Identification of Plant Diseases Using Gray-Level and Color-Based Features of Leaf Images

Ibrahim Isah

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Electrical and Electronic Engineering

Eastern Mediterranean University
July 2020
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Prof. Dr. Ali Hakan Ulusoy
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science in Electrical and Electronic Engineering.

Assoc. Prof. Dr. Rasime Uyguroğlu
Chair, Department of Electrical and Electronic Engineering

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Master of Science in Electrical and Electronic Engineering.

Assoc. Prof. Dr. Önsen Toygar
Supervisor

Examining Committee

1. Assoc. Prof. Dr. Önsen Toygar

2. Assoc. Prof. Dr. Rasime Uyguroğlu

3. Asst. Prof. Dr. Mehtap Köse Ulukök

ABSTRACT

The detection of plant diseases is a vital factor in agricultural production worldwide, which if ignored, can lead to tremendous losses of plant products and revenue. Farmers and researchers from many centuries ago have learnt to identify some plant disease manually by inspection, but presently, technological development have advanced cultivation to an industrial scale, therefore detection of plant diseases has also become a great issue of concern as the farmers may be unable to identify the diseases, their point of origin or even the infected plants early enough. This can lead to a disease outbreak. Early detection of plant diseases can immensely reduce or avoid massive potential losses as it will provide the opportunity for active and cautionary measures. In view of the aforementioned issue, carrying out researches on different ways and methods to curb this problem is a vital necessity.

This thesis study employs the application of computer vision and image processing techniques for plant disease identification. Principal Component Analysis (PCA), Local Binary Patterns (LBP) and Completed Local Binary Patterns (CLBP) feature extraction methods are used for the extraction of texture-based and appearance-based image features. Disease symptoms are analyzed and identified from four different plant leaves to evaluate the performance of the proposed method. We propose a method that incorporates Feature-Level Fusion of the gray level features and color based features using PCA and LBP methods to create a robust system. The proposed method has proven to be more robust compared to the individual systems using LBP and CLBP.

Experiments are conducted on PlantVillage dataset due to its diversified collection of plant leaves. Furthermore, two classifiers are used for classification purposes namely k-Nearest Neighbor (k-NN) and Support Vector Machine (SVM). At the end of the empirical evaluations, a comparative study is presented.

Keywords: plant disease identification, leaf images, color spaces, Feature-Level Fusion, feature extraction, texture-based features, appearance-based features.

ÖZ

Bitki hastalıklarının tespiti, dünya çapındaki tarımsal üretimde hayati bir faktördür ve göz ardı edilirse çok büyük çapta bitkisel ürün ve gelir kaybına yol açabilir. Çiftçiler ve araştırmacılar yüzyıllardan beri bazı bitki hastalıklarını gözlemleyerek tanımlamışlardır. Ancak günümüzde teknolojik ilerlemelerle bitki yetiştirme endüstriyel bir ölçekte yapılmaktadır. Böylece bitki hastalıklarının tespiti büyük ilgi uyandırmıştır çünkü çiftçilerin bitki hastalıklarını tanımlaması, başlangıç noktasını bulmaları ve hastalıklı bitkileri tespit etmeleri mümkün olmayabilir. Bu da bir salgın hastalığa sebep olabilir. Bitki hastalıkların erken tespiti, aktif ve uyarıcı tedbirlere fırsat verdiği için büyük potansiyel kayıpları etkili bir şekilde azaltacak veya yok edecektir. Bahsedilen görüş doğrutusunda farklı yaklaşımlar ve yöntemlerle bu problemi çözmek için araştırmalar yapmak hayati bir zorunluluktur.

Bu tez çalışması, bilgisayarla görü ve görüntü işleme uygulamalarını kullanmaktadır. Ana Bileşenler Analizi (PCA), Yerel İkili Örüntü (LBP) ve Tamamlanmış Yerel İkili Örüntü (CLBP) öznitelik çıkarma yöntemleri, dokuya-bağlı ve görünüm tabanlı özniteliklerin saptanmasında kullanılmıştır. Hastalık semptomları, önerilen yöntemin performansını değerlendirmek için dört farklı çeşit bitkinin yapraklarından analiz edilmiş ve tanımlanmıştır. Önerilen yöntemde, PCA ve LBP kullanılarak elde edilen renksiz ve renkli öznitelikler birleştirilerek güçlü bir yöntem yaratılmıştır. Önerilen yöntemin LBP ve CLBP kullanan bireysel sistemlere göre daha güçlü olduğu ispatlanmıştır.

Deneyler, çeşitli bitki yaprakları koleksiyonundan oluşan PlantVillage veri kümesi üzerinde yapılmıştır. Ayrıca, sınıflandırma yöntemi olarak k-En Yakın Komşu (k-NN) ve Destek Vektör Makinesi (SVM) sınıflandırıcıları kullanılmıştır. Deney sonuçlarının değerlendirilmesinden sonra karşılaştırma çalışması da sunulmuştur.

Anahtar kelimeler: bitki hastalıklarının tanımlanması, yaprak görüntüleri, renk uzayları, Karar-Seviyesi Kaynaşımı, öznitelik çıkarımı, dokuya-bağlı öznitelikler, görünüm tabanlı öznitelikler.

DEDICATION

This work is dedicated to my Family for their continuous love, support and encouragement.

ACKNOWLEDGMENT

First and foremost, I have to thank all mighty Allah for protecting, guiding and giving me the opportunity to be successful in all my endeavors.

I would like to wholeheartedly thank my parents, my sisters, and my entire family for their love, support and inspiration throughout my existence.

I would like to sincerely thank my supervisor Assoc. Prof. Dr. Önsen Toygar for her guidance, support, patience, and understanding and especially for her confidence in me throughout this study.

To all my friends, my colleagues, both at home and around the globe, listing all the names would possibly leave no space for this thesis, thank you for your understanding, encouragement, inspiration and for making all the chaos and hard times become a fun memory worth remembering in the future.

Finally, I would like to say thank you to the whole of Electrical Engineering, Computer Engineering and Eastern Mediterranean University at large for providing a friendly environment equipped with all the required resources to prepare me and thousands of others for this day.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CLBP	Completed Local Binary Pattern
HIS	Hue, Saturation and Intensity Channels
k-NN	k-Nearest Neighbor
LBP	Local Binary Patterns
LBPH	Local Binary Patterns Histogram
LBP_{u2}	Local Binary Patterns uniform pattern
LDA	Linear Discriminant Analysis
PCA	Principal Component Analysis
RGB	Red, Green, and Blue Channels
SVM	Support Vector Machine
YCbCr	Luma, Blue difference and Red difference Chroma

Chapter 1

INTRODUCTION

The Food and Agriculture Organization (FAO) on 2 December 2019 in Rome, launched the United Nations' International Year of Plant Health (IYPH) for 2020, aimed at raising global awareness regarding how plant health protection can help reduce hunger, combat poverty, protect the environment and accelerate economic growth. Plants make up 80% of the food we consume, and they account for about 98% of the oxygen we ingest. However, the interminable impact of pests and diseases are persistently on the rise.

Each year, plant pests and diseases are the cause of 40 percent of global food crops loss. This leads to more than $220 billion in annual agricultural trade losses, which in turn exposes millions of people to hunger, and severely damages agriculture – the primary source of income for poor rural communities [1].

This is why policies, research and actions to promote plant health are fundamental for reaching the Sustainable Development Goals.

Plant disease have the ability to evolve in different parts of a plant including roots, stems and leaves. Their surface area and color make them most visible to the human eyes. Farmers and experts exploit this feature to access and analyze the health conditions of their crops [2]. Today, with the development of computer vision and

robustness of computer algorithms in various aspects of life, innovative and exciting methods are used to classify, track and diagnose various crop diseases. These techniques and methods are affordable, convenient, more efficient and more accurate than the conventional methods.

This thesis aims to identify and compare the performance of the proposed method and existing research on computer aided early detection of diseases using various image texture-based algorithms for detection and classification of plant diseases using PlantVillage dataset [3]. PlantVillage dataset is an online platform that provides an avenue for researchers, farmers or any interested entity to get knowledge and tools on crop health. It has over 50,000 professionally collected images of over 150 crops with more than 1,800 diseases. The dataset comprises of numerous plant leaves with specific symptoms of a disease. In this thesis the plants studied are apple, corn, grape and tomato.

1.1 Statement of the Problem

Plant diseases have in several ways caused tremendous losses to humans. A popular example is the Irish famine caused by late blight disease of potatoes (by infestans of Phytophthora) which entered Europe in1845, spreading quickly from Belgium to other mainland European countries, and eventually to Great Britain and Ireland. The populace in Ireland at that time, was heavily reliant on potato for their livelihood as compared to other parts of Europe (Bourke, 1964; Reader, 2009), the severity of the epidemic reached disastrous levels. The subsequent Great Famine killed around 1 million people, and an additional million were forced to leave the island [4].

On the other hand, the near annihilation of the American chestnut by chestnut blight (caused by Cryphonectria parasitica) almost lead to the loss of a valued resource. An additional example of a direct impact on the economy is that of southern corn leaf blight disease (caused by Cochliobolus maydis, anamorphous Bipolaris maydis) which approximated to one billion dollars loss in a year by American corn growers [5]. Plant diseases when considered individually may cause less drastic losses annually throughout the world but when collectively analyzed, constitute sizable losses to farmers and can reduce the aesthetic values of landscape plants and home gardens.

As such, research shows that from the time that the first farmers began to cultivate plants, people have been concerned with reducing the crop losses caused by plant diseases [6]. Through the evolution of our crop production technologies over the last ten thousand years, the principles of plant disease detection have been woven into the fabric of our civilization. The decisions of when, where, and what to plant and the development of specific cultural practices have been based on countless generations

of trial and error. Without a doubt, the successful farming methods depended upon being able to detect the presence of plant diseases early enough so as to provide necessary care and attention to the plant and eliminate the pest or pathogens that cause this damage.

1.2 Relevance of the Study

This study uses different techniques of computer vision and image processing to detect and classify plant disease using images of the leaves. Accurately detecting these diseases and identifying the infected plants by their leaves can greatly minimize the amount of loss of farm produce to disease infections and also prevent a spread in the case of contagious diseases. This would immensely reduce the pressure on experts and farmers who would have had to carry out a manual inspection. It demonstrates the achievements of the application of computer vision in plant disease detection [7].

