Mustansiriyah Journal of Pure and Applied Sciences, Print ISSN: 2957-9910. Online ISSN: 2957-9929, Vol. 3, No. 1, 2025, Pages 64-81
DOI: 10.47831/mjpas.v3i1.121

MJPAS

MUSTANSIRIYAH JOURNAL OF PURE AND APPLIED SCIENCES

Journal homepage*:*
https://mjpas.uomustansiriyah.edu.iq/index.php/mjpas

RESEARCH ARTICLE ***- Computer Science***

Towards smart farming approach for plants disease detection and classification by using HoG feature extraction and PNN algorithm

Juliet Kadum [1*], Jamal Mustafa Al-Tuwaijar [2], Amaal Kadum [3] , Sadam Hameed Rasheed[4]

[1,2] Computer Science Department, College of Science/University of Diyala, Diyala, Iraq

[3] Computer Science Department, College of Education/ University of Mustansiriyah, Baghdad, Iraq

[4] General Directorate for Education of Diyala, Diyala, Iraq

[*] Corresponding author E-mail: julietkadum@uodiyala.edu.iq

Article Info.

Article history:

Received
09 February 2024

Accepted
28 April 2024

Publishing
30 January 2025

Abstract

The advancement of computerized automated diagnostic technologies ensures that health plants are more effective by detecting diseases early. The damage caused by them can be minimized. Recently, many scientists have been trying to use methods and algorithms in image processing and extracting features based on color, texture, and shape and employing them in building smart systems capable of detecting plant diseases in their early stages. In this paper, proposed a smart farming aims to automated remove shadow and segmentation the pepper fruit diseases region, then extracted strong features from diseases region, and finally detected and diagnosis pepper fruit disease using PNN algorithm. In addition, in this work we used, (Diyala Pepper), which relies on a dataset of pepper plants named Diyala Pepper. The suggested system's implementation was done on a real dataset of pepper fruits. Experimental results of the proposed method showed 100% accuracy in the training phase and 81.82% in the phase of testing.

The official journal published by the College of Education at Mustansiriya University

Keywords*:* 2D Log Chromaticity Image, PNN algorithm, HoG Feature Extraction, Evaluation Measures.

1. Introduction

Agricultural practices are fundamental to human civilization. Not only does it provide food, but it also benefits the economy. When plant grown, several diseases can infect plant leaves or crops. The illnesses put a stop to the development of those species. Additional harm to the plants might be prevented if the illnesses could be identified and classified early and precisely. There are now major issues with the diagnosis and categorization of these disorders. The usual method that farmers use to forecast and categorize plant leaf diseases is often tedious and inaccurate. Manually trying to forecast the kinds of diseases could lead to problems. Destroying crop plants due to a lack of rapid disease detection and classification could lead to a drastic drop in crop yields. Computerized image processing techniques can help farmers enhance output while decreasing losses [1] , [2]. Automated identification and tracking of diseases may promote selective and timely management of diseases, contributing to enhanced yields, enhanced crop quality and a decrease in the use of pesticides. Lower processing prices, less sensitivity to chemicals for field staff, regulators, and improved biodiversity are further benefits. For each disease and seed, the signs are special, and each plant can suffer from several threats. A focused integrated framework and algorithms for the detection of diseases are therefore required [3].Algorithms for machine learning enable one to interpret and evaluate complex data [4]. To categorize and identify illnesses in agricultural commodities, common tactics have included techniques centered on the field of machine learning, recognition of images, and classification. Different image recognition approaches accompanied by different labeling procedures have been used in the current methods for disease detection [5]. Pepper is a major crop of peppers planted in the world's tropics and subtropics areas. This is attributed to the strong biological importance of fruits (high dry matter quality, minerals, essential oils, vitamin B and C complexes, and carotenoids, etc.) and their diverse forms of usage in various countries the culinary and food sectors [6]. Pepper diseases induced by biological (infectious) and biological (non-infectious) pathogens interfere with the development of pepper. Fungi, bugs, nematodes, and viruses are biotic agents for pepper disease. Abiotic illnesses are among there are some adverse cultural or environmental elements, for example excessive sunlight, dietary deprivation, and hot temperatures.

Diseases infect pepper plants in each of their forms, such as the leaves, branches, roots, berries, and small seedlings. Several of symptoms are triggered by fungi and bacteria [7]. Black Fruit infection contributed to the emergence of this disease is the pepper fruit disease environmental factors, for example a drop in temperature [8]. Fruit rot symptoms initially appeared as soft spots that darkened to become black form. Usually, a fruit's calyx end will develop a lesion, which will then extend to the fruit's sides [9]. Insect Damage Peppers are invaded by several insect forms, among the most significant pests are aphids, flea beetles, cutworms, hornworms, pepper maggot, and pepper weevil [8].

This research aims to identify diseases based on the shadow replicas of the data, solve the problems with the dataset such as replication data, and determine the accuracy of the classification the pepper plant diseases. It uses a new technique to pick the best images, feature extraction using HoG, and standardized the features using Z-Score. This method will motivate clever farming by farmers and enable them to take the necessary preventive and corrective measures for pepper crops, thereby taking a step toward making better yield decisions.

The structure of the article is as follows: Section 2 describes the work linked (related work).The related techniques and resources of this work are presented in Section 3. The suggested scheme is illustrated in Section 4 and Section 5 discusses the proposed process of simulation and the experimental findings and discussion. Section 6 completes the proposed work.

2. Related works

Processing by Hlaing, et al, a set of surface attributes is suggested for the organization of various ailments of foliage of plants. Photographs are obtained from various cell sensors. A “SIFT” Scale-invariant transform feature is used as a form part and is symmetric for scaling, hinge, and commotion, and lighting. The SIFT vector is too tough to be organized and it demands a high computational period. 10-fold cross validation is used, in which 10 separate samples are used for research, then randomly selecting an unusual sample for each test run. Although the suggested feature will effectively simulate the SIFT feature and be used to detect plant diseases, it must try to enhance the proposed feature by acknowledging and working with other picture analysis techniques [10]. According to the previous study by Sharif, et al, there is a modern method for the diagnosis and therapy of citrus diseases in fruits. There are two main phases in the recommended approach. The first phase is the identification of spot lesions on citrus plants and fruits. The second phase consists of detecting citrus diseases. This research examined citrus lesion spots utilizing an adaptive, weighted segmentation technique conducted on an improved input image. The characteristic functions are forwarded to the Multi-Class Support Vector Machine (M-SVM) for the final diagnosis of citrus disease. This image recognition program is tested on the Citrus Diseases Database and the Citrus Database. The software learns from a series of pictures and the consumer requires access to a wide collection of photographs in a single area. In the citrus disease picture dataset, the proposed approach performs well and achieves 97% accuracy. In the combined citrus disease image and symptom dataset, the proposed approach performs 95% accuracy, whereas the current approach performs 68% accuracy. In the local symptom dataset, the proposed approach performs 90.4% accuracy on the local dataset [11].

