DEEP LEARNING-BASED APPROACH IN PLANT SPECIES IDENTIFICATION

TAN JING WEI

FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2018

DEEP LEARNING-BASED APPROACH IN PLANT SPECIES IDENTIFICATION

TAN JING WEI

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

INSTITUTE OF BIOLOGICAL SCIENCES
FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2018

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **Tan Jing Wei**

Matric No: **SGR160036**

Name of Degree: **Master of Science**

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

Deep Learning-Based Approach in Plant Species Identification

Field of Study: **Bioinformatics**

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge, nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every right in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date:

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:

Designation:

DEEP LEARNING-BASED APPROACH IN PLANT SPECIES IDENTIFICATION

ABSTRACT

Plant species identification and classification is one of the main tasks for botanists as well as a matter of interest for public. An automated plant species identification system could help the botanists and the layman to identify plant species in a more structured and speedy manner. Conventional machine learning techniques are widely used in the development of automated identification system in various fields including in biology and biodiversity. Deep learning is an emerging area in the machine learning approach. It has been considered as one of the powerful approaches for feature extraction as compared to the conventional approaches due to its superiority in providing deeper information of an image rather than the surface information. In this research, a total of 1290 leaf samples were collected in the University of Malaya (UM), Malaysia from 43 species of tropical trees with 30 samples for each species. The leaf images were pre-processed based on the feature extraction approaches which included the removal of background noises, segmentation of region of interest (ROI) and conversion of RGB images into grey-scaled images. The features were then extracted by using one of the deep learning approaches which is Convolutional Neural Network (CNN). Based on the literature review, this is one of the first few studies, which has applied CNN in tropical tree species identification, by using both leaf morphometric and venation pattern approaches. Three CNN-based models were used for feature extraction which are pre-trained AlexNet, fine-tuned pre-trained AlexNet and a newly proposed CNN model – D-Leaf model. A conventional morphometric method was employed for benchmarking purposes, which computed the morphological measurements based on the Sobel segmented veins. These features were classified by using four machine learning techniques, namely, Support Vector Machine

(SVM), Artificial Neural Network (ANN), k-Nearest Neighbour (k-NN), Naïve Bayes (NB) and Convolutional Neural Network (CNN). The fine-tuned AlexNet model performed slightly better (testing accuracy = 95.54%) than the D-Leaf (testing accuracy = 94.88%) models and AlexNet (testing accuracy = 93.26%). However, the execution time of D-Leaf model was 7 times faster than AlexNet and fine-tuned AlexNet models, respectively. The CNN models obtained a much higher performance than the vein morphometric measurement model which obtained only 66.28% in testing accuracy. In addition, ANN classifiers have achieved much better performance than SVM, k-NN, NB and CNN. In this research, D-Leaf can be a more efficient and effective automated tool for plant species identification with a high accuracy and shorter execution time than AlexNet and the fine-tuned AlexNet models as the CNN models performed better than the conventional morphometric measurements model. The conventional morphometric measurements method was less desirable in extracting features as compared to the CNN approach. The CNN extracted features are found to be fitted well with the ANN classifier as compared to other classifiers.

Keywords: tropical plants, deep learning, Convolutional Neural Network, Artificial Neural Network

PENDEKATAN BERASASKAN PEMBELAJARAN MENDALAM UNTUK IDENTIFIKASI SPECIES TUMBUHAN

ABSTRAK

Identifikasi dan klasifikasi spesies tumbuhan merupakan salah satu tugas utama untuk ahli botani serta memberi kepentingan kepada masyarakat. Sistem identifikasi spesies tumbuhan automatik boleh membantu ahli botani dan orang awam untuk mengenal pasti spesies tumbuhan dengan cepat dan lebih sistematik. Teknik pembelajaran mesin konvensional telah digunakan secara meluas dalam sistem pengenalan automatik dalam pelbagai bidang termasuk biologi dan biodiversiti. Pembelajaran mendalam merupakan kaedah terkini dalam pendekatan pembelajaran mesin. Ia dianggap sebagai salah satu pendekatan yang mampu mengekstrak ciri-ciri berbanding dengan pendekatan konvensional kerana keberkesanaannya dalam menyediakan maklumat yang lebih mendalam tentang imej dan bukan hanya maklumat permukaan sahaja. Di dalam kajian ini, sejumlah 1290 sampel daun dikumpulkan di Universiti Malaya (UM), Malaysia yang terdiri daripada 43 spesies pokok tropika. 30 sampel telah dikumpul untuk setiap spesies pokok. Imej daun telah diproses terlebih dahulu dengan penyingkiran objek-objek yang tidak penting, segmentasi region of interest (ROI) dan penukaran imej RGB kepada imej skala kelabu berdasarkan pendekatan pengekstrakan ciri. Ciri-ciri tersebut kemudiannya diekstrak dengan menggunakan salah satu pendekatan pembelajaran mendalam iaitu Rangkaian neural konvolusi (CNN). Berdasarkan tinjauan kajian, ini adalah salah satu daripada penyelidikan pertama yang menggunakan CNN dalam klasifikasi spesies pokok tropika dengan menggunakan pendekatan pola morfometrik dan vena daun. Tiga CNN model telah digunakan untuk pengekstrakan ciri, iaitu pre-trained AlexNet, fine-tuned AlexNet dan satu model baharu yang dicadangkan – D-Leaf. Kaedah morfometrik konvensional yang mengukur morfologi berdasarkan vena segmentasi Sobel digunakan untuk tujuan penanda aras. Ciri-ciri ini diklasifikasikan dengan menggunakan empat teknik pembelajaran mesin, iaitu Support Vector Machine (SVM), Rangkaian neural buatan (ANN), k-Nearest Neighbour (k-NN), Naïve Bayes (NB) dan Rangkaian neural convolutional (CNN). Model fine-tuned AlexNet lebih baik (ketepatan pengujian = 95.54%) daripada model D-Leaf (ketepatan pengujian = 94.88%) dan AlexNet (ketepatan pengujian = 93.26%). Walau bagaimanapun, masa penglaksanaan model D-Leaf adalah 7 kali lebih cepat daripada model AlexNet dan fine-tuned AlexNet masing-masing. Model-model CNN mendapat prestasi yang lebih tinggi daripada morfometrik tradisional, yang hanya memperolehi 66.28% dalam ketepatan pengujiannya. Di samping itu, klasifier ANN telah mencapai prestasi yang lebih baik daripada SVM, k-NN, NB dan CNN. Dalam kajian ini, D-Leaf adalah satu sistem yang lebih cekap dan berkesan dengan ketepatan yang tinggi dan masa pelaksanaan yang lebih pendek. Model-model CNN lebih baik daripada model pengukuran morfometrik konvensional. Kaedah pengukuran morfometrik konvensional adalah tidak sebaik jika dibandingkan dengan pendekatan CNN dalam mengekstrak ciri-ciri. Ciri-ciri yang diekstrak oleh CNN dengan menggunakan ANN adalah lebih baik berbanding teknik klasifikasi yang lain.

Kata Kunci: pokok tropika, pembelajaran mendalam, rangkaian neural konvolusi, rangkaian neural buatan

ACKNOWLEDGEMENTS

First and foremost, I wish to express my sincere gratitude and appreciation to my supervisors Dr. Chang Siow Wee and Prof. Datin Dr. Sameem Binti Abdul Kareem, for sharing expertise, and sincere and invaluable guidance and encouragement extended to me until this research is fully completed.

I would like to take this opportunity to express gratitude to Dr. Yong Kien Thai for his help and contributions towards accomplishment of this research project.

Next, my sincere thanks go to my beloved parents, Mr. Tan Gie Lim and Ms. Law Swee Ai for the unceasing encouragement, support and attention. I am grateful to my uncle and aunty, Mr. Lim Kin Seng and Ms. Law Swee Hoon who supported me throughout this period. To my sisters and brothers, thanks for their motivation and understanding. From the bottom of my heart, I have owned them so much.

Finally, thanks to all my friends and peers for their encouragement and help. I would like to thank to one and all, who directly or indirectly, have lent their hand to me in my project.

THANK YOU FOR ALL YOUR SUPPORT.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

--	:	Not Applicable
ANN	:	Artificial Neural Network
C	:	Convolutional Layer
CNN	:	Convolutional Neural Network
CS	:	Convolutional Stage
Den.	:	Density
DTC	:	Dewan Tunku Canselor Hall
FC	:	Fully Connected Layer
FS	:	Faculty of Science
Ft	:	Fine-tuned
GLCM	:	Gray-Level Co-Occurrence Matrix
HOG	:	Histogram of Oriented Gradient
k-NN	:	k-Nearest Neighbour
LBP	:	Local Binary Pattern
Lgth	:	Length
Max.	:	Maximum
Min.	:	Minimum
ML	:	Main Library
MLP	:	Multilayer Perceptron
MSD	:	Morphological Shape Descriptors
NB	:	Naïve Bayes
No.	:	Number
PCA	:	Principal Component Analysis
PS	:	Pooling Stage
RBM	:	Restricted Boltzmann Machines
ROC	:	Receiver Operating Characteristic
SD	:	Standard Deviation
SVM	:	Support Vector Machine
VL	:	Varsity Lake

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

1.1 Background

Plants, the organisms that exist everywhere in this world, play a vital role in biological diversity as well as in the economic sector. Plant help to optimize the ecosystem, for example, they help in maintaining the quantity of oxygen and carbon dioxide via the photosynthesis process. Furthermore, plants are important natural resources for foodstuff, furniture, medicine and so on. However, recently, quite a number of plants are at the risk of extinction. The main reasons that caused this problem are the human activities such as pollutions, global warming, greenhouse effect, deforestation and others. Hence, it is necessary to preserve and conserve those endangered species as well as others which are not at the risk of extinction.

The number of the plant species is extremely huge all over the world with about 391,000 vascular plant species (Willis, 2017). Hence, it is impossible and not practical for a botanist or expert to memorize and recognize all of the species (Carvalho et al., 2007). Moreover, some plants may have high similarity between each other, which require botanists or experts to spend a lot of time in differentiating these species. Therefore, it is necessary to develop a computerized or automated system in order to resolve these matters.

Most of the automated plant identification systems are developed based on the leaves due to the vast availability (Aakif & Khan, 2015; Beghin et al., 2010; Caglayan et al., 2013; Cope et al., 2010; Danti et al., 2012; Kadir et al., 2013; Kadir et al., 2014; Lee et al. 2016; Lin et al., 2008; Murat et al., 2017; Sharma et al., 2014; Wang et al., 2005). Most of the leaves are available throughout the whole year, unlike flowers, seeds and fruits which may be available at certain seasonal periods in a year. Consequently, leaves are the most practical organ that are commonly employed in the automated identification systems. Leaves are able to provide us with information such as shape, textures, and veins, colour.

Figure 1.1 shows the general approach of an automated plant identification system. Firstly, the leaf images would be collected or acquired either by using digital camera, scanner or other equipment. The leaf images are then pre-processed by image enhancement, image segmentation and so on. Next, the features of each image would be obtained by feature extraction approach. Lastly, the classification methods will be used to recognize the leaf. In an identification system, the most important stage is the feature extraction stage. An optimum feature set could help in improving the accuracy of plant species identification.

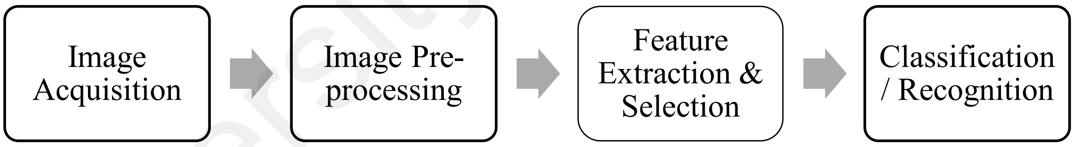


Figure 1.1: General approach of automated plant identification system.

The classification or identification tasks would be performed by using traditional statistical analysis or machine learning approach. With the advancement of science and technology, machine learning had been widely employed in many domains especially in the biological domain (Larranaga et al., 2006; Libbrecht & Noble, 2015; Sommer & Gerlich, 2013). Machine learning is a type of artificial intelligent techniques which mainly perform pattern identification role which are useful or practical for the development of leaf identification system (e.g.: Artificial Neural Network, Support

Vector Machines, etc.). Machine learning approaches, feature selection and classification methods were implemented in the automated system to ensure high accuracy of plant identification via the image features of leaves.

Currently, deep learning, a modern sub-discipline of artificial intelligence (AI) has emerged as a favored and well-known method that provides robust supervised learning model. It has been widely applied in various disciplines such as medical, biology, speech recognition, image recognition and others (Al-Angari et al., 2012; Benard et al., 2014; Camurri et al., 2003; Jarasch et al., 2017; Kim et al., 2013; Nasir et al., 2013; Lien et al., 1998; Oskouie et al., 2017). It can be further classified into a few different architectures like Convolutional Neural Network based methods, Restricted Boltzmann Machines based methods and others (Guo et al., 2016).

The superiority of deep learning is its ability in extracting a more detailed and deepen data in contrast with those conventional feature extraction techniques. It can be applied to images, sounds, videos, and others directly for feature extraction (Goswami et al., 2014; Lévy & Jain, 2016; Ngiam et al., 2011; Wu et al., 2015).

1.2 Problem Statement

The plant species all over the world is extremely huge with about 391,000 vascular plant species according to the report from the Royal Botanic Gardens, Kew, England (Millis, 2017). The application of deep learning techniques in plant identification system has been done by Grinblat et al. (2016), Lee et al. (2015, 2016, 2017), Sladojevic et al. (2016) and Sünderhauf et al. (2014). However, to date, there is no related study regarding the application of deep learning approach in tropical plant species. A recent study that has developed an automated Malaysian shrub classification with the use of conventional feature extraction methods was done by Murat et al. (2017). Thus, we believe, it is necessary to investigate the deep learning-based approaches in automated plant species classification using tropical plant species.

Furthermore, there is a need to include all the features of a leaf such as shape, texture, vein and colour in order to make an accurate plant species identification. Many studies have applied conventional feature extraction methods, which used only one type of the leaf features. Thus, these studies may be less effective in automated plant species identification (Danti et al., 2012; Murat et al., 2017; Hati et al., 2013; Sharma et al., 2014; Wang et al., 2005). It has been proven in Beghin et al. (2010), Chaki et al. (2015), Lin et al. (2008) and Wang et al. (2016) that the identification results were more accurate when using more than one type of features of a leaf.

The traditional plant identification approach is to train the taxonomists to examine the specimen and allocate taxonomic labels to each of them. However, there are some problems faced by using the traditional approach, namely, there is a shortage on related field experts and also it is time consuming and costly to use the traditional approach (Leishman et al., 1992; Medin et al., 1997). Furthermore, it is impossible, even for an expert to know all the plant species. Hence, the development of an automated plant species identification system could help the experts identify the plant species quickly and efficiently at low cost.

Besides that, it is difficult for novices to get a botanist or an expert in helping them to identify an unknown leaf. Hence, an automated plant identification tool could help the novices identify an unknown leaf sample in a more convenient way. In addition, this automated plant tool can help in raising the interest toward plant preservation and conservation. This could also aid in educating the public about plant knowledge with lower cost.

1.3 Research Objectives

The objectives of this study are:

1. To extract leaf features from the selected tropical plant species using deep learning-based approach.
2. To compare the performance of extracted features by using deep learning and conventional approaches.
3. To identify the optimum leaf features in deep learning-based plant species identification.
4. To develop an automated plant species identification system using deep learning-based approach.

1.4 Scope of the Study

In this research, a leaf dataset was locally collected in the University of Malaya, Kuala Lumpur, Malaysia. This dataset consists of 43 tropical tree species. Due to time and budget constraint, only 43 common tropical tree species were selected and included in this dataset. In addition, the sampling sites were restricted within the campus of University of Malaya only.

Leaves were selected as the research samples due to their vast availability as compared to fruits, flowers or other parts of a plant. Commonly, the botanists would categorize and identify the plant species by using the external traits of the plants. However, such elements would only appear on a specific period of a year. For example, the flowers or fruits of a specific plant would appear once or twice per year. Due to this matter, it was difficult for sampling to be carried out. Leaf is the most common element which is used in many plant-related researches because it is available for sampling throughout the year. However, the individuals of same species might share the same visual properties with different varieties or cultivars.

1.5 Significance of the Study

Based on our literature review, this is one of the first studies have been applied the deep learning approach in extracting the features of tropical plant for species identification. An optimum leaf features set was obtained using the deep learning approach. A comparison was made to find out the optimum tool for developing a deep learning-based plant species identification system. Thus, the model with more deepen, and detailed features help in predicting the plant species with higher accuracy.

This system may assist botanists, taxonomists or other scientists to identify or recognise an unknown sample in a shorter time. Besides, the system can help to raise the interest of laymen especially students in the botanical study to conserve and protect the existing plant species in the world.

1.6 Chapter Organization

This thesis is organized as follows:

- Chapter 1 provides the introduction of the proposed research including the problem statements, research objectives, scope of the study and significance of the study.

- Chapter 2 presents previous studies in plant identification, introduction to deep learning-based approach, convolutional neural network-based approach and classifiers.
- Chapter 3 provides a detailed methodology used in the field sampling, leaf image acquisition, image pre-processing, deep learning-based feature extraction – Convolution Neural Network and classification methods.
- Chapter 4 reveals the results, benchmarking, validation methods and developed a prototype of graphical user interface of the proposed model.
- Chapter 5 discusses and compares the results.
- Chapter 6 concludes the presented research and proposes some future works.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

Automated identification systems are widely implemented in various domains which include medical, biological, engineering, architecture, and others (Al-Angari et al., 2012; Benard et al., 2014; Camurri et al., 2003; Jarasch et al., 2017; Kim et al., 2013; Lien et al., 1998; Nasir et al., 2013; Oskouie et al., 2017). It is essentially referred as computerized pattern recognition based on the text, images, sounds and videos. Image recognition is one of the common types of recognition system that has been employed for face recognition, fingerprint recognition, medical images recognition, word recognition and so on (Benard et al., 2014; Hrechak & McHugh, 1990; Lien et al., 1998; Nasir et al., 2013). There is a need to develop such systems in order to ease the daily tasks or works to make them more efficient and effective.

Feature extraction has a huge influence on an identification system. If the features of an object are not extracted effectively, it may affect the entire identification process. Thus, feature extraction is considered as the most important stage in the development of an automated identification system. Features can be either extracted by employing conventional methods (Choras, 2007; Nakanishi et al., 1995) or by advance deep learning-based methods (Han et al., 2017; Hu et al., 2015; Jin et al., 2014; Sermanet et al., 2012).

Conventional methods refer to the methods which require a series of steps and consume high computational time. Shape features can be computed by different approaches such as morphological shape descriptors (MSD), Histogram of oriented gradients (HOG), Zernike Moment, Hu Moments, and Polar Fourier Descriptor (Sariyanidi et al., 2013; Kale et al., 2014; Murat et al., 2017). The common statistic techniques used to compute the texture of the leaf are local binary pattern (LBP) operator, grey-level co-occurrence matrix (GLCM), Grey Tone Spatial Dependency Matrix (GTSDM) and Gabor filters (Ou et al., 2014; Pantic et al., 2015).

Deep learning – an advance method, a high level spectrum of machine learning, is robust in executing classification tasks directly from the text, sound or images. It is employed commonly in various domains including computer vision and robotics (Goodfellow et al., 2016). It is well suited particularly in the recognition systems such as speech recognition, text translation, face recognition and so on (Hinton et al., 2012; Wang, 2015; Wang et al., 2016).

2.2 Plant Species Identification

Plants are commonly identified based on their external morphology or appearance such as leaf, flower, fruits and bark. Among these plant organs, leaves contribute more useful information for species identification (Du et al., 2006; Du et al., 2007; Wang et al., 2005). Commonly, botanists and taxonomists identify an unknown specimen by comparing it with herbarium specimens which had been identified properly by referring to the published plant descriptions, images and illustrations or using keys which is discussed in detail in the next section. Traditionally, the leaf specimens are collected and stored as herbarium for unknown specimen identification later. The collected samples are compressed by storing between magazines or newspaper while being dried out quickly as shown in Figure 2.1.

Due to the huge number of plant species all over the world, botanists and taxonomists have limitations in identifying and recognising all plant species (Carvalho et al., 2007). Computational methods are more advantageous nowadays due to their availability in providing more useful information and features to identify a plant species. Additionally, computational approach could store and identify large number of plant species in more appropriate manner.