1.3 Plant Disease Signs and Symptoms

Plant-affected diseases emanate from non-living or living factors. The abiotic components known as non-living components include climatically induced situations, contaminated water sources, deficient or surplus liquid for plant use, air and soil chemical substances, nutritional deficiencies, while the Biotic components (living) are usually influenced by unwanted weeds, pests, and pathogenic organisms or microorganisms (chromists, fungi, viruses, nematodes, phytoplasms, and bacteria). Therefore, by critically analyzing factors such as ease of spread, ability to exist in diverse forms and complexity in extermination, biotic-borne pathogens continue to present a deleterious threat to the lifespan of plants survival and harvest quality. The three most commonly considered defects for identifying and classifying plant diseases are spots (caused either by fungi or bacteria), mildew, and rust [8].

Every living organism on earth exhibit or react in a particular way when in a condition or situation that deviates from the normal state of being. For example, when the human skin goes red or develops a rash could be due to an allergic reaction or an early indication of an underlying ailment. Plants are not excluded; the leaves in many instances serve as our gateway for diagnosing a lot of diseases in plants. For example, the early blight disease of tomato leads to the appearance of small dark spots that expand into circular plaques made up of rings that circumnutates on the leaves [9]. This in turn, results in premature defoliation of the leaves and heavy losses in yield. The Figure 1 shows a healthy tomato leaf on the left and a diseased leaf affected by early blight on the right.

a. Healthy Tomato leaf b. Early Blight Disease of Tomato Leaf

Figure 1: Healthy and Diseased Tomato Leaves

The fungus, Guignardia bidwellii is responsible for the black rot disease of grapes. It is usually common in regions of wet, warm and humid climate as this provides a conducive situation for spore germination and infection. This disease spreads when the spores are carried by wind or splashed by rain onto the surfaces of developing plant tissue. This goes on for as long as the environmental conditions remain suitable. Black rot can be identified when round, tan plaques with dark purple to brown edges are spotted on the leaves. Critical infections may result in leaf deformity, wilting of the

leaves [10]. Figure 2 shows a healthy grape leaf on the left and a diseased leaf affected by Black rot on the right.

a. Healthy Grape Leaf

b. Black Rot of Grape Leaf

Figure 2: Healthy and Diseased Grape Leaves

Puccinia Sorghi Schwein is the pathogen (fungus) responsible for the popular disease of maize known as Common Rust. Early plaques mostly occur in clusters and are circular. But as the plaques ripen, the fungus protrudes through the foliage surface and the plaques elongate with time.

The characteristic symptom observed on the maize leaves are Brownish-red oblong pustules, plaques of common rust occur on both the upper and lower surfaces of the leaves and are spread sporadically along the leaves. Spores are transported by wind with new infections ensuing weekly or bi-weekly. One plaque is capable of producing both brownish-red urediniospores and black teliospores, yet lastly, only black teliospores will be seen within the plaque [11]. Figure 3 shows a healthy and a diseased corn leaf.

a. Healthy Corn leaf

b. Common Rust of Corn leaf

Figure 3: Healthy and Diseased Corn Leaves

Cedar apple rust is a member of the Pucciniaceae family; a class of fungi with several species that typically need two or more host to complete their life cycle. Member of this class are known as rust, seeing as at some stage in their evolution, many are orange or reddish in color. The fungus spreads through the leaves and develops aecia beneath the leaves. The aecia produces aeciospores which are wind-blown back to the redcedars. They afterwards germinate and begin gall formation which produce telial horns to restart the process. A heavily infested apple tree can take on a yellowish cast from multiple plaques on the leaves. The figures below show a healthy apple leaf and a cedar apple rust infested leaf. Figure 4 shows a healthy and a diseased apple leaf.

a. Healthy apple leaf

b. Cedar apple rust infected leaf

Figure 4: Healthy and diseased Apple leaves

1.4 Commercial Crops, Food Crops and Fruits

Commercial crops are crops that are cultivated predominantly for the purpose of generating revenue. Cash crops like coffee, cacao, wheat, maize, soybeans, rice and several more, provide direct source of revenue as commodity for export and employment prospects to the rural economy and nations at large. Additionally, farmers get the opportunity to generate funds for better maintenance and innovation. Cash crops facilitate the growth and development of institutions that allow more commercialization [12] while food crops, on the other hand, are crops which contribute to the essential food supply of the entire world. They include grains, which serve as the principal source of food for the bulk of the population as they are consumed as the daily meal. This includes corn, wheat, barley, seeds and nuts, which serve as the powerhouse of energy and a concentrated form of food; herbs, which are mainly comprised of leaves, stems, flowers, roots and seeds; fruits such as oranges, grapes, apples; and vegetables.

1.5 Review of Previous Work

Sabrol and Satish [13] in 2016 proposed a method using classification tree, to classify five types of tomato diseases, namely tomato late blight, Septoria spot, bacterial spot, bacterial canker, tomato leaf curl and healthy tomato using the plant leaf and stem images. The classification was conducted by extracting color, shape and texture features from healthy and unhealthy tomato plant images. The plant images were segmented, afterwards, feature extraction process was carried out on the segmented images. The extracted features were then fed to classification tree. Finally, the disease classification was based on these six different class types. The classification of the six types of tomato images yielded overall 97.3% of classification accuracy.

Kulkarni et al. [14] in 2012 presented a method for detection of plant diseases, using Artificial Neural Network (ANN) and image processing techniques. They were able to achieve this objective, by capturing, filtering and segmenting the images using Gabor filter. Then, color and texture features were extracted from the segmented images. Training was then carried out by ANN on the features that could accurately distinguish between the healthy and diseased samples. Experimental results showed that classification performance by ANN was 91%.

Singh and Misra [15] proposed an algorithm for automatic detection and classification of plant leaf diseases through image segmentation technique using Genetic Algorithm. Color co-occurrence method was used for feature extraction to enable them consider both the texture and color of the image. It also covered survey on different diseases classification techniques that can be used for plant leaf disease detection. The average accuracy of classification of proposed algorithm was 97.6%.

In 2018, Maheswari et al. [9] proposed a method for identification of early blight fungal disease of tomato caused by the pathogen Alternaria solani using fold scope and quadratic Support Vector Machine (SVM). It is quoted that a fold scope is a paper microscope which was brought in to existence by an engineer, Manu Prakash and his colleagues. The images of the tomato leaves were acquired using this set up of using a high resolution mobile phone to give a magnified image. These images were classified with the help of several machine learning algorithms which include LDA, k-NN and linear SVM. The proposed method yielded an accuracy of 91% and 89% for training and prediction respectively as compared to the other methods.

1.6 Thesis Contribution

This study involves the supervised extraction of texture-based and appearance-based image features. Agricultural productivity is so important in countries in which their economy is highly dependent on agriculture. It is estimated that every year, 30 to 40 percent of crops are wasted through the production chain [16]. Losses from diseases also have a serious economic effect, causing a decline in income for crop farmers, consumers and distributors. Several experiments have been carried out in different locations under altered environmental conditions to quantify the losses due to various diseases [16]. In this respect, it is vital to detect and identify plant diseases in their initial stage with the help of computer vision and image processing techniques.

This thesis comprises an investigation on various plant diseases that are observed on corn, grape, orange and tomato leaves and how we can make use of supervised extraction of features using appearance-based and texture-based methods on gray-level and colored leaf images to implement an identification and classification algorithm. RGB, HSI and YCbCr color spaces are used for color image experiments. The experiments are conducted on PlantVillage dataset of leaf images. Comparative results on gray-level and colored leaf images are presented. Eighteen diseases in total are analyzed against the healthy counterpart of the four plants selected for this thesis. The experiments are carried out in four phases, with the first phase being a collection of images from the database which is already stratified as stated earlier. The second phase involves pre-processing the images. This is done with the help of some MATLAB commands to crop, resize or convert to grayscale as required [17]. The third phase is the extraction of relevant features and finally, classification of the diseases based on the analyzed features.

The algorithms utilized for extracting the appearance-based and texture-based features from the images in this thesis include Principal Component Analysis (PCA) [18], Local Binary Patterns (LBP) [19] and Completed Local Binary Patterns (CLBP) [20].

The remaining part of the thesis is ordered as follows. Chapter 2 discusses the general structure of the plant disease detection algorithm. Chapter 3 sets out methodologies implemented for the feature extraction. It also contains details about the strategies used to ascertain the classification performance. Chapter 4 offers the proposed model structural framework in detail and Chapter 5 is devoted to empirical analysis and results. Finally, Chapter 6 encompasses conclusions and inferences.

Chapter 2

GENERAL STRUCTURE OF PLANT DISEASE DETECTION ALGORITHM

2.1 Introduction

In the following section, we discuss the mainstream method employed in plant disease identification of an image-based computer aided system. The general structure of plant disease identification is shown in Figure 5.

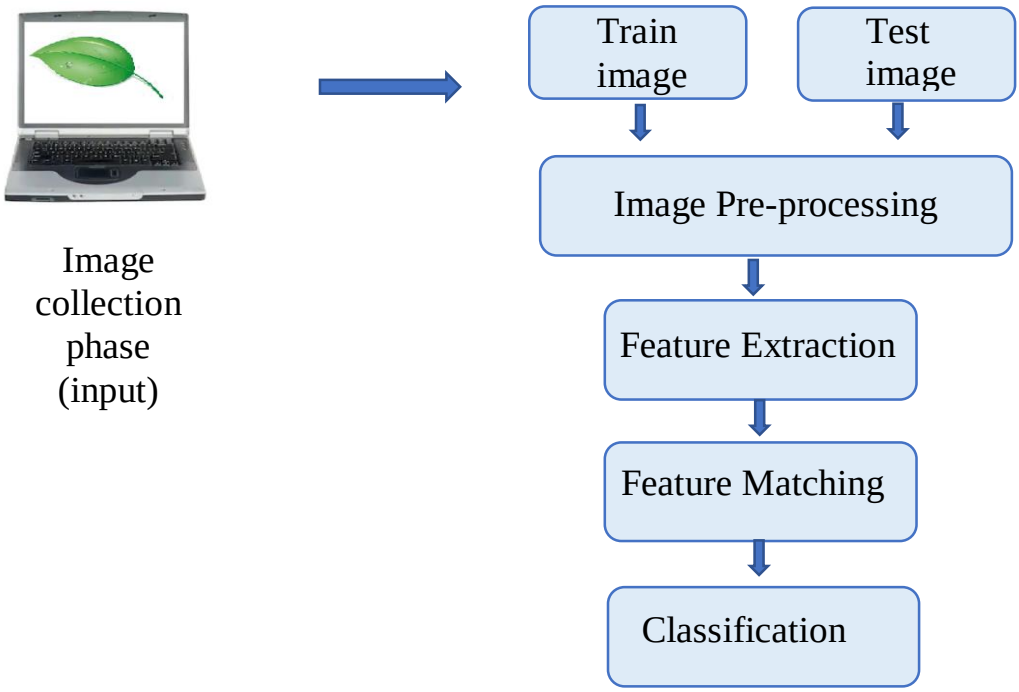

Figure 5: General Structure of Plant Disease Identification Algorithm.

