as the proposed model being only intended for two classifications and not being able to classify multiple diseases at the same time. It is suggested to check the readings on other equipment such as temperature and humidity devices to ensure a more accurate result. Future research should focus on enhancing the model's ability to segment extremely small targets, classify numerous diseases, and broaden the usage of the superpixel clusters method for image segmentation. Thermal images can also be introduced to increase the quality of performance and confirm the type of disease. Finally, a mobile system to automate the detecting procedure will be looked at in the future.

ACKNOWLEDGEMENT

The authors thank Dr. Ahmed I. ElManawy for his guidance and technical assistance and Eng. Yahya Ejeez for his assistance during obtaining the data.

5. REFERENCES

- Bandamaravuri, K. B., Nayak, A. K., Bandamaravuri, A. S., and Samad, A. (2020).** "Simultaneous detection of downy mildew and powdery mildew pathogens on *Cucumis sativus* and other cucurbits using duplex-qPCR and HRM analysis". *AMB Express*, 10(1).
- Daunde, A. T., Baghele, R. D., and Khandare, V. S. (2020).** "Management of Prevalent Diseases of Cucumber (*Cucumis sativus*) through Integrated Approach". *International Journal of Current Microbiology and Applied Sciences*, 9(7):3022–3028.
- ElManawy, A. I., Sun, D., Abdalla, A., Zhu, Y., and Cen, H. (2022).** "HSI-PP: A flexible open-source software for hyperspectral imaging-based plant phenotyping". *Computers and Electronics in Agriculture*, 200:107248.
- FAOSTAT. (2020).** "Cucumber, production quantity for Egypt". Available at <<https://www.fao.org/faostat/en/#data/QCL>> (Accessed September 2022).
- Han, D. (2013).** "Comparison of Commonly Used Image Interpolation Methods". *Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering (ICCSEE 2013)*, Hangzhou, China, 1556–1559.
- Jiang, F., Lu, Y., Chen, Y., Cai, D., and Li, G. (2020).** "Image recognition of four rice leaf diseases based on deep learning and support vector machine". *Computers and Electronics in Agriculture*, 179.
- Khan, M. A., Alqahtani, A., Khan, A., Alsubai, S., Binbusayyis, A., Ch, M. M. I., ... and Cha, J. (2022).** "Cucumber Leaf Diseases Recognition Using Multi Level Deep Entropy-ELM Feature Selection". *Applied Sciences (Switzerland)*, 12(2).
- Lebeda, A., and Cohen, Y. (2011).** "Cucurbit downy mildew (*Pseudoperonospora cubensis*)-biology, ecology, epidemiology, host-pathogen interaction and control". *European Journal of Plant Pathology*, 129(2):157–192.
- Lin, K., Gong, L., Huang, Y., Liu, C., and Pan, J. (2019).** "Deep learning-based segmentation and quantification of cucumber powdery mildew using convolutional neural network". *Frontiers in Plant Science*, 10:155.
- Ma, J., Du, K., Zheng, F., Zhang, L., Gong, Z., and Sun, Z. (2018).** "A recognition method for cucumber diseases using leaf symptom images based on deep convolutional neural network". *Computers and Electronics in Agriculture*, 154:18–24.
- Mahmud, M. S., Ki Chang, Y., and Esau, T. (2018).** "Detection of Strawberry Powdery Mildew Disease in Leaf Using Image Texture and Supervised Classifiers". In *Proceedings of the CSBE/SCGAB 2018 Annual Conference*, Guelph, ON, USA, 22-25.
- Mia, M. J., Maria, S. K., Taki, S. S., and Biswas, A. A. (2021).** "Cucumber disease recognition using machine learning and transfer learning". *Bulletin of Electrical Engineering and Informatics*, 10(6):3432–3443.
- Moghimi, A., Yang, C., and Marchetto, P. M. (2018).** "Ensemble Feature Selection for Plant Phenotyping: A Journey from Hyperspectral to Multispectral Imaging". in *IEEE Access*, 6:56870-56884, doi: 10.1109/ACCESS.2018.2872801.