Figure 2.1: Compress the leaf samples between drawing boards.

2.2.1 Taxonomic Approach

In previous studies, several methods have been employed to identify a plant species such as asking an expert who was knowledgeable in plant, comparing the unknown sample to images or sketches in book, and etc. (Cope et al., 2012; Randler, 2008). Among these methods, a taxonomic key is a simple and the most common approach that used for identifying an unknown object (Krimmel et al., 2014).

Dichotomous keys, one of the taxonomic key, is a very structured approach for identifying an unknown plant. It works based on a series of paired alternative choices such as margin structure of a leaf, leaf arrangement and type of the root system (Kroening et al., 2007) as shown in the Figure 2.2 (To-Anun et al., 2011). Yet, the use of the dichotomous key would require some botanical knowledge and experience.

1. Trees 2
1. Herbs 3
2. Palmate venation .... Species A
2. Pinnate venation .... Species B
3. Ovaries interior .... Species C
3. Ovaries superior 4
4. Fruits nutlets .... Species D
4. Fruits capsules .... Species E

Figure 2.2: The example of Dichotomous keys.

Another taxonomic tool with the name of polyclave keys is used for identifying the unknown species. Interactive computer programs are being used to generate the polyclavve keys. It is basically employed as an elimination process to identify an unknown object (Krimmel et al., 2014) and it allows the users select the features or characteristics, then, take the selections from the character set (Morse, 1971). An elimination process is repeated until a preliminary identification is made. Table 2.1 shows the example of the Polycave keys (Tyrl, 2010).

Table 2.1: The example of Polyclave keys.

	Species A	Species B	Species C	Species D	Species E
Trees	-	-	+	-	+
Herbs	+	+	-	+	-
Plamate venation	+	-	+	-	+
Pinnate venation	-	+	-	+	-
Ovaries interior	+	+	+	-	-
Ovaries superior	-	-	-	+	+
Fruits nutlets	-	-	+	-	-
Fruits capsules	+	+	-	+	+

2.2.2 Automated Computational Approach

A plant species is commonly recognized by the inspections on their bark, flowers, fruits, and leaves. However, the leaf is the most popular and practical organ due to its availability throughout the year while other parts of the tree are only seasonally available.

In the automated computational approach, leaf features are extracted from their leaf images (Cope et al., 21012; Gwo et al., 2013; Murat et al., 2017). The most common leaf features used in developing a plant identification system are textures, shape, veins and colour.

Leaf shape is the most common feature that used in plant identification systems. The leaf shape features that can be obtained from a leaf are compactness, area, major and minor axis length and aspect ratio. Another significant feature that can be acquired from the leaf images are the vein features. The venation architectures of each plant species are distinctly different from each other. The leaf venation can be classified into three categories which are primary veins, secondary veins and tertiary veins (Cope et al., 2012) as shown in Figure 2.3. Texture refers to the surface structure of a leaf which can be used to represent a species significantly. For example, some plant species may have tiny and soft hair on the surface of their leaf. Different species leave can be differentiated

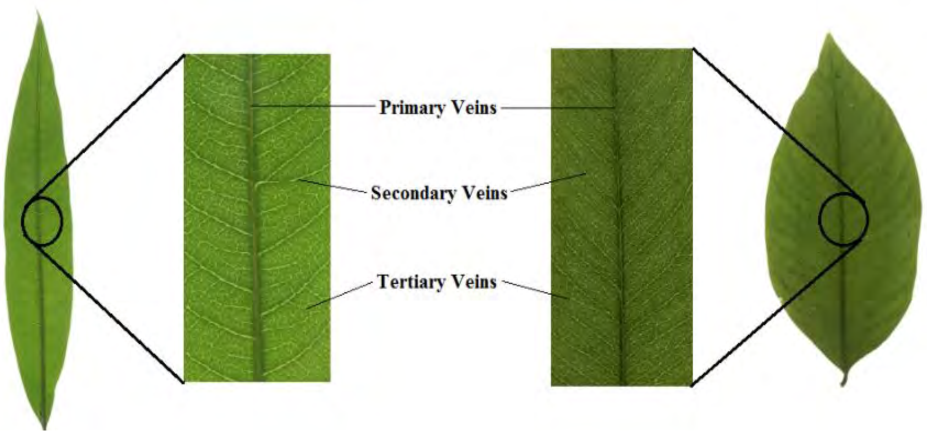


Figure 2.3: Leaf venation.

according to their leave colour, although this cannot be done efficiently with the naked eyes. Some plant from the same species may have different colour of leaves. As an example, *Cinnamomum inners* have young leaves which are red and green in colours as shown in Figure 2.4.

Figure 2.4: Young leave of *Cinnamomum inners*.

2.3 Automated Plant Identification

Several leaf image datasets are publicly available for plant-related studies. Numerous algorithms and approaches can be applied in plant identification. These algorithms and approaches can be further divided into conventional methods and deep learning methods (Caglayan et al., 2013; Kadir et al., 2013; Lin & Peng, 2008; Lee at al., 2017; Pawara et al., 2017).

2.3.1 Plant Image Datasets

The leaf images are commonly acquired either scanned by scanner or captured by camera. There are several leaf image datasets which are available publicly for analysis, such as, Leafsnap dataset, Swedish Leaf dataset, Flavia dataset, MalayaKew dataset and Foliage Dataset.

2.3.1.1 Leafsnap Dataset

Leafsnap dataset consists of 185 tree species from the North-eastern United States. The leaf images are categorised into two groups which are "Lab" images and "Field" images. The "Lab" images refer to those high-quality images which were acquired from the Smithsonian Collection, with 23,147 images. Whereas, the "Field" images consist of mobile phone taken images from the outdoor, which composed of shadows, blur, noises and illumination patterns. Figure 2.6 shows leaf samples of each tree species in the Leafsnap dataset (Kumar et al., 2012).

Figure 2.5: The leaf sample of each species in Leafsnap dataset.

2.3.1.2 Swedish Leaf Dataset

Swedish leaf dataset, which made up of 15 Swedish trees classes with 75 samples per species as shown in Figure 2.6, are collected in Sweden (Söderkvist, 2001). It is a part of the collaboration project between Linkoping University and Swedish Museum (Söderkvist, 2001).

Figure 2.6: The leaf samples of each species in Swedish Leaf dataset.

2.3.1.3 Flavia Dataset

Flavia dataset is a dataset that is commonly employed for plant identification studies. The Flavia dataset consists of 32 plant species collected from the YangTze Delta, China (Wu et al., 2007). Each plant species consists of 50 samples in this dataset. Figure 2.7 shows the leaf samples of each plant species in the Flavia dataset.

Figure 2.7: The leaf samples of each species in Flavia dataset.

2.3.1.4 MalayaKew Dataset

Malayakew dataset consists of 44 leaf classes which gathered at the Royal Botanic Gardens, Kew, England (Lee et al., 2015). The leaf sample of all species in the MalayaKew dataset are showed in Figure 2.8. However, the background of the leaf images in this dataset are black in colour.

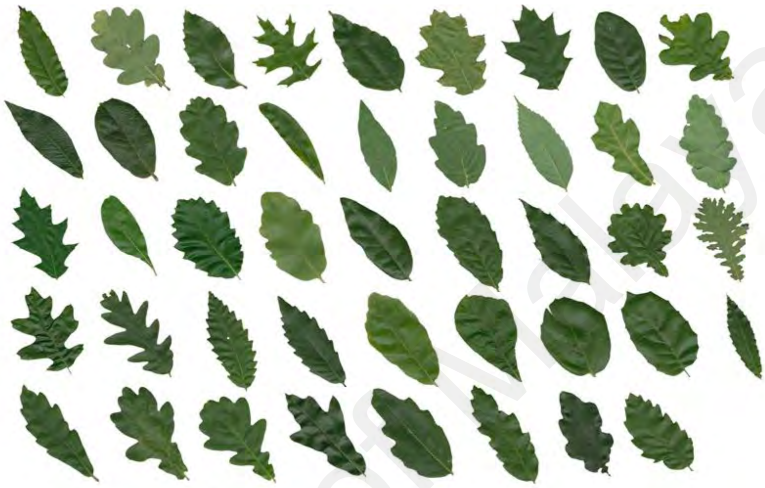


Figure 2.8: The leaf sample of each species in MalayaKew dataset.

2.3.1.5 ICL Leaf Dataset

ICL leaf dataset has 17,032 leaf images which were sampled from 220 plant species. The leaf samples were collected at the Botanical Garden of Hefei, Anhui Province, China, by members of Intelligent Computing Laboratory (ICL). The number of leaf images for each species is ranging between 26 and 1078. The leaf sample of each plant species from the ICL leaf dataset is as shown in Figure 2.9 (Wang et al., 2017).

Figure 2.9: The leaf sample of each species in ICL dataset.

2.3.2 Conventional Methods in Plant Identification

Conventional methods require a series of process to obtain a set of optimum features, which are pre-processing, segmentation and feature extraction (Murat et al., 2017; Wang et al. 2016). Firstly, the images have to be pre-processed to remove all the noises and enhance the images. Each image is then segmented to obtain the ROIs followed by feature extraction. The extracted features are then fed into the classifiers for training and testing. If both the shape and vein features are used, two separate segmentation processes have to be carried out (Chaki et al., 2015; Kadir et al., 2014). Thus, the conventional feature extraction methods would result in a high execution time (Murat et al., 2017).

Segmentation is a process that is used for obtaining the regions of interest. There are numerous types of segmentation method such as edge detection, surface fitting, region splitting, etc. (Bhanu & Lee, 1994). The most commonly used method is the edge detection which employs the grayscale images for transforming the images into binary images. There are several common types of edge detection such as Sobel, Prewitt, Roberts and so on (Zaitoun & Aqel, 2015).

Sobel is an edge detection approach, which measures the 2-D spatial gradient of the image and then highlight the edges corresponding regions with high spatial gradients. It is basically used to search for the relative magnitude of absolute gradient in each point of a greyscale image. As compared to other edge detection approach, Sobel can help to smooth the random noises and enhance the edge elements (Gupta & Mazumdar, 2013). Thus, Sobel can be used as a segmentation approach to segment or extract the vein architectures of the leaf samples since it's able to detect the edge of an object in a simple and effective manner (Gupta & Mazumdar, 2013; Vincent & Folorunso, 2009).

According to earlier reports, feature extraction is the key process in conventional methods of plant identification. The common features that are employed in plant identification system are the shape (Aakif & Khan, 2015; Caglayan et al., 2013; Murat et al., 2017; Sharma et al., 2015; Wang et al., 2005), vein (Cope et al., 2010; Kadir et al., 2013; Kadir et al., 2014), colour (Danti et al., 2012; Kadir et al., 2013; Kadir et al., 2014) and texture (Beghin et al., 2010; Lin et al., 2008; Kadir et al., 2013; Kadir et al., 2014). Various approaches can be applied to extract the leaf features such as Zernike Moment, Histogram of Oriented Gradient (HOG), Hu's moment and others. Table 2.2 summarizes some of the previous studies which have employed conventional feature extraction methods.

Table 2.2: Summary of convention feature extraction methods used in the previous studies.

Feature	Authors	Dataset	Feature Extraction	Results
Shape	Murat et al., 2017	myDAUN	Hu invariant moments, morphological shape descriptors, Zernike moments, Histogram of Oriented Gradients	98.23%
Shape	Gwo et al., 2013	13 fresh plant species	Leaf contour	92.70%
Shape	Sharma & Gupta, 2015	16 Classes	Morphological Features	> 90%
Shape	Aakif et al., 2015	Own Dataset, Flavia, ICL	Morphological features, Shape-defining feature, Fourier descriptor	96.50%
Shape	Hati et al., 2013	20 kinds of plants	Morphological Features	92%
Shape	Du et al., 2007	20 Species	Digital morphology feature extraction	91%
Shape	Wang et al., 2005	20 Species	Morphological Features	92.40%
Shape	Wu et al., 2007	32 Species	Basic Geometric Features, Digital Morphological Features, Vein features	90.31%
Texture	Arun et al., 2013	5 Medicinal Plant Species	Grey textures, Grey Tone Spatial Dependency Matrices Based Textures, LBP	94.70%
Texture	Yadav et al., 2013	25 Wood Species	GLCM	92.60%

Table 2.2, continued.

Feature	Authors	Dataset	Feature Extraction	Results
Texture	Prasvita et al., 2013	30 Medicinal Plant Species	LBP	56.33%
Texture	Ehsanirad et al., 2010	13 Species	GLCM & PCA	98%
Texture	Sulc et al., 2014	Austrian Federal Forest (AFF), Flavia, Foliage, Swedish, Middle European Woods (MEW)	Ffirst	> 99%
Texture	Cope et al., 2010	32 Species of leave	Gabor Co-Occurrences	79.69%
Colour	Danti et al., 2012	10 Species of Indian Leafy Vegetables	--	92-100%
Vein	Cope et al., 2010	N/A	Genetic algorithms (GA), Ant Colony	Combination of both GA and Ant Colony performed better.
Shape, Colour	Caglayan et al., 2013	Flavia	Morphological features, Colours	96.30%
Shape, Texture	Chaki et al., 2015	31 Classes	Gabor Filter, GLCM, Curvelet transform coefficients, Invariant Moments	97.60%
Shape, Texture	Beghin et al., 2010	Royal Botanic Garden, Kew, UK	--	81.10%

Table 2.2, continued.

Feature	Authors	Dataset	Feature Extraction	Results
Shape, Texture	Lin et al., 2008	30 Broad-Leaved Tress Species	--	98.30%
Shape, Texture	Wang et al., 2016	Flavia, ICL, MEW2012	Entropy sequence, Zernike moments, Hu's invariants	96.67%
Shape, Colour, Vein, Texture	Kadir et al., 2013	Flavia	--	93.75%
Shape, Colour, Vein, Texture	Kadir et al., 2014	Flavia, Foliage	--	97.19%

-- Not Applicable

Most of the studies had collected their local leaf samples as the dataset of their dataset. In addition, the only tropical dataset was collected by Murat et al. (2017), involving with tropical shrub species only. Therefore, this research proposed an image dataset with tropical tree species, which is considered the first of the same kind dataset.

2.3.3 Deep Learning-based Methods in Plant Species Identification

Deep learning is a subfield of the machine learning approaches which is able to learn high-level features in data by employing hierarchical structure (Guo et al., 2016). Numerous deep learning studies have been proposed to overcome the traditional artificial intelligent problems (Guo et al., 2016). It is difficult for a computer or machine to understand those raw sensory data or the pixels of an image. It is very complicated when a function is used to map the identity of an object from those pixels. Thus, deep learning helps to split those complicated mapping into a series of nested mapping layers to resolve the difficulty (Goodfellow et al., 2016). It is a powerful feature extraction method since it can extract more detailed information as compared to those conventional feature extraction methods (Goodfellow et al., 2016). It is also robust in dealing with huge image datasets. Furthermore, it is capable of extracting all features of an image such as shape, colour, texture and veins at the same time.

Deep learning model is basically a multilayer perceptron (MLP) which is built from three basic layers which are input, hidden and output layers as other neural networks as shown in Figure 2.10 (Goodfellow et al., 2016). MLP consists of one or more hidden layers in between input and output layer, in which a function maps input values to output values mathematically. Besides that, every node of the previous and next layers are connected to each other. It is referred as a feed forward neural network (Gardner & Dorling, 1998).

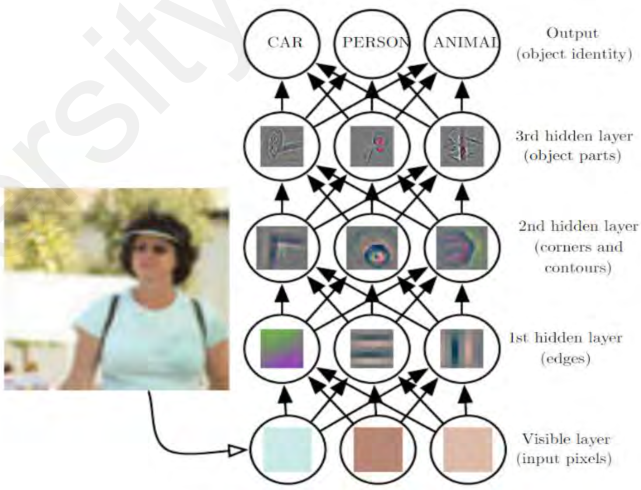


Figure 2.10: General illustration of a deep learning.

Deep learning approaches can be divided into 4 categories, namely, Convolutional Neural Networks (CNN), Autoencoder, Restricted Boltzmann Machines (RBM) and Sparse Coding (Guo et al., 2016). Figure 2.11 illustrates the categorization of deep learning approaches along with their representative works (Guo et al., 2016).

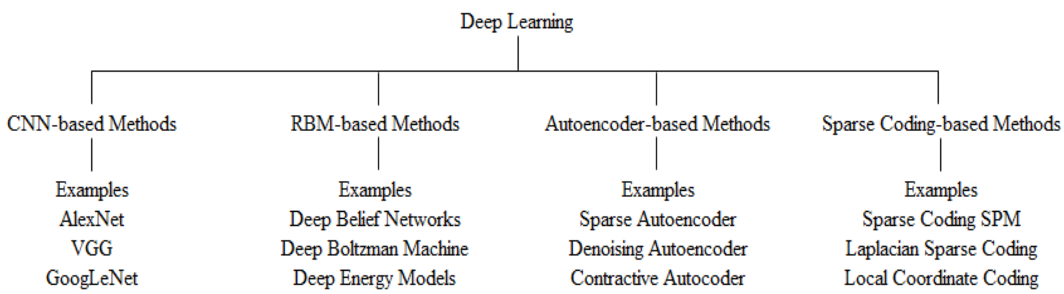


Figure 2.11: Categorization of deep learning approaches and their representative examples.

CNN-based method, the most notable and common of deep learning approaches, is a robust technique which has been used extensively in the field of image and video feature extraction recent. CNN basically learns the image features based on more than one perceptron layer. Thus, it can be concluded that CNN is another type of multilayer perceptron (MLP) in machine learning (Goodfellow et al., 2016). There are some representative models which has been derived from the CNN-based method, such as, AlexNet, VGG, and GoogLeNet.

RBM-based method is referred as a generative stochastic neural network which was proposed by Hinto & Sejnowski et al. (1986). RBM is an alternative Boltzmann machine which restrict the visible and hidden units to generate a bipartite graph. This restriction helps in improving the efficiency of the training algorithms (Guo et al., 2016). The representative works for RBM-based method are Deep Belief Network (DBN), Deep Boltzmann Machine (DBM), and Deep Energy Models (DEM).

The autoencoder-based method, on the other hand, is a special type of artificial neural network (ANN), which basically learns efficient encodings. An autoencoder is trained to construct its own specific inputs, instead of training the given inputs to predict their targets directly (Guo et al., 2016). Sparse Autoencoder, Denoising Autoencoder and

Constractive Autocoder are some of the examples of the autoencoder-based deep learning method.

Sparse coding-based method is used to learn over a complete basic function set to characterize the inputs (Guo et al., 2016). Examples are Sparse Coding SPM, Laplacian Sparse Coding and Local Coordinate Coding.

2.4 Deep Learning

A general CNN model consists of 3 main neural layers which are convolutional layer, fully connected layer and classification layer as shown in Figure 2.12.

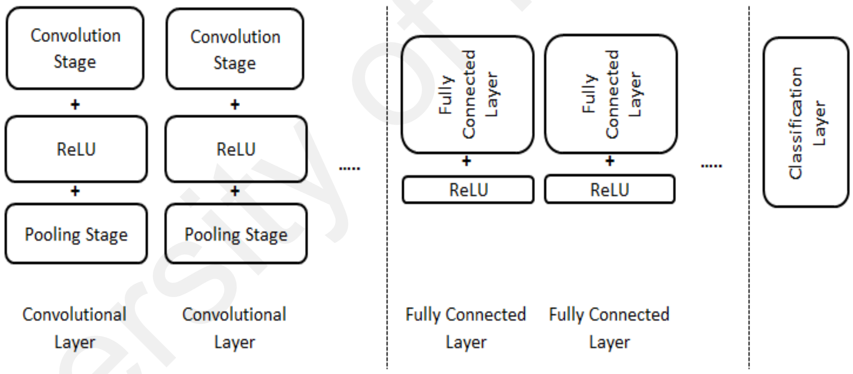


Figure 2.12: General Architecture of a CNN Model.