The pipeline starts with image collection. This can be executed by the use of a high resolution camera or an ordinary mobile phone camera to create a database of plant leaves for analysis. But for the purpose of this research, PlantVillage dataset is used. The rest of the steps are described in the following subsections.

2.2 Image Database and Collection of Training and Testing Samples

The database (PlantVillage dataset) comprises of images of apple, cherry, corn, grape, orange, peach, pepper, potato, raspberry, soybean, squash, strawberry and tomato leaves grouped as colored, gray-scale and segmented images. However, in this thesis, images of tomato, grape, corn and apple are considered. For the three plants namely apple, corn and grape; three diseases are chosen along with a healthy counterpart making a total of 4 groups and 100 pictures are chosen for training. Another set of 100 pictures corresponding to the aforementioned category are chosen for testing the model. But in the case of tomato, 9 tomato diseases are considered along with its healthy counterpart with 100 images for each category as in the former. Tables 1 to 4 illustrate how the selected plants are stratified in the data collection phase for this analysis.

Table 1: Apple Leaf Image Data

Apple Leaf Dataset	Train set	Test set
Sample set A1 (Apple scab)	100	100
Sample set A2 (Black rot)	100	100
Sample set A3 (Cedar apple rust)	100	100
Sample set A4 (Healthy)	100	100
Sample set A5 (Total)	400	400

Table 2: Corn Leaf Image Data

Corn Leaf Dataset	Train set	Test set
Sample set C1 (Cercospora leaf spot)	100	100
Sample set C2 (Northern leaf blight)	100	100
Sample set C3 (Common Rust)	100	100
Sample set C4 (Healthy)	100	100
Sample set C5 (Total)	400	400

Table 3: Grape Leaf Image Data

Grape Leaf Dataset	Train set	Test set
Sample set G1 (Leaf Blight)	100	100
Sample set G2 Ecsa (Black measles)	100	100
Sample set G3 (Black Rot)	100	100
Sample set G4 (Healthy)	100	100
Sample set G5 (Total)	400	400

Table 4: Tomato Leaf Image Data

Tomato Leaf Dataset	Train set	Test set
Sample set T1 (Yellow leaf curl)	100	100
Sample set T2 Ecsa(Tomato Mosaic Virus)	100	100
Sample set T3 (target Spot)	100	100
Sample set T4 (Spider Mites)	100	100
Sample set T5 (Septoria leaf spot)	100	100
Sample set T6 (Leaf Mold)	100	100
Sample set T7 (Late Blight)	100	100
Sample set T8 (Early Blight)	100	100
Sample set T9 (Bacterial spot)	100	100
Sample set T10(Healthy)	100	100
Sample set T11 (Total)	1000	1000

2.3 Data Pre-processing

Data pre-processing is a mandatory step to achieve meaningful results in designing a detection model. As a result of the different methods used in image collection, some modalities call for special pre-processing strategies (e.g. intensity normalization,

segmentation, image resizing, cropping, etc.) to get rid of foreign artifacts that are not initially part of the picture or area of interest.

In this research, our preprocessing phase is peculiar to those mandatory for clear identification of leaves which include resizing the image to a specific scale or dimension in order to efficiently extract the pertinent features from the images as plant leaves are not limited to a specific shape or size. Some algorithms may further require that colored images be converted to grayscale to work efficiently. This is also implemented where applicable.

2.4 Feature Extraction

Feature extraction is an approach used to obtain distinctive characteristics from a picture in such a manner that these characteristics delineate a highly-descriptive information in a lower dimensional structure. These features may be global, local or a combinational representation of the picture such as color, shape, regions or edges.
Feature fusion is a method that incorporates discrepant relevant features to produce a more refined descriptor of a picture. There are different approaches and algorithms purposely for feature extraction, but for the purpose of this thesis, Principal Component Analysis [18], Local Binary Pattern [19] and Completed Local Binary Pattern [20] are implemented as they can easily be fine-tuned mathematically by vector manipulations for efficient problem solving in computer vision.

These features afterwards serve as input to various types of machine learning algorithms (e.g. Support Vector Machines (SVM), k-Nearest-Neighbor (k-NN), K-means Clustering, Linear Discriminant Analysis (LDA), Fuzzy Logic, Neural Networks, Logistic regression, Naïve Bayes, etc.) for matching and classification.

2.5 Classification and Matching

In supervised learning, the algorithm is trained on datasets that contain examples associated with items which are correctly labeled. This enhances its capability to come to an accurate inference when provided with a new set of data. In this research, the plant diseases are considered as the class labels of the images. Such classifiers fall under supervised, unsupervised or semi-supervised techniques of machine learning [21]. k-Nearest-Neighbor and Support Vector Machine classifiers are used in this thesis for classification by a measure of Manhattan distance between the features.

The k-Nearest Neighbors algorithm is an instance-based, or a lazy method of machine learning. It has been considered as one of the easiest of all machine learning algorithms [22]. The justification of k-NN is that samples with some level of similarity belonging to the same class is highly probable, while the main concept of k-NN algorithm is to first select k closest neighbors to each test observation, followed by using the learnt k nearest neighbors to predict this test observation. As such, k-NN algorithm is perhaps conceived of as an algorithm, which needs no explicit training [23].

Chapter 3

METHODOLOGY

3.1 Introduction

This chapter deals with the methodology adopted for the extraction of features by the algorithms in this thesis. The algorithms are Principal Component Analysis [18], Local Binary Patterns [19] and Completed Local Binary Patterns [20].

3.2 Principal Component Analysis

Principal Component Analysis is a multivariate methodology that evaluates data in which many interrelated quantitative dependent variables explain the observations. Its objective is to extract relevant data from the statistical data to convey it as a set of new orthogonal variables known as principal components, and to view the similarity pattern between observations and variables as points in spot maps. Mathematically, PCA relies on the eigen-decomposition of positive semi-definite matrices and on the Singular Value Decomposition (SVD) of rectangular matrices. It is determined by eigenvectors and eigenvalues which are numbers and vectors associated to square matrices; collectively they provide the eigen-decomposition of a matrix, which analyzes the structure of that matrix such as correlation, covariance or cross-product matrices.

The directions with the most variation of data is the path of projection which Principal Component Analysis takes. The largest eigenvalues corresponding to eigen vectors determines the course of projection. The variance of the information along the eigen vector’s direction is in correspondence to the magnitude of eigen values.

3.2.1 Principal Component Analysis Algorithm

Performing PCA is quite simple in practice. We assume that leaf image I (x, y) is the size of M x N gray image, then each column are linked together to form a D = M x N columns vector. The dimensions of the leaf image vector is D, the dimensions of the image space is also D. Then, we assign the number of training observations n and x_i to represent the column vector of the i^{th} leaf image, the entire training sample's covariance matrix (*S*) can be calculated with the following equation:

$$s = \sum_{i=1}^{N} (X_i - m)\,(X_i - m)^T \tag{3.1}$$

$$and \quad m = \frac{1}{N}\sum_{i=1}^{N} X_i \tag{3.2}$$

Where m is the mean image of the training set and X_i is each sample of the training set. Let A = [x_1-m, x_2-m, …. x_n-m], then s = AA^T. The eigenvalues and eigenvectors are then calculated from the covariance matrix. Let p = (p_1,p_2,….p_r) (r>N) be the r normalized eigenvectors corresponding to r largest eigenvalues. Each leaf image of the training set is projected into the eigenspace to obtain its corresponding eigenspace based feature Y_k, which is defined by the equation as:

$$Y_k = p^T \varsigma_k \; ; \quad k = 2, \ldots . N \tag{3.3}$$

where p^T is the eigen vectors and ς_k is the scaled data. PCA extracts the global grayscale features of a whole image and the global features are useful and important [24]. It was found that there were many fascinating applications of PCA, out of which in our day to day life, knowingly or unknowingly multivariate data analysis and image compression are being used as a substitute for either [25].

3.3 Local Binary Patterns (LBP)

Local Binary Patterns approach is a popular descriptor in texture analysis which was discovered by Ojala et al. [19] in 1996. The various variations of this algorithm makes it efficient when used for feature extraction. Using the neighboring pixels, every pixel

is represented by binary code. LBP uses only sign component [26]. Assuming the lightening angle varies from one picture to another, LBP will still effectively extract the features. Its simplicity also allows for models to be adjusted or combined with other descriptors to adapt to different types of classification algorithms.

3.3.1 Local Binary Patterns Algorithm

The pixel values of eight neighbors are defined by the value of the center pixel, after which the threshold binary values are weighted by powers of 2 and afterwards summed to obtain the LBP code of the center pixel. Figure 6 shows an illustration of the LBP operator. Let's assume g_c and g_p signify accordingly the gray values of the center and its eight-neighbor pixels, then the LBP code for the center pixel with coordinate (x, y) is calculated as illustrated in equation (3.3).

$$\mathrm{LBP}(x, y) = \sum_{p=0}^{7} s(g_c - g_p) \tag{3.4}$$

where s(z) is the threshold function as

$$s(z) = \begin{cases} 1, & z \geq 0 \\ 0, & z < 0 \end{cases} \tag{3.5}$$

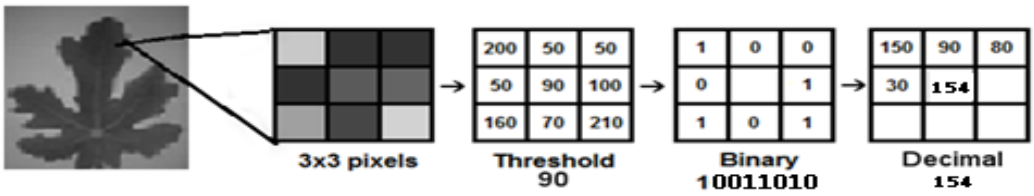

Figure 6: LBP Algorithm

LBP histogram (LBPH) is used in such a way that, the LBP codes of all pixels for an input picture are compiled into one histogram to form a texture descriptor. Figure 7 illustrates this phenomenon. Equation (3.5) is used to obtain the statistical representation of LBPH.

$$\text{LBPH(i)} = \sum_{x,y} \delta\{\text{i}, \text{LBP(x, y)}\}, \text{i} = 0, \cdots 27 \quad (3.6)$$

where δ (.) is the Kroneck product function.