- Mostafa, Y. S., Hashem, M., Alshehri, A. M., Alamri, S., Eid, E. M., Ziedan, E. S. H. E., and Alrumman, S. A. (2021).** "Effective Management of Cucumber Powdery Mildew with Essential Oils". *Agriculture (Switzerland)*, 11(11).
- Ni, L., and Punja, Z. K. (2019).** "Management of Fungal Diseases on Cucumber (*Cucumis sativus* L.) and Tomato (*Solanum lycopersicum* L.) Crops in Greenhouses Using *Bacillus subtilis*". In *Bacilli and Agrobiotechnology: Phytostimulation and Biocontrol*, 2:1–28.
- Ni, L., and Punja, Z. K. (2021).** "Management of powdery mildew on greenhouse cucumber (*Cucumis sativus* L.) plants using biological and chemical approaches". *Canadian Journal of Plant Pathology*, 43(1):35–42.
- Pawar, P., Turkar, V., and Patil, P. (2016).** "Cucumber disease detection using artificial neural network". In *2016 international conference on inventive computation technologies (ICICT)*, Coimbatore, India, 3:1-5, IEEE.
- Pujari, J., Yakkundimath, R., and Byadgi, Abdulmunaf. S. (2016).** "SVM and ANN Based Classification of Plant Diseases Using Feature Reduction Technique". *International Journal of Interactive Multimedia and Artificial Intelligence*, 3(7):6-14.
- Savory, E. A., Granke, L. L., Quesada-Ocampo, L. M., Varbanova, M., Hausbeck, M. K., and Day, B. (2011).** "The cucurbit downy mildew pathogen *Pseudoperonospora cubensis*". *Molecular Plant Pathology*, 12(3):217–226.
- Sharma, A., Katoch, V., and Rana, C. (2016).** "Important Diseases of Cucurbitaceous Crops and Their Management". In *Handbook of Cucurbits: Growth, Cultural Practices, and Physiology (Issue February)*, 301–323.
- Shoukry, M., al Gazar, T., and EL-Sheshtawi, M. (2021).** "Ability of Some Antagonistic Fungi for Controlling Cucumber Downy Mildew Disease Caused by *Pseudoperonospora cubensis*". *Journal of Plant Protection and Pathology*, 12(1):67–69.
- Titus, J., and Geroje, S. (2013).** "A Comparison Study on Different Interpolation Methods Based on Satellite Images". *International Journal of Engineering Research & Technology (IJERT)*, 2(6):82–85.
- Vakilian, K. A., and Massah, J. (2013).** "An artificial neural network approach to identify fungal diseases of cucumber (*Cucumis sativus* L.) plants using digital image processing". *Archives of Phytopathology and Plant Protection*, 46(13):1580–1588.
- Verhaar, M. A. (1998).** "Studies on biological control of powdery mildew in cucumber (*Sphaerotheca fuliginea*) and rose (*S. pannosa*) by means of mycoparasites". *Wageningen University and Research*, 3-5.
- Wspanialy, P., and Moussa, M. (2016).** "Early powdery mildew detection system for application in greenhouse automation". *Computers and Electronics in Agriculture*, 127: 487–494.
- Ying, G., Miao, L., Yuan, Y., and Zelin, H. (2009).** "A study on the method of image pre-processing for recognition of crop diseases". *International Conference on Advanced Computer Control*, Singapore, 2009:202-206. doi: 10.1109/ICACC.2009.10.

- Yousuf, S., and Dar, G. H. (2016).** "Prevalence and Status of Major Fungal Foliar Diseases of Cucumber in Western Himalayan State of Jammu & Kashmir, India". *International Journal of Current Microbiology and Applied Sciences*, 5(7):550–557.
- Zhang, S., Zhu, Y., You, Z., and Wu, X. (2017).** "Fusion of superpixel, expectation maximization and PHOG for recognizing cucumber diseases". *Computers and Electronics in Agriculture*, 140:338–347.
- Zhou, B., Xu, J., Zhao, J., Li, A., and Xia, Q. (2015).** "Research on Cucumber Downy Mildew Detection System based on SVM Classification Algorithm". In 3rd international conference on material, mechanical and manufacturing engineering (IC3ME 2015), Guangzhou, China, 1681-1684. Atlantis Press.