An early detection technique utilizing spectral reflexes in pepper plants has been developed by Kerim Karadag and his colleagues [12]. Reflectance spectra from plants in the field are easier and more cost-effective. Inputs have been described using K-Nearest Neighbor, Naive Bayes, and Artificial Neural Networks. The overall performance rate for the individual algorithms was estimated at 100 percent for KNN, 97.5 percent for ANN and 90 percent for NB. If KNN, Ann and NB will render the same, these rates can be measured as 100%, 88,125% and 82% respectively.

A. Adeel, et al , invested into creating a technique for detecting and segmenting of grape leaf diseases to promote the safe and reliable control of grape leaf diseases [13]. A CNN review for classifying plant diseases was presented by J.Lu, et al, they assessed the key issues with CNN's classification of plant diseases and their remedies and the relevance of DL criteria. They found that further investigation using more complicated datasets was needed to produce a more desirable outcome [14]. Our solution will possibly consist of a four stage phase. In the first stage, a local contrast haze reduction (LCHR) strategy is applied that improves evident contrast at the presentation of the object of interest. In stage two, LAB color transition has initial picture thresholding of the color channel and the better channel is chosen depending on the pixel details that will later be used in stage three of the LAB color conversion. Using a canonical correlation analysis (CCA), we may analyze the color, shape, and geometric characteristics of an entity. The proposed approach has achieved an overall rate of segmentation accuracy 90 percent and an accuracy rate of more than 92 percent, which is better than the other subsequent methods.

Many issues arose in previous research, such differences in symptoms of a specific disease among various plants, a problem where several plants have the same symptoms, a problem where symptoms for a particular disease vary in different places, and many different automated disease detection systems.

3. Research methods and material

This section will be reviewed the methods and techniques that used in this research.

3.1 2D log chromaticity image

A method of expanding the recovery of a grey scale picture and that is without shadow effects to provide recovery of color details is given. See the following information in the 2D Log Chromaticity Image:

3.1.1 Production of the Color Input Image's 2D Log Chromaticity

The RGB image input's R, G, and B channels in equations (1, 2, 3) are first normalized by the following formula [12]:

$$R' = R/(R+G+B) \tag{1}$$

And;

$$G' = G/(R+G+B) \tag{2}$$

Finally;

$$B' = B/(R+G+B) \tag{3}$$

R', G', and B' are the invariant images of the three color bands, where R, G, and B are the input picture color channels. After that, in equation (4) must find value of (p) which is represents the log of each normalized channels Since ρ is not 2D, we have to find a projector U that can project ρ on to the 2D plane.

$$\chi = U\rho \tag{4}$$

3.1.2 1D illumination invariant picture receiving

The strongest invariant representation of 1D (Iinv) illumination is produced by projecting the representations of 2D log chromaticity, $x1$ and $x2$, to the projection direction q [15].

$$\text{Iinv}= x1\,cos(\theta) + x2\,sin(\theta) \quad (5)$$

It is likely to discover the finest projection direction θ, so to find θ an entropy minimization method is used. For θ_k in 1...180, the projection angle θ is selected to be the value of θ_k that has the least entropy. Project the data into 1D space using equation (6) and calculate the entropy η_k in each case as follows

$$\eta = -\Sigma\,(Pi(\text{ Iinv}) * log\,Pi(\text{ Iinv})) \quad (6)$$

3.1.3 Acquisition of 2D illumination-invariant chromaticity images

This process can be carried out to create a chromaticity image without shadows for better output image display [13]. Firstly, calculate the average of χ1 or χ2, then compute the median for both χ1 and χ2, (m1, m2) and then reconstruct the chromaticity picture using the 1D invariant image, Iinv.

3.2 HoG feature extraction

Dalal and Triggs proposed the attribute identifier HoG, which has gained popularity for recognition via history, facial recognition, and machine learning. It splits the image into tiny square cells, counts the histogram components, and records the histogram unit's value in the image information [16].

3.2.1 Directed Gradient Extraction from Histogram

At this level, gradients are identified from each pixel, and they are employed to create a texture function vector known as an optical histogram of gradients. First locate the horizontal and vertical derivative at pixel (K,L), then compute the magnitude of the gradient[17].

$$G(K,L) = \sqrt{G_K(K,L)^2 + G_L(K,L)^2} \quad (7)$$

The gradient's direction is :

$$\propto 0\ (K,L) = \tan{-1}\left[\frac{G_K(K,L)}{G_L(K,L)}\right] \quad (8)$$

Where $G_K(K,L), G_L(K,L)$ are at pixel (K,L), the derivative along the axes horizontal and vertical, respectively., α0 € [-π/2, π/2] .

3.2.2 The HoG Description

The gradient in the image is the focus of the HoG description. Next, it is separated into blocks of similar duration. The horizontal direction spectrum is quantized into 9 coordinate intervals (bins). To create a feature vector of length 9, the gradient histogram of each cell's pixels is produced, which describes the positions of all of the pixels in the cell. Next, 4 blocks are combined into a superblock and each function vector is around the map. Finally, by scanning a high-density file, The 36 D vectors of all the super blocks are concatenated to create the image's HoG function [17].

3.3 Probabilistic neural network (PNN) classification algorithm

In 1989, Dr. D.F. Speeht was the first to propose the probabilistic neural network. It refers to a network of feed forwards and is a branch of a network of radial bases. It offers the following advantages: quick process of learning, a faster pace of preparation, more detailed classification, reasonable tolerance for defects, and so on. It is a member of a supervised network classifier with a focus on pattern categorization that is based on Bayesian low-risk parameters [18]. Actually, the main drawback of PNNs is the need for a good computer with enough memory and fast processing [19]. Probabilistic neural networks generally have four layers, which are as follows: the input layer, secret layer, competition layer, and convolution layers. The vector feature is sent to the network via the input layer as well, and the quantity represents all sample characteristics together; the hidden level is linked by the reference weight to the input layer, used to measure the similarity between the vector feature input and each mode in the training collection, to acquire the hidden level's output, you feed its distance into the Gaussian function. The Gaussian function's expression is:

$$G(x) = \frac{1}{\sqrt{2\pi}\ n\sigma_p} \sum_{i=1}^{n} \quad exp\ exp\ \left[-(\frac{\|x - x_i\|^2}{2\sigma^2})\right]^{[22]} \qquad (9)$$

Where σ represents the standard variance, n indicates the number of instances of preparation, p is how many input nodes there are, x is the random variable vector, and the vector of i^{th} is x_i [19]. The number of vectors in the input sample equals the number of secret layer neurons. The rivalry layer is in charge of connecting the pieces of each class's pattern layer. The number of neurons in this level corresponds to the number of sample types. The performance layer in the summation layer is responsible for generating the maximum score. The relationship between the input layer and the hidden layer is achieved using a Gaussian function to determine the degree of similarity between each neuron in the subsequent level and each neuron in the input level. The appropriate score of each class is then added together, and the resulting average is used to determine the class of the input sample [18].