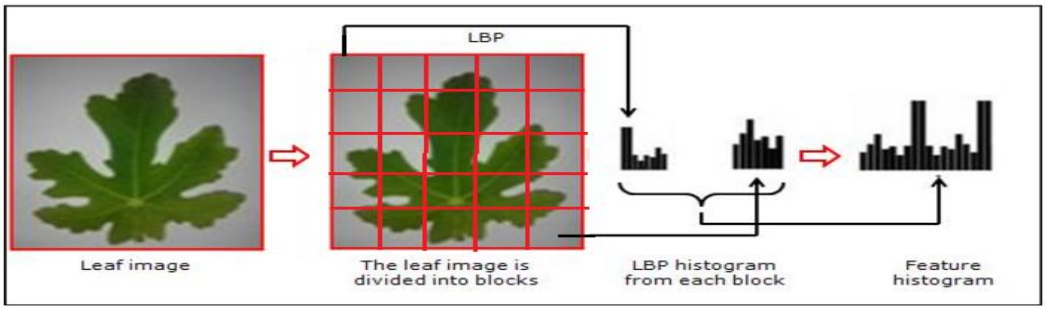

Figure 7: LBP Histogram (LBPH).

The use of neighborhoods of varying sizes is one of the extensions of LBP [27]. The extension will be capable of taking any radius and neighbors around a centered pixel denoted by $LBP_{P,R}$ using a circular neighborhood and bilinear interpolation where the sampling point is not located in the middle.

Example of these extensions include $LBP_{16,2}$ which implies 16 neighbors in a neighborhood of radius 2, and LBP_{u2} implies uniform patterns if it contains at most two bitwise transitions from 0 to1 or the reverse when a circular binary string is considered. For LBPH measurements, the uniform patterns are used in such a way that every uniform pattern has a single bin and all non-uniform patterns are assigned to a standalone bin, so, with 8 neighbors, the regular number of bins for LBPH is 256 and 59 for uniform patterns. With 16 neighbors, the number of bins are 65,536 and 243 respectively. Figure 8 shows the different radii and neighbor configuration of LBP algorithm.

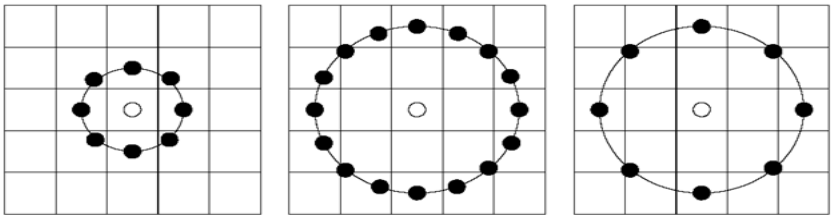

Figure 8: An Example with Different Radii and Neighbors.

3.4 Completed Local Binary Patterns (CLBP)

Completed Local Binary Patterns [28] is an extension of LBP. It incorporates three units measurement to extract features from an image. The first one is CLBP-Sign (CLBP-S) which indicates the local disparity between the gray levels. The second is CLBP-Magnitude (CLBP-M) which indicates the magnitude or slope descriptor. Finally the last unit is CLBP-Center (CLBP- C) which contains the central pixel gray level value. The equations below illustrate how it operates:

$$\mathrm{CLBP}_{\mathrm{Sp,\,r}} = \sum_{n=0}^{p-1} s(g_n - g_c)2^n \tag{3.7}$$

$$\mathrm{CLBP}_{\mathrm{Mp,\,r}} = \sum_{n=0}^{p-1} s(D_n - \mathrm{T})2^n \tag{3.8}$$

$$\mathrm{CLBP}_{\mathrm{Cp,\,r}} = \mathrm{s}(g_c - g_n) \tag{3.9}$$

where,

$$\mathrm{D_n} = (g_n - g_c) \tag{3.10}$$

$$\mathrm{T} = \frac{1}{N}\sum_{i=0}^{n-1} \quad \frac{1}{p}\sum_{n=0}^{p-1}(g_n - g_c) \tag{3.11}$$

$$\mathrm{G_{N,\,r}} = \frac{1}{n}\sum_{i=0}^{N-1} g_i \tag{3.12}$$

where $\mathrm{D_n}$ is the difference between a given pixel g_n and its neighbors g_c, T is the threshold to be determined adaptively, $\mathrm{G_{N,r}}$ is the local intensity variant.

3.4.1 Completed Local Binary Patterns Algorithm

Completed Local Binary Patterns is a combination of value of center gray level (CLBP-C) and a Local Difference Sign-Magnitude Transform (LDSMT). The pixel in

the center depicts the value of gray level. The LDSMT comprises of two other values which are CLBPS and CLBPM. Equation (3.7) to (3.10) illustrates how CLBP_S and CLBP_M are combined. Figure 9 is a graphical illustration of CLBP and how its components are combined.

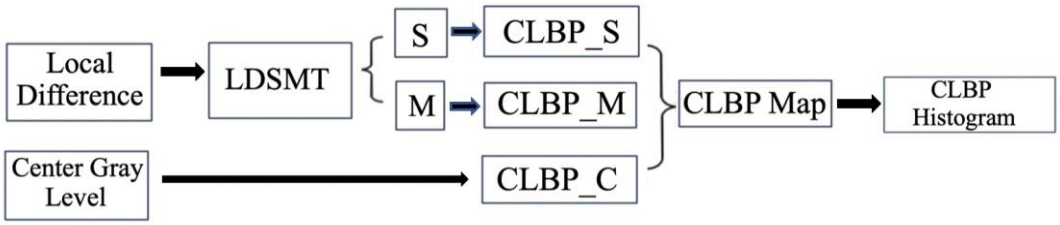

Figure 9: Framework of Completed Local Binary Patterns Algorithm

d_r is the local difference vector and can be calculated as follows:

$$d_r = S_r * a \begin{cases} s(d_r) \\ m_r = |d_r| \end{cases} \quad (3.13)$$

where s_r is the sign component, m_r is the magnitude component and s_p is determined as follows:

$$S_p = \begin{cases} 1, & d_r \geq 0 \\ -1, & d_r < 0 \end{cases} \quad (3.14)$$

where the local difference vector d_r=[d0,……. dr−1] which can be constructed from the following vectors:

$$\text{sign vector } S_r = [S_{0,\ldots\ldots} \; S_{r-1}] \quad (3.15)$$

$$\text{magnitude vector } m_r = [m_{0,\ldots\ldots} \; m_{r-1}] \quad (3.16)$$

The major alteration between LBP and CLBP is that LBP descriptor uses only the sign component for feature comparison. In many cases, considering sign component alone is not sufficient to design a robust matching system. In contrast, the CLBP uses

traditional structure magnitudes along with the sign characteristic. This is a more refined approach to maximize the algorithm's accuracy.

Chapter 4

PROPOSED MODEL FRAMEWORK

The proposed method is a combination of two or more methods with the sole objective of achieving a higher performance of the system. Feature-Level Fusion has been proven to give higher performance accuracy and provide a more robust recognition system [29]. In this thesis, PCA features are fused or concatenated with LBP features of the Cb component of the YCbCr color space and this method provides better accuracy compared to using any of the individual methods alone.

4.1 Model Overview

The proposed model consists of pre-processing, feature extraction, normalization, Feature-Level Fusion, matching and classification stages. A plant image is firstly applied pre-processing and then, the features are extracted using PCA and LBP methods. Then the features are normalized and concatenated with Feature-Level Fusion strategy. Afterwards, matching and classification stages are applied to obtain the final decision of the model. These stages are shown in Figure 10 as the proposed model's framework and each stage is described in the following subsections.

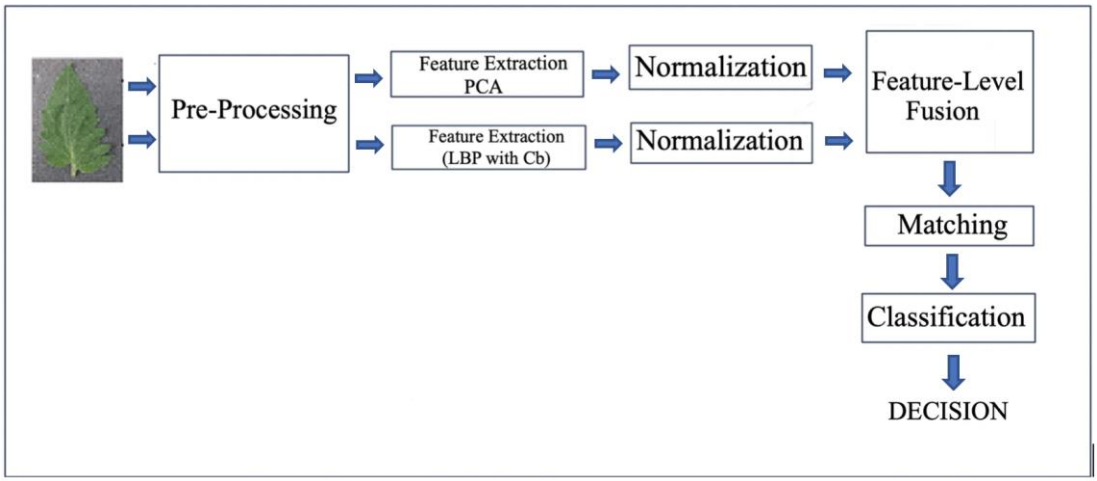

Figure 10: The Proposed Model's Framework.

4.2 Pre-processing Phase of the Proposed Model

The pre-processing phase is carried out on the leaf images of apple, corn, grape and tomato plants. Colored segmented images are resized to a uniform dimension and converted to grayscale for PCA while the original colored version is used for LBP [30]. The images are further divided into 5 by 5 blocks of pixels to be used in LBP. This pre-processing phase is followed by the feature extraction process.