الكشف عن أمراض البياض في الخيار باستخدام تقنية معالجة الصور

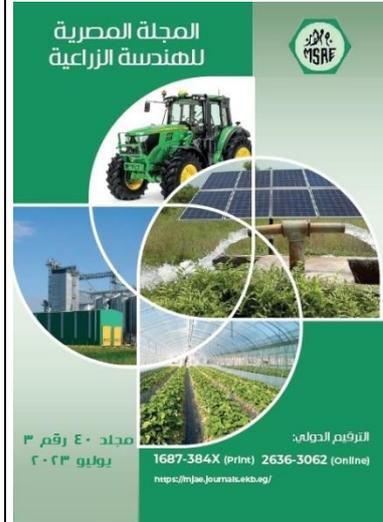
آلاء علاء الدين ابراهيم الدسوقي^١ ، هبه محمد عبدالنبي^٢ ، السيد عويس عمران^٣

- ^١ إخصائي زراعي - الإدارة العامة للحجر الزراعي ببورسعيد وشمال سيناء - بورسعيد - مصر.
^٢ استاذ امراض النبات - قسم امراض النبات - كلية الزراعة - جامعة قناة السويس - الاسماعيلية - مصر.
^٣ استاذ الاراضي والمياه والإستشعار عن بعد - قسم الاراضي والمياه - كلية الزراعة - جامعة قناة السويس - الاسماعيلية - مصر.

الملخص العربي

تساهم زراعة الخيار في الاقتصاد المصري كونه أحد أكثر الخضروات شيوعاً وأهمية، ولأنه يصاب بالعديد من الأمراض فإن ذلك يشكل خطراً على الإنتاج، كما أن الطرق التقليدية في الكشف تكلفنا الوقت والمال، لذا تقدم هذه الدراسة معرفة بآلية الكشف المبكر عن أمراض البياض الدقيقي والزرغبي في الخيار باستخدام التعلم الآلي من خلال تقييم أداء خوارزميات الكشف المختلفة وتصميم برنامج يعكس مهامها، وتحتوي هذه الدراسة على خمس خطوات، بداية من عملية الحصول على الصورة والمعالجة المبدئية واستخراج الصفات ومعالجات أخرى نهائية إلى تصنيف الصور.

وقد تم استخدام ٩٣١ صورة واستخراج افضل الصفات لتدريب وتقييم الأساليب المقترحة وكشفت النتائج عن خمس خواص اعطوا أفضل أداء. ثم تم إجراء مقارنة بين النماذج التقليدية المختلفة للتعلم الآلي مثل التحليل المميز ونظام الدعم الآلي و الجيران الأقرب، وللحصول علي دقة أكبر تم استخدام نهج (Fine Gaussian) في مصنف (نظام الدعم الآلي) الذي اعطي أعلى كفاءة في الأداء بقيمة تساوي ٩٦٪ والتي كانت تقاربها في الأداء نهج (Fine) في مصنف (الجيران الأقرب) بقيمة تساوي ٨٠.٩٥٪، بينما أعطي نهج (Quadratic) في مصنف التحليل المميز أداءً اقل بقيمة تساوي ٨٠.٩٢٪ ومن خلال مقارنة الطريقة المقترحة بالطرق الحديثة الأخرى في المراجع السابقة تبين أنها تعطي دقة أفضل، بالإضافة إلى أن الطريقة المقترحة لها تطبيق عملي يتيح لنا الكشف التلقائي عن مرض البياض الدقيقي والزرغبي من خلال أجهزة الكمبيوتر المختلفة، وعلى الرغم من أن التركيز الحالي على أوراق الخيار، لكننا نرى أنها ستكون مفيدة في الكشف المبكر على الأمراض النباتية الأخرى وتتبع تطور الأمراض وذلك مع الدراسات المستقبلية.



© المجلة المصرية للهندسة الزراعية

الكلمات المفتاحية:

الخيار؛ البياض الزرغبي؛ البياض الدقيقي؛ التعلم الآلي؛ معالجة الصور.