A typical convolutional layer is composed of the convolutional stage, Rectified Linear Unit (ReLU) and the pooling stage (Goodfellow et al., 2016; Guo et al., 2016; Grinblat et al., 2016; Lee et al., 2015; Lee et al., 2016; Lee et al., 2017; Liu et al., 2017; Sladojevic et al., 2016).

The purpose of the convolutional stage is to extract the common patterns from the local area of the input images. This is done by undergoing convolution operation with the use of filters and kernels over the pixels of the input image and then calculating the inner product of the filter at every image region to create a feature map (Zeiler & Fergus, 2013). Each feature map would be implemented with the ReLu, which is a non-linear activation function. This function would remain to linear closely (piecewise linear function). Thus, this will preserve many of the properties that optimize and generalize the linear models with gradient based methods easily (Goodfellow et al., 2016). The last stage (pooling stage) would simplify and summarize all the features of each individual layer with pooling function. There are two conventional types for pooling which are average and maxpooling. For average pooling, every element in a pooling region is considerable, even there are many elements with low magnitude (Zeiler & Fergus, 2013). Whereas, max-pooling applies a factor of K_x and K_y (K indicates a kernel of size) via each direction, producing position invariance over larger local regions and down-sampling the input image (Nagi et al., 2011).

The extracted features of each convolutional layer are compacted and summarized as outputs and then passed to next layer as the input as shown in the Figure 2.13. After convolutional layers, the features are then passed to the fully connected layers. The features would be condensed and compacted again in the fully connected layers. The extracted features from fully connected layers would be then used to train the classifiers.

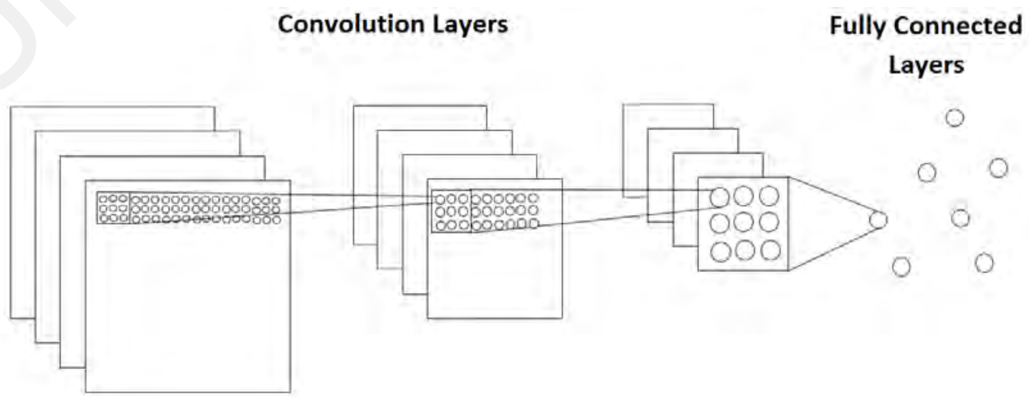


Figure 2.13: The extracted features are compacted and passed to the next layer.

Instead of spending time to build a new CNN model, various researches have implemented pre-trained CNN models. For example, Lee et al. (2015) proposed plant identification by using DeCAF; Minaee et al. (2016) proposed an iris recognition with VGG; Hu et al. (2015) proposed AlexNet, CaffeNet, VGGNet and PlacesNet for remote sensing images recognition. These pre-trained models have been trained with a large set of images which might be different from the current study. There are several types of pre-trained model which include AlexNet, VGGNet, CaffeNet and others. Each pre-trained model is different in architecture and number of convolutional layers, fully connected layers and some other parameters. In order to get a more desirable set of features, a pre-trained model can be fine-tuned to fit with the proposed dataset.

2.4.1 AlexNet

AlexNet is one of the CNN-based pre-trained model proposed by Krizhevsky et al. (2012). This model was trained to identify the images from 1000 classes in the ImageNet LSVRC-2010 contest with eight learned layers (5 convolutional layers + 3 fully connected layers) as illustrated in Figure 2.14. AlexNet has been employed in various field of studies since its capability to perform excellently. For example, Han et al. (2017) used AlexNet for remote sensing images recognition; Nguyen et al. (2015) used AlexNet for unrecognizable images predictions and Lévy and Jain (2016) used AlexNet for mammogram recognition.

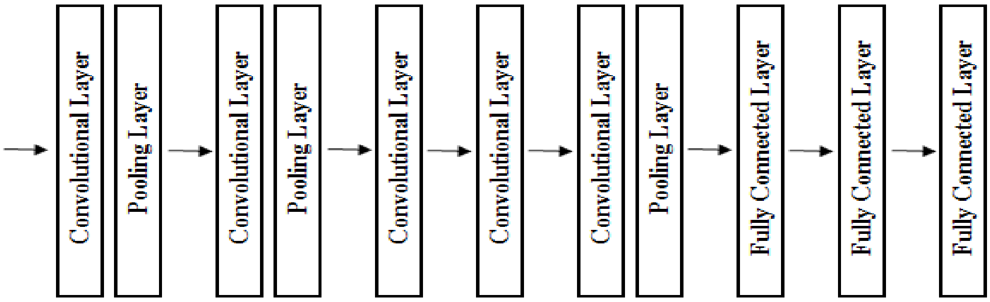


Figure 2.14: The architecture of AlexNet and CaffeNet models.

2.4.2 CaffeNet

The CaffeNet (Convolutional Architecture for Fast Feature Embedding) model was developed and maintained by the Berkely Vision and Learning Center (BVLC) (Jia et al., 2014) with the same architecture as AlexNet model (Figure 2.14) except that the data augmentation is excluded in the training and the order of the pooling and normalization layers are exchanged (Hu et al., 2015). The ImageNet dataset is employed for training, validating and testing purposes which achieved a comparable performance with AlexNet (Hu et al., 2015). CaffeNet has been employed in remote sensing images recognition (Hu et al., 2015) and plant species identification (Lee et al., 2017; Sladojevic et al., 2016).

2.4.3 VGGNet

Another CNN-based method, VGG is developed by Simonyan and Zisserman (2014). There are several architectures of VGG models which composed of different number in the convolution layers (8-16 layers) with 3 fully connected layers.

Figure 2.15 shows one of the VGG architecture (VGG-16) which consists of sixteen learned layers with thirteen convolution layers and three fully connected layers. VGGNet model has been employed popularly as a pre-trained model in various studies such as scene recognition (Wang et al., 2015), iris recognition (Minaee et al., 2016) and breast lesion classification (Hadad et al., 2017).

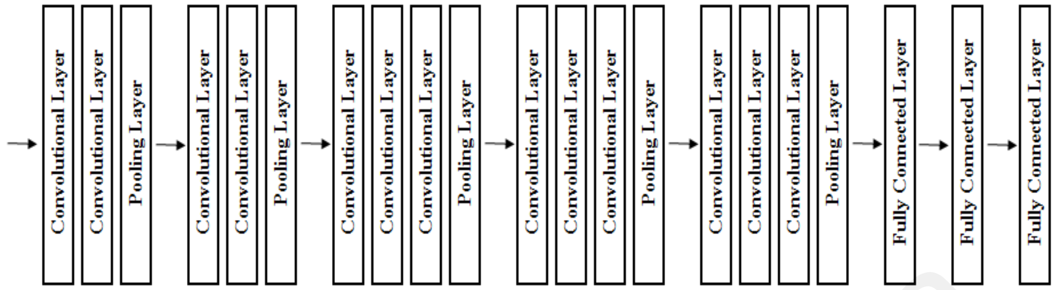


Figure 2.15: The architecture of VGG-16 model.

2.4.4 GoogLeNet

Szegedy et al. (2015) proposed another CNN-based model, namely, GoogLeNet. It is the first model which introduced deep learning with the inception module, which dropped off the great number of network training parameters (Pawara et al., 2017). The inception module made up of three convolution layers (1 x 1, 3 x 3, 5 x 5) and a pooling layer. Basically, a GoogLeNet model consists of 22 deep network layers, which include 9 inception modules as shown in Figure 2.16. It is widely used in handwritten Chinese

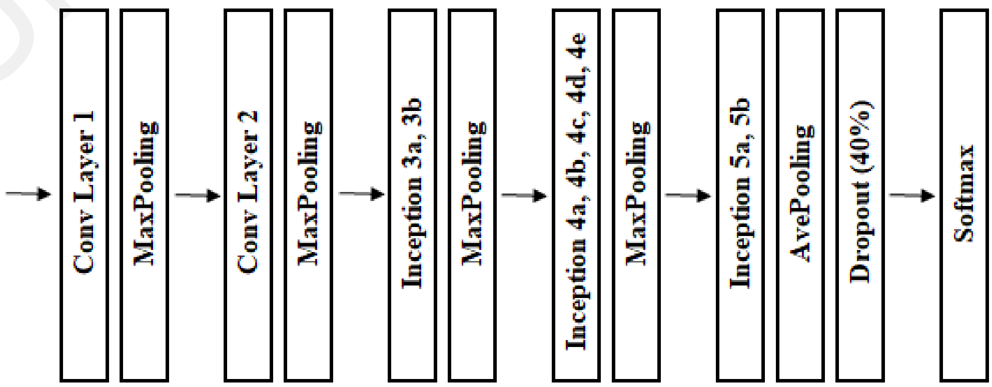


Figure 2.16: The architecture of GoogLeNet model.

character recognition (Zhong et al., 2015), plant recognition (Pawara et al., 2017) and pulmonary tuberculosis classification (Lakhani and Sundaram, 2017).

2.4.5 CNN-based Applications

This section reviews some of the CNN-based studies in various fields. The reviewed studies are summarised and listed in Table 2.3, such as, facial expression recognition (Matsugu et al., 2003; Choi et al., 2018), iris recognition (Minaee et al., 2016), lung pattern classification (Anthimopoulos et al., 2016), pulmonary tuberculosis recognition (Lakhani & Sundaram, 2017), Alzheimer's disease classification (Sarraf & Tofighi, 2016), breast lesion classification (Hadad et al., 2017), vision-based hand gesture recognition (Nagi et al., 2011), handwritten Chinese character recognition (Zhong et al., 2015), house numbers digit classification (Sermanet et al., 2012), traffic sign recognition (Jin et al., 2014), remote sensing images classification (Han et al., 2017; Hu et al., 2015), scene recognition (Wang et al., 2015), gas classification (Peng et al., 2018) and plant identification (Grinblat et al., 2016; Lee et al., 2015; Lee et al., 2016; Lee et al., 2017; Liu et al., 2017; Pawara et al., 2017; Sladojevic et al., 2016).

Table 2.3: CNN-based identification studies.

Type of Identification	Methods	Accuracy	Reference
Facial Expression	Proposed	97.60%	Matsugu et al., 2003
Facial Expression	Proposed	93.95%	Choia et al., 2018
Iris	VGG	99.40%	Minaee et al., 2016
Lung Pattern	Proposed	85.50%	Anthimopoulos et al., 2016
Pulmonary Tuberculosis	GoogLeNet, AlexNet	0.99 (AUC)	Lakhani & Sundaram, 2017

Table 2.3, continued.

Type of Identification	Methods	Accuracy	Reference
Alzheimer's Disease	LeNet-5	96.85%	Sarraf & Tofighi, 2016
Breast Lesion	VGG	>90.00%	Hadad et al., 2017
Vision-based Hand Gesture	Max-Pooling CNN	96.00%	Nagi et al., 2011
Handwritten Chinese Character	HCCR-GoogLeNet	>96.00%	Zhong et al., 2015
House Numbers Digit	ConvNet	95.10%	Sermanet et al., 2012
Traffic Sign	ConvNets	>97.00%	Jin et al., 2014
Remote Sensing Images	AlexNet, CaffeNet, VGGNet, PlacesNet	> 96.00%	Hu et al., 2015
Remote Sensing Images	AlexNet	> 90.00%	Han et al., 2017
Scene	VGG	50%-82%	Wang et al., 2015
Gas	Proposed - GasNet	95.30%	Peng et al., 2018
Plant	VGG	> 97.00%	Lee et al., 2015
Plant	VGG	54% - 71%	Lee et al., 2016
Plant	CaffeNet	>91.00%	Lee et al., 2017
Plant	Proposed	96.90%	Grinblat et al., 2016
Plant	GoogLeNet, AlexNet	>75.00%	Pawara et al., 2017
Plant Disease	CaffeNet	>95.00%	Sladojevic et al., 2016
Plant Disease	Proposed	97.62%	Liu et al., 2017

2.4.6 Plant Species Identification with CNN

Currently, there are very few published articles on studies that have applied CNN in plant species identification. Most of the plant identification studies employed pre-trained

CNN models for feature extraction rather than proposing a new CNN model. This can help to reduce the time consuming in training a new CNN model.

Lee et al. (2015) proposed a CNN model to identify 44 plant species acquired from the Royal Botanic Gardens of Kew, England. A set of whole leaf images and another set of manually cropped leaf patch images were tested in this study by employing a pre-trained CNN model for feature extraction and deconvolutional network (DL) for unique features filtration. The extracted features were then classified with a Multilayer Perceptron (MLP) and SVM. This proposed approach obtained an accuracy of greater than 97%.

Lee et al. (2016) proposed another study which employed a high-level fusion CNN model with the combination of features from species and organs. A comparison study was performed between the fine-tuned VGG-16, a pre-trained model, and the proposed method. However, the VGG models (species features only) outperformed their proposed methods (combination of species and organ features). The proposed approach obtained an accuracy of only 54.4% and 68.9% for non-augmented and augmented models respectively, while the VGG models achieved an accuracy of 56.4% and 71.2% for non-augmented and augmented models respectively.

Recently, Lee et al. (2017) investigated in another study in which the Caffe framework was employed as training model. Deconvolutional network (DN) was used in this study for image visualisation on the extracted features. Two different datasets were used, the dataset of whole images and dataset of leaf patches. Both datasets achieved an accuracy of more than 97% using MLP and the SVM. Furthermore, the researchers had combined both local and global features together and achieved more than 91% accuracy.

Another study based on leaf vein morphological patterns and using deep learning technique for plant identification was proposed by Grinblat et al. (2016). The authors trained the CNN models with different number of layers which ranged from 2 layers (1 convolutional layer + 1 Softmax layer) to 6 layers (5 convolutional layers + 1 Softmax layer). The CNN 5-layers model which combined veins with three different scale factors (100%, 80% and 60%) performed the best with an average accuracy of 96.9%.

Sünderhauf et al. (2014) proposed a study that employed CNN to extract the features from the images of LifeCLEF Plant Task, and classified with Extremely Randomized Tree classifier. For feature extraction, a pre-trained CNN model was employed and applied on 7 categories of images which were entire plant, flower, leaf, stem, fruit, branch and leafScan. The leafScan refers to the leaf images with homogenous background, in which the leaf was well centred. However, the other images were compromised of various background and the object was not well centred. The best performance was achieved by the leafScan categories with an accuracy of more than 50%. Furthermore, the authors made a comparison between the features extracted from the first and the second fully connected layers of CNN. From the results, it was notable that the performance of the models according to extracted features from both the fully-connected layers was competitive.

Furthermore, the CNN was employed by Sladojevic et al. (2016) for plant diseases identification. This study was tested on 13 classes of plant by using an open source pre-trained network model known as the CaffeNet model, which consisted of a set of weights trained by using ImageNet. This model consists of 5 convolutional layers and 3 fully-connected layers out of 8 learning layers. The performances were improved from accuracy of 95.8% (before fine-tuning) to 96.3% (after fine-tuning) with 100 training iterations.

A study has been done by Liu et al. (2017) which proposed a CNN model to identify diseases of apple leaves based on the AlexNet model. A total of 13,689 images were collected as the dataset for this research. The first part of the proposed model was built based on the standard AlexNet which was named as the AlexNet precursor. Then, followed by Cascade Inception which included two Inception structures and two max-pooling layers in order to extract an optimal feature set. This model achieved 97.62% of overall accuracy. The execution time for the proposed CNN model is comparable with AlexNet, however, the performance of the proposed CNN model outperformed AlexNet (91.19%) and other models.

2.5 Machine Learning

Classification is a technique that trains the classifier by data to recognize the specific traits of each group of objects and assists the new observation to seek for the group that it belongs to. It can be performed either by using machine learning (Adeniyi et al., 2016; Bijalwan et al., 2014) or statistical approaches (Hefner & Ousley, 2014; Sawaf et al., 2001). The machine learning classification are categorised into supervised and unsupervised approaches. Supervised machine learning refers to an approach which trains a set of data that provided with its corresponding desired or identified output data. However, unsupervised machine learning known as clustering, is normally implemented to train the data without their identified output.

Classification approaches are widely used in various fields of studies. The examples are, Chang et al. (2013) and Tan et al. (2016) in oral cancer prognosis, Mathioulakis et al.

(2018) in the modelling of air-to-water heat pumps, Guo et al. (2018) in protein interaction sites prediction, Adeniyi et al. (2016) in web usage data mining, and others. Among various classifiers, the most commonly used classifiers are Artificial Neural Network (ANN), Support Vector Machine (SVM), Naïve Bayes (NB), Decision Tree (DT) and k-Nearest Neighbour (k-NN).

2.5.1 Support Vector Machine (SVM)

SVM is a machine learning algorithm that has been popularly implemented in various identification system. In general, SVM is a powerful classification method capable of dealing with high dimensional space and data points which are not linearly separated (Caglayan et al., 2013). SVM is associated with those learning algorithms which involves data analysis, regression support and pattern recognition. Additionally, SVM is capable of dealing with various types of data impressively such as categorical, multiple as well as continuous data. The execution of linear SVM on the feature mapped data could be speed up with low storage and improved in the performance (Sulc & Matas, 2014).

Maldonado-Bascón et al. (2007) had proposed a road-sign recognition study based on SVM; Guyon et al. (2002) had used SVM in classifying cancer; Tzotsos et al. (2008) had applied SVM in analysing object-based images and Hartley et al. (2017) had proposed a strong gravitational lens classification based on SVM.

2.5.2 Artificial Neural Network (ANN)

The development of ANN is inspired by the structure and operation of our human's brain. It is constructed by using 3 layers which are input, hidden layer and output. These interconnected nodes are implemented to train the input and predict the output. Firstly, the inputs would be multiplied with the weights and then computed an activation function to predict the outputs as shown in the Figure 2.17 (Gerhenson, 2003). Different number of neurons and hidden layers may have impact on the accuracy rate of the identification. Thus, neural network is commonly applied in those complex categorization and pattern recognition. The architectures of neural network can be different in term of mathematical function, topology, training approaches and information flow (Krenker et al., 2011).

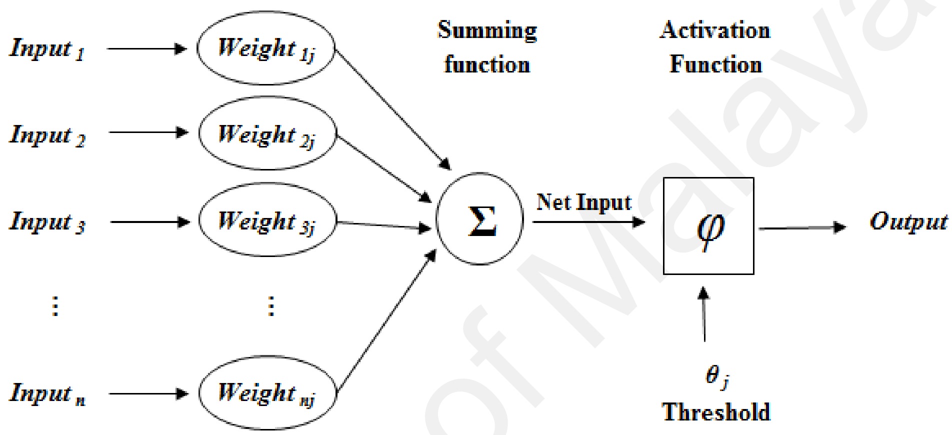


Figure 2.17: General architecture of ANN.

ANN had been widely applied in various fields of studies, such as cancer diagnosis prediction (Khan et al., 2001); in aquatic insect species prediction (Park et al., 2003); in bankruptcy prediction (Wilson & Sharda, 1994) and in weather type classification (Chen et al., 2011).

2.5.3 k-Nearest Neighbour(k-NN)

k-Nearest Neighbour(k-NN), a classification approach which classifies a sample based on the majority vote of its neighbours (Pornpanomchai et al., 2011). In another word, a sample would be classified to the class which was most common with its k-nearest neighbours (Pornpanomchai et al., 2011). The k (number of neighbour) is defined based on the rate of error. In k-NN, it consisted of numerous kinds of distance metrics, such as, city block, cosine and Euclidean.