4.3 Feature Extraction and Normalization Phases of the Proposed Model

In the feature extraction phase, feature extraction is done individually for both PCA and LBP. Principal components are first extracted using PCA algorithm after the images have been converted to grayscale. The extraction of LBP features of the Cb component from YCbCr in the 5 by 5 blocks of pixels follows. We then proceed this action by a standard normalization where the mean of the features is calculated, then subtracted from each feature value and then divided by the mean value.

4.4 Feature-Level Fusion

Feature-Level Fusion at this phase is done by simply horizontally concatenating the previously extracted PCA feature vector with LBP feature vector of the respective images. At this point, a new characteristic vector is generated which is a combination of the two individual algorithms, then the matching and classification commences.

4.5 Matching and Classification Phases of the Proposed Model

For the proposed model, SVM classifier is used in the matching and classification phase as it shows a better performance when compared to the k-Nearest Neighbor classifier which will be presented in Chapter 5 in empirical configuration I and II. SVM works by creating a hyper plane that classifies the training vectors into different classes and the hyper plane that gives the maximum margin of separation between the classes is chosen as the best option because this gives a higher classification accuracy.

Chapter 5

EXPERIMENTS AND RESULTS

5.1 Introduction

This section discusses our model's assessment criteria, empirical results and analysis. The experimental setup described in the section below is the quantitative measure of the model with regards to other related works. Findings are extracted and presented on the selection of appropriate configurations, after evaluating the performance of the model on adopted experimentation.

5.2 Empirical Configuration I (with k-NN)

In this thesis, Tables 1 to 4 in Chapter 2 depict the organization of the database illustrating how a total of four plants with a total of 4,400 images were categorized and distributed between training and testing. For simplicity of identification, the plants are represented with letters A1 to A5 for apple, C1 to C5 for corn, G1 to G5 for grape and T1 to T10 for tomato. All the categories corresponding to 1 (i.e. A1, C1, G2, and T1) represent healthy leaves while the rest of the categories are diseased. Therefore just as illustrated in Tables 1 to 4 in Chapter 2, for Train data, the plants of apple, corn and grape have a total of 400 images each, and when distributed between their four categories, each category has 100 images. However, for tomato, there are a total of 1000 images and 10 categories, therefore, each category has 100 images, respectively. The same configuration is maintained for Test data as well.

The first set of results from this section is obtained from application of Principal Component Analysis approach using the aforementioned configuration. The experimental images are imported into Matlab in a vector form, after which the mean image is calculated. This is followed by Normalization of the data. The mean image is then subtracted from the entire data dimensions. This produces a dataset whose mean is zero. Then the covariance matrix is computed after which we also obtain Eigen values and vectors where the maximum variance is computed giving us an idea of the relationship between the magnitude and the variance, based on this relationship, we select the best components to form our feature vector. From this point, Manhattan distance is used to measure the distance between the features after which the features with the least distance corresponding to the images are chosen as a match.

The second set of results is obtained by the application of Local Binary Patterns feature extraction approach. In LBP, conversion of images to grayscale is performed and then both the train and test images are divided into 5x5 of pixels for a smoother operation. LBP algorithm is applied to the individual blocks in order to enumerate the algorithms using circular symmetric pattern for each pixel using (P, R) method. The parameters employed in carrying out these experiments are neighbors with P = 8 and radius with R = 1. Uniform pattern LBPu2 is used to obtain the descriptors of the Local Binary Patterns. After this phase, the 59 column feature is obtained that will be passed to the classifier for prediction.

The third phase involves the use of Completed Local Binary Patterns approach. CLBP is a variation of LBP with 3 units. The first unit is CLBP-Sign (CLBP-S) representing the local disparity among gray levels. The second unit is CLBP-Magnitude (CLBP-M), representing the magnitude or known as slope descriptor. Finally, the last unit is

CLBP-Center (CLBP- C) which carries the central pixel gray-level values. For this case, two of the components CLBP_S and CLBP_M are combined for better performance. Neighbors P = 8 and radius R = 1 are considered for carrying out this experiment and uniform pattern LBPu2 is adopted to obtain the descriptors of the Local Binary Patterns. After this phase, the 59 column feature is obtained that will be passed to the classifier for prediction. At the end of the experimental process, the performance of the model is calculated using the following equation:

$$\text{Accuracy} = \frac{\text{Number of correctly classified images}}{\text{Total number of images}} \times 100 \quad (5.1)$$

5.2.1 Results from Empirical Configuration I

In this section, the experiments are carried out using the three algorithms namely PCA, LBP and CLBP on gray-scale images and the comparative results are presented in Tables 5 to 8.

Table 5: Accuracy (%) for PCA, LBP, and CLBP on Grayscale Apple Dataset.

	Method		
Apple Leaf Dataset	**PCA**	**LBP**	**CLBP**
Dataset A1 (Apple scab)	92.00	67.00	70.00
Dataset A2 (Black rot)	86.14	84.00	88.00
Dataset A3 (Cedar apple rust)	84.00	92.00	89.00
Dataset A4 (Healthy)	82.00	60.00	55.00
Dataset A5 (Total)	86.53	75.75	75.50

Table 6: Accuracy (%) for PCA, LBP, and CLBP on Grayscale Corn Dataset.

	Method		
Corn Leaf Dataset	**PCA**	**LBP**	**CLBP**
Dataset C1 (Cercospora leaf spot)	94.00	80.00	84.00
Dataset C2 (Northern leaf blight)	93.00	56.00	55.00
Dataset C3 (Common Rust)	97.00	99.00	96.00
Dataset C4 (Healthy)	95.00	93.00	87.00
Dataset C5 (Total)	94.00	82.00	80.00

Table 7: Accuracy (%) for PCA, LBP, and CLBP on Grayscale Grape Dataset.

	Method		
Grape Leaf Dataset	**PCA**	**LBP**	**CLBP**
Dataset G1 (Leaf Blight)	96.00	91.00	86.00
Dataset G2 Ecsa (Black Measles)	95.00	75.00	71.00
Dataset G3 (Black Rot)	91.00	77.00	81.00
Dataset G4 (Healthy)	96.00	95.00	96.00
Dataset G5 (Total)	94.00	84.50	83.00

Table 8: Accuracy (%) for PCA, LBP, and CLBP on Grayscale Tomato Dataset.

	Method		
Tomato Leaf Dataset	**PCA**	**LBP**	**CLBP**
Dataset T1 (Yellow Leaf Curl)	91.00	70.00	68.00
Dataset T2 Ecsa (Tomato Mosaic Virus)	97.00	57.00	61.00
Dataset T3 (Target Spot)	92.00	54.00	51.00
Dataset T4 (Spider Mites)	99.00	87.00	83.00
Dataset T5 (Septoria Leaf Spot)	95.00	48.00	56.00
Dataset T6 (Leaf Mold)	97.00	56.00	58.00
Dataset T7 (Late Blight)	96.00	47.00	52.00
Dataset T8 (Early Blight)	99.00	57.00	53.00
Dataset T9 (Bacterial Spot)	97.00	93.00	93.00
Dataset T10(Healthy)	98.00	84.00	91.00
Dataset T11 (Total)	96.10	65.30	66.60

5.3 Empirical Configuration II (with k-NN)

In this section, the experiments are carried out using the three algorithms PCA, LBP and CLBP on colored images in RGB, HIS and YCbCr color spaces and k-NN classifier is employed. The comparative results are tabulated in the following subsections.

5.3.1 Results from Empirical Configuration II

The experimental results on apple, corn, grape and tomato datasets using LBP feature extraction method on HSI color space are presented in Tables 9 to 12.

Table 9: Accuracy (%) for LBP on Segmented HSI Color Space for Apple Dataset.

	Method: LBP			
Apple Leaf Dataset	**Hue channel**	**Saturation channel**	**Intensity channel**	**H+S+I**
Dataset A1 (Apple Scab)	84.00	72.00	83.00	75.00
Dataset A2 (Black Rot)	69.00	81.00	77.00	81.00
Dataset A3 (Cedar Apple Rust)	94.00	98.00	94.00	93.00
Dataset A4 (Healthy)	88.00	63.00	65.00	82.00
Dataset A5 (Total)	83.75	78.50	79.75	84.00

Table 10: Accuracy (%) for LBP on Segmented HSI Color Space for Corn Dataset.

	Method : LBP			
Corn Leaf Dataset	**Hue channel**	**Saturation channel**	**Intensity channel**	**H+S+I**
Dataset C1 (Cercospora leaf spot)	75.00	87.00	85.00	75.00
Dataset C2 (Northern leaf blight)	62.00	52.00	55.00	61.00
Dataset C3 (Common Rust)	98.00	95.00	93.00	95.00
Dataset C4 (Healthy)	96.00	99.00	95.00	98.00
Dataset C5 (Total)	82.75	83.25	82.00	82.25

Table 11: Accuracy (%) for LBP on Segmented HSI Color Space for Grape Dataset.

	Method: LBP			
Grape Leaf Dataset	**Hue channel**	**Saturation channel**	**Intensity channel**	**H+S+I**
Dataset G1 (Leaf Blight)	92.00	85.00	91.00	93.00
Dataset G2 Ecsa (Black measles)	76.00	78.00	74.00	80.00
Dataset G3 (Black Rot)	85.00	80.00	88.00	86.00
Dataset G4 (Healthy)	94.00	94.00	95.00	99.00
Dataset G5 (Total)	86.75	84.25	87.00	89.50

Table 12: Accuracy (%) for LBP on Segmented HSI Color Space for Tomato Dataset.

	Method: LBP			
Tomato Leaf Dataset	**Hue channel**	**Saturation channel**	**Intensity channel**	**H+S+I**
Dataset T1 (Yellow leaf curl)	74.00	67.00	67.00	67.00
Dataset T2 Ecsa(Tomato Mosaic Virus)	91.00	83.00	84.00	82.00
Dataset T3 (target Spot)	76.00	78.00	68.00	74.00
Dataset T4 (Spider Mites)	85.00	80.00	77.00	85.00
Dataset T5 (Septoria leaf spot)	58.00	55.00	53.00	53.00
Dataset T6 (Leaf Mold)	72.00	60.00	55.00	65.00
Dataset T7 (Late Blight)	33.00	46.00	46.00	40.00
Dataset T8 (Early Blight)	45.00	47.00	46.00	43.00
Dataset T9 (Bacterial spot)	79.00	94.00	94.00	91.00
Dataset T10(Healthy)	88.00	85.00	87.00	89.00
Dataset T11 (Total)	70.10	69.50	67.70	68.90

Accuracy of the experiments conducted on apple, corn, grape and tomato datasets using LBP feature extraction method on RGB color space are shown in Tables 13 to 16.