4. Proposed system

Fig. 1 illustrated the block diagrams for a proposed system to classification the pepper fruit diseases from the dataset, It is divided into five major phases: - acquiring the dataset (loading the chosen dataset), preparing the image (2D log chromaticity image, adaptive threshold, and crop), HoG approach for feature extraction, z-score normalization, and PNN algorithm for classification. Each stage in the proposed system discussed in details as following:

4.1 Dataset stage

The dataset (Diyala Pepper), includes (78) images of peppers and fruits that are both healthy and harmful [20]. Fig. 2 clarifies examples of pepper fruits dataset.

4.2 Preprocessing stage

This preprocessing of pepper fruit images aims to the determined region of interested ROI and remove removed shadow from it, this stage includes four sub-steps as following:

4.2.1 2D log chromaticity image

Shadows distort artifacts because of an obstruction from an obstruction that sometimes contributes to incorrect readings of the form, colour, or position of an image. Shadowing has a detrimental impact on the accuracy of classifiers built using training data. Additionally, this phase tries to create an illumination-invariant grayscale picture from one color image that is free of shadows by utilizing the 2D log chromaticity technique to eliminate the shadow from input color images.

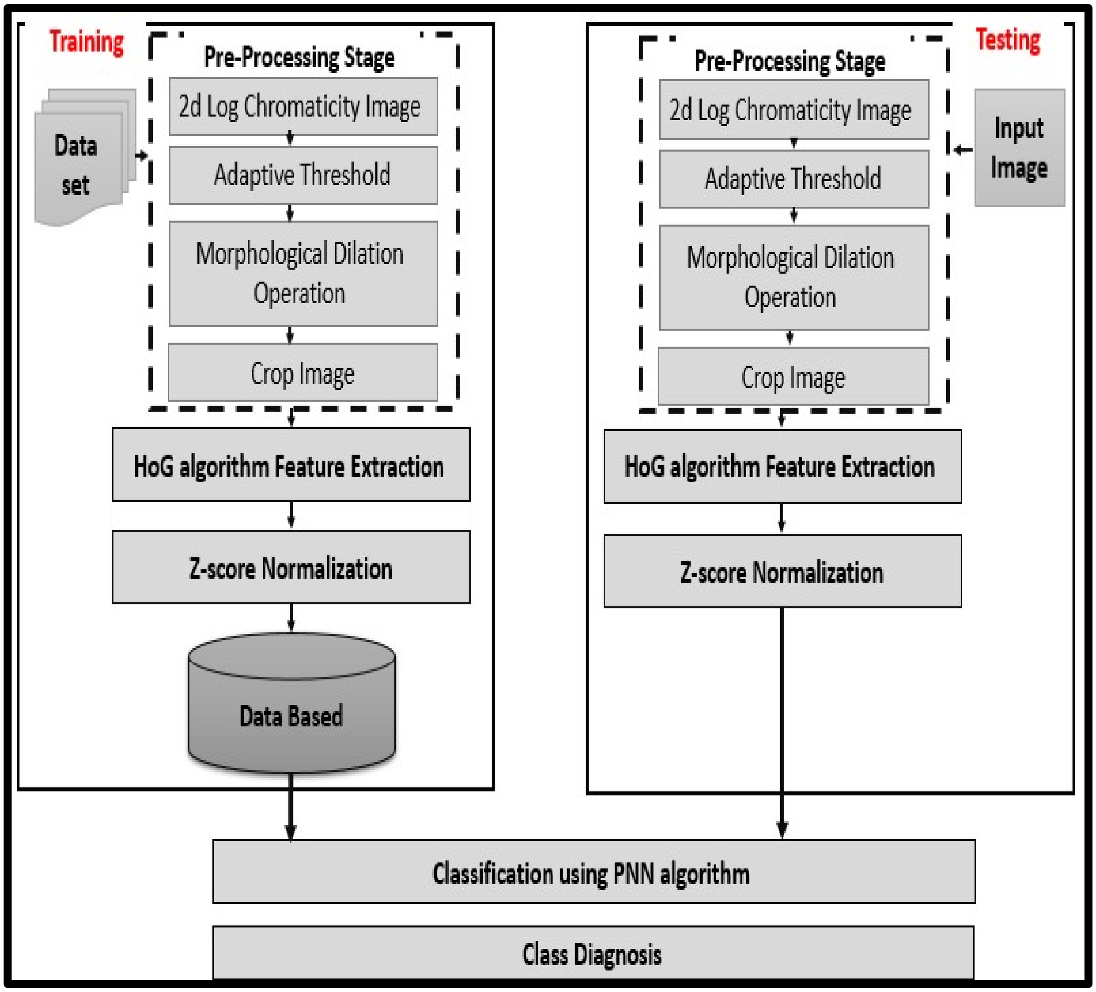


Fig.1. General block diagrams of proposed system.

Fig. 2. Samples from the pepper fruit dataset for all classes: a) fruits with peppers are healthy; b) black fruit lesions are illnesses; c) rots of fruit ; d) insect harm diseases.

4.2.2 Adaptive threshold technique

By selecting a threshold value, global thresholding applies to the whole image, the pixels are allocated to one area and those that fell below the mark are allocated to another (adjoin) region [21] , [22].

In this step, convert pepper fruits image into binary image without losing the important details of the image by using the proposed adaptive thresholding technique. To find the best threshold based on the proposed adaptive threshold, first, determine the likelihood that an image pixel will be white (P(W)) and the probability of black pixel of the image (P(B)), find the new threshold value = ((P(W)+P(B))/2 , this threshold is the optimal threshold when the old and new thresholds are equal.

4.2.3 Morphology dilation operation

Morphology is an image processing methodology focused on shape [23] . Things can dilate or increase in mass thanks to dilation. The choice of the structuring component affects both the total and how they evolve. It would not make much more sense to dilate before defining the structural feature than it would to try to filter the image without knowing what filter to use, his process seeks to increase the visibility of weak pixels. The description of dilation of A to B is [11]:

$$D(A,B) = A \oplus B = \cup_{\beta \in B} (A + B) \quad (10)$$

4.2.4 Crop image

The preceding stage produces a binary picture in which the white region represents the object or ROI, and the black region, indicates the backdrop. Given that the images dataset utilized for this study are indoor photographs and that one trait of indoor images is that the backdrop's color white, Therefore, by downsizing the image, for example, the crop phase seeks to shorten its completion time, if the size of the input image is 255*255 then in crop step can be resizing the image to 128*128. Additionally, by altering the color value of a pixel in the original image such that its coordinates match those of the white pixel in the cropped image, this phase seeks to distinguish between significant and unimportant portions in the image explain in result Fig. 8.