Bijalwan et al. (2014) and Moldagulova & Sulaiman (2017) had used k-NN in classifying textual documents; Shirvan and Tahami (2011) had applied k-NN in voice analysis for Parkinson's disease detection and Li et al. (2014) had employed in developing a new intrusion detection system in wireless sensor network.

2.5.4 Naïve Bayes (NB)

Bayes classifier is a statistical classifier which makes prediction on the class which an unknown object belongs to based on probability. It computes the frequency and combinations of values of a data set and assumes that all features of the samples are unrelated to each other (McCallum & Nigam, 1998; Patil & Sherekar, 2013). Each feature can be learned individually, and larger number of features would help to simplify the process of learning (McCallum & Nigam, 1998). However, the conditional independence of Bayes theorem may cause the decrease in the performance (Caglayan et al., 2013).

There are various studies which have been conducted based on the NB classifier, such as text classification (McCallum & Nigam, 1998); development of intrusion detection system (Amor et al., 2004) and sentiment analysis (Tan et al., 2009).

2.5.5 Random Forest (RF)

Random forest is a tree-based classification approach, which aggregates the prediction of multiple trees of a data set. Bootstrap samples are then used to grow each tree of the forest. Each tree would contribute to the classification results the trees in the forest attempt to employ their votes for target class. The forest would select the class with highest number of votes. It is able to deal with large data efficiently and perform high accuracy rate (Caglayan et al., 2013).

RF was applied by Díaz-Uriarte and De Andres (2006) for gene selection and microarray data classification; Pal (2005) used RF for remote sensing classification; Svetnik et al. (2003) employed RF for compound classification and Rodriguez-Galiano et al. (2012) used RF for land-cover classification.

2.5.6 Decision tree (DT)

Decision tree is a classifier which classifies a dataset by partitioning the dataset based on the decision framework defined by the tree (Friedl & Brodley, 1997). It is a non-parametric supervised approach, which predicts the value of a considered variable via the simple decision pattern, obtained from the data features (Arun et al., 2013). DT is efficient in handling nonlinear relations between classes and features in both categorical and numeric data, and it is robust to missing data (Friedl & Brodley, 1997). It is capable of processing data that is measured and computed at different scales, and considering the data frequency distribution without any assumptions (Tooke et al., 2009).

DT is widely used in various studies, e.g., Friedl and Brodley (1997) had conducted a study to classify the remotely sensed land cover by using DT; Tooke et al. (2009) had applied DT in classifying urban vegetation characteristics and Tarter (1990) had evaluated and treated adolescent substance abuse by employing a DT method.

2.5.7 The Application of Classifiers in Plant Species Identification

The summary of implementation of classifiers in previous plant species identification studies is as shown in Table 2.4.

Table 2.4: Summary of classifiers that implemented in previous plant identification studies.

Classifiers	Authors	No. of Classes	Accuracy
Neural Network (NN)	Aakif & Khan, 2015	14	68 - 96%
	Chaki et al., 2015	31	41 - 86%
	Danti et al., 2012	10	92 - 100%
	Hati & Sajeevan, 2013	20	92.00%
	Kadir et al., 2013	32	> 95%
	Lee et al., 2015	44	> 97%
	Lee et al., 2017	44	> 97%
	Lin & Peng, 2008	30	98.30%
	Murat et al., 2017	45	98.23%
	Prasvita et al., 2013	30	56.33%
	Sharma & Gupta, 2015	16	> 90%
	Sladojevic et al., 2016	13	91 - 98%
	Wang et al., 2005	20	92.20%
	Wu et al., 2007	32	> 90%
	Yadav et al., 2013	25	88 - 93%
Support Vector Machines (SVM)	Arun et al., 2013	5, 32, 60, 15, 153	58 - 92%
	Caglayan et al., 2013	32	71 - 87%
	Lee et al., 2015	44	> 98%
	Lee et al., 2017	44	> 98%
	Murat et al., 2017	45	32 - 85%
	Sulc & Matas, 2014	5	> 99%
	Wang et al., 2016	32, 220, 153	> 83%

Table 2.4, continued.

Classifiers	Authors	No. of Classes	Accuracy
k-Nearest Neighbour (kNN)	Arun et al., 2013	5	83 -100
	Caglayan et al., 2013	32	81 -93%
	Cope et al., 2010	32	> 79%
	Lin & Peng, 2008	30	> 90%
	Murat et al.,2017	45	82 -92%
	Wang et al., 2005	20	92 - 93%
Bayes' Theorem	Caglayan et al.,2013	32	79 - 89%
	Gwo & Wei, 2013	13	70 -93%
	Kadir et al., 2014	32, 60	95 -97%
Random Forest (RF)	Arun et al., 2013	5	81 - 93%
	Caglayan et al.,2013	32	86 -94%
	Murat et al.,2017	45	> 83%
Hypersphere	Du et al., 2007	20	91%
	Wang et al., 2005	20	92.20%
Extra Trees	Arun et al., 2013	5	70 - 88%
	Sünderhauf et al., 2014	--	< 60%
Stochastic Gradient Descent	Arun et al., 2013	5	92 -100%
	Grinblat et al., 2016	--	> 77%
Others			
Decision Trees	Arun et al., 2013	5	75 - 93%
Eigenspace	Ehsanirad & Sharath Kumar, 2010	13	98%
Incremental Classification	Beghin et al., 2010	18	81.10%
Linear Discriminant Analysis	Murat et al.,2017	45	37 - 83%
Neuro-Fuzzy Controller	Chaki et al., 2015	31	97.60%

As shown in Table 2.4, Artificial neural network (ANN) is the most common and frequent classifier that have been employed in numerous plant identification researches. Murat et al. (2017) and Hati and Sajeevan (2013) applied the basic ANN as a classifier obtaining high accuracies of 98.23% and 92% respectively. One of theANN architecture, which is the back-propagation ANN, has been frequently employed in plant identification studies by researchers Yadav et al. (2013), Wang et al. (2005), Danti et al. (2012), Chaki et al. (2015) and Lin and Peng (2008). The ANN model with this architecture design can recognise and classify the images efficiently since it can perform a high accuracy rate (>80%) in most cases. However, the identification rate is highly dependent on the extracted features. If irrelevant or less optimum features are extracted, it may cause inaccurate rate of identification and decrease the performance. For example, in the study by Aakif & Khan (2015), the extracted morphological features alone resulted 68.3% accuracy but when combined with the extracted features from Fourier descriptor and shape-defining features, the accuracy increased up to around 96%. In another study by Chaki et al. (2015), the performance of the back-propagation ANN with shape features was only 41.6%, but, after they were combined with texture features, the classification rate increased to 85.6%.

Probabilistic Neural Network (PNN), another alternative method to the back-propagation neural network which derived from Radial Basis Function (RBF) (Prasvita & Herdiyeni, 2013; Sharma & Gupta, 2015), which is famous for its fast training speed and robustness to noise. Prasvita and Herdiyeni (2013) had tested the extracted texture features with PNN, but, the accuracy obtained was 56.33% only due to the low quality of acquired images. Wu et al. (2007), Lin and Peng (2008) and Kadir et al. (2013) had achieved a high identification rate of more than 90%.

In classification using SVM, Murat et al. (2017) had obtained an accuracy which ranged from 32-85% with different set of extracted features, Caglayan et al. (2013) obtained an accuracy which ranged from 72-93%, Arun et al. (2013) obtained an accuracy that ranged from 58-92% in a different category, Sulc and Matas (2014) obtained more than 99% of accuracy and Wang et al. (2016) achieved more than 83% of identification rate. Murat et al. (2017) has employed another version of SVM known as the directed acyclic graph multiclass least squares twin support vector machine (DAG MLSTSVM) in their study, achieving more than 85% accuracy. DAG MLSTSVM used the directed acyclic graph to choose and rebuild the classifier in "one-versus-one' approach (Murat et al., 2017).

Other than that, the studies that employed k-NN as a classifier achieved a good performance with more than 80% of identification rate. Murat et al. (2017) obtained an accuracy that ranged from 82.99% to 91.96% with different sets of shape features and Caglayan et al. (2013) achieved more than 81% of accuracy with different extracted features. Arun et al. (2013) obtained more than 88.0% in different combination of extracted features, whereas Cope et al. (2010) achieved an identification rate of 79.69% with texture features. In addition, Wang et al. (2005) employed two different k-NN as classifier, which were 1-NN and 4-NN, obtaining 92.6% and 92.3% accuracy respectively. Lin and Peng (2008) had obtained more than 90% of accuracy as well with 1-NN and 4-NN.

Caglayan et al. (2013), Gwo et al. (2013) and Kadir et al. (2014) had implemented Bayes classifier in their identification system and the identification rate had exceeded 80% with different extracted features included in the classification process.

As for the Random Forest classifier, Murat et al. (2017) achieved a good performance which was more than 83% of accuracy by using shape features in Random Forest classifier. Whilst, Caglayan et al. (2013) had tested in different combination of extracted features and achieved more than 86% accuracy. Random forest was employed by Arun et al. (2013) achieved a high accuracy range in the data without pre-processing which was 81% to 93%.

Hypersphere classification algorithm was also used to implement plant species identification model. The identification rate was 91% in Du et al. (2007) study and 92.2% in Wang et al. (2005) study respectively which was slightly lower than other methods such as 1-NN, 4-NN and Back-Propagation Neural Network.

Extra Tree is a classifier which use a meta-estimator and attempts to adjust extra-tree on diverse sub-objects of dataset. It employs the averaging to increase the identification accuracy and regulate the over-fitting problems (Arun et al., 2013). A moderately high performance ranging 70% to 88% was performed in the study by Arun et al. (2013). However, the identification rate in the works of Sünderhauf et al. (2014) was less than 60% by using extra tree.

Stochastic Gradient Descent (SGD) classifier is capable of performing at a high accuracy as shown in study by of Arun et al. (2013) with up to 94.7% for the non-pre-processed data. A test had been done by Arun et al. (2013) which used SGD on 5 different plant species. It was able to achieve 92% to 100% of accuracy with less than 0.1 of error rates. Grinblat et al. (2016) employed SGD as a classifier to classify the CNN features and obtained more than 77% of accuracy.

There are other classifiers that had been employed in some previous studies such as Decision Trees (DT), Eigenspace, Extra Trees, Incremental Classification, Linear Discriminant Analysis and Neuro-Fuzzy Controller. By using DT classifier, the performance of the study by Arun et al. (2013) ranged from 82% to 93% accuracy for species identification in the data without any pre-processing. Whilst, in the study by Ehsanirad et al. (2010) which the team used Eigenspace algorithm to build the leaves identification system, successfully classified the texture features by using PCA, obtaining with 92% accuracy. However, this algorithm only achieved an accuracy of 78% when the texture features were extracted by using GLCM.

Incremental classification had been proposed by Beghin et al. (2010) to identify leaves. By using this classifier, the achieved identification rate was up to 81.1% in average. However, LDA achieved a poor performance as compared to other classifiers with an accuracy ranging from 37% to 83% in Murat et al. (2017). Instead of exclusive classification, Neuro-Fuzzy Controller integrates the advantages of fuzzy classifier and neural network together (Chaki et al., 2015). In this case, the fuzzy classifier provides the probabilities of several classes of an observation that it may belong to (Chaki et al., 2015). In the study by Chaki et al. (2015), the accuracy achieved up to 97.60% by using this classifier.

CHAPTER 3: METHODOLOGY

3.1 Introduction

Various feature extraction techniques, including deep learning-based approaches and classification algorithms which were applied to plant images research have been reviewed in the previous chapter. The promising results from deep learning-based approaches were also presented. The methodology adopted for this research, is discussed in the following sections.

Basically, there are four components which are explained in detail in this chapter. The four components are:

(i) Field sampling of leaf samples and acquisition of leaf images.

(ii) The application of pre-processing on the leaf image dataset.

(iii) The application of deep learning and convention methods for leaf feature extraction.

(iv) The evaluation of extracted features using classification methods.

3.2 Proposed Architecture

Basically, this research proposed four main steps for developing an automated plant species identification based on deep learning technique as shown in Figure 3.1. The four main steps included sampling, image pre-processing, feature extraction and classification. Firstly, the leaf samples were collected, and images were acquired. The acquired leaf images were then pre-processed and important information were retrieved from the images through feature extraction. Lastly, the extracted features were fed into the classifiers for training and testing purposes.

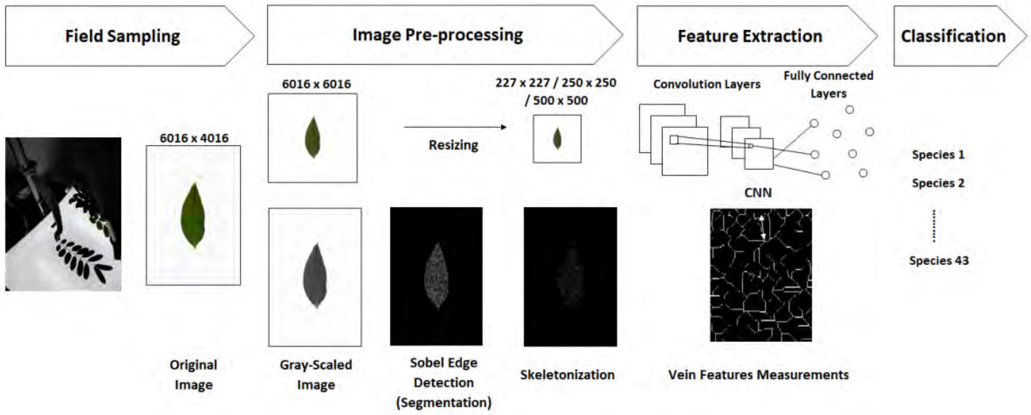


Figure 3.1: Proposed Architecture of the Research.

3.3 Software and Hardware

Workstation with Intel® Xeon® CPU E5-1603 v3 @ 2.80GHz processor, 32GB of RAM and Nvidia Quadro K2200 4GB was used in the development and execution of this research as listed in Table 3.1.

Table 3.1: Specification of hardware used in this research.

Specifications	**Details**
Processor	Intel® Xeon® CPU E5-1603 v3 @ 2.80GHz
Graphic Processor Unit	Nvidia Quadro K2200 4GB
Operating System	Window 7 Professional (64-bit)
RAM	32GB

Adobe Photoshop was used for image pre-processing and Neural Network toolbox of MATLAB version 2016a (for feature extraction and classification) and 2017a (for visualisation) were used in feature extraction and classification. Adobe Photoshop is a powerful graphic editor developed by Adobe system which can be employed to edit, enhance, create, etc. image as well as video. Whilst, MATLAB is developed by Math

Works Inc., which stands for Matrix Laboratory. It is a multi-paradigm numerical computing environment which performs scientific computations and visualization. Its capability in analysing various scientific problems, flexibility and powerful graphics makes it a very useful software package.

3.4 Field Sampling

The leaf samples of this research were collected from four locations in the University of Malaya, Kuala Lumpur, Malaysia. These locations were the areas around the main library (ML), Varsity Lake (VL), Dewan Tunku Canselor hall (DTC) and the Faculty of Science (FS) as shown in Figure 3.2 (Google Map). The leaf samples of this research were mainly collected from the common tropical plant species, which can be obtained simply in the University of Malaya and anywhere else in Malaysia. Leaf is the part that is always chosen as the studied part due to its availability all year round, instead of choosing flowers, fruits or other parts of the plant. 43 species of tropical plants with 30 samples per species were collected as the dataset of this study. Consequently, a total 1290

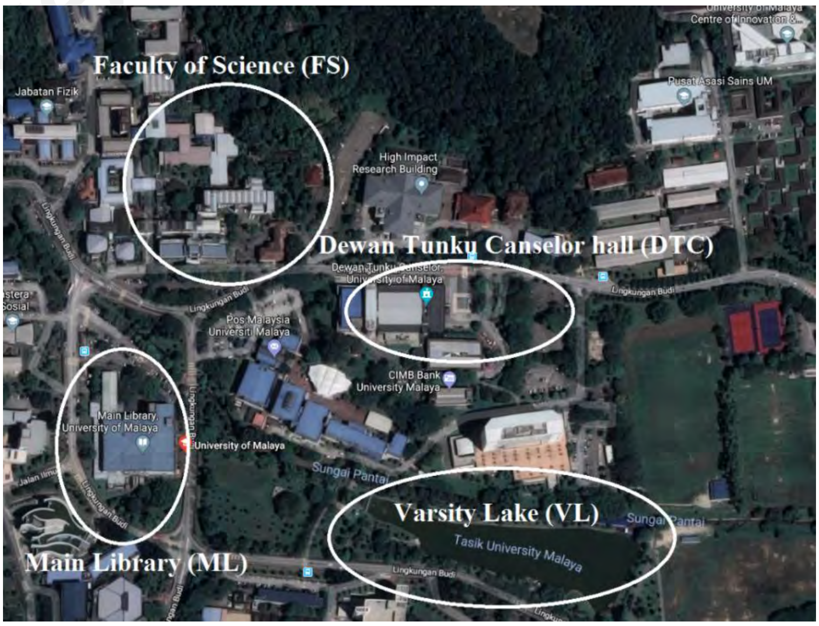


Figure 3.2: The locations for sample collection (Google Map, 2018).

leaf images were acquired for this research. The chosen species with their scientific name, common name and location are shown in Table 3.2.

Table 3.2: The details of the selected plant species.

No.	Species Name	Common Name	Location
1	*Acacia auriculiformis* Benth.	Earleaf Acacia, Common Acacia	DTC
2	*Filicium decipiens* (Wight & Arn.) Thwaites	Fern Tree	DTC
3	*Alstonia scholaris* (L.) R.Br.	Indian Pulai, Pulai Tree	ML
4	*Barringtonia racemosa* (L.) Spreng.	Putat Kampung	ML
5	*Bucida molinetii* (M.Gómez) Alwan & Stace	Spiny Black Olive	ML
6	*Dryobalanops aromatica* C.F.Gaertn.	Kapur, Borneo Camphor	ML
7	*Hopea odorata* Roxb.	White Thingan, Cengal Pasir	ML
8	*Syzygium aqueum* (Burm.f.) Alston	Water Apple, Jambu Air	ML
9	*Adenanthera pavonina* L.	Red Bead tree, Red Sandalwood, Saga	VL
10	*Albizia saman* (Jacq.) Merr.	Rain Tree	VL
11	*Aquilaria malaccensis* Lam.	Agarwood, Gaharu	VL
12	*Artocarpus integer* (Thunb.) Merr.	Chempedak	VL
13	*Bauhinia blakeana* Dunn	Hong Kong Orchid Tree	VL
14	*Cassia fistula* L.	Golden Rain Tree	VL
15	*Cinnamomum iners* Reinw. *ex* Blume	Wild Cinnamon, Clove Cinnamon	VL
16	*Cynometra malaccensis* Meeuwen	Belangkan, Kekatong	VL
17	*Delonix regia* (Hook.) Raf.	Royal Poinciana, Flame Tree	VL
18	*Dipterocarpus grandiflorus* (Blanco) Blanco	Keruing Belimbing	VL
19	*Erythrina variegata* L.	Indian Coral Tree	VL

Table 3.2, continued.

No.	Species Name	Common Name	Location
20	*Eucalyptus alba* Reinw. *ex* Blume	Ridge Gum	VL
21	*Cyrtophyllum fragrans* (Roxb.) DC.	Ironwood, Tembusu	VL
22	*Ficus microcarpa* L.f.	Indian Laurel, Malayan Banyan	VL
23	*Hura crepitans* L.	Sandbox Tree	VL
24	*Hymenaea courbaril* L.	West Indian Locust	VL
25	*Khaya senegalensis* (Desv.) A.Juss.	African Mahogany	VL
26	*Lagerstroemia floribunda* Jack	Kedah Bungor	VL
27	*Mangifera indica* L.	Mango	VL
28	*Melaleuca cajuputi* Powell	Tea Tree, Cajuput Tree, Gelam	VL
29	*Mesua ferrea* L.	Indian Rose Chestnut	VL
30	*Mimusops elengi* L.	Spanish Cherry, Tanjong Tree	VL
31	*Plumeria rubra* L.	Bunga Kubur	VL
32	*Polyalthia longifolia* (Sonn.) Thwaites	False Ashoka Tree, Mast Tree	VL
33	*Pterocarpus indicus* Willd.	Narra, Angsana Tree	VL
34	*Saraca thaipingensis* Prain	Yellow Saraca	VL
35	*Senna surattensis* (Burm.f.) H.S.Irwin & Barneby	Scrambled Egg Bush	VL
36	*Spathodea campanulata* P.Beauv.	African Tulip Tree, Pancut-pancut	VL
37	*Sterculia foetida* L.	Hazel Sterculia, Wild Almond	VL
38	*Swietenia macrophylla* King	Sky Fruit, Big-leaf Mahogany	VL
39	*Syzygium myrtifolium* Walp.	Kelat Paya, Red Lip, Kelat Oil	VL
40	*Tabebuia rosea* (Bertol.) Bertero *ex* A.DC.	New World Trumpet, Trumpet Tree	VL
41	*Terminalia catappa* L.	Ketapang, Indian Almond	VL
42	*Theobroma cacao* L.	Cocoa Tree	VL

Table 3.2, continued.

