Ministry of Higher Education and
Scientific Research
University of Diyala
College of Science
Department of Computer Science

Classification of Iraqi Pepper Disease using PNN Technique

A Thesis

Submitted to the Department of Computer Science\ College of Sciences\ University of Diyala

In a Partial Fulfillment of the Requirements for the Degree of Master in Computer Science

By

Rasha Basim Issa

Supervised By

Asst. Prof. Dr. Jamal Mustafa Al-Tuwaijari

2021 A.D. **Iraq** **1442 A.H.**

بسم الله الرحمن الرحيم

﴿ وَلَسَوْفَ يُعْطِيكَ رَبُّكَ فَتَرْضَىٰ ﴾

صدق الله العلي العظيم

(سورة الضحى : الاية 5)

Acknowledgement

My thanks are above all to ***God Almighty****, who guided my steps towards the path of knowledge and without His help and blessing, This thesis would not be advanced or will see the light.*

I wish to express my thanks to my supervisor, ***Assist. Prof. Dr. Jamal Mustafa Al-Tuwaijari*** *for supervising this thesis and for the generosity, patience and continuous guidance throughout the work.My thanks to the academic and administrative staff at the department of the computer sciences.*

Last and not least, thanks a lot go to ***my family****,* ***my friends****, and anyone who helped me in one way or another.*

Rasha Basim Issa

Dedication

I would like to dedicate this work to:

the soul of both my father and my cousin,I would not wish you were with me because I am sure you are with me.

To my mother, she patiently supported me to see my success.

To my husband for his unlimited love, support, endurance and encouragement.

Supervisor's Certification

We certify that this thesis entitled "***Classification of Iraqi Pepper Disease using PNN Technique***" was prepared by "**Rasha Basim Issa**" Under our supervisions at the University of Diyala, Faculty of Science, Department of Computer Science, as partial fulfillment of the requirement needed to award the degree of Master of Science in Computer Science.

Signature:

Name: **Assistant Prof. Dr. Jamal Mustafa Al-Tuwaijari**

Date : / /2021

Approved by the University of Diyala, Faculty of Science Department of Computer Science.

Signature:

Name: **Assistant Prof. Dr. Bashar Talib Al-Nuaimi**

Date: / /2021

(Head of Computer Science Department)

Scientific Certification

I certify that the thesis entitled “***Classification of Iraqi Pepper Disease using PNN Technique***” was prepared by " ***Rasha Basim Issa*** " has been evaluated scientifically; therefore, it is suitable for debate by examining committee.

Signature:

Name : Assistant Prof. Dr. Rawaa Dawoud Hassan

Date : / / 2021

Linguistic Certification

This is to certify that this thesis entitled "***Classification of Iraqi Pepper Disease using PNN Technique*** "was prepared by "***Rasha Basim Issa***" at the University of Diyala/ Department of Computer Science, is reviewed linguistically. Its language was amended to meet the style of the English language.

Signature:

Name :

Date : / / 2021

Examination Committee Certification

We certify that we have read the thesis entitled "***Classification of Iraqi Pepper Disease using PNN Technique***" and an examination committee, examined the student "***Rasha Basim Issa***" in the thesis content and that in our opinion, it is adequate as fulfill the requirement for the Degree of Master in Computer Science at the Computer Science Department, College of Science University of Diyala.

(**Chairman**)

Signature:

Name: **Prof. Dr. Dhahir Abdulhadi Abdulah**

Date**: / / 2021**

Signature:

Name: **Assist. Prof. Dr.Taha Mohammed Hassan** **(Member)**

Date**: / / 2021**

Signature:

Name: **Assist. Prof. Dr. Amel H. Abbas** **(Member)**

Date**: / / 2021**

Signature:

Name: **Asst. Prof. Dr. Jamal Mustafa AL-Tuwaijari** **(Supervisor)**

Date**: / / 2021**

Approved by the **Dean** of College of Science, University of Diyala

(**The Dean**)

Signature:

Name: **Prof. Dr. Tahseen H. Mubarak**

Date**: / / 2021**

Abstract

Quality agriculture production is the essential trait for any nation's economic growth. The agricultural sector has been facing great challenges to feed the increasing number of population living in the world. In the future, it will be very difficult to rely on traditional farming to produce food. In plants, pepper is used as a major source of nutrients throughout the world. However, peppers diseases badly affect the production and quality of pepper plants. Image processing and machine learning techniques have been widely used in the agriculture for detection and classification of diseases in plants. In this thesis, we propose a method for classification of diseases in peppers plants. The proposed method consists of several stages, which are include, the image acquisition stage, image pre-processing, features extraction, features normalization, and the classification stage. In the pre-processing stage, 2-D log chromaticity image, image thresholding, morphological dilation operation and, crop image were used. For feature extraction the Histogram of Oriented Gradients (HOG) used and Z-score for normalization, Forty-five features were extracted from each peppers plant image, and probabilistic neural networks algorithm was used to perform the classification process. The proposed technique is tested on the new dataset for pepper plant images, the dataset named (**DiyalaPepper)** and contains (244) samples of healthy and unhealthy images for both peppers fruits and leaves collectively. The total number of leaves pepper images in dataset equal to (166) and (78) images for pepper fruits. Accuracy rate obtained from the peppers fruits image was 81.82 % , and 94.11% for peppers leaves image.

Table of Contents

List of Abbreviations

Abbreviations	Meaning
ACC	Accuracy
ANN	Artificial Neural Networks
B'	Blue normalization
CCA	Canonical Correlation Analysis
CCV	Color Coherence Vector
ERR	Error Rate
FAR	False Accept Rate
FN	False Negatives
FRR	False Reject Rate
G'	Green normalization
GUI	Graphical User Interface
HER	Electronic Health Records
HOG	Histogram of Oriented Gradients
IF	Intelligent Farming
I_{inv}	1D Illumination Invariant Image
IoT	Internet of Things
KNN	K-nearest Neighbor
LAB	Luminance and B Components
LBP	Local Binary Pattern

LCHR	Local Contrast Haze Reduction
Log	Logarithm
ML	Machine Learning
M-SVM	Multi- Class Support Vector Machine
NB	Naive Bayes
NCA	Neighborhood Component Analysis
PCA	Principal Component Analysis
R'	Red normalization
SNS	Sensitivity
SPC	Specificity
SVM	Support Vector machine
T	Threshold
TN	True Negatives
TP	True Positives
TP	False Positives

Table of Symbols

Symbol	Meaning
Θ	Theta
η	Eta
$\oplus$	Circled plus / oplus - xor
$\ominus$	Circled minus /ominus
α	Alpha
$\|\|x\|\|$	Norm
σ	Standard Deviation

List of Figures

List of Table

List of Pseudo Code no.

Chapter one

General Introduction

Chapter One
General Introduction

1.1 Introduction

Agriculture remains a vital sector for most countries. It presents the main source of food for the population of the world [1]. Plants are susceptible to many diseases and pests. Researchers have shown that almost one-third of the world's agricultural crops are being destroyed by these harmful factors. Diseases causing product losses around 60–70% or even close to 100% have been a common concern of humankind. For this reason, they have searched for ways to fight against these diseases and weeds. In order to reduce economic losses and environmental pollution from agricultural chemicals, damage caused by diseases should be handled appropriately [2]. Hence, it is extremely important to switch from traditional agricultural methods to modern agriculture. Smart management consists of collecting, transmitting, selecting and analyzing data. As the amount of agricultural data increases significantly, robust analytical techniques capable of processing and analyzing large amounts of data to obtain more reliable information and much more accurate predictions are essential[1].

Automation of disease detection and monitoring can facilitate targeted and timely disease control, which can lead to increased yield, improved crop quality and reduction in the quantity of applied pesticides. Further advantages are reduced production costs, reduced exposure to pesticides for farm workers and inspectors and increased sustainability. Symptoms are unique for each disease and crop, and each plant may suffer from multiple threats. Thus, a dedicated integrated disease-detection system and algorithms are required [3].

Machine learning algorithms help us better understand and analyze complex data[4]. Popular methods have utilized machine learning, image processing and classification based approaches to identify and detect the diseases on agricultural products. The existing techniques for disease detection have utilized various image processing methods followed by various classification techniques[5].

In this thesis, due to real dataset for pepper plant (leaves and fruits) image used, probabilistic neural networks is used in classification problems because tolerant of noisy inputs, instances can be classified by more than one output and adaptive to changing data[6].

1.2 Pepper Plant

Pepper is an important vegetable crop grown in temperate and tropical regions of the world. This fact is due to the high biological value of the fruits (high content of dry substance, vitamin C and B-complex, minerals, essential oils, carotenoids, etc.) and their various kinds of utilization in the culinary and food industry of different countries [7].

1.3 Diseases of Pepper Plant

Diseases of pepper caused by biotic (infectious) and abiotic (non-infectious) agents interfere with the production of pepper. Biotic agents of disease of pepper include fungi, bacteria, nematodes, and viruses. Abiotic disorders include a number of unfavorable cultural or climatic conditions, such as sunlight, nutrient deficiency, and temperature excesses. Diseases affect all parts of the pepper plant including the foliage, stems, roots, fruit, and young seedlings. Fungi and bacteria cause a variety of symptoms such as leaf and fruit spotting, wilting and plant death [8].