4.3 HoG feature extraction

When a disease region is found in a fruit pepper image, the HoG feature extraction method is used to extract 45 features from each image at this step. Furthermore, the suggested system computed sigma for each characteristic before storing it in the database. The sigma stands for Standard Deviation from the mean is a measure of how far data deviates from the mean, and it plays a crucial role in improving classification accuracy.

4.4 Z -score normalization

This procedure comprises initially determining the average values, and the standard deviation of the features, and finally normalization is accomplished using the following formula [24], [20]:

$$z = \frac{f_i - \mu}{\sigma_f} \quad (11)$$

Where f represents feature values, μ represents feature mean, and σ represents standard deviation. In order to prevent feature class overlapping and improve the suggested system's precision for classification, Z-score normalization seeks to increase the similarity between features and sigma that belong to the same class associated while simultaneously isolating them from another class.

5 Classification stage

The database that is employed in the suggested approach for Classification of Pepper Fruit Diseases Using PNN Algorithm is split into two parts: training 70% and testing 30%.

5.1 Training Step

The suggested approach also uses the pre-trained network method, which saves time and effort by allowing the network to be trained once, and then used later. For each form of plant disease, the data are labeled during the training stage, and training is carried out based on these labels. Epoch and shuffle are the two main training alternatives needed for the training stage. Epoch is used to divide the whole amount of data into smaller groups or batches, which are then successively fed into the system to train it to make decisions. The batch is referred to as an epoch when all data is added to the system exactly once. Multiple epochs may be thought of as the addition of various batches of data into the system. Before each epoch during the training of a neural network, it is crucial practice to shuffle the training data in every class so that each batch of data is different.

5.2 Testing Step

Based on the PNN algorithm, we designed a testing step categorizes the pepper plant as either healthy or diseased and assigns a disease class to each category. Fig. 3 shows clarifies the an architecture for a PNN that can identify the class K = {1,..,t}, where t is the pepper classes number for the pepper fruit image that are equal to 4.

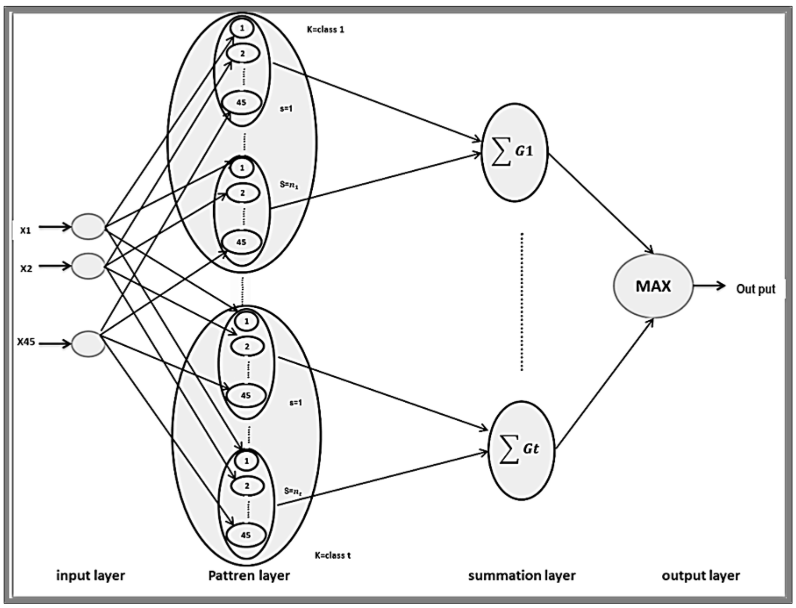


Fig. 3. Architecture for a PNN.

The input level comprises (45) nodes, all of the nodes in the pattern layer are branched out at each function input node in order that every pattern node obtains the complete capability of the x input. The pattern nodes are divided into categories, and each group has various training samples (nk) for the K classes, along with (45) training components for the classes. Equation (11) states that each training sample node in the class k category corresponds to a Gaussian function depending on its associated feature. All values with a Gaussian function center are summed separately and for all K groups totaled in the summation layer, respectively. Finally, attempt to evaluate the class with a maximum value (maximum a posteriori MAP value) in the output layer to mark instance x accordingly.

6. Implementation and results

The suggested system was applied on a computer with a processor: 11th Gen Intel(R) Core(TM) i5-1135G7 @ 2.40GHz 2.42GHz, RAM (8.00 GB), 64-bit (O.S) operating system, and is applied in Windows 10 through version Microsoft Visual Studio 2019 Visual throughout the implementation and testing.

6.1 The Dataset

The dataset includes (6) images of pepper leaves in both good and bad conditions, as indicated in Table 1. Each image has a resolution of 1600 x 1066 pixels. A DSLR (Nikon D3300 sensor size is 18-55mm) with a CMOS sensor and a resolution of (width 1600 pixels - height of 1066 pixels) was used to gather the data. Each image in the JPEG combo has its own RGB color range.

Table 1. Pepper disease dataset attributes.

Fruit(peppers)	
Dataset Attribute	**No. of Images**
Black fruit lesion	10
Fruit rots	35
Insect injury	10
Fruit Healthy	23
Number of all the fruits image	78
Number of all class	4

6.2 Results of implementation of the suggested system

The configuration pepper leaf dataset is the first stage of the proposed system; it seeks to remove duplicate or similar pictures from the dataset, which increases classification accuracy. The dataset contains the (78) pepper images and (4) class. Fig. 4 displays an illustration of the outcomes of selecting the best sample picture using the threshold value of 80.

The results of the second are preprocessing stage, which includes four main steps, which are (2D log chromaticity image, adaptive (threshold), operation of dilation, and crop), Fig .5, Fig .6, and Fig .7 illustrated the results of the 2D log chromaticity image. The color image is divided into three separate color channels, R, G, and B, as the initial stage in creating a 2D log chromaticity image.

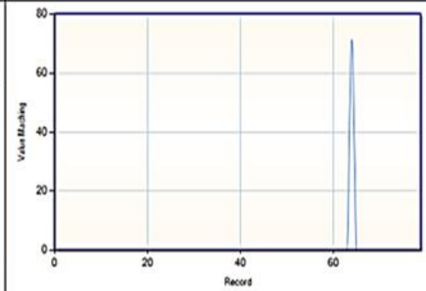
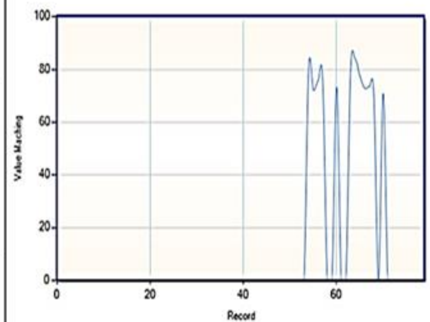
Class	Input Image	Histogram of matching features	Max matching	Threshold <80
1			71.43	Don't Add to training dataset
2			83.99	Add to training dataset

Fig. 4. Results of the configuration training dataset.