No.	Species Name	Common Name	Location
43	*Tristaniopsis whiteana* (Griff.) Peter G.Wilson & J.T.Waterh.	River Tristania	VL

VL- Varsity Lake, ML - Main Library, DCT - Dewan Tunku Canselor hall

The leaf images were captured by using a D750 Nikon DSLR camera. During image acquisition, in order to capture good quality images with standard background, the leaf samples were pasted on a white background box with fluorescent light at the bottom of the box as shown in Figure 3.3. It helped to reduce the formation of shadow and glaze on the leaf image, as the lighting of this setup is illuminated from the bottom of the sample.

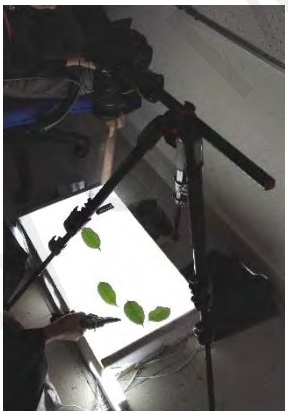


Figure 3.3: The setup for image acquisition.

The leaf sample of each species are illustrated as in Figure 3.4.

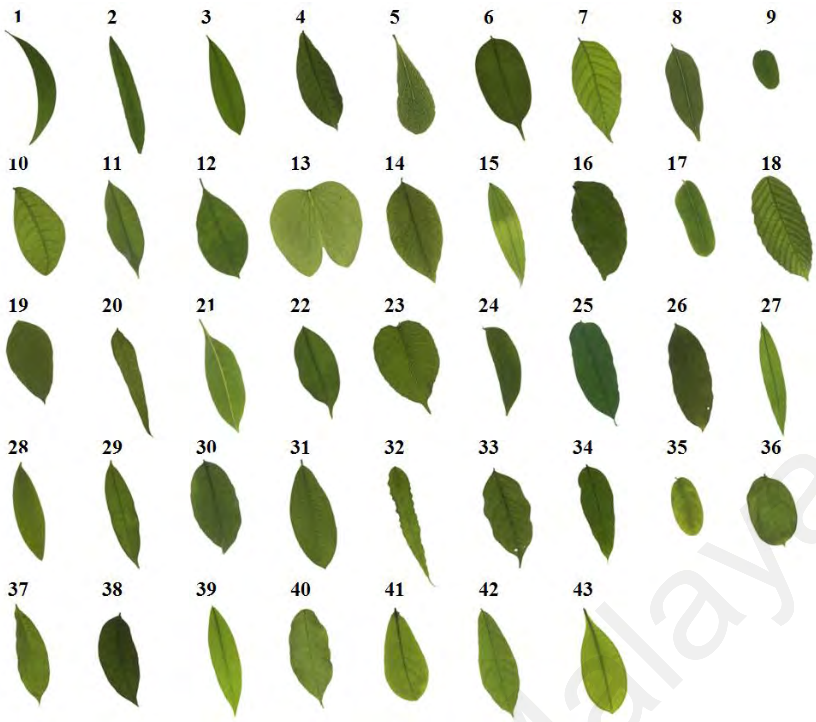


Figure 3.4: Leaf sample of each species.

From Figure 3.4, it can be seen that some species in the proposed dataset possessed a high similarity in their colour and shape. For example, Species 29 and Species 39 are highly similar in shape.

3.5 Image Pre-processing

Raw images were in the original format of the camera which was not proper for study purposes, thus, there was a need to convert the raw images into the processed format, such as, jpg, jpeg and tiff. The acquired raw images in this research were stored in a Nikon camera format which is known as the Nikon Electronic File (NEF) with 6016 x 4016 of resolution. Therefore, Adobe Photoshop was used to convert the raw images into Tagged Image File Format (TIFF) and reduce the background noises of the images. Background noises were referred to pixel values which does not represent the true intensities of an image during the image acquisition (Malladi & Sethian, 1996). Background noise removal could help to enhance and highlight the important information of the leaf images. After that, the leaf images were subjected to two different pre-processing stages, namely, image reconstruction for CNN and edge detection method – Sobel.

3.5.1 Image Reconstruction for CNN

Image reconstruction is necessary since only the images with square dimension (m x m) are acceptable as the input in CNN. The original leaf images with 6016 x 4016 resolution were pre-processed by adding 1000 zero padding at both sides to produce a square dimension images with 6016 x 6016 as shown in the Figure 3.5. This process was performed before resizing as to ensure the ratio of major axis length to minor axis length could be maintained. Then, the leaf images were resized into several sizes, which were, 227 x 227, 250 x 250, 500 x 500 and 750 x 750 resolution.

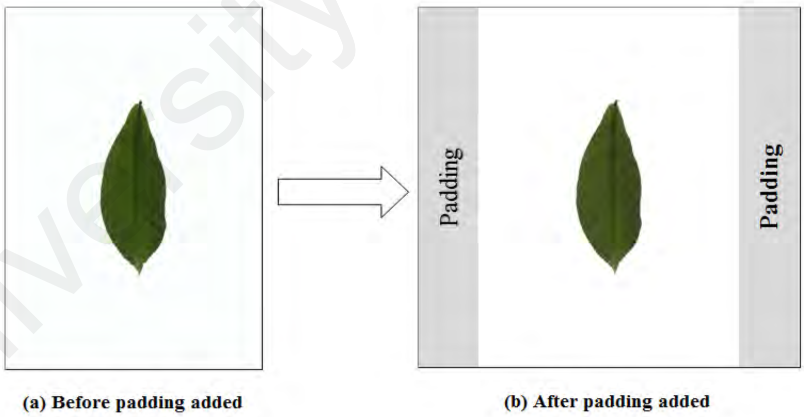


Figure 3.5: The image before and after padding was added.

3.5.2 Edge Detection Method – Sobel

Firstly, the RGB images were converted into greyscale images. Then, Sobel was employed to segment the region of interest (ROI) from the greyscale images, which was the vein architecture in this research. The segmented images were then subjected to the post-processing and skeletonising to ensure a clean vein architecture could be acquired. The process of vein architecture extraction is as shown in the Figure 3.6.

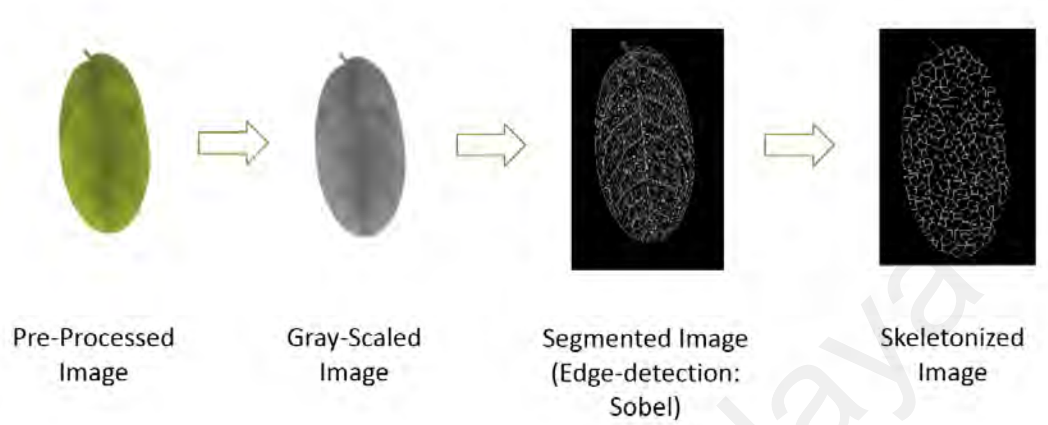


Figure 3.6: Process of vein architecture extraction.

3.5.3 Leaf Patch

Whole leaf of each sample was cropped into two different types of patches which were centre patch and side patch. The centre patches were cropped automatically based on the centroid of the leaf samples by using MATLAB at the centre of the leaf which generally consisted of the primary vein, whilst and the side patches were cropped automatically at

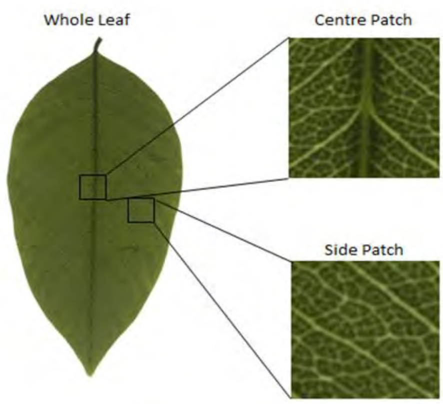


Figure 3.7: Centre leaf patch and side leaf patch.

the side of the leaf which generally consist of the veins other than the primary veins as shown in Figure 3.7.

3.6 Feature Extraction

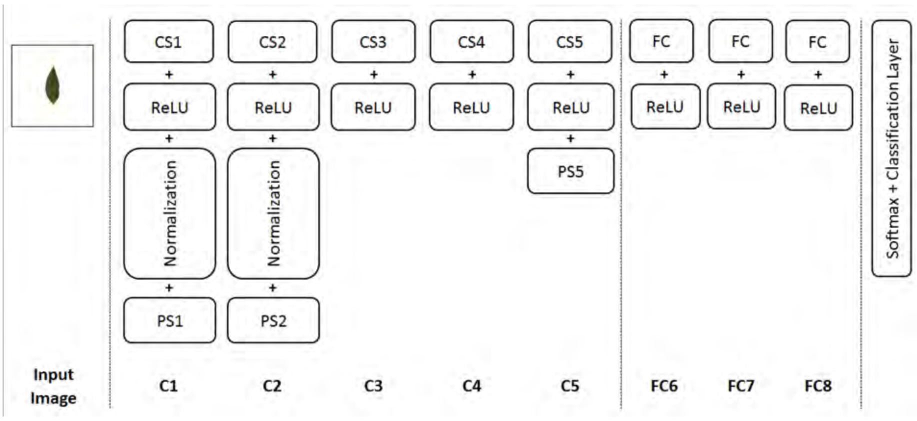
In this research, Convolutional Neural Network, a deep learning algorithm, was used for feature extraction. In addition, a conventional approach was investigated and compared by computing the morphological features of the leaf venation which was segmented by Sobel edge detection approach.

3.6.1 Deep learning – Convolutional Neural Network

Three different CNN models were proposed for feature extraction in this research to find out the most optimal model for plant identification system. The aim was extracting the most optimal features within the shortest time and minimal costs for development of an automated plant identification system. The proposed CNN models were pre-trained AlexNet, fine-tuned AlexNet and D-Leaf.

3.6.1.1 Pre-trained AlexNet Model

As mentioned in Chapter 2, AlexNet model is a pre-trained CNN model which was trained with the images from 1000 different classes (Krizhevsky et al., 2012). The details of AlextNet architecture which was employed in this research is shown in Figure 3.8. After five convolution layers and three fully connected layers, the AlexNet model would be added with a softmax classification layer as illustrated in Figure 3.8. Softmax classification layer is referred to the softmax function which yield the predicted probability of each group and is fully-connected to the final convolutional layer.



CS - Convolutional Stage, C - Convolutional Layer, PS - Pooling Stage, FC - Fully-Connected Layer

Figure 3.8: The AlexNet Architecture for this research.

In AlexNet, the input images were in 227 x 227 x 3 of resolution. The last three fully-connected layers in this model were set with 4096, 4096 and 1000 neurons respectively. This was followed by the softmax classification layer. The parameters of the AlexNet model are as shown in Table 3.3.

Table 3.3: Parameters of AlexNet model.

CNN Layer	Filter Size	No. of Kernel	Size of Stride[1]
CS1	11 x 11	96	[4 4]
PS1	3 x 3	---	[2 2]
CS2	5 x 5	256	[1 1]
PS2	3 x 3	---	[2 2]
CS3	3 x 3	384	[1 1]
CS4	3 x 3	384	[1 1]
CS5	3 x 3	256	[1 1]
PS5	3 x 3	---	[2 2]
FC6	---	4096	---
FC7	---	4096	---
FC8	---	1000	---

--- Not Applicable, [1]Distance between the receptive field centers of neighbouring neurons (Krizhevsky et al., 2012)

3.6.1.2 Fine-tuned AlexNet Model

In order to make the pre-trained model more specific with the proposed dataset, the pre-trained AlexNet model was fine-tuned. Yet, the architecture of the fine-tuned AlexNet model remained the same as the original AlexNet model with five convolutional layers, three fully connected layers and a softmax classification layer as shown in Figure 3.8. Some layers of AlexNet were fine-tuned as shown in Table 3.4 in order to avoid the overfitting problem, as the proposed dataset is small as compared to the original training dataset of AlexNet.

Table 3.4: Parameters of fine-tuned AlexNet model.

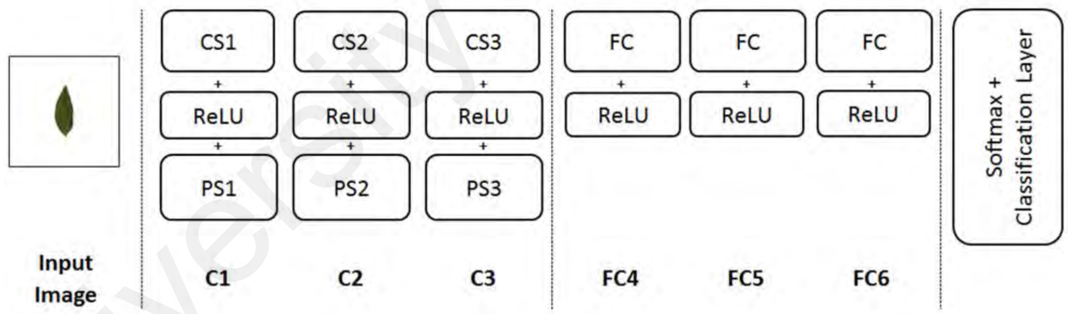
CNN Layer	Filter Size	No. of Kernel	Size of Stride[1]
CS1	7 x 7	96	[2 2]
PS1	3 x 3	---	[2 2]
CS2	5 x 5	256	[1 1]
PS2	3 x 3	---	[2 2]
CS3	3 x 3	384	[1 1]
CS4	3 x 3	384	[1 1]
CS5	3 x 3	256	[1 1]
PS5	3 x 3	---	[2 2]
FC6	---	1290	---
FC7	---	1290	---
FC8	---	43	---

--- Not Applicable, [1]Distance between the receptive field centers of neighbouring neurons (Krizhevsky et al., 2012)

The input size of the image remained the same as in the original AlexNet model which is 227 x 227 x 3. The original AlexNet would cause aliasing effects on the second layer and create "dead" kernels for the first layer. Thus, the filter size and stride of the first convolutional stage (CS1) were reduced into 7 x 7 and [2 2] respectively in order to solve the two issues and to improve the performance. Additionally, the last layer of fully connected layer was set with the number of the plant species. So, the three fully connected layers were fine-tuned into 1290, 1290 and 43 neurons. However, the other parameters remained unchange as in the original AlexNet model.

3.6.1.3 Proposed Model – D-Leaf

In this research, a CNN model with the name D-Leaf is proposed for leaf feature extraction with a total of seven layers. It is a simpler model than the AlexNet model as the AlexNet model consists of more number of layers and takes a longer execution time. This model consists of three convolution layers, three fully connected layers, and a softmax classification layer. Each convolution layer consists of a convolution stage, a ReLU layer and a pooling stage as depicted in the Figure 3.9.



CS - Convolutional Stage, C - Convolutional Layer, PS - Pooling Stage, FC - Fully-Connected Layer

Figure 3.9: The architecture of D-Leaf.

There were three different image inputs, with the size of 250 x 250 x 3, 500 x 500 x 3 and 750 x 750 x 3, were tested to obtain the foremost model for automated plant species identification development. The parameters of convolutional layer and fully connected layers of the D-Leaf model were kept consistent even different sizes of input images were used. The parameters of each D-Leaf layer are shown in Table 3.5.

Table 3.5: Parameters of D-Leaf model.

CNN Layer	Filter Size	No. of Kernel	Size of Stride*
CS1	11 x 11	64	[4 4]
PS1	2 x 2	---	[2 2]
CS2	5 x 5	96	[2 2]
PS2	2 x 2	---	[2 2]
CS3	4 x 4	256	[1 1]
PS3	2 x 2	---	[2 2]
FC4	---	1290	---
FC5	---	1290	---
FC6	---	43	---

--- Not Applicable, 1 Distance between the receptive field centers of neighbouring neurons (Krizhevsky et al., 2012)

The input leaf images were filtered by the first convolutional layer (C1) with the first convolutional stage (CS1) which was made up of a filter size of 11 x 11, 64 kernels and a stride in [4 4] size. This was followed by a pooling stage consisting of a 2 x 2 filter and [4 4] in the stride size. Each of the convolutional layers was followed by a pooling stage using the same parameters. Next, the feature maps from the C1 was convolved with 96 filters of size 5 x 5 and [2 2] stride followed by a pooling stage. The pooled feature maps were passed to the next stage – the third convolutional layer (C3). In the C3, the convolutional stage consisted of a 4 x 4 filter, 256 of kernels and a [1 1] stride followed by a pooling stage. The output of C3 was fed into three fully connected layers which consisted of 1290, 1290 and 43 neurons, respectively.

3.6.2 Vein Morphometric Measurements

As mentioned in the previous section of this chapter, an edge detection method-Sobel was employed to segment the vein architecture from the leaf images. The first step was the conversion of RGB or coloured images to grey-scaled images and followed by the vein architecture segmentation. After that, the vein architecture would be post-processed by using skeletonizing method (convert the objects of the leaf image into lines, without changing the essential structure of the vein architecture) to ease the vein morphometric measurement. a Finally, features were extracted by computing and measuring the morphology of the vein architecture, with a threshold of 0.05 to compute the magnitude of gradient. Based on the vein architecture that has been segmented by using Sobel method, the vein features were extracted by computing or measuring the morphological features of the veins. A total of 62 features were extracted from vein, such as, number of vein branches, number of ending points, number of branching points, number of areoles and others as listed in Table 3.6.

Table 3.6: List of vein features.

1	Leaf Area	32	Mean Areoles Perimeter
2	No. of Branching Point	33	Median Areoles Perimeter
3	No. of Ending Point	34	Max. Areoles Perimeter
4	No. of Branches	35	Min. Areoles Perimeter
5	Total Lgth of Branches	36	Total Areoles Convex Area
6	Total Area of Branches	37	Mean Areoles Convex Area
7	Mean of Branch Area	38	Median Areoles Convex Area
8	Median of Branch Area	39	Max. Areoles Convex Area
9	Max. Branch Area	40	Min. Areoles Convex Area
10	Min. Branch Area	41	Total Areoles Solidity
11	Mean of Branch Lgth	42	Mean Areoles Solidity
12	Median of Branch Lgth	43	Median Areoles Solidity
13	Max. Branch Lgth	44	Max. Areoles Solidity
14	Min. Branch Lgth	45	Min. Areoles Solidity
15	Mean Branch Width	46	Mean Areoles Major Axis Lgth
16	Median Branch Width	47	Median Areoles Major Axis Lgth

Table 3.6, continued.