This thesis will focus on the following diseases for the leaves and fruits. In addition to healthy ones , that can be describe bellow:

A- For the Leaves Disease:

1- **Noninfectious Diseases (Environmental)**

Symptoms of the disease are wrapped in leaves and wilts, and the edges of leaves and buds are colored brown [9].

2- Noninfectious Diseases (Nutrient Deficiencies)

Such as lack of nitrogen, magnesium, potassium and iron. Symptoms of nitrogen deficiency are the general yellowing of old leaves and its gradual rise to the upper leaves [10]. Potassium deficiency leads to yellowing of the old leaves, with the ends tanning, burning edges and stains, as the fleshy parts show necrosis at the ends [11]. Symptoms of a magnesium deficiency begin with areas of patchy yellowing between the veins [12]. Iron deficiency causes the yellowing of modern leaves, while their veins remain green, and later develop into general yellowing and shortening of the color of the leaves with the appearance of necrotic spots [11,12].

3- Leaf Curl:

The disease syndrome includes shortening of stem internodes, interveinal yellowing, and upward rolling of the leaf blade, accompanied by fruit discoloration and size reduction[13].

4- Leaf Spots:

Symptoms on leaves included small, irregularly shaped, brown lesions with yellow halos and marginal necrosis [14].

5- Leaf Worm or Cut Worm

Cutworms attacking pepper, principally feltia subterranean (fabricus) and peridroma saucia ,are stout, softbodied smoth larvae , colored dull gray, brown , or black sometimes spotted or striped . they cut off the young plants near the soil surface. Some cutworms climb plants and cut off the leaves and buds. They are found wherever pepper are grown [15].

B- For the Fruits Diseases:

1- Black Fruit lesion

The causes that led to the appearance of this disease are environmental conditions, such as a decrease in temperature [15].

2- Fruit Rots

Symptoms began as soft lesions that turned dark brown to black. Lesions usually originated at the calyx end of the fruit and extended down the sides [16].

3- **Insect Injury**

Peppers are attacked by many kinds of insects , the most important are aphids, cutworms, flea ,beetles, hornworms, the pepper weevil, and the pepper maggot. Some of these insects are widely distributed; others are serious pest in limited area only [15].

1.4 Related Works

Several researchers have shown their interest in classification of plant diseases. The following are some of the published works that are relevant to the current work:

- **Manisha Bhange, et al. , 2015[17]** : propose a web based tool that helps farmers for identifying fruit disease by uploading fruit image to the system. The system has an already trained dataset of images for the pomegranate fruit. Input image given by the user undergoes several processing steps to detect the severity of disease by comparing with the trained dataset images. First the image is resized and then its features are extracted on parameters such as color, morphology, and CCV and clustering is done by using k-means algorithm. Next, SVM is used for classification to classify the image as infected or non-infected. An intent search technique is also provided which is very useful to find the user intension. Out of three features extracted him got best results using morphology. Experimental results display different accuracy levels of disease detection based on the input image quality and the stages of the disease. The overall system accuracy is measured to be 82%.
- **Thomas J. Hirschauer,et al. ,2015[18]**: In this paper, a comprehensive computer model is presented for the diagnosis of f Parkinson’s disease (PD) based on motor, non-motor, and neuroimaging features using the recently-developed enhanced probabilistic neural network (EPNN). The model is tested for differentiating PD patients from those with scans without evidence of dopaminergic deficit (SWEDDs) using the Parkinson’s Progression Markers Initiative (PPMI) database, an observational, multi-center study designed to identify PD biomarkers for diagnosis and disease progression. The results are compared to four other commonly-used machine learning algorithms: the

probabilistic neural network (PNN), support vector machine (SVM), k-nearest neighbors (k-NN) algorithm, and classification tree (CT). The EPNN had the highest classification accuracy at 92.5 % followed by the PNN (91.6 %), k-NN (90.8 %) and CT (90.2 %). The EPNN exhibited an accuracy of 98.6 % when classifying healthy control (HC) versus PD, higher than any previous studies.