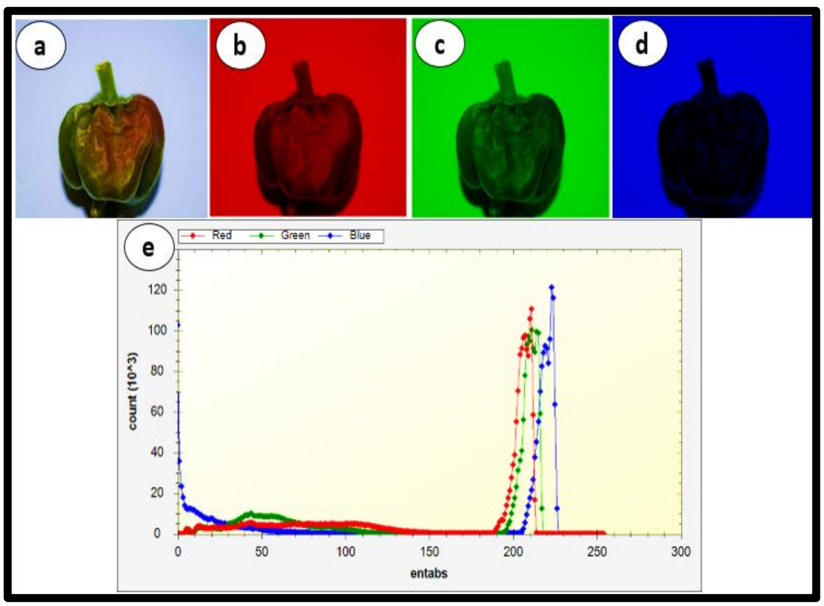


Fig. 5. Image of a pepper fruit with its R, G, and B color channels split apart and displayed; a)The original image ; b,c,d) Color channels in that order: red, green, and blue ; e) An R, G, and B channel histogram.

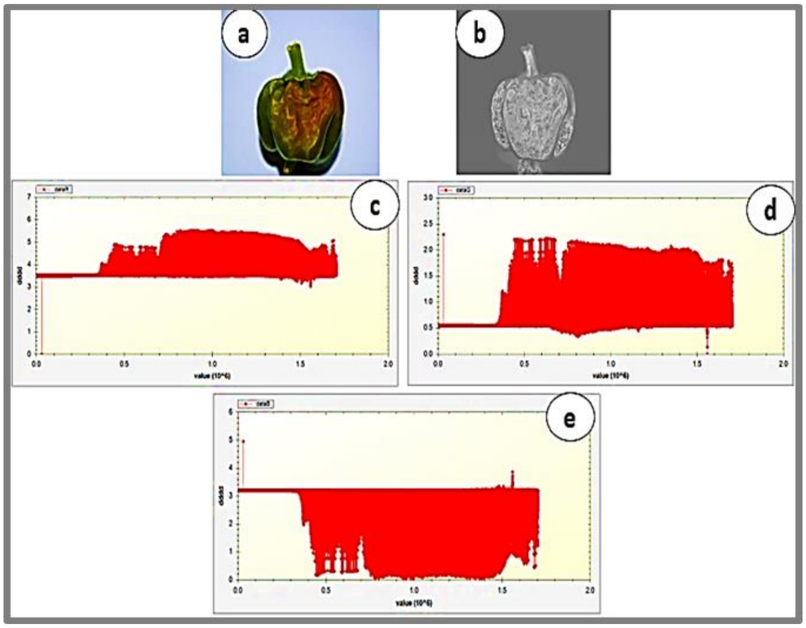


Fig. 6. Results of the 2d log chromaticity technique for removing shadows from chromaticity images when θ = 70; a) The original image ;b) Invariant grayscale; c,d,e) Histogram of R,G,B.

Fig. 6 illustrates the results of removing shadow from pepper leaf image depending on the θ=70 as a value. The results of the transformation of the pepper fruit picture to two dimensions (x1,x2) are presented in Fig. 7 histogram.

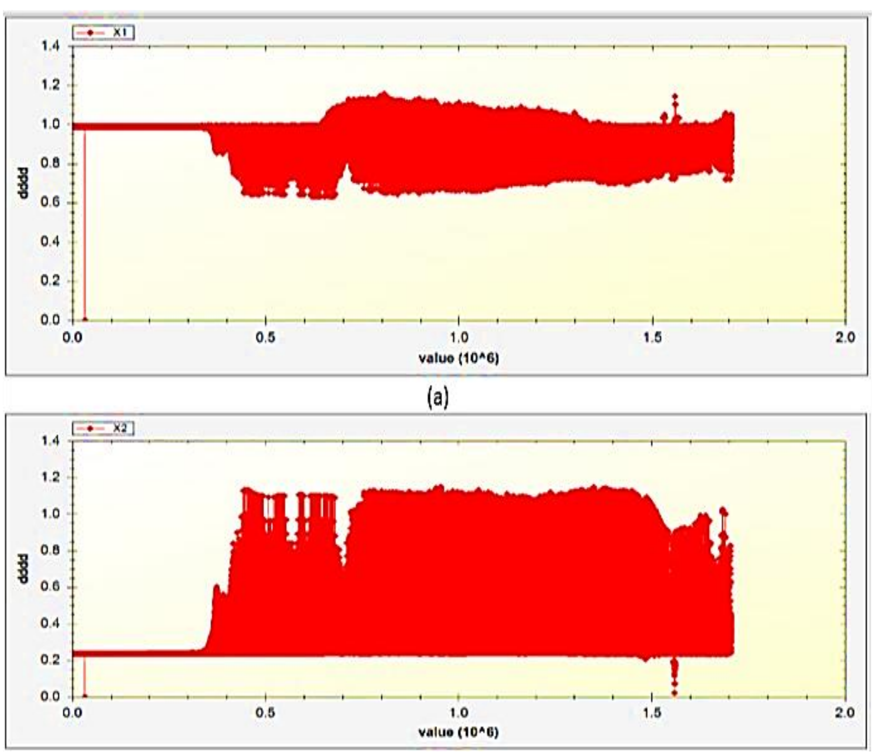


Fig. 7. Results the transformation to two dimensions (x1,x2); a,b) Histogram of value x1,x2 respectively .

Fig. 8 shows the results of the adaptive threshold technique, dilation operation depending on the size of the kernel [3*3], and crop image which represents the ending step in preprocessing the Fruit image. In this step, to get only the ROI from the image by determining the min (x,y) and max (x,y).
Fig. 9 represents the results of the HoG feature extraction algorithm, where Fig. 9.a represents 45 pepper fruit images of 4 classes; Fig 9.b represents their sigma.
Fig.10 depicted the effect of z-score normalization on HoG features, where features belonging to one class are more closely connected yet separated from the other class to reduce feature class overlapping and boost the precision of the suggested method during the classification phase. Results from the two classification stages (training and testing) for the disease of pepper leaves are included in the phase of classification using the PNN algorithm,70% of the database training in this phase, and 30% of the database was testing. The features that remain when normalization and sigma parameters are maintained in the database for use in the testing procedure since the training database underwent normalization to compute the z-score for each class. A sample of the computation sum for each class by the PNN classification method is presented in Fig.11. a, along with the histogram for that class. The probability of summation (Psum) for one class is presented in Fig.11.b, and the largest value of (Psum) is used to categorize the classes of pepper leaves. As shown in Fig.11.b the maximum probability of summation = 0.262500349091923then PNN can be classified a class of the input leaf image which is belongs to class=3.