17	Max. Branch Width	48	Max. Areoles Major Axis Lgth
18	Min. Branch Width	49	Min. Areoles Major Axis Lgth
19	Den. of Vein	50	Mean Areoles Minor Axis Lgth
20	Den. of Branching Point	51	Median Areoles Minor Axis Lgth
21	Den. of Ending Point	52	Max. Areoles Minor Axis Lgth
22	Complexity	53	Min. Areoles Minor Axis Lgth
23	Feature Points	54	Mean Areoles Eccentricity
24	Number of Areoles	55	Median Areoles Eccentricity
25	Total Areoles Area	56	Max. Areoles Eccentricity
26	Mean Areoles Area	57	Min. Areoles Eccentricity
27	Median Areoles Area	58	Mean Areoles Equivdiameter
28	Max. Areoles Area	59	Median Areoles Equivdiameter
29	Min. Areoles Area	60	Max. Areoles Equivdiameter
30	SD of Areoles Area	61	Min. Areoles Equivdiameter
31	Total Areoles Perimeter	62	Den. of Areoles

No. – Number, Lgth – Length, Max. – Maximum, Min. – Minimum, Den. – Density, SD – Standard Deviation

3.7 Classification

In the last stage of an automated plant species classification system, the extracted features from the leaf images were then fed into a classifier for training and recognising. Five classifiers were proposed in this research, namely, Support Vector Machines, Artificial Neural Network, k-Nearest Neighbours, Naïve Bayes and Convolutional Neural Network.

3.7.1 Support Vector Machines (SVM)

SVM consists of four basic kernel functions which are linear, polynomial, radial basis function (RBF) and sigmoid. The formulae for the kernel function are stated as below (Hsu et al., 2003):

i. Linear: $K(x_i , x_j)= x_i^T x_j$ (3.1)

ii. Polynomial: $K(x_i , x_j)=(\gamma x_i^T x_j + r)^d , \gamma > 0$ (3.2)

iii. RBF: $K(x_i , x_j) = exp(-\gamma || x_i - x_j ||^2) , \gamma > 0$ (3.3)

iv. Sigmoid: $K(x_i , x_j)= tanh(\gamma x_i^T x_j + r)$ (3.4)

where $K(x_i , x_j)= \emptyset (x_i)^T \emptyset(x_j)$ is called the kernel function, γ, r, and d are kernel parameters.

A linear SVM was employed for leaf feature classification since this is a multi-class dataset research. The scheme of "one versus all" was used in SVM architecture.

3.7.2 Artificial Neural Network (ANN)

A typical ANN consists of an input layer, hidden layer and output layer. The output of ANN, h_i is formulated as below:

$$h_i = \sigma \left(\sum_{j=1}^{N} v_{ij} x_j + T_i^{hid}\right) \quad (3.5)$$

where $\sigma()$ is known as the activation function, N is the input neuron number, v_{ij} is the weight, x_j is the inputs to input neurons and T_i^{hid} is the threshold (Wang, 2003).

A feed forward neural network with a single layer which composed of 10 to 100 neurons was employed in this research as illustrated in Figure 3.10. The training process was set to the default function known as the Scaled Conjugate Gradient function and the stopping criterion was set to the minimum gradient achieved as shown in Figure 3.11.

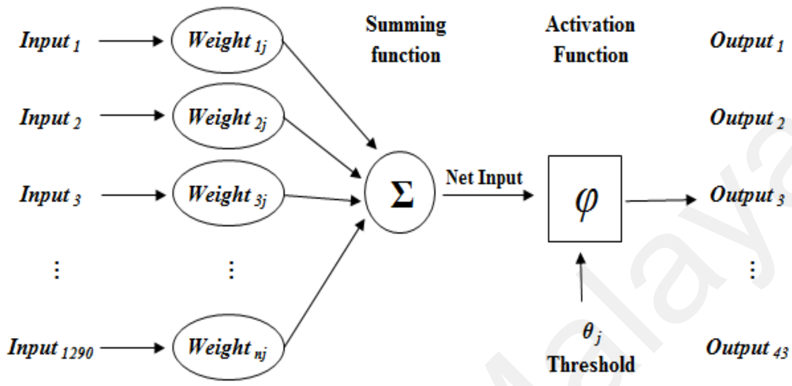


Figure 3.10: Proposed ANN architecture.

Figure 3.11: Training function and stopping criteria of the proposed ANN.

3.7.3 k-Nearest Neighbour (k-NN)

The performance of the k-NN classifier depends on the inputs, decision rule, "k" value and distance metrics (Rizwan & Anderson, 2014). There are four types of distance metrics in k-NN, namely, Euclidean, city block, cosine, correlation and Chebyshev. The formula of each distance metric is shown in Equations 3.6 to 3.10 (Rizwan & Anderson, 2014):

i. Euclidean: $d = \sqrt{\sum_{j=1}^{n}(x_{sj} - x_{tj})^2}$ (3.6)

ii. City block: $d = \sum_{j=1}^{n} | x_{sj} - x_{tj} |$ (3.7)

iii. Cosine: $d = 1 - \frac{\sum_{j=1}^{n} x_{sj}\, x_{tj}}{\sqrt{\sum_{j=1}^{n} x_{sj}\, x_{sj}}\sqrt{\sum_{j=1}^{n} x_{tj}\, x_{tj}}}$ (3.8)

iv. Correlation: $d = 1 - \frac{(x_s - \tilde{x}_s)(x_t - \tilde{x}_t)'}{\sqrt{(x_s - \tilde{x}_s)(x_s - \tilde{x}_s)'}\sqrt{(x_t - \tilde{x}_t)(x_t - \tilde{x}_t)'}}$ (3.9)

v. Chebyshev: $d = max_j \{ \left| x_{sj} - x_{tj} \right| \}$ (3.10)

where x represents the vector of feature in m dimensional space, s represents the query point, whereas t represents the point from the instance space, $\tilde{x}_s = \frac{1}{m}\sum_{j=1}^{m} x_{sj}$ and $\tilde{x}_t = \frac{1}{m}\sum_{j=1}^{m} x_{tj}$.

The performance of the k-NN classifier depends on the inputs, decision rule, distance metric and the 'k' value (Rizwan & Anderson, 2014). In this research, several values of nearest neighbours have been tested with different distance metric. The models were tested with 1 to 10 nearest neighbours with city block distance metric due to the high efficiency and high accuracy.

3.7.4 Naïve Bayes (NB)

NB, is a classification method which classifies and identifies the samples based on Bayes theorem. The basic idea of Naive Bayes theorem is formulated as shown in Equation 3.11:

$$P(C_i|X) = \frac{p(X|C_i)\,P(C_i)}{P(X)} \tag{3.11}$$

where X represents a data sample which has no class label, C indicates a specified class, $P(C_i|X)$ is the highest conditional probability, and i=1, 2, …, k (Catal et al., 2011).

A multiclass Naive Bayes model was employed in this study to train and classify leaf features based on the probability. NB was set at the default parameters in the MATLAB which the prior class probability distribution is considered as the relative frequency distribution of the species.

3.7.5 Convolutional Neural Network (CNN)

In this research, the CNN was also used as a classifier. Two CNN models, namely, a model that used the same architecture and parameters of fine-tuned AlexNet (Figure 3.8 and Table 3.4) and a model that used the same architecture and parameters of the D-Leaf model (Please refer to the Figure 3.9 and Table 3.5) were used as the CNN classifier to classify the extracted plant features.

3.7.6 Performance Evaluation

The dataset was partitioned into training set and testing set with a ratio of 80:20. The performance of the proposed models was evaluated based on the accuracy metrics as stated in Equations 3.12 and 3.13.

$$\text{Training Accuracy} = \frac{\text{Number of training samples that has been correctly classified}}{\text{Total number of training samples}} \text{X } 100\% \tag{3.12}$$

$$\text{Testing Accuracy} = \frac{\text{Number of testing samples that has been correctly classified}}{\text{Total number of testing samples}} \text{X } 100\% \tag{3.13}$$

3.7.7 Parameter Setting

For the performance evaluation, the data was partitioned into training and testing set with a ratio of 80:20. 24 leaf images were randomly selected for the training set and 6 leaf images for the testing set out of 30 leaf images of each plant species. In total, 1032 of leaf images were included in the training set and 258 leaf images in the testing set.

In order to obtain the best identification model, classifiers such as the Artificial Neural Network and k-Nearest Neighbours were tested with several parameter sets. The parameter setting of the model with the best performance was then employed throughout the whole study without changes. Whereas, the parameter of SVM was set at the scheme of “one versus all” and NB classifier was using the default setting in the MATLAB. While, the CNN used the parameters as stated in Table 3.5.

3.8 Cross-validation (CV)

Cross validation refers to a common method which is used for algorithms selection. Typically, CV is employed for splitting data, either once or several times, to estimate the risk of each algorithm (Arlot & Celisse, 2010). It helps to avoid overfitting problem by keeping the training data independent from the validation data (Arlot & Celisse, 2010). The most popular CV approaches which are commonly used are k-fold CV and leave-one-out. The main concept of k-fold CV procedure is that the dataset is partitioning the dataset into k subsets randomly. Then, one of the k subsets is used as the testing set and the remaining subsets are used as a training set each time (Chen et al., 2013). While, leave-one-out CV is that one sample is leave out for testing in every iteration and others are used for training (Refaeilzadeh et al., 2009).

In the proposed dataset, each plant species consisted of 30 samples, thus cross-validation approach was applied on the dataset partition to reduce the overfitting. 5-fold and 10-fold CV approaches were conducted to evaluate the D-Leaf approach with an ANN classifier.

3.9 Summary

In this chapter, the proposed methodology in current study was discussed. Four stages (sampling, image pre-processing, feature extraction and classification) were proposed for this study. First, the leaf samples were collected and then the leaf images were acquired. Second, three types of CNN-based methods were used as feature extraction methods, namely, AlexNet pre-trained, AlexNet fine-tuned and D-Leaf (proposed) CNN-based methods. The extracted features were tested and evaluated using five classification methods, namely, the SVM, ANN, k-NN, NB and CNN. The proposed method was benchmarked and compared with conventional feature extraction method, known as the vein morphometric measurement.

CHAPTER 4: RESULTS

4.1 Introduction

This chapter presents the results and performance evaluation in this research. The performances of three different CNN models (AlexNet, fine-tuned AlexNet and proposed D-Leaf models) and conventional feature extraction method (vein morphometric measurement) are evaluated and revealed in the following sections. The proposed approach – D-Leaf is then benchmarked by using cross-validation and validated with other plant datasets. A graphical user interface is developed and illustrated in last section of this chapter.

4.2 Parameter Setting

The ANN classifier was tested with CNN features at different number of neurons which ranged from 10 to 100 as listed in Table 4.1.

Table 4.1: Performance of ANN classifier with different number of neurons.

No. Of Neuron	Accuracy	
	Training	Testing
10	90.68	65.74
20	98.71	63.18
30	99.56	63.68
40	99.70	64.34
50	99.69	65.74
60	99.75	64.22
70	99.84	65.54
80	**99.89**	**66.28**

Table 4.1, continued.

No. Of Neuron	Accuracy	
	Training	Testing
90	99.89	65.97
100	99.85	65.85

It was found that the ANN classifier with 80 neurons performed the best. As the number of neurons increased, both the training and testing accuracy were increased as well. However, the accuracy dropped after 80 neurons. Hence, this parameter was fixed and remained the same in the ANN classifier throughout the whole study.

In the case of the k-NN, several numbers of nearest neighbours were tested. The k-NN classifier was tested with 1 to 10 nearest neighbours. The k-NN classifier with 1 nearest neighbour achieved the best performance as shown in Table 4.2. Therefore, the k value in k-NN classifier was fixed to one nearest neighbour throughout the whole research.

Table 4.2: Performance of k-NN with different number of nearest neighbour.

Number of Nearest Neighbour	Accuracy	
	Training	Testing
1	**100.00**	**82.17**
2	98.26	81.94
3	96.15	79.42
4	97.77	81.78
5	95.78	79.25
6	95.63	79.19

Table 4.2, continued.

Number of Nearest Neighbour	Accuracy	
	Training	Testing
7	95.52	79.35
8	94.98	79.92
9	94.90	79.84
10	94.17	78.80

4.3 Pre-trained CNN Model

Two CNN models were implemented to extract the leaf features from the leaf images which were the AlexNet and the fine-tuned AlexNet models. The extracted features were tested and validated by using different classifiers in order to obtain the most desirable and suitable feature extraction method for plant species identification.

4.3.1 Pre-trained AlexNet Model

A pre-trained model, known as the AlexNet, was employed to extract the leaf features. A total of 4096 leaf features were extracted from FC7 layer of pre-trained AlexNet as shown in Section 3.6.1.1. The extracted features were then fed into four different classifiers, namely SVM, ANN, k-NN and NB.

4.3.1.1 Performance of the AlexNet Model

The AlexNet was executed for 10 times in order to obtain the average accuracy. Table 4.3 shows the performance of different classifiers using AlexNet extracted features.

Table 4.3: Performance of the pre-trained AlexNet model.

Model	Accuracy	
	Training*	**Testing***
AlexNet-SVM	84.20	79.40
AlexNet-ANN	**100.00**	**93.26**
AlexNet-kNN	100.00	85.60
AlexNet-NB	90.10	83.33

* Average accuracy of 10 runs

(a) AlexNet-SVM

AlexNet-SVM, the SVM classifier with extracted features form the AlexNet model, was the only model which achieved less than 90% of training accuracy and less than 80% of testing accuracy. It had performed 84.20% and 79.40% of training accuracy and testing accuracy, respectively. AlexNet-SVM is the worst model as compared to the other models with much lower accuracy as revealed in Table 4.3.

(b) AlexNet-ANN

As shown in Table 4.3, the AlexNet-ANN achieved a 100% training accuracy and 93.26% of testing accuracy. This model outperformed the other model in this experiment, making it the best model for this part of the research.

(c) AlexNet-kNN

The k-NN classifier achieved a 100% in the training accuracy and 85.60% in testing accuracy.

(d) AlexNet-NB

The AlexNet-NB model obtained only 90.10% of training accuracy and a testing accuracy of 83.33%, which was lower than 90%. As compared to the AlexNet-kNN classifier, the AlexNet-NB's performance is slightly poorer performance in both training and testing.

(e) Comprehensive Comparison

As shown in Table 4.3, only the AlexNet-ANN and AlexNet-k-NN classifiers achieved 100% in the training accuracy, while, AlexNet-SVM and AlexNet-NB performed 84.20% and 90.10% in the training accuracy, respectively. Furthermore, the other classifiers only obtained a testing accuracy which is lower than 90% except for the AlexNet-ANN model which obtained a testing result of 93.26%.

In general, the best model is the AlexNet-ANN classifier with the extracted features of the AlexNet model (100% of training accuracy and 93.26% of testing accuracy), while, the worst model is the AlexNet-SVM model (84.20% of training accuracy and 79.40% of testing accuracy). This can be concluded that ANN is more compatible with the extracted features of AlexNet as compared to the other classifiers due to the similar development concept in both ANN and CNN models.

4.3.1.2 Execution Time of AlexNet Model

Table 4.4 shows the execution time of the pre-trained AlexNet models with different classifiers.

Table 4.4: Execution time of pre-trained AlexNet models.

Model	Time * (Minutes)
AlexNet-SVM	33.86
AlexNet-ANN	**37.26**
AlexNet-kNN	36.87
AlexNet-NB	35.73

* Average execution time for 10 runs

The execution time of the AlexNet models which included feature extraction and classification was from 33 to 37 minutes. Although the AlexNet-ANN model was taken a slightly longer time than the other models, it achieved the best performance. However, SVM performed with lower accuracy, but in shorter time compare to other classifiers.

4.3.2 Fine-tuned AlexNet Models

The AlexNet was also fine-tuned to get fitted with the proposed dataset and implemented for leaf feature extraction. Fine-tuned AlexNet (Ft-AlexNet) was capable to extract 4096 leaf features. The extracted features were then classified with five different classifiers, namely, SVM, ANN, k-NN, NB and CNN.

4.3.2.1 Performance of Fine-tuned AlexNet Model

Each of these models were executed for 10 runs in order to get an average accuracy due to the slight difference in the extracted features per run. Table 4.5 shows the performance of different classifiers using fine-tuned AlexNet extracted features.

Table 4.5: Performance of fine-tuned pre-trained AlexNet model.

Model	Accuracy	
	Training*	Testing*
Ft-AlexNet-SVM	100.00	87.79
Ft-AlexNet-ANN	**100.00**	**95.54**
Ft-AlexNet-kNN	97.41	87.33
Ft-AlexNet-NB	99.22	87.33
Ft-AlexNet-CNN	99.88	88.30

* Average accuracy for 10 runs

(a) Ft-AlexNet-SVM

Ft-AlexNet-SVM model achieved a training accuracy of 100%. However, the testing accuracy of the Ft-AlexNet-SVM was 87.79%.

(b) Ft-AlexNet-ANN

The fine-tuned AlexNet extracted features, which were classified with ANN, achieved 100% of training accuracy. Furthermore, this model obtained a testing accuracy of 95.54%. The Ft-AlexNet-ANN achieved the best performance as compared to the other models.

(c) Ft-AlexNet-k-NN

The Ft-AlexNet-kNN achieved 97.41% of training accuracy which was lower than the other models. The testing result of Ft-AlexNet-kNN was 87.33% which was the lowest in this experiment.

(d) Ft-AlexNet-NB

The Ft-AlexNet-NB achieved a better training performance than Ft-AlexNet-kNN with accuracy of 99.22%. It obtained a comparable testing performance with the Ft-AlexNet-kNN model, with 87.33% of accuracy. However, the testing accuracy of both Ft-AlexNet-NB and the Ft-AlexNet-kNN models were the lowest.

(e) Ft-AlexNet-CNN

The training performance of the Ft-AlexNet-CNN achieved 99.88% of accuracy. However, the testing accuracy achieved 88.30%. The Ft-AlexNet-CNN was the second highest model among the classifiers.

(f) Comprehensive Comparison

The classifiers of fine-tuned pre-trained AlexNet models achieved a training accuracy ranging from 97% to 100%. Only the SVM and ANN classifiers achieved a 100% training accuracy among the fine-tuned AlexNet models. Besides that, ANN was the only classifier which performed at a testing accuracy of more than 90%.

Generally, ANN performed the best with 100% of training accuracy and 95.54% of testing accuracy. Whilst, the lowest accuracy was achieved by the k-NN classifier with 97.41% of training accuracy and 87.33% of testing accuracy. It is notable that the extracted features of fine-tuned AlexNet are more compatible with ANN rather than the other classifiers. This is because the fundamental concept of both ANN and CNN are similar to each other.

4.3.2.2 Execution Time of Fine-Tuned AlexNet Model

The time durations for the execution of the fine-tuned AlexNet models are shown in Table 4.6.

Table 4.6: Execution time of fine-tuned pre-trained AlexNet models.

Model	Time * (Minutes)
Ft-AlexNet-SVM	37.55
Ft-AlexNet-ANN	36.75
Ft-AlexNet-kNN	37.55
Ft-AlexNet-NB	38.55
Ft-AlexNet-CNN	37.59

* Average execution time for 10 runs

The time as illustrated in Table 4.6 was the average execution time of the fine-tuned AlexNet model for 10 runs. The time duration included the feature extraction and classification. However, extra execution time was needed for re-training the fine-tuned model. The computational time for each of the models were ranging between 36 to 38 minutes.

4.4 D-Leaf Model

D-Leaf is the proposed CNN-based model used to extract the leaf features for plant species identification. Since D-Leaf is the proposed method, several experiments were conducted in order to obtain the most optimum model for the leaf feature extraction. In this section, performance of different size of input images, fully connected layers, leaf patches and classifiers on each species were carried out to define the proficiency of D-Leaf.

4.4.1 Input Size Evaluation

Proposed D-Leaf model was tested by several images with different sizes which were 250 x 250, 500 x 500, 750 x 750 and 6016 x 6016 (original size). These were then evaluated with the CNN in order to select the optimum parameter settings for further analysis. However, the input size of 750 x 750 and 6016 x 6016 failed to execute due to the huge computer memory requirement which could not be supported by the used computer. Thus, only the input size of 250 x 250 and 500 x 500 were employed in this research.

4.4.1.1 Input Size of 250 x 250

The results obtained from the proposed D-Leaf models, using images with size of 250 x 250 are shown in Table 4.7.The training accuracy of images with the size of 250 x 250 ranged between 82% and 96% and the testing accuracy ranged about 69% to 79%. The model with maximum 60 epochs achieved a slightly better performance at 96.80% of training accuracy and 79.03% of testing accuracy as compared to the other models. As the number of maximum epoch increased, the training and testing performance increased. However, for the input size of 250 x250, the accuracy started to drop at 80 maximum epochs from 96.80% of training accuracy and 79.03% of testing accuracy to 95.50% of training accuracy and 78.72% of testing accuracy. The number of maximum epochs rose proportionally with the execution time. These models took about one minute to six minutes to complete the execution.