Table 13: Accuracy (%) for LBP on Segmented RGB Color Space for Apple Dataset.

	Method: LBP			
Apple Leaf Dataset	**Red channel**	**Green channel**	**Blue channel**	**R+G+B**
Dataset A1 (Apple scab)	80.00	74.00	69.00	89.00
Dataset A2 (Black rot)	67.00	79.00	81.00	80.00
Dataset A3 (Cedar apple rust)	95.00	94.00	95.00	89.00
Dataset A4 (Healthy)	64.00	63.00	66.00	70.00
Dataset A5 (Total)	76.50	77.50	77.75	82.00

Table 14: Accuracy (%) for LBP on Segmented RGB Color Space for Corn Dataset.

	Method : LBP			
Corn Leaf Dataset	**Red channel**	**Green channel**	**Blue channel**	**R+G+B**
Dataset C1 (Cercospora leaf spot)	88.00	86.00	84.00	82.00
Dataset C2 (Northern leaf blight)	57.00	58.00	53.00	58.00
Dataset C3 (Common Rust)	94.00	91.00	91.00	95.00
Dataset C4 (Healthy)	96.00	97.00	95,19	96.00
Dataset C5 (Total)	83,95	83.21	80.99	82.25

Table 15: Accuracy (%) for LBP on Segmented RGB Color Space for Grape Dataset.

	Method: LBP			
Grape Leaf Dataset	**Red channel**	**Green channel**	**Blue channel**	**R+G+B**
Dataset G1 (Leaf Blight)	91.00	92.00	89.00	85.00
Dataset G2 Ecsa(Black measles)	72.00	73.00	71.00	71.00
Dataset G3 (Black Rot)	90.00	90.00	88.00	84.00
Dataset G4 (Healthy)	96.00	93.00	93.00	93.00
Dataset G5 (Total)	87.25	87.00	85.25	83.25

Table 16: Accuracy (%) for LBP on Segmented RGB Color Space for Tomato Dataset.

	Method: LBP			
Tomato Leaf Dataset	**Red channel**	**Green channel**	**Blue channel**	**R+G+B**
Dataset T1 (Yellow leaf curl)	68.00	66.00	68.00	62.00
Dataset T2 Ecsa(Tomato Mosaic Virus)	77.00	75.00	74.00	81.00
Dataset T3 (target Spot)	73.00	63.00	70.00	75.00
Dataset T4 (Spider Mites)	84.00	79.00	83.00	89.00
Dataset T5 (Septoria leaf spot)	50.0	56.00	50.00	52.00
Dataset T6 (Leaf Mold)	62.00	59.00	58.00	56.00
Dataset T7 (Late Blight)	44.00	44.00	45.00	47.00
Dataset T8 (Early Blight)	37.00	41.00	47.00	44.00
Dataset T9 (Bacterial spot)	96.00	93.00	93.00	87.00
Dataset T10(Healthy)	84.00	88.00	82.00	87.00
Dataset T11 (Total)	67.50	66.40	67.00	68.00

Results in terms of accuracy on apple, corn, grape and tomato datasets using LBP feature extraction method on YCbCr color space are shown in Tables 17 to 20.

Table 17: Accuracy (%) for LBP on Segmented YCbCr Color Space for Apple Dataset.

	Method: LBP			
Apple Leaf Dataset	**Y**	**Cb**	**Cr**	**Y+Cb+Cr**
Dataset A1 (Apple scab)	76.00	69.00	84.00	81.00
Dataset A2 (Black rot)	79.00	64.00	66.00	74.00
Dataset A3 (Cedar apple rust)	93.00	95.00	90.00	93.00
Dataset A4 (Healthy)	59.00	70.00	90.00	92.00
Dataset A5 (Total)	76.75	74.50	82.50	85.00

Table 18: Accuracy (%) for LBP on Segmented YCbCr Color Space for Corn Dataset.

	Method : LBP			
Corn Leaf Dataset	**Y**	**Cb**	**Cr**	**Y+Cb+Cr**
Dataset C1 (Cercospora leaf spot)	83.00	76.00	91.00	75.00
Dataset C2 (Northern leaf blight)	59.00	55.00	44.00	67.00
Dataset C3 (Common Rust)	91.00	68.00	82.00	94.00
Dataset C4 (Healthy)	95.00	99.00	99.00	96.00
Dataset C5 (Total)	82.00	74.50	79.00	83.00

Table 19: Accuracy (%) for LBP on Segmented YCbCr Color Space for Grape Dataset.

	Method: LBP			
Grape Leaf Dataset	**Y**	**Cb**	**Cr**	**Y+Cb+Cr**
Dataset G1 (Leaf Blight)	92.00	89.00	87.00	92.00
Dataset G2 Ecsa(Black measles)	73.00	58.00	53.00	66.00
Dataset G3 (Black Rot)	87.00	75.00	78.00	80.00
Dataset G4 (Healthy)	93.00	98.00	99.00	99.60
Dataset G5 (Total)	86.25	80.00	79.25	84.50

Table 20: Accuracy (%) for LBP on Segmented YCbCr Color Space for Tomato Dataset.

	Method: LBP			
Tomato Leaf Dataset	**Y**	**Cb**	**Cr**	**Y+Cb+Cr**
Dataset T1 (Yellow leaf curl)	63.00	82.00	42.00	61.00
Dataset T2 Ecsa(Tomato Mosaic Virus)	74.00	51.00	46.00	81.00
Dataset T3 (target Spot)	75.00	73.00	48.00	75.00
Dataset T4 (Spider Mites)	82.00	68.00	68.00	90.00
Dataset T5 (Septoria leaf spot)	54.00	53.00	38.00	61.00
Dataset T6 (Leaf Mold)	59.00	69.00	60.00	63.00
Dataset T7 (Late Blight)	47.00	44.00	49.00	45.00
Dataset T8 (Early Blight)	43.00	46.00	41.00	48.00
Dataset T9 (Bacterial spot)	95.00	84.00	89.00	88.00
Dataset T10(Healthy)	88.00	81.00	70.00	89.00
Dataset T11 (Total)	68.00	65.10	55.10	70.10

The experimental results on apple, corn, grape and tomato datasets using CLBP feature extraction method on HSI color space are presented in Tables 21 to 24.

Table 21: Accuracy (%) for CLBP on Segmented HSI Color Space for Apple Dataset.

	Method: CLBP			
Apple Leaf Dataset	**Hue channel**	**Saturation channel**	**Intensity channel**	**H + S + I**
Dataset A1 (Apple scab)	84.00	80.00	80.00	75.00
Dataset A2 (Black rot)	85.00	85.00	80.00	81.00
Dataset A3 (Cedar apple rust)	92.00	95.00	92.00	93.00
Dataset A4 (Healthy)	91.00	69.00	64.00	82.00
Dataset A5 (Total)	88.00	82.25	79.00	82.75

Table 22: Accuracy (%) for CLBP on Segmented HSI Color Space for Corn Dataset.

	Method : CLBP			
Corn Leaf Dataset	**Hue channel**	**Saturation channel**	**Intensity channel**	**H + S + I**
Dataset C1 (Cercospora leaf spot)	72.00	80.00	85.00	75.00
Dataset C2 (Northern leaf blight)	70.00	61.00	57.00	61.00
Dataset C3 (Common Rust)	100	93.00	94.00	95.00
Dataset C4 (Healthy)	99.00	94.00	95.00	98.00
Dataset C5 (Total)	85.25	82.00	82.75	82.25

Table 23: Accuracy (%) for CLBP on Segmented HSI Color Space for Grape Dataset.

	Method: CLBP			
Grape Leaf Dataset	**Hue channel**	**Saturation channel**	**Intensity channel**	**H + S + I**
Dataset G1 (Leaf Blight)	94.00	88.00	86.00	93.00
Dataset G2 Ecsa(Black measles)	75.00	72.00	69.00	80.00
Dataset G3 (Black Rot)	78.00	85.00	85.00	86.00
Dataset G4 (Healthy)	97.00	91.00	94.00	99.00
Dataset G5 (Total)	86.00	83.50	83.50	89.50

Table 24: Accuracy (%) for CLBP on Segmented HSI Color Space for Tomato Dataset.

	Method: CLBP			
Tomato Leaf Dataset	**Hue channel**	**Saturation channel**	**Intensity channel**	**H+S+I**
Dataset T1 (Yellow leaf curl)	59.00	62.00	64.00	67.00
Dataset T2 Ecsa(Tomato Mosaic Virus)	93.00	77.00	79.00	82.00
Dataset T3 (target Spot)	83.00	74.00	70.00	74.00
Dataset T4 (Spider Mites)	84.00	84.00	89.00	85.00
Dataset T5 (Septoria leaf spot)	62.00	56.00	54.00	53.00
Dataset T6 (Leaf Mold)	59.00	60.00	55.00	65.00
Dataset T7 (Late Blight)	41.00	43.00	42.00	40.00
Dataset T8 (Early Blight)	45.00	46.00	47.00	43.00
Dataset T9 (Bacterial spot)	85.00	98.00	91.00	91.00
Dataset T10(Healthy)	85.00	81.00	85.00	89.00
Dataset T11 (Total)	69.80	70.20	67.60	68.90

Accuracy of the experiments conducted on apple, corn, grape and tomato datasets using CLBP feature extraction method on RGB color space are shown in Tables 25 to 28.

Table 25: Accuracy (%) for CLBP on segmented RGB Color Space for Apple Dataset.

	Method: CLBP			
Apple Leaf Dataset	**Red channel**	**Green channel**	**Blue channel**	**R+G+B**
Dataset A1 (Apple scab)	87.00	81.00	82.00	89.00
Dataset A2 (Black rot)	80.00	81.00	79.00	80.00
Dataset A3 (Cedar apple rust)	91.00	89.00	92.00	89.00
Dataset A4 (Healthy)	68.00	66.00	68.00	70.00
Dataset A5 (Total)	81.50	78.25	80.25	82.00

Table 26: Accuracy (%) for CLBP on Segmented RGB Color Space for Corn Dataset.