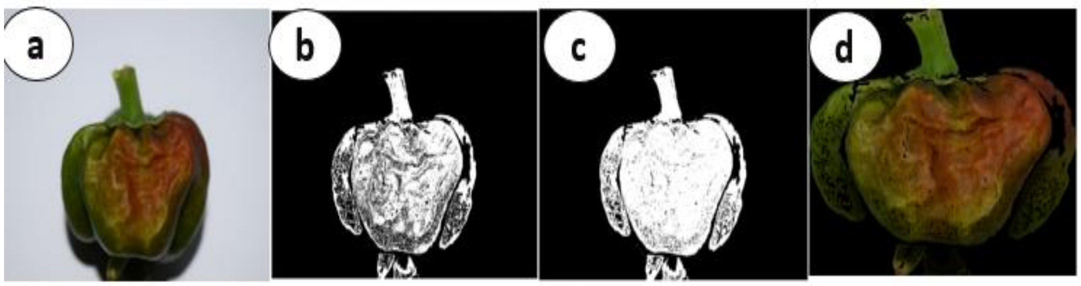
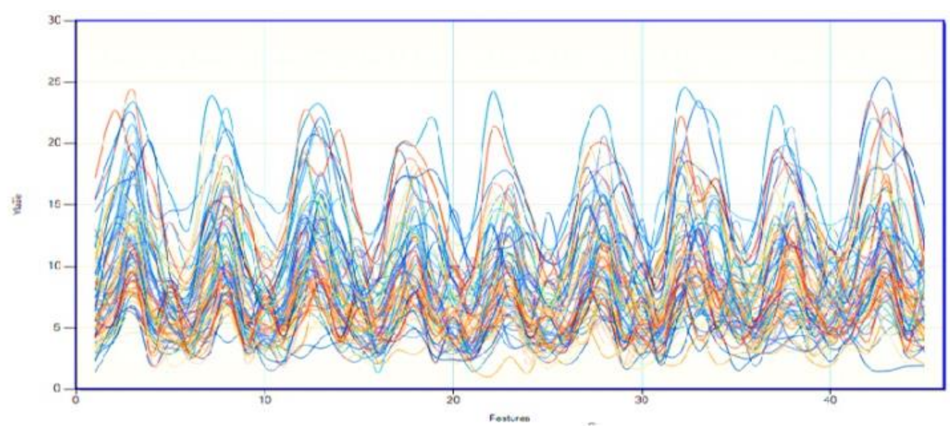
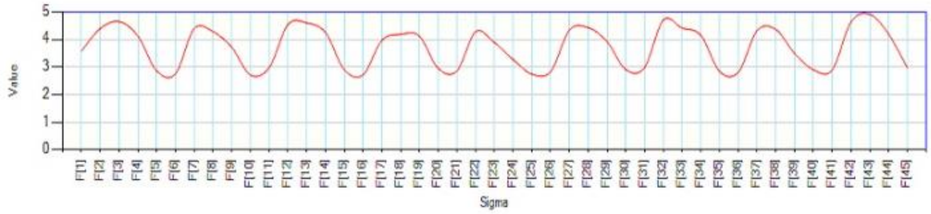


Fig. 8. Results of the adaptive threshold, dilation operation with kernel size3*3, and crop fruit image operation; a) original image; b) Adaptive threshold ;c)Dilation image; d) Crop image.



(a)



(b)

Fig. 9. Results of HoG feature extraction algorithm; a) original HoG features; b) Sigma values for 45 HoG features

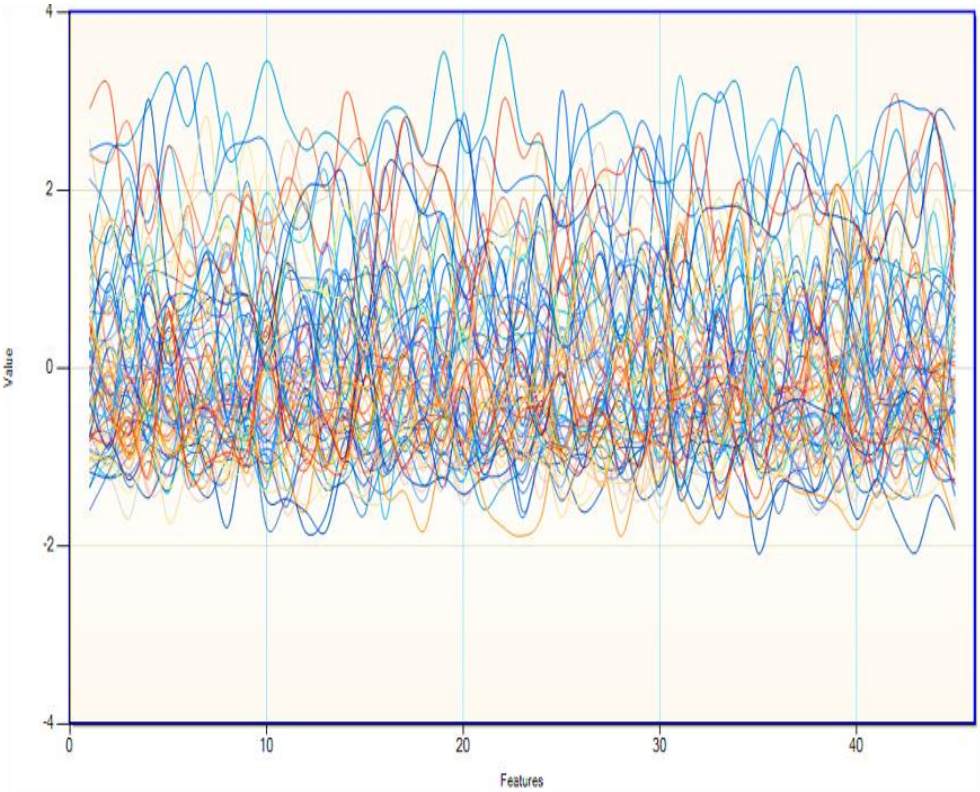
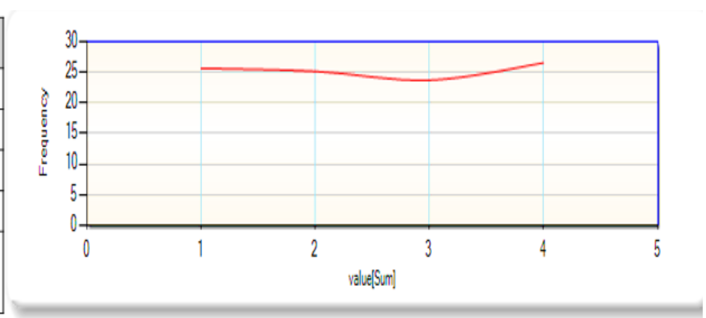