Table 4.7: Performance of D-Leaf models with 250 x 250 of input size.

Input Size	Maximum Epochs	Accuracy		Time Taken (Minutes)
		Training*	Testing*	
250 x 250	20	82.20	69.34	1.54
250 x 250	40	95.00	77.02	2.94
250 x 250	**60**	**96.80**	**79.03**	**4.31**
250 x 250	80	95.50	78.72	5.55

* Average accuracy for 10 runs

4.4.1.2 Input Size of 500 x 500

Table 4.8 shows the results of the D-Leaf models with an input size of 500 x 500.

Table 4.8: Performance of D-Leaf models with 500 x 500 of input size.

Input Size	Maximum Epochs	Accuracy		Time (Minutes)
		Training*	Testing*	
500 x 500	20	62.60	53.29	7.70
500 x 500	40	70.60	59.15	15.08
500 x 500	60	70.20	60.89	22.30
500 x 500	**80**	**73.10**	**62.83**	**28.66**

* Average accuracy for 10 runs

In general, with an input size of 500 x 500, the training accuracy and testing accuracy of the models ranged from 62% to 73% and from 53% to 63%, respectively. The increase in the number of maximum epochs caused the accuracy and execution time to rise concurrently. The time durations for executing these models were between 7 minutes to 29 minutes.

4.4.1.3 Comprehensive Comparison

The performances of the CNN models with the 250 x 250 input size were better than the images with the size of 500 x 500. Furthermore, the execution time for the models with input size of 250 x 250 was shorter than that of the models with input size of 500 x 500.

In short, the best performance was achieved with the input size of 250 x 250 and at a maximum epoch of 60. Thus, the parameters of this model were subsequently used throughout the research.

4.4.2 Fully Connected Layer 4 (FC4)

According to the proposed D-Leaf architecture which is demonstrated in Figure 3.9, a total of 1290 leaf features were extracted from the FC4. The extracted features from FC4 were then further experimented by using five different classifiers, namely, SVM, ANN, k-NN, NB and CNN. The average performance of each FC4 models are shown in Table 4.9.

Table 4.9: Performance of FC4 of D-Leaf models.

Classifiers	**Accuracy**		**Time (Minutes)**
	Training*	**Testing***	
SVM	100.00	83.53	5.87
ANN	**100.00**	**91.36**	**5.27**
k-NN	87.52	73.45	5.62
NB	93.69	81.43	5.07
CNN	96.20	78.22	5.24

* Average accuracy for 10 runs

As depicted in Table 4.9, the SVM classifier achieved a 100% of training accuracy and 83.53% of testing accuracy. k-NN, NB and CNN obtained a training accuracy less than 100%, at87.52%, 93.69% and 96.20%, respectively. Whilst, k-NN obtained 73.45% of testing accuracy which was the lowest for this part of the research followed by CNN, which obtained only 78.22% testing accuracy. NB and SVM performed at more than 80% of testing accuracy, at 81.43% and 83.53%, respectively. However, the ANN performed the best with 100% of training accuracy and 91.36% of testing accuracy compared to the other classifiers.

4.4.3 Fully Connected Layer 5 (FC5)

FC5 was one of the fully connected layers of the proposed model D-Leaf as shown in Figure 3.9. 1290 of leaf features were extracted from FC5 and fed into SVM, ANN, k-NN, NB and CNN for classifying. Table 4.10 shows the results of average accuracy of different classifiers, fed by FC5 extracted features which has been obtained from 10 iterations.

Table 4.10: Performance of FC5 of D-Leaf models.

Classifiers	Accuracy		Time (Minutes)
	Training*	Testing*	
SVM	100.00	82.75	4.40
ANN	**100.00**	**94.88**	**4.69**
k-NN	100.00	82.44	4.72
NB	98.40	81.86	4.44
CNN	96.80	79.03	4.31

* Average accuracy for 10 runs

The execution time for every classifier with the features of FC5 was approximately equivalent, which was about 5 minutes. SVM, ANN and k-NN had obtained 100% of training accuracy, whereas, NB and CNN obtained 98.40% and 96.80%only. Furthermore, only the ANN successfully classified testing data according to the features of FC5 with more than 90% of testing accuracy, at 94.88%. Nevertheless, the other classifiers such as SVM, k-NN and NB obtained a testing accuracy which ranged between 81% and 83% with the exception of the CNN which achieved only a testing accuracy of 79.03%. The testing accuracy of SVM, k-NN and NB was 82.75%, 87.44% and 81.86%, respectively.

It can be concluded that the ANN had achieved the best performance among the classifiers with 100% of training accuracy and 94.80% of testing accuracy. Yet, the CNN obtained a training accuracy of 96.80% and testing accuracy of 79.03% which was the model with the lowest accuracy for this part of the study.

4.4.4 Fully Connected Layer 6 (FC6)

Only 43 leaf features were extracted from the FC6 and the performance of FC6 was evaluated by feeding the extracted features into five classifiers, which were SVM, ANN, k-NN, NB and CNN. The performance of FC6 was as showed in the Table 4.11

Table 4.11: Performance of FC6 of D-Leaf models.

Classifiers	Accuracy		Time (Minutes)
	Training*	Testing*	
SVM	100.00	80.93	3.98
ANN	**100.00**	**93.84**	**4.25**
k-NN	100.00	81.12	4.22

Table 4.11, continued.

Classifiers	Accuracy		Time (Minutes)
	Training*	**Testing***	
NB	98.38	81.12	4.03
CNN	96.10	79.42	4.43

* Average accuracy for 10 runs

As tabulated in Table 4.11, all the classifiers had achieved 100% of training accuracy except for NB and CNN which performed at only 98.38% and 96.10% respectively. All classifiers achieved more than 80% of testing accuracy except for CNN which obtained only 79.42% accuracy. SVM achieved 80.93% of testing accuracy while k-NN and NB obtained 81.12%. ANN was the only classifier which obtained more than 90% of testing accuracy (93.84%). The time for the FC6 models to complete each execution was in average four minutes. Generally, the ANN classifier outperformed the other classifiers while CNN had the lowest performance.

4.4.5 Leaf Patches

Another experiment was conducted by applying D-Leaf on the centre leaf patches and side leaf patches as shown in Figure 3.7. The centre patches and side patches were cropped automatically at the centre and side of the leaf sample, respectively. The aim of this part was to examine the performance of the leaf patches in plant species identification. Hence, D-Leaf was employed to extract the leaf features and then classified by SVM, ANN, k-NN, NB and CNN separately.

4.4.5.1 Centre Leaf Patches

The performance of the centre leaf patches is shown in Table 4.12.

Table 4.12: Performance of D-Leaf model with centre leaf patches.

Classifiers	Accuracy	
	Training*	**Testing***
SVM	100.00	74.92
ANN	**100.00**	**91.63**
k-NN	86.43	63.88
NB	94.59	71.24
CNN	94.40	65.39

* Average accuracy for 10 runs

The training results of SVM and ANN classifiers with the extracted features from the centre patches were 100%. However, k-NN, NB and CNN can only obtain 86.43%, 94.59% and 94.40% in training performance, respectively. Among the centre patch models, the only classifier that achieved more 90% of testing accuracy was the ANN at 91.63%. On the other hands, the other classifiers had obtained a lower range of testing accuracy which ranged from 63% to 75%. SVM and NB obtained about 75% and 71% of testing accuracy, respectively, whereas k-NN and CNN obtained 63.88% and 65.39%. The ANN classifier was the best model with 100% of training accuracy and 91.63% of testing accuracy. Yet, the k-NN only obtained about 86% and 54% of training accuracy and testing accuracy, respectively, was the model with the lowest performance.

4.4.5.2 Side Leaf Patches

Table 4.13 depicts the results of the side leaf patches by using the D-Leaf model.

Table 4.13: Performance of D-Leaf models with side leaf patches.

Classifiers	Accuracy	
	Training*	**Testing***
SVM	100.00	58.68
ANN	**100.00**	**86.40**
k-NN	78.90	47.05
NB	82.45	52.29
CNN	92.90	51.86

* Average accuracy for 10 runs

Using side leaf patches, k-NN, NB and CNN achieved a training accuracy which ranged from 78% to 93% while SVM and ANN achieved 100% testing accuracy. However, only the ANN classifier achieved a testing accuracy of more than 86%, at 86.40%. The other classifiers performed poorer, ranged from between 47% to 59% of testing accuracy. Among these classifiers, k-NN achieved the lowest testing accuracy of 47.05%

.

4.4.6 Performance of Classifiers in Each Species

As discussed in Section 3.7 of Methodology Chapter, the experiments continued by selecting six samples of each species as a testing set. These plant samples (whole leaf images) were subsequently extracted using the D-Leaf approach and different classifiers and the results shown in Table 4.14.

Table 4.14: Number of testing samples which were classified correctly.

Species	SVM	ANN	NB	k-NN	CNN
1	6	6	4	4	5
2	5	6	5	5	0
3	6	6	6	6	6
4	5	6	6	6	6
5	4	6	5	5	4
6	5	5	4	4	4
7	4	6	3	3	4
8	4	5	4	4	4
9	6	6	6	6	6
10	6	6	6	6	5
11	4	6	5	5	6
12	3	5	4	3	3
13	4	6	4	5	3
14	5	5	4	4	4
15	4	5	5	6	5
16	5	6	6	6	6
17	6	6	5	5	6
18	4	5	5	4	4
19	5	6	6	5	6
20	5	5	4	5	5
21	6	6	5	5	5
22	4	6	4	5	5
23	6	6	6	6	6
24	6	6	6	6	5
25	5	5	5	5	5
26	5	6	6	6	6
27	3	5	4	3	4
28	5	6	6	6	6
29	4	5	5	5	3

Table 4.4, continued.

Species	SVM	ANN	NB	k-NN	CNN
30	5	5	3	3	2
31	6	6	6	6	6
32	4	6	6	5	5
33	4	5	2	2	3
34	5	6	6	6	6
35	4	6	5	6	4
36	5	6	4	4	4
37	4	6	5	5	0
38	6	6	6	6	6
39	4	6	5	5	4
40	5	6	5	6	5
41	5	6	6	6	6
42	4	6	5	5	5
43	5	6	5	6	6
Total Number of Correctly Classified Sample	206	246	213	215	199

From here on, we use the term "misclassified" to denote when only 1-3 of the samples were correctly classified and "cannot classify" when none of the samples were correctly classified.

Generally, the ANN classifier was capable of correctly classifying up to 5 and 6 testing samples for all species. In the case of the SVM classifier, the performance was moderate. It could only classify half of the samples correctly in Species 12 (*Cinnamomum iners*) and Species 27 (*Melaleuca cajuputi*).

The NB classifier misclassified in Species 7 (*Erythrina variegata*), Species 30 (*Plumeria rubra*) and Species 33 (*Saraca thaipingensis*). While, the k-NN misclassified Species 7 (*Erythrina variegata*), Species 12 (*Cinnamomum iners*), Species 27 (*Melaleuca cajuputi*), Species 30 (Plumeria rubra) and Species 33 (*Saraca thaipingensis*).

The CNN classifier performed the worst (199 samples out of 258 samples) in classifying the testing samples correctly. The CNN cannot classify the testing samples of Species 2 and Species 37. In addition, there were low number of sample in certain plant species which were classified wrongly by CNN classifier as well which are Species 12 (*Cinnamomum iners*), Species 29 (*Mimusops elengi*), Species 30 (*Plumeria rubra*) and Species 33 (*Saraca thaipingensis*).

4.5 Benchmarking

The D-Leaf model was benchmarked with the use of conventional feature extraction method namely vein morphometric measurements and cross-validation approach.

4.5.1 Vein Morphometric Measurements

The vein morphometric measurements were employed as a benchmark, in which 62 vein morphological features based on Sobel-based segmented vein architecture (as listed in Table 3.6) were extracted and fed in ANN classifier with different number of neuron which ranging from 20 to 100 neurons for performance evaluation. Table 4.15 shows the performance of the vein features.

Table 4.15: Performance of morphometric measurements.

No. of Neuron	Accuracy	
	Training	Testing
20	98.71	63.81
40	99.7	64.34
60	99.75	64.22
80	**99.89**	**66.28**
100	99.85	65.97
120	99.93	66.09
140	99.93	66.02

* Average accuracy for 10 runs

The morphometric measurement models achieved a training accuracy ranging between 98.71% and 99.89%. As the number of ANN's neuron increased, the training accuracy rose as well, except for the model with 100 neurons. The performance of the models started to drop from 99.89% to 99.85% when the number of neurons were 100 neurons. However, the testing accuracy increased as the number of neurons increased, for the models with the number of neurons less than 100. The performance of the model with 100 neurons decreased (65.97%) and the models of more than 100 neurons increased slightly. The best performance was achieved by the models with 80 with a testing accuracy of 66.28%.

4.5.2 Cross-Validation (CV)

The potential of D-Leaf approach with an ANN classifier was further validated with 5-fold and 10-fold CV approaches. The results of the CV are shown in the Table 4.16.

Table 4.16: Validation results with cross validation.

Data Partition	Accuracy	
	Training*	**Testing***
5-CV	100.00	93.15
10-CV	100.00	93.31

* Average accuracy for 10 runs

Both the CV models achieved up to 100% of training accuracy and performed comparatively for the testing accuracy. The 5-fold CV model performed at 93.15% of testing accuracy while the 10-fold CV model obtained a testing accuracy of 93.31%. The testing results of both the 5-fold CV and 10-fold CV models were comparable to the result with the D-Leaf model without the employment of CV (94.88%).

4.6 Validation

Lastly, the proposed CNN method – D-Leaf was further evaluated and validated against other datasets along with the ANN classifier. D-Leaf method was evaluated by using three publicly available datasets which were MalayaKew, Flavia and Swedish Leaf Dataset in order to test the performance of the proposed D-Leaf. These three datasets are selected for validation as their sample size are similar with D-Leaf dataset.

The results of the D-Leaf method with the ANN classifier using the MalayaKew, Flavia and Swedish dataset is as shown in Table 4.17.

Table 4.17: Validation results with MalayaKew, Flavia and Swedish dataset.

Dataset	Accuracy	
	Training*	Testing*
MalayaKew	100.00	90.38
Flavia	100.00	94.63
Swedish	100	98.09

* Average accuracy for 10 runs

The model using the extracted leaf features from the MalayaKew dataset achieved 100% and 90.38% in training accuracy and testing accuracy, respectively. D-Leaf method successfully identified the plant species of the Flavia dataset with a performance 100% training and 94.63% testing. As seen from Table 4.17, the D-Leaf method achieved 100% of training accuracy and 98.09% of testing accuracy.

4.7 D-Leaf Plant Species Identification System

A prototype of the graphic user interface (GUI) based on the D-Leaf CNN approach for feature extraction and the ANN classifier for classification was developed. The user requires to upload the leaf image into the graphical user interface (GUI) of the prototype. The system would read the uploaded leaf image and extract the leaf features by using the D-Leaf approach. The extracted features are then classified and identified by using the ANN classifier. The output shows top five most related species to the user. The user is able to select the species from the list in order to view the whole tree, leaf, bark, fruit and flower images of the selected species. Figure 4.1 shows the GUI of the D-Leaf identification system before the leaf image is loaded. While, Figure 4.2 shows the results and output of the D-Leaf identification after the leaf image was loaded.

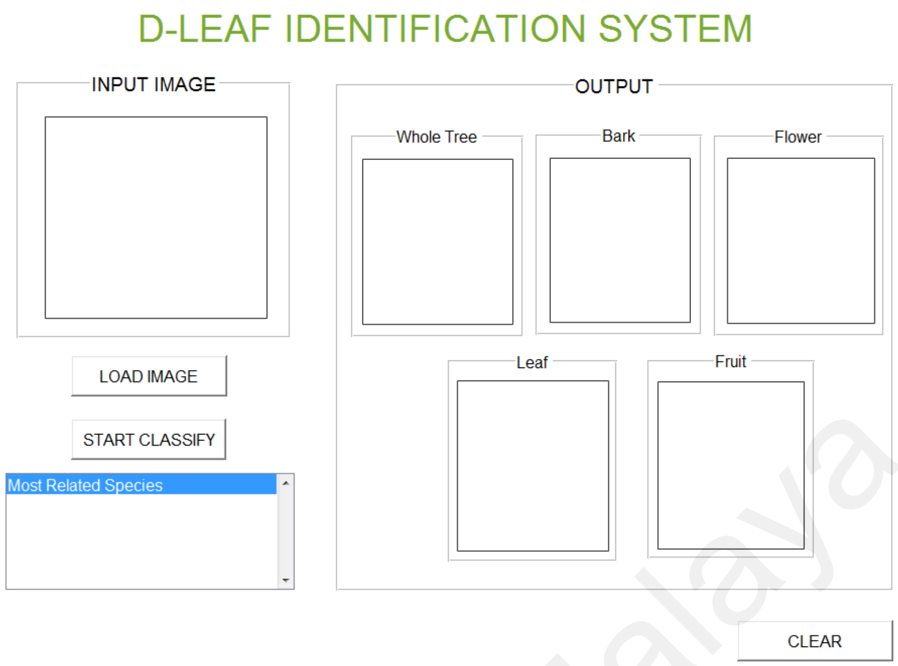


Figure 4.1: Graphic user interface of D-Leaf pant identification system before the leaf image is loaded.

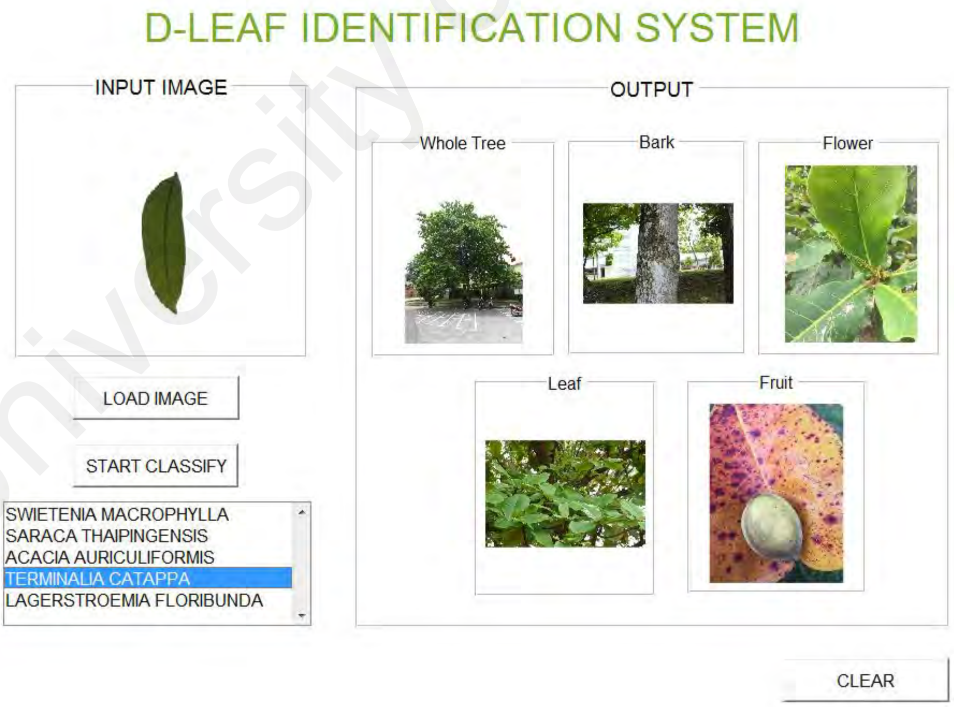


Figure 4.2: Graphic user interface of D-Leaf plant identification system after the leaf image is loaded.

The proposed D-Leaf Plant Identification system was further tested with four independent samples, which include two samples with white background and 2 samples with complex background. Figure 4.3 shows the top 5 results based on the images from independent samples.