	Method : CLBP			
Corn Leaf Dataset	**Red channel**	**Green channel**	**Blue channel**	**R+G+B**
Dataset C1 (Cercospora leaf spot)	84.00	83.00	81.00	82.00
Dataset C2 (Northern leaf blight)	56.00	62.0	58.00	58.00
Dataset C3 (Common Rust)	94.00	95.00	95.00	95.00
Dataset C4 (Healthy)	96.00	95.00	95.00	96.00
Dataset C5 (Total)	82.50	83.75	82.25	82.75

Table 27: Accuracy (%) for CLBP on Segmented RGB Color Space for Grape Dataset.

	Method: CLBP			
Grape Leaf Dataset	**Red channel**	**Green channel**	**Blue channel**	**R+G+B**
Dataset G1 (Leaf Blight)	87	87.00	90.00	85.00
Dataset G2 Ecsa(Black measles)	71.00	72.00	76.00	71.00
Dataset G3 (Black Rot)	83.00	82.00	86.00	84.00
Dataset G4 (Healthy)	92.00	96.00	92.00	93.00
Dataset G5 (Total)	83.25	84.25	86.00	83.25

Table 28: Accuracy (%) for CLBP on Segmented RGB Color Space for Tomato Dataset.

	Method: CLBP			
Tomato Leaf Dataset	**Red channel**	**Green channel**	**Blue channel**	**R+G+B**
Dataset T1 (Yellow leaf curl)	61.00	65.00	65.00	62.00
Dataset T2 Ecsa(Tomato Mosaic Virus)	84.00	81.00	76.00	81.00
Dataset T3 (target Spot)	75.00	77.00	76.00	75.00
Dataset T4 (Spider Mites)	89.00	91.00	88.00	89.00
Dataset T5 (Septoria leaf spot)	49.00	50.00	54.00	52.00
Dataset T6 (Leaf Mold)	61.00	58.00	53.00	56.00
Dataset T7 (Late Blight)	47.00	50.00	43.00	47.00
Dataset T8 (Early Blight)	41.00	45.00	46.00	44.00
Dataset T9 (Bacterial spot)	87.00	86.00	92.00	87.00
Dataset T10(Healthy)	86.00	86.00	83.00	87.00
Dataset T11 (Total)	68.00	68.90	67.60	68.00

Results in terms of accuracy on apple, corn, grape and tomato datasets using CLBP feature extraction method on YCbCr color space are shown in Tables 29 to 32.

Table 29: Accuracy (%) for CLBP on Segmented YCbCr Color Space for Apple dataset.

	Method: CLBP			
Apple Leaf Dataset	**Y**	**Cb**	**Cr**	**Y+Cb+Cr**
Dataset A1 (Apple scab)	78.00	83.00	86.00	81.00
Dataset A2 (Black rot)	75.00	77.00	78.00	74.00
Dataset A3 (Cedar apple rust)	89.00	91.00	92.00	93.00
Dataset A4 (Healthy)	62.00	75.00	97.00	92.00
Dataset A5 (Total)	76.00	81.50	88.25	85.00

Table 30: Accuracy (%) for CLBP on Segmented YCbCr Color Space for Corn Dataset.

	Method : CLBP			
Corn Leaf Dataset	**Y**	**Cb**	**Cr**	**Y+Cb+Cr**
Dataset C1 (Cercospora leaf spot)	84.00	82.00	78.00	75.00
Dataset C2 (Northern leaf blight)	64.00	64.00	57.00	67.00
Dataset C3 (Common Rust)	95.00	89.00	94.00	94.00
Dataset C4 (Healthy)	95.00	99.0	99.60	96.00
Dataset C5 (Total)	84.50	83.50	82.25	83.00

Table 31: Accuracy (%) for CLBP on Segmented YCBCR Color Space for Grape Dataset.

	Method: CLBP			
Grape Leaf Dataset	**Y**	**Cb**	**Cr**	**Y+CB+CR**
Dataset G1 (Leaf Blight)	84.00	87.00	94.00	92.00
Dataset G2 Ecsa(Black measles)	68.00	80.00	72.00	66.00
Dataset G3 (Black Rot)	85.00	75.00	81.00	80.00
Dataset G4 (Healthy)	96.00	96.00	99.60	92.00
Dataset G5 (Total)	83.25	84.50	86.75	84.50

Table 32: Accuracy (%) for CLBP on Segmented YCbCr Color Space for Tomato Dataset.

	Method: CLBP			
Tomato Leaf Dataset	**Y**	**Cb**	**Cr**	**Y+Cb+Cr**
Dataset T1 (Yellow leaf curl)	65.00	74.00	71.00	61.00
Dataset T2 Ecsa(Tomato Mosaic Virus)	83.00	88.00	83.00	81.00
Dataset T3 (target Spot)	72.00	83.00	72.00	75.00
Dataset T4 (Spider Mites)	92.00	77.00	84.00	90.00
Dataset T5 (Septoria leaf spot)	53.00	64.00	65.00	61.00
Dataset T6 (Leaf Mold)	59.00	66.00	72.00	63.00
Dataset T7 (Late Blight)	45.00	44.00	45.00	45.00
Dataset T8 (Early Blight)	51.00	45.00	35.00	48.00
Dataset T9 (Bacterial spot)	91.00	73.00	87.00	88.00
Dataset T10(Healthy)	85.00	86.00	80.00	89.00
Dataset T11 (Total)	69.60	70.20	69.40	70.10

5.3.2 Empirical Results III (With SVM)

In this section, the experiments are carried out on grayscale, colored and segmented images using three algorithms namely PCA, LBP and CLBP with SVM classifier and the comparative results are presented in Tables 33 to 35.

Table 33: Accuracy (%) of PCA, LBP and CLBP on All Images with SVM Classifier

TOTAL	PCA with SVM	LBP With SVM	CLBP With SVM
Apple grayscale	99.00	N/A	N/A
Apple colored	N/A	67.33	74.33
Apple segmented	N/A	43.50	46.50
Corn grayscale	98.75	N/A	N/A
Corn colored	N/A	65.75	66.00
Corn segmented	N/A	72.66	66.00
Grape grayscale	99.00	N/A	N/A
Grape colored	N/A	40.66	41.33
Grape segmented	N/A	56.66	48.00
Tomato grayscale	97.30	N/A	N/A
Tomato colored	N/A	51.25	42.00
Tomato segmented	N/A	55.60	41.75

Table 34: Accuracy (%) for LBP on Three Color Space with SVM Classifier.

	Method : LBP with SVM Classifier			
Dataset	**red channel**	**green channel**	**blue channel**	**R+G+B**
Apple (colored)	67.33	67.00	66.00	67.33
Apple (segmented)	46.50	44.50	36.25	41.50
Corn (colored)	65.50	65.00	65.50	65.75
Corn (segmented)	72.33	72.33	70.33	72.00
Grape (colored)	40.33	30.75	32.25	41.33
Grape (segmented)	57.33	56.33	58.00	57.00
Tomato (colored)	44.80	44.60	43.80	56.75
Tomato (segmented)	57.60	57.00	45.80	58.66
	Hue channel	**Saturation channel**	**Intensity channel**	**H + S + I**
Apple (colored)	67.00	66.00	67.33	65.75
Apple (segmented)	46.5	44.5	36.25	49.40
Corn (colored)	65.5	65.00	65.50	65.75
Corn (segmented)	72.33	72.33	70.33	51.25
Grape (colored)	40.33	30.75	32.25	41.33
Grape (segmented)	57.33	56.33	58.00	42.25
Tomato (colored)	44.80	44.60	43.80	46.00
Tomato (segmented)	57.80	57.20	45.66	53.20
	Y channel	**Cb channel**	**Cr channel**	**Y+Cb+Cr**
Apple (colored)	67.00	66.00	67.33	62.25
Apple (segmented)	46.50	44.50	36.25	49.50
Corn (colored)	65.50	65.00	65.50	63.50
Corn (segmented)	72.33	72.33	70.33	59.50
Grape (colored)	40.33	30.75	32.25	58.75
Grape(segmented)	57.33	56.23	58.00	54.75
Tomato (colored)	44.80	44.60	43.80	38.88
Tomato (segmented)	57.80	57.20	45.60	42.50

Table 35: Accuracy (%) for CLBP on Three Color Space with SVM Classifier.

	Method : CLBP with SVM Classifier			
Dataset	**red channel**	**green channel**	**blue channel**	**R+G+B**
Apple (colored)	69.33	75.00	75.33	69.67
Apple (segmented)	49.50	47.25	46.00	66.50
Corn (colored)	64.25	65.55	65.75	76.25
Corn (segmented)	72.33	73.00	70.66	72.00
Grape (colored)	41.00	41.66	42.00	48.00
Grape (segmented)	51.75	57.33	58.33	73.25
Tomato (colored)	41.83	42.66	37.00	41.66
Tomato (segmented)	52.50	52.33	52.00	52.66
	Hue channel	**Saturation channel**	**Intensity channel**	**H + S + I**
Apple (colored)	69.33	68.66	74.00	66.00
Apple (segmented)	53.50	47.75	40.25	67.25
Corn (colored)	76.75	67.75	65.75	74.00
Corn (segmented)	72.33	73.00	70.66	51.25
Grape (colored)	54.5	44.33	41.00	63.00
Grape (segmented)	61.00	57.33	58.33	73.25
Tomato (colored)	41.83	42.66	37.00	43.71
Tomato (segmented)	52.50	53.66	52.00	49.50
	Y channel	**Cb channel**	**Cr channel**	**Y+Cb+Cr**
Apple (colored)	74.66	68.66	69.33	80.50
Apple (segmented)	47.50	53.50	61.00	72.75
Corn (colored)	65.50	64.75	76.25	78.50
Corn (segmented)	72.33	73.00	70.66	60.00
Grape (colored)	41.33	58.00	58.00	63.25
Grape(segmented)	56.37	57.16	63.00	72.75
Tomato (colored)	41.83	42.66	37.00	37.33
Tomato (segmented)	52.50	57.33	47.50	53.80

Table 36 presents a comparative result of using PCA, LBP and CLBP algorithm with two classifiers involved in this thesis which are k-NN and SVM on colored, segmented and grayscale images.

Table 36: Accuracy (%) for LBP and CLBP on segmented images Comparing k-NN with SVM Classifier.