Fig.10. Z-Score normalization on pepper fruits database features

Class	sum
1	25.6789024420508
2	25.1669597996644
3	23.7282997824256
4	26.54339366869
Total Sum	101.117555692831



(a)

class	Psum
0	0.253950980777826
1	0.248888134480971
2	0.23466053564928
3	**0.262500349091923**
MAX Psum	**0.262500349091923**

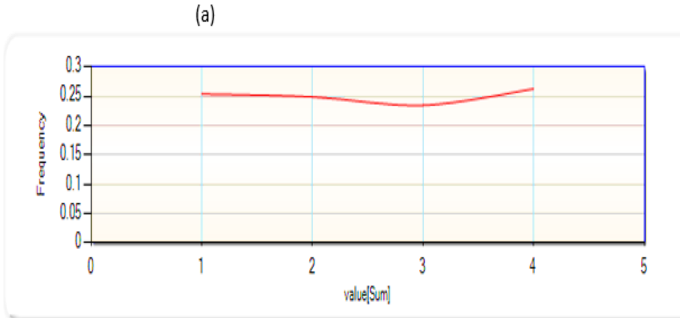


Fig. 11. Example of PNN classification algorithm; a) summation of the samples for one class using PNN, b) probability summation (Psum).

6.3 Performance Evaluation Measures

Calculating the number of accurately identified classes allows one to assess the classification's accuracy examples (true positives [TP]), the total sum of correctly recognized examples that do not fit the category (true negatives [TN], and examples that were either wrongly assigned to the class (false positives [FP]) or not taken seriously as class examples (false negatives [FN]). Collection of widely used output assessment indicators as seen in equations (12, 10 ,17) [25], [26], [27] would be used in this work.

$$Accuracy(ACC) = \frac{TP + TN}{TP + FN + FP + TN} \qquad (12)$$

And;

$$Error\ rate(ERR) = \frac{FP + FN}{TP + TN + FP + FN} \qquad (13)$$

And;

$$False\ Accept\ Rate(FAR) = \frac{wrongly\ accepted\ sample}{Total\ no.of\ wrong\ match} \qquad (14)$$

And;

$$False\ Reject\ Rate(FRR) = \frac{wrongly\ rejected\ sample}{Total\ no.of\ corrected\ match} \qquad (15)$$

And;

$$Recall\ (Sensitivity)(SNS) = \frac{TP}{TP + FN} \qquad (16)$$

Finally;

$$Specificity(SPC) = \frac{TN}{FP + TN} \qquad (17)$$

To acquire the best outcome accuracy, the experiment involved choosing a varied number of training epochs. The suggested algorithm used epoch and shuffling techniques. By testing the best accuracy ratio that getting when the value of epoch [0 to10] , accuracy value = 83.92857 as Fig.12 illustrates how the sigma histogram changes as the training epoch changes.

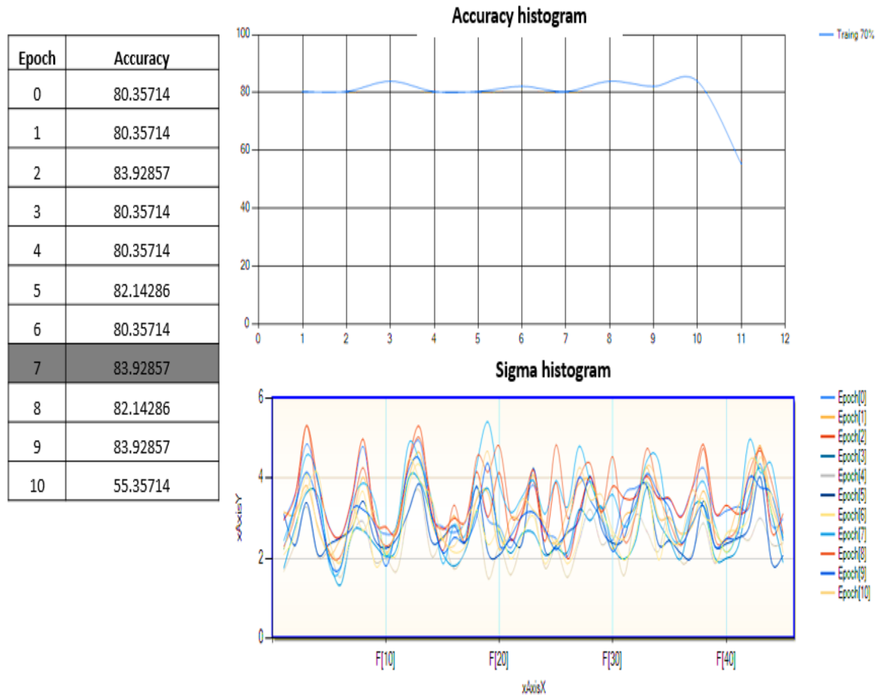


Fig. 12. Accuracy and sigma change against training Epoch.

Table 2 shown value performance evaluation measurements using confusion matrix, these measures are true positive [TP], true negative [TN], false positive [FP],false negative [FN], Accuracy of single[ACC],Error of single [ERR],FAR,FRR, sensitivity [SNS],then lastly Specificity [SPC].As displayed in Table 2 the suggested system has ability to achieve the optimal values of performance evaluation metrics with class 1 where value of ACC=1,ERR=0,FAR=0,FRR=0,SNS=1,and finally SPC=1.

Table 2. Results of performance evaluation measurements.

The class	TP	FP	FN	TN	ACC	ERR	FAR	FRR	SNS	SPC
1	3	0	0	18	1.0000	0.0000	0.0000	0.0000	1.0000	1.0000
2	4	6	1	11	0.6818	0.3182	0.3529	0.2000	0.8000	0.6471
3	5	1	0	16	0.9545	0.0455	0.0588	0.0000	1.0000	0.9412
4	2	1	7	12	0.6364	0.3636	0.0769	0.7778	0.2222	0.9231
					0.8182	0.1818	0.1222	0.2444	0.7556	0.8778

6 Conclusion

For the crop to be successfully grown, effective plant disease detection and categorization are crucial. In this research, the collection of real pepper fruit disease data set has several problems and challenges such as shadow and replication images that affected on accuracy of the classification algorithm. Because of this, it has become possible to identify and categorize the health or sickness of peppers and fruits using contemporary technology and sturdy design. In the configuration pepper fruit dataset stage, the proposed system was used new techniques (configuration training data set) to isolate the samples with the best features to remove weak or repetitive samples that affect the accuracy of class classification through the proposed system. Implementation of the suggested technique using actual data from pepper fruits. The data divided 30% of the data were for testing, and 70% were for training. In the training and testing stages, the suggested system achieved an accuracy of (100%) and (81.82%), respectively.