- **Yasha Zeinali, et al. , 2017 [19]**: This paper presents a novel PNN algorithm, the competitive probabilistic neural network (CPNN). In the CPNN, a competitive layer ranks kernels for each class and an optimum fraction of kernels are selected to estimate the class-conditional probability. Using a stratified, repeated, random subsampling cross-validation procedure and 9 benchmark classification datasets, CPNN is compared to both traditional PNN and the state of the art (e.g. enhanced probabilistic neural network, EPNN). These datasets are examined with and without noise and the algorithm is evaluated with several ratios of training to testing data. In all datasets (225 simulation categories), performance percentages of both CPNN and EPNN are greater than or equivalent to that of the traditional PNN; in 73% of simulation categories, the CPNN analyses show modest improvement in performance over the state of the art.
- **Muhammad Sharif, et al. , 2018[20]** : In this article, are propose a hybrid method for detection and classification of diseases in citrus plants. The proposed method consists of two primary phases; (a) detection of lesion spot on the citrus fruits and leaves; (b) classification of citrus diseases. The citrus lesion spots are extracted by an optimized weighted segmentation method, which is performed on an enhanced input image. Then, color, texture, and geometric features are fused in a codebook. Furthermore, the best features are selected by implementing a hybrid feature selection method, which consists of PCA score, entropy, and skewness-based covariance vector. The selected features are fed to Multi- Class Support

Vector Machine (M-SVM) for final citrus disease classification. The proposed technique is tested on Citrus Disease Image Gallery Dataset, Combined dataset (Plant Village and Citrus Images Database of Infested with Scale), and own collected images database. It used these datasets for detection and classification of citrus diseases namely anthracnose, black spot, canker, scab, greening, and melanose. The proposed technique outperforms the existing methods and achieves 97% classification accuracy on citrus disease image gallery dataset, 89% on combined dataset and 90.4% on local dataset. In future, it would like to construct a deep model and apply the model on the selected citrus datasets, as the deep learning performed significantly well in the field of computer vision. However, it need a big citrus dataset for it.

- **Dr. Shaik Asif Hussain, et al., 2018[21]** : To detect plant diseases Image processing steps are used to extract features from the images of plant leaves. The images are classified to define the disease detected through graphical user interface (GUI) and it also calculates the affected region and it is shown as percentage of disease detection. The algorithms used are K-means clustering and Support vector machine for comparison and percentage of disease detection. The analysis of the work is carried through objective values such as mean, Entropy, variance, kurtosis, skewness, contrast and homogeneity. However contrast enhancement is also done for low intensity images. The resultant values are future enhanced with a hardware placed to monitor the status and update the information in Internet of things (IOT) for effective management of the plant disease detection.
- **Balakrishna K.,et al.,2019[22]:**In this article, the authors proposed two methods for identification and classification of healthy and unhealthy tomato leaves. In the first stage, the tomato leaf is classified as healthy or unhealthy using the KNN approach. Later, in the second stage, they classify the unhealthy tomato

leaf using PNN and the KNN approach. The features are like GLCM, Gabor, and color are used for classification purposes. Experimentation is conducted on the authors own dataset of 600 healthy and unhealthy leaves. The experimentation reveals that the fusion approach with PNN classifier outperforms than other methods, when the training Percentage of dataset is 70% , PNN classification results is 85.89 and 70.37 for KNN.

- **Velamakanni Sahithya, et al. , 2019[23]** : In this paper, ladies finger plant leaves are chosen and examined to find an early stage of various diseases such as yellow mosaic vein, leaf spot, powdery mildew etc. Leaf images are captured, processed, segmented, features extracted, and classified to know if they are healthy or unhealthy. Due to practical limitations in climatic conditions and other terrain regions, noisy image data sets are also created and taken into consideration. K-means clustering is used for segmentation and for classification, SVM and ANN are used. This work uses PCA to reduce the feature set. Results show that, the average accuracy of detection in SVM and ANN are 85% and 97% respectively. Without noise they are observed to be 92% and 98% respectively.
- **Kerim Karadag, et al., 2019[24]**: In this work, spectral reflections were used for early detection of fusarium disease in pepper plants. It is faster and more cost-effective to use the reflectance spectra from the plants in the laboratory. Pepper disease detection takes place in two stages. In the first step, the feature vector is obtained. In the second step, the feature vectors of the input data are classified. The feature vector consist of the coefficients of wavelet decomposition and the statistical values of these coefficients. Artificial Neural Networks (ANN), Naive Bayes (NB) and K-nearest Neighbor (KNN) were used for classification. In detection the health case of pepper, the average success rates of different classification algorithms for the first two groups (diseased and healthy peppers)

were calculated as 100% for KNN, 97.5% for ANN and 90% for NB. Likewise, these rates for the classification of all groups were calculated as 100% for KNN, 88.125% for ANN and 82% for NB. Overall, the results have shown that leaf reflections can be successfully used in disease detection.