	Method: LBP and CLBP with k-NN and SVM Classifier			
Dataset	k-NN		SVM	
	LBP	CLBP	LBP	CLBP
Apple (colored)	75.75	75.50	67.33	69.67
Apple (segmented)	82.00	82.00	41.50	66.50
Apple (PCA)	86.53		87.33	
Corn (colored)	82.00	80.00	65.75	76.25
Corn (segmented)	82.25	82.75	72.00	72.00
Corn (PCA)	94.00		85.75	
Grape (colored)	84.50	83.00	41.33	48.00
Grape (segmented)	83.25	83.25	57.00	73.25
Grape (PCA)	94.00		80.667	
Tomato (colored)	66.30	66.00	56.75	41.66
Tomato (segmented)	68.00	68.00	78.66	52.66
Tomato (PCA)	96.10		71.25	

5.4 Empirical Configuration IV (Proposed Method)

The empirical configuration for the proposed method involves a Feature-Level Fusion of features. That is, the horizontal concatenation of the feature vectors of PCA and the feature vectors of the blue difference component (Cb) of YCbCr color space with LBP which are extracted from colored segmented images. This concatenation produces a new feature vector that contains all the relevant information that is then passed to the Support Vector Machine Classifier for classification.

Combining three color spaces in RGB, HSI and YCbCr is not improving the accuracy whereas each channel separately achieves better results compared to the combination of three channels for each color space. In general, gray images demonstrate better results compared to color images. However, if we want to consider color information for the leaf images which is rich in information, each channel can be used separately

to obtain better plant disease identification results as compared to the combination of the three channels of each color space.

5.4.1 Results from Empirical Configuration IV (Proposed Method)

Proposed method results compared with PCA, LBP and CLBP methods are presented in Table 37. The results are in terms of accuracy for each plant type namely apple, corn, grape and tomato. All types of plant diseases for the aforementioned plants are included in this experiment and the results are presented for each plant type separately.

Table 37: Accuracy (%) for the Proposed Method and Comparison with the Other Methods

	PCA	LBP	CLBP	Proposed Method [PCA +LBP-Cb]
Apple A5 (total)	94.00	41.50	66.50	99.00
Corn C5 (total)	97.00	72.00	72.00	98.75
Grape G5 (total)	95.75	57.00	73.25	99.00
Tomato T11 (total)	95.58	57.60	52.66	97.30

According to the comparison results, the proposed method demonstrates better results with respect to each feature extraction method for all types of plants. Therefore, using Feature-Level Fusion of PCA features extracted from gray level images and LBP features extracted from Cb color channel achieves the best results for plant disease classification compared to the existing PCA, LBP and CLBP features extracted from plant leaf images.

Chapter 6

CONCLUSION

In this work, we proposed a Feature-Level Fusion approach on a texture-based and an appearance-based approach to analyze and classify plant diseases based on the signs and symptoms visible on the plant leaves. This was done by employing the use of three feature extractors namely, Principal Component Analysis (PCA), Local Binary Patterns (LBP) and Completed Local Binary Patterns (CLBP). The three feature extraction algorithms were used for extracting the features from the leaf images. The newly proposed approach in this thesis involves the use of Feature-Level Fusion to integrate the features from PCA and LBP extracted from the Cb (blue-difference) component of YCbCr color space. This gives rise to a new feature that is rich in useful information, hence leading to better classification accuracy. The experiments were done on PlantVillage dataset which is a repository of various professionally collected plant leaves from different plants. The proposed system has proven to be efficient in that, it presents a better performance score as compared to the individual performances of LBP and CLBP. As a future work, we intend to explore the possibility of simultaneously identifying and detecting multiple diseases from a single leaf image. In addition, feature extraction can be done in parallel as this will significantly reduce the computation time. We believe that this will yield performance and gains over the current approach and also bring about the possibility of deploying the disease detection system to easily accessible devices like mobile phones, tablets or smart devices with a camera and cloud computing capabilities.

REFERENCES

[1] FAO launches 2020 as the UN's International Year of Plant Health - World. (2019, December 2). Retrieved from https://reliefweb.int/report/world/fao-launches-2020-un-s-international-year-plant-health

[2] Kaur, S., Pandey, S., & Goel, S. (2019). Plants Disease Identification and Classification Through Leaf Images: A Survey. *Archives of Computational Methods in Engineering*, 26(2), 507-530. doi:10.1007/s11831-018-9255-6.

[3] Hughes, D., & Salathé, M. (2015). An open access repository of images on plant health to enable the development of mobile disease diagnostics. *arXiv preprint arXiv:1511.08060*.

[4] Yoshida, K. (2013). The rise and fall of the Phytophthora infestans lineage that triggered the Irish potato famine. Retrieved from https://elifesciences.org/articles/00731 DOI: 10.7554/eLife.00731

[5] Plant disease management (2011). retrieved from https://www.apsnet.org/edcenter/disimpactmngmnt/topc/EpidemiologyTemporal/Pages/default.aspx

[6] Arneson, P. A. (2006). *Plant disease epidemiology: temporal aspects. the plant health instructor. DOI: 10.1094*. PHI-A-2001-0524-01. http://dx. doi. org/10.1094/PHI-A-2001-0524-01.

[7] Barbedo, J. G., Koenigkan, L. V., & Santos, T. T. (2016). Identifying multiple plant diseases using digital image processing. *Biosystems Engineering*,147, 104-116. doi:10.1016/j.biosystemseng.2016.03.012

[8] Pujari, D., Yakkundimath, R., & Byadgi, A. S. (2016) SVM and ANN based classification of plant diseases using feature reduction technique. *Int J Interact Multimed Artif Intell* 3(7), 6–14.

[9] Maheswari, P., Raja, P., & Ghangaonkar, N. M. (2018, December). Intelligent Disease Detection System for Early Blight of Tomato Using Foldscope: A Pilot Study. In *2018 IEEE 4th International Symposium in Robotics and Manufacturing Automation (ROMA)* (pp. 1-6). IEEE.

[10] Madeiras, A. (2019, September 27). Grape IPM- Black Rot. Retrieved June 29, 2020, from https://ag.umass.edu/fruit/fact-sheets/grape-ipm-black-rot

[11] Jackson, T. A. (2017). Rust disease of corn in Nebraska. Univ of Nebraska-Lincoln Extension. *Journal Intitute of Agriculture and Natural Resources. www. ianrpubs. unl. edu/.../publicationD. js. Nebraska Publication. Diakses pada*, *5*.

[12] Achterbosch, T. J., van Berkum, S., Meijerink, G. W., Asbreuk, H., & Oudendag, D. A. (2014). *Cash crops and food security: Contributions to income, livelihood risk and agricultural innovation* (No. 2014-15). LEI Wageningen UR.

[13] Sabrol, H., & Satish, K. (2016, April). Tomato plant disease classification in digital images using classification tree. In *2016 International Conference on Communication and Signal Processing (ICCSP)* (pp. 1242-1246). IEEE.

[14] Kulkarni, A. H., & Patil, A. (2012). Applying image processing technique to detect plant diseases. *International Journal of Modern Engineering Research*, *2*(5), 3661-3664.

[15] Singh, V., & Misra, A. K. (2017). Detection of plant leaf diseases using image segmentation and soft computing techniques. *Information processing in Agriculture*, *4*(1), 41-49.

[16] Dhingra, G., Kumar, V., & Joshi, H. D. (2018). Study of digital image processing techniques for leaf disease detection and classification. *Multimedia Tools and Applications*, *77*(15), 19951-20000.

[17] Matlab. [online].https//www.mathworks.com/products/matlab.html

[18] Moon, H. (1998). Analysis of PCA-based face recognition algorithms. *Empirical Evaluation Techniques in Computer Vision*, *1998*, 57-71.

[19] Ojala, T., Pietikäinen, M., & Harwood, D. (1996). A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, *29*(1), 51-59.

[20] Ahmed, F., Hossain, E., Bari, A. H., & Shihavuddin, A. S. M. (2011, November). Compound local binary pattern (CLBP) for robust facial expression recognition. In *2011 IEEE 12th International Symposium on Computational Intelligence and Informatics (CINTI)* (pp. 391-395). IEEE. doi: 10.1109/CINTI.2011.6108536.

[21] Kaur, S., Pandey, S., & Goel, S. (2019). Plants disease identification and classification through leaf images: A survey. *Archives of Computational Methods in Engineering*, *26*(2), 507-530. doi:10.1007/s11831-018-9255-6.

[22] Cover, T., Hart, P.: Nearest neighbor pattern classification. *IEEE Transactions on Information Theory* 13(1), 21–27 (1967)

[23] Cheng, D., Zhang, S., Deng, Z., Zhu, Y., & Zong, M. (2014, December). kNN algorithm with data-driven k value. In *International Conference on Advanced Data Mining and Applications* (pp. 499-512). Springer, Cham.

[24] Zhai, B. & Li, Z. (2018). Face Recognition Based on Fusion Feature of LBP and PCA with KNN. *DEStech Transactions on Computer Science and Engineering*. 10.12783/dtcse/cmsam2018/26592.

[25] Mishra, S. P., Sarkar, U., Taraphder, S., Datta, S., Swain, D. P., Saikhom, R., & Laishram, M. (2017). Multivariate statistical data analysis-principal component analysis (PCA). *International Journal of Livestock Research*, *7*(5), 60-78.

[26] Schaefer, G., & Doshi, N. (2017, July). LBP vs. LBP variance for texture classification. In *International Conference on Data Mining and Big Data* (pp. 156-164). Springer, Cham.

[27] Ojala, T., Pietikainen, M., & Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on pattern analysis and machine intelligence*, *24*(7), 971-987.

[28] Guo, Z., Zhang, L., & Zhang, D. (2010). A completed modeling of local binary pattern operator for texture classification. *IEEE transactions on image processing*, *19*(6), 1657-1663.

[29] Awang, S., Yusof, R., Zamzuri, M. F., & Arfa, R. (2013, December). Feature level fusion of face and signature using a modified feature selection technique. In *2013 International Conference on Signal-Image Technology & Internet-Based Systems* (pp. 706-713). IEEE. doi: 10.1109/SITIS.2013.115.

[30] Yan-Li, A. (2015, June). Introduction to digital image pre-processing and segmentation. In *2015 Seventh International Conference on Measuring Technology and Mechatronics Automation* (pp. 588-593). IEEE.