Acknowledgement
The author's thanks the "Department of Computer Science", "Collage of Science", " Diyala University", for supporting this work.

Reference

[1] K. Kansotia *et al.*, "Role of Modern Technologies in Plant Disease Management: A Comprehensive Review of Benefits, Challenges, and Future Perspectives," *International Journal of Environment and Climate Change*, vol. 13, no. 12, pp. 1325–1335, 2023.

[2] W. B. Demilie, "Plant disease detection and classification techniques: a comparative study of the performances," *J Big Data*, vol. 11, no. 1, p. 5, 2024.

[3] N. Schor, S. Berman, A. Dombrovsky, Y. Elad, T. Ignat, and A. Bechar, "Development of a robotic detection system for greenhouse pepper plant diseases," *Precis Agric*, vol. 18, pp. 394–409, 2017.

[4] A. A. Maryoosh and S. Pashazadeh, "Leukemia Detection using Machine Learning Algorithms: Current Trends and Future Directions," Mustansiriyah Journal of Pure and Applied Sciences, vol. 2, no. 3, pp. 33-49, 2024.

[5] M. K. Tripathi and D. D. Maktedar, "Recent machine learning based approaches for disease detection and classification of agricultural products," in *2016 international conference on computing communication control and automation (ICCUBEA)*, IEEE, 2016, pp. 1–6.

[6] S. Hussain, R. Hasan, and S. J. Hussain, "Classification and detection of plant disease using feature extraction methods," *International Journal of Applied Engineering Research*, vol. 13, no. 6, pp. 4219–4226, 2018.

[7] V. Sahithya, B. Saivihari, V. K. Vamsi, P. S. Reddy, and K. Balamurugan, "GUI based detection of unhealthy leaves using image processing techniques," in *2019 International Conference on Communication and Signal Processing (ICCSP)*, IEEE, 2019, pp. 818–822.

[8] S. Kohombange, H. K. L. K. Gunasekera, and J. P. Eeswara, "Effect of Nitrobenzene on Yield of Bell Pepper (Capsicum annuum L.) and Sweet Cucumber (Cucumis sativus L.) under Green House Condition," *Cutting-edge Research in Agricultural Sciences Vol. 9*, no. May 2018, pp. 23–32, 2021, doi: 10.9734/bpi/cras/v9/1986f.

[9] "2023/2024 New Jersey Commercial Tree Fruit Production Guide," 2024.

[10] C. S. Hlaing and S. M. M. Zaw, “Plant diseases recognition for smart farming using model-based statistical features,” in *2017 IEEE 6th global conference on consumer electronics (GCCE)*, IEEE, 2017, pp. 1–4.

[11] M. Sharif, M. A. Khan, Z. Iqbal, M. F. Azam, M. I. U. Lali, and M. Y. Javed, “Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection,” *Comput Electron Agric*, vol. 150, pp. 220–234, 2018.

[12] K. Karadağ, M. E. Tenekeci, R. Taşaltın, and A. Bilgili, “Detection of pepper fusarium disease using machine learning algorithms based on spectral reflectance,” *Sustainable Computing: Informatics and Systems*, vol. 28, p. 100299, 2020.

[13] A. Adeel *et al.*, “Diagnosis and recognition of grape leaf diseases: An automated system based on a novel saliency approach and canonical correlation analysis based multiple features fusion,” *Sustainable Computing: Informatics and Systems*, vol. 24, p. 100349, 2019.

[14] J. Lu, L. Tan, and H. Jiang, “Review on convolutional neural network (CNN) applied to plant leaf disease classification,” *Agriculture*, vol. 11, no. 8, p. 707, 2021.

[15] A. Krishnan, P. Jayadevan, and J. V. Panicker, “Shadow removal from single image using color invariant method,” in *2017 International Conference on Communication and Signal Processing (ICCSP)*, IEEE, 2017, pp. 583–587.

[16] A. Challa, “Automatic Handwritten Digit Recognition On Document Images Using Machine Learning Methods.” 2019.

[17] Z. H. Rasool and M. A. A. Amir, "Comprehensive Image Classification using Hybrid CNN-LSTM Model with Advanced Feature Extraction on Coco Dataset," Mustansiriyah Journal of Pure and Applied Sciences, vol. 2, no. 2, pp. 28-47, 2024.

[18] M. Turkoglu and D. Hanbay, “Recognition of plant leaves: An approach with hybrid features produced by dividing leaf images into two and four parts,” *Appl Math Comput*, vol. 352, pp. 1–14, 2019.

[19] S. Kuroda, H. Okuda, W. Ishida, and S. Koseki, “Modeling growth limits of Bacillus spp. spores by using deep-learning algorithm,” *Food Microbiol*, vol. 78, pp. 38–45, 2019.

[20] R. Basim, J. M. Al-Tuwaijari, and T. S. A. Al-Hayali, “An Employed of PNN algorithm in Smart Farming System for Early Plant Disease Detection and Diagnosis,” *Solid State Technology*, vol. 63, no. 6, pp. 18771–18789, 2020.

[21] P. Roy, S. Dutta, N. Dey, G. Dey, S. Chakraborty, and R. Ray, “Adaptive thresholding: A comparative study,” in *2014 International conference on control, Instrumentation, communication and Computational Technologies (ICCICCT)*, IEEE, 2014, pp. 1182–1186.

[22] M. Nixon and A. Aguado, *Feature extraction and image processing for computer vision*. Academic press, 2019.

[23] N. T. Gadare and S. A. Ladhake, “Image Enhancement and Background Detection Using Morphological Transformation,” *Int. Journal of Engineering Research and Applications*, vol. 4, no. 2.

[24] X. xiao Wang, L. Dong, S. yang Liu, Y. Hao, and B. Wang, “A fault classification method of photovoltaic array based on probabilistic neural network,” in *2019 Chinese Control And Decision Conference (CCDC)*, IEEE, 2019, pp. 5260–5265.

[25] M. Gösgens, A. Zhiyanov, A. Tikhonov, and L. Prokhorenkova, “Good classification measures and how to find them,” *Adv Neural Inf Process Syst*, vol. 34, pp. 17136–17147, 2021.

[26] G. Stegmayer, D. H. Milone, S. Garran, and L. Burdyn, “Automatic recognition of quarantine citrus diseases,” *Expert Syst Appl*, vol. 40, no. 9, pp. 3512–3517, 2013.

[27] S. Sunny and M. I. Gandhi, “An efficient citrus canker detection method based on contrast limited adaptive histogram equalization enhancement,” *Int. J. Appl. Eng. Res*, vol. 13, no. 1, pp. 809–815, 2018.