- **Alishba Adeel, et al., 2019[25] :** In this work, it proposed an automated system for segmentation and recognition of grape leaf diseases. The proposed system comprises of four main steps. In first step, a local contrast haze reduction (LCHR) enhancement technique is proposed for increasing the local contrast of symptoms. Thereafter, LAB color transformation is held in the second step and the best channel is selected based on the pixels information that is later utilized into thresholding function. Color, texture, and geometric features are extracted and fused by canonical correlation analysis (CCA) approach. At the time of features fusion, a noise is added in the form of irrelevant and redundant features that are removed by Neighborhood Component Analysis (NCA). The classification of final reduced features is then performed by M-class SVM. The introduced system is assessed on Plant Village dataset of three types of grape leaf diseases such as black measles, black rot, and leaf blight including healthy. The proposed method acquired an average segmentation accuracy rate of 90% and classification accuracy is above 92% which is superior in contrast of existing techniques. The computational cost of the proposed system is also calculated which is minimized after selection approach. However, the proposed system also contains few limitations such as degrade accuracy for complex images.
- **Mustafa, M. S., et al., 2020[26] :** This research has been developed a system for recognizing the species and detecting the early disease of the herbs using computer vision and electronic nose, which focus on odour, shape, color and texture extraction of herb leaves, together with a hybrid intelligent system that are

involved fuzzy inference system, naıve Bayes (NB), probabilistic neural network (PNN) and support vector machine (SVM) classifier. These techniques were used to perform a convenient and effective herb species recognition and early disease detection on ten different herb species samples. The species recognition accuracy rate among ten different species using computer vision and electronic nose is archived 97% and 96%, respectively, in SVM, 98% and 98%, respectively, in PNN and both 94% in NB. In the early disease detection, the detection rate among ten different herb's species using computer vision and electronic nose are 98% and 97%, respectively, in SVM, both 98% in PNN, 95% and 94%, respectively, in NB. Integrated three machine learning approaches have successfully achieved almost 99% for recognition and detection rate. An integrated approach can contribute high accuracy in species recognition and early disease detection, but unfortunately there is an increase in computational time in the testing phase.

1.5 Problem Statement

Several problems emerged when building a system about the classification of plant diseases, could be classified as follows:

A. shadow problem which has a negative effect of classifier performance.

B. Symptom segmentation:

Most pepper plant disease symptoms have no well-defined edges, and they fade on plants slowly because of which there will not be a proper segmentation, which will affect final result.

C. Different disorders with similar symptoms:

Many pepper plant disorders have similar symptoms such as diseases, nutritional deficiencies, pests, excessive cold or heat. It's a challenge to differentiate and identify the disorders by automatic plant disease detection techniques.

D. Similarity of texture pattern for infection and healthy region.

1.6 Aim of The Thesis

The main aim of this thesis is to design and implement an efficient system for classification of pepper plant diseases (leaves and fruits) based on image processing and machine learning techniques, by using real dataset for pepper (leaves and fruits) images , and identify the problems in this dataset that effect on accuracy of classification. This system take one step towards promoting the farmers to do the smart farming, and allowing them to take decisions for a better yield by making them capable to take the necessary preventive and corrective action on their pepper crop.

1.7 Contributions

Create new dataset for pepper plant (DiyalaPepper)[78], which consist of 244 image (78 for fruits and 166 for leaves) and face some challenges, including :

A. Collection of data set

To acquire images of pepper plant, one has to travel to different places. Data collection will be a challenging since variety of pepper plant diseases may not be available at some farms and diseases occur only during certain seasons.

B. Image capture conditions

Automatic plant disease detection systems give steady and efficient results, only if all the images are captured under same conditions. Capturing images under same conditions is possible only inside laboratories. It’s a challenge to working on images captured under different conditions in the any ordinary room because of uncontrollable environment.

1.8 Thesis Organization

Beside this chapter, the remaining parts of this thesis include the following chapters:

Chapter Two: ***Theoretical Background***

It presents an overview of machine learning . Also, it illustrates the basic principles and the scientific theories used when building a system about the classification of plant diseases.

Chapter Three: ***Design And Implementation Of The Proposed System***

This chapter introduces the steps of the proposed system, describes the developed algorithms to execute the system.

Chapter Four: ***Experimental Results***

This chapter presents the experiments and the results which are obtained from the system running and the performance measures of the tests results.

Chapter Five: ***Conclusions and Suggestions for Future Works***

This chapter gives a list of conclusions derived from the results of the presented work and some suggestions for future works.