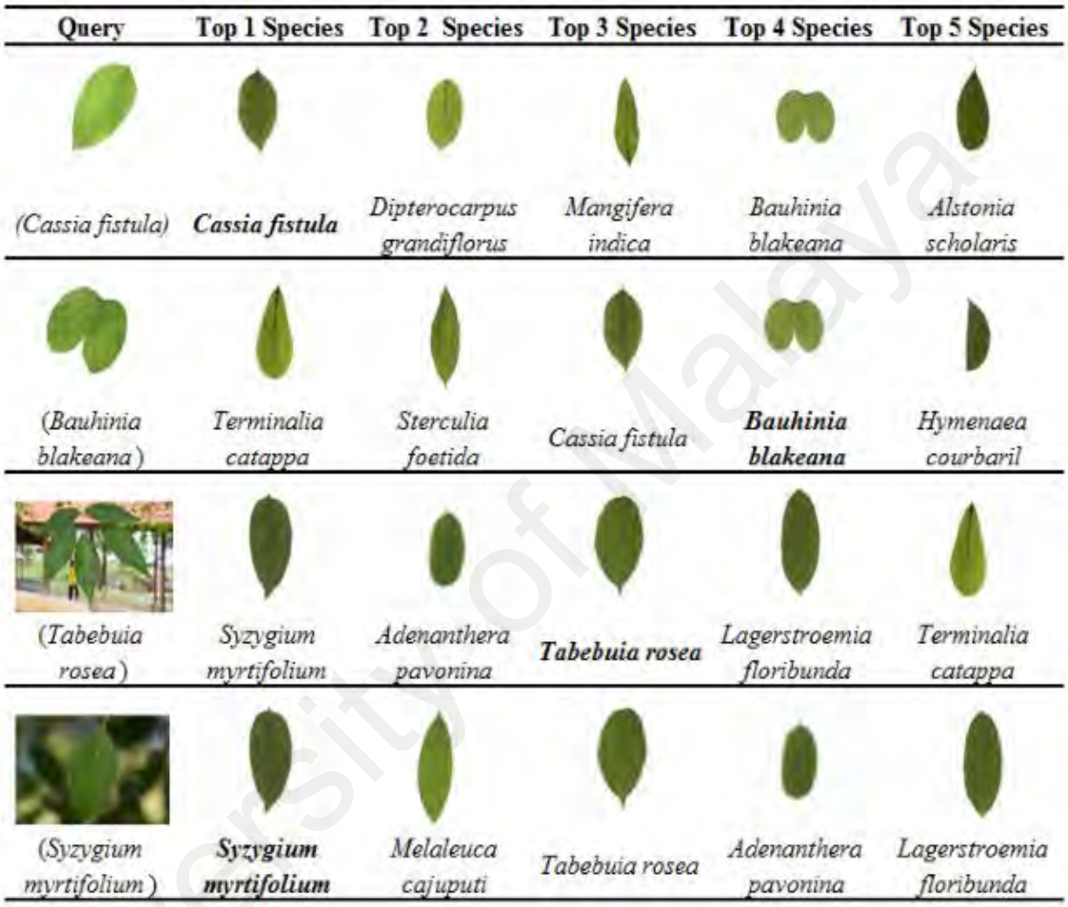


Figure 4.3: The top 5 results that returned by the D-Leaf GUI system.

The first and second queries with white background were successfully identified as Top 1 and Top 4 species, respectively. The D-Leaf Identification system also correctly identified the third queries as the Top 3 species and the fourth queries as Top 1 species even though they had a complex background.

4.8 Summary

The proposed D-Leaf model performed a comparable performance as the AlexNet and fine-tuned AlexNet models with a shorter execution time. Additionally, the D-Leaf model outperformed the conventional feature extraction method – vein morphometric measurement. D-Leaf model with 5-CV and 10-CV achieved a comparative accuracy as the D-Leaf model without cross validation. Next, the D-Leaf model was also further tested and validated using other leaf datasets, achieving more than 90% accuracy. Finally, a prototype of the D-Leaf Plant Species Identification System consisting of a graphical user interface was developed based on the D-Leaf model in order to ease and assist the experts and laymen to identify plant species.

CHAPTER 5: DISCUSSION

5.1 Introduction

This chapter summarises and discusses the performance of the proposed model to identify tropical plant species.

The components that were discussed and compared in this chapter are:

(i) The performance of three different CNN-based models, which are the AlexNet model, fine-tuned AlexNet model and the proposed CNN-based model, namely, the D-Leaf.

(ii) The selection of optimum parameters used in the D-Leaf model

(iii) The extracted features from three different fully connected layers.

(iv) The illustration of feature maps in the D-Leaf model.

(v) The optimum classifier for the proposed model.

(vi) The performance of the whole leaf, centre leaf patches and side leaf patches.

(vii) The comparison between feature extraction methods which are CNN-based methods and conventional methods, namely, vein morphometric measurement.

(viii) The application of cross-validation methods.

(ix) The validation of the proposed D-Leaf model against other leaf datasets.

(x) The development of the tropical plant species identification system.

5.2 CNN Models

A pre-trained CNN-based model, known as AlexNet, was investigated in the first part of this research, for extracting the leaf features to be fed into the classifiers. AlexNet was chosen for this experiment rather than other CNN-based models such as VGG and CaffeNet due to the lowest computation memory consumption and execution time (Liu et al., 2017).

Furthermore, in order to make AlexNet compatible with the proposed dataset, the AlexNet model was fine-tuned in the first convolutional layer by reducing the filter size in order to avoid aliasing effects on the second layer which might cause kernels of the first layer to be "dead". In addition, the fully-connected layers were fine-tuned into 1290, 1290 and 43 neurons as mentioned in Section 3.6.1.2 of the Methodology chapter.

AlexNet took about 36 minutes for a complete execution while fine-tuned AlexNet took about 36.75 minutes. However, despite the computational time, both pre-trained and fine-tuned AlexNet models took a long time for the completion of the feature extraction and classification process as shown in Table 5.1. Thus, D-Leaf as a new and simpler CNN-based model D-Leaf was proposed in this study in order to shorten the execution time while maintaining the performance.

Table 5.1: Comparison of layer number and average execution time for CNN models.

CNN Models	Number of Layers	Average Execution Time (Minutes)	Testing Accuracy
AlexNet	8	35.93	93.26
Fine-tuned AlxeNet	8	36.75	95.54
Proposed D-Leaf	**6**	**4.70**	**94.88**

The D-Leaf model performed comparably with the pre-trained AlexNet and fine-tuned AlexNet models as stated in Table 5.1. The fine-tuned AlexNet model (95.54%) performed slightly better than D-Leaf (94.88%) and AlexNet models (93.26%). Generally, the D-Leaf model is more feasible as a feature extraction approach than the

AlexNet and the fine-tuned AlexNet approaches for the development of an automated plant species identification system due to its execution time (about 5 minutes) and its performance.

In terms of the CNN architecture, both AlexNet models consisted of 8 layers while D-Leaf consisted of 6 layers. Thus, the AlexNet model is more complicated than the D-Leaf model. A model with more CNN layers will definitely require a longer time of execution. The D-Leaf has two convolutional layers lesser than AlexNet models. As shown in Table 5.1, the classification of pre-trained and fine-tuned AlexNet features took about 35.93 minutes and 36.75 minutes, respectively. However, D-Leaf model took only about 4.70 minutes for a complete execution which was 7 times faster than AlexNet and fine-tuned AlexNet models.

There were some drawbacks and limitations in the selected CNN models. The input images had to be in a square dimensional in order to feed into the CNN model for feature extraction purposes. Besides that, the execution of CNN required highthe computational memory or the RAM, causing the workstation to slow down and unable to perform other tasks. However, since the D-Leaf model composed of fewer layers, the workstation executed much faster in the D-Leaf models than the AlexNet models.

5.3 Parameters of D-Leaf

The ideal parameter setting that produced the best D-Leaf model for feature extraction was found through trial and error. Several sizes of images including 250 x 250, 500 x 500, 750 x 750 and 6016 x 6016 were used as inputs. Theoretically, the images with the larger size which consisted of more pixels were sharper, clearer and contained detailed information than the images with smaller size. However, the input images with larger size required higher computational memory. The 750 x 750 and 6016 x 6016 input images could not be processed as they were not supported by the used workstation RAM. Hence, only the 250 x 250 and 500 x 500 input images were successfully processed and analysed.

Theoretically, the 500 x 500 images contain more detailed information than the 205 x 250 images, however, as shown in Table 4.3 of the Results section, the models with the input size of 250 x 250 outperformed the models with the input size of 500 x 500, This indicated that the resolution of an image is not a significant trait for the feature extraction stage of the CNN models.

Besides that, increasing the image resolution requires longer execution time. For the 250 x 250 images, the execution time was 4 times faster than the 500 x 500 images. In short, the CNN models were able to work well with smaller resolution images in extracting the leaf features even if the images were less sharp, less clear and hence, consisted of lesser information.

Another parameter that had been considered in the D-Leaf model was the number of maximum epochs. The D-Leaf models with higher maximum epochs performed better than those models with lower epochs. However, for the D-Leaf models with 250 x 250 of image input size, the performance began to decrease as the number of maximum epochs was higher than 60. Moreover, as the number of maximum epochs increased, the execution time was also increased. Thus, the optimum parameters of D-Leaf were achieved within 60 maximum epochs, using 250 x 250 images.

5.4 Fully Connected Layers of D-Leaf

The leaf features were extracted from the fully connected layers of the CNN models. Since there were three fully connected layers in the D-Leaf model, three sets of leaf features were extracted from each of the respective layers. The performance of each set of leaf features were tested by feeding into the classifiers for learning and classifying.

In term of performance, the three extracted feature sets of the three fully connected layers performed comparably. Concisely, the leaf features of the FC5 layer performed slightly better than the FC4 (91.36%) and the FC6 (93.84) with 94.88% of testing accuracy as shown in Table 5.2.

Table 5.2: Performance of each fully connected layer in D-Leaf model.

Fully Connected Layer	Number of Extracted Features	Testing Accuracy	Time Taken (Minutes)
FC4	1290	91.36	5.27
FC5	**1290**	**94.88**	**4.69**
FC6	43	93.84	4.25

As expected, the results from the extracted features of the FC4 layer were not as good as that of the FC5 and FC6 layers. This was because of the extracted features of the FC4 layer which were less condensed and detailed than the FC5 and FC6 layers since it was the first layer of the fully connected layer. Supposedly, the FC6 feature set should be made up of the most condensed and detailed leaf information than the FC4 and FC5 layers, however, the features that were extracted from the FC5 layer produced better results than that of the FC6 layer. This can be easily explained by the fact that the model with the higher number of features produce better results as shown in the performance of the FC5 and FC6 layers.

In short, the FC5 layer was the optimum fully connected layer for extracting the leaf features in this research with the highest accuracy and in the optimum execution time.

5.5 Visualization of Feature Maps in the D-Leaf Model

The features that were extracted from each layer of the D-Leaf models can be visualized. The visualization of the features of each layer of the D-Leaf models are illustrated in Figure 5.1.From the feature maps, it could be deduced that the D-Leaf learnt the features based on the colour pixels for identifying the plant species. The feature map of each individual D-Leaf layers showed significant difference from each other. Besides that, the feature maps from each species have its own significant traits that could be used for differentiating.

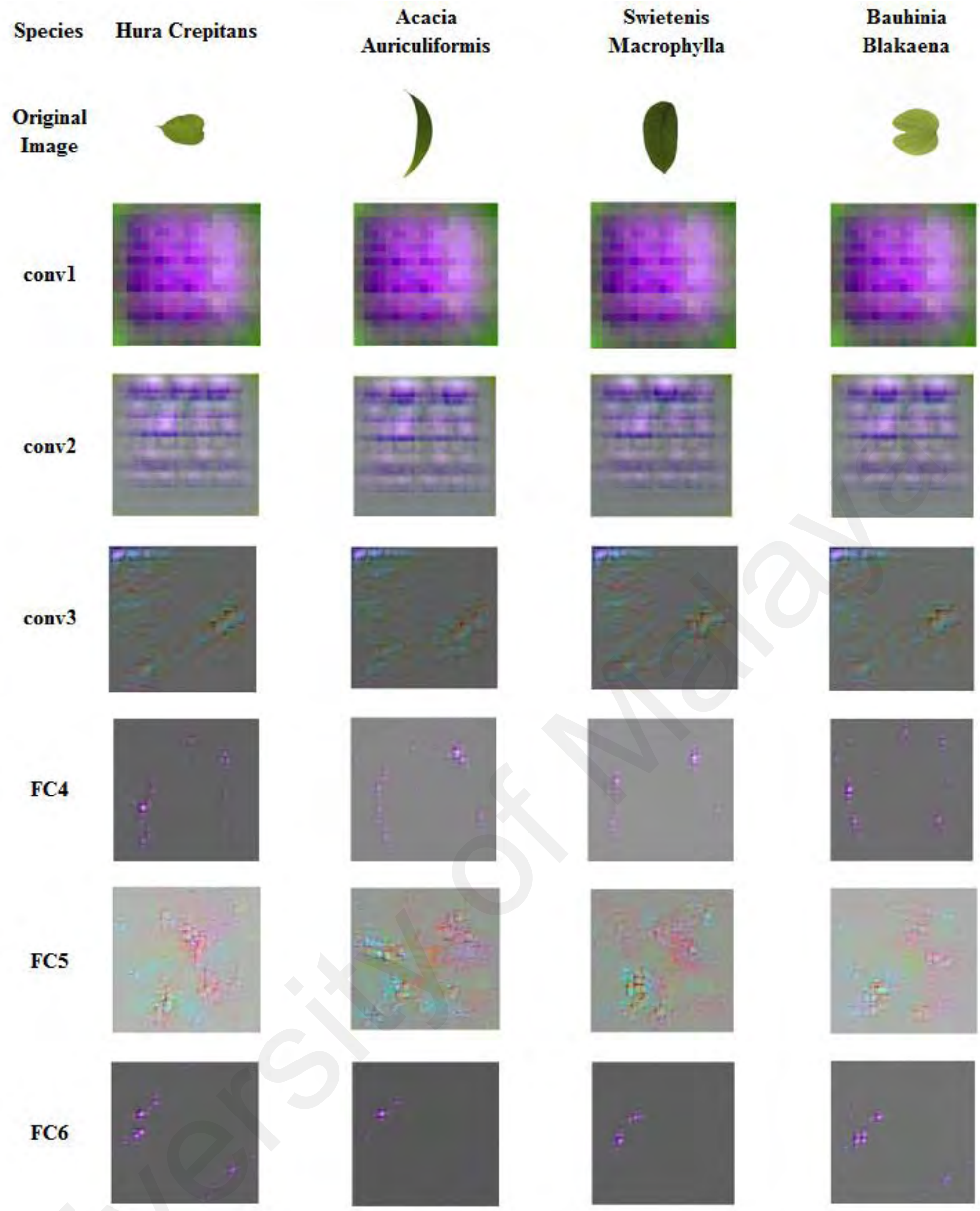


Figure 5.1: Visualization of feature maps of each D-Leaf layer.

5.6 Classifiers

In this study, five classifiers were employed to test and examine the feasibility of the extracted features from different models. The five classifiers were SVM, ANN, k-NN, NB and CNN. The performance of each classifier with different CNN models is summarised in the Table 5.3.

Table 5.3: Performance of the classifiers in different CNN models.

Classifiers	Testing Accuracy		
	AlexNet	Fine-tuned AlexNet	D-Leaf
SVM	79.40	87.79	82.75
ANN	**93.26**	**95.54**	**94.88**
k-NN	85.60	87.33	82.44
NB	83.33	87.33	81.86
CNN	---	88.30	79.03

--- Not Applicable

By comparing the performance of the different classifiers, the ANN classifier outperformed the other classifiers in all the experiment settings due to the stability of the ANN algorithm. ANN classifier had performed more than 93% of accuracy in all three different CNN models. It can be claimed that the extracted leaf features from the CNN were highly compatible with the ANN classifier. This may be because, just like the ANN, the fundamental architecture of the CNN models is also inspired from the neural network as ANN. As stated in the literature review, similar to the ANN classifier, the CNN has three basic layers, which are the input, hidden and output layers. The CNN is made up of the multilayer perceptron as well for learning and extracting the features from images.

The testing performance of the SVM, NB, k-NN and CNN were less than 89% accuracy. The SVM performed in an accuracy that ranged between 79% and 88%, while, the k-NN performed in a range between 82% and 87%. In addition, the NB and CNN achieved a result ranging between 81% to 87% and 79% to 88%, respectively. This can be explained by the fact that these classifiers were less compatible with the leaf features that were extracted from the CNN models due to the differences of the algorithms of the SVM, k-NN and NB from the CNN. SVM is based on the maximum margin hyperplane concept, and NB basically a statistical approach while k-NN is executed based on the nearest neighbour. These types of classifier may not be compatible with the neural network architecture of the CNN.

Other than that, the AlexNet cannot be applied as a classifier to classify its own extracted features. This is because the last fully connected layer of AlexNet is fixed at 1000 neurons as shown in the Section 3.6.1.1 of the Methodology Chapter whereas the number of the classes in this study is 43. If AlexNet was employed as a classifier, some fine-tuning processes were needed. The last fully connected layer of fine-tuned AlexNet was changed to 43 neurons and the fine-tuned AlexNet was subsequently employed as a classifier. As shown in Table 5.3, the CNN classifier seemed to be less efficient in classifying plant species. Hence, CNN is more suitable to be a feature extractor rather than a classifier in this case.

The Species 12 (*Cinnamomum iners*), Species 27 (*Melaleuca cajuputi*), Species 30 (*Plumeria rubra*) and Species 33 (*Saraca thaipingensis*) were always being misclassified even though their leaf shape were not really similar to each other (Figure 5.2). Yet, the misclassification of these species occurred only in SVM, NB, k-NN and CNN classifiers. Conversely, the ANN classifier has no problem in classifying those species. Again, these confirmed that the ANN classifier is the best classifier to use together with the features extracted using the D-Leaf model.

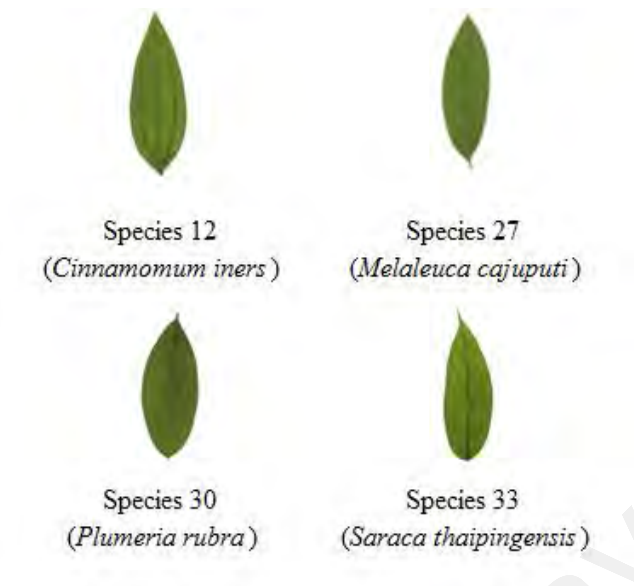


Figure 5.2: The misclassified species.

5.7 Whole Leaf and Leaf Patches

Experiments were conducted on the performance of whole leaf and leaf patches by using the ANN classifier since the ANN classifier outperformed the other classifiers.

The centre patches and side patches did not produce results comparable to the whole leaf images. The images with the whole leaf seemed to be preferable for identifying plant species than the leaf patches. The whole leaf images achieved up to 94.88% of testing accuracy, whereas, the centre patches and side patches obtained testing accuracy of 91.63% and 86.40% only, respectively.

The images of whole leaf were composed of finer and more detailed leaf information than the leaf patches. Since the leaf patches are a small part that have been cropped from the whole leaf, thus the patches composed only a small portion of leaf information. These might be caused by the loss of some important leaf features after the patches were cropped from the whole leaf image. An important morphological feature that was missed in the centre and side patches was the whole leaf shape. Our findings show that the leaf shape features are the dominant features in plant species identification system as supported by previous studies (Murat et al., 2017; Caglayan et al., 2013; Sharma et al., 2014; Hati et al., 2013; Du et al., 20117).

Nonetheless, the centre patches outperformed the side patches as shown in Table 4.12 and Table 4.13. This might be because the centre patches consisted of more important information than the side patches such as primary leaf vein.

In addition, since the centre patches and side patches were cropped automatically, some of the patches might be cropped inaccurately or incorrectly as shown in Figure 5.3. As illustrated in Figure 5.3, the centre patch supposed to include the main centre primary vein, but the image in Figure 5.3 shows that the primary vein is absent in the centre patch. Furthermore, the side patch should consist the secondary and tertiary vein only, but the image in Figure 5.3 shows that it contained the main centre primary vein. These issues could have been happened because the leaf was not placed in the centre of the image perfectly or the leaf was not aligned properly during image acquisition and as a result, some important leaf features and information were lost. In addition, the patches which were cropped manually may incur some bias during feature extraction process. For example, the leaf patches will be cropped purposely on the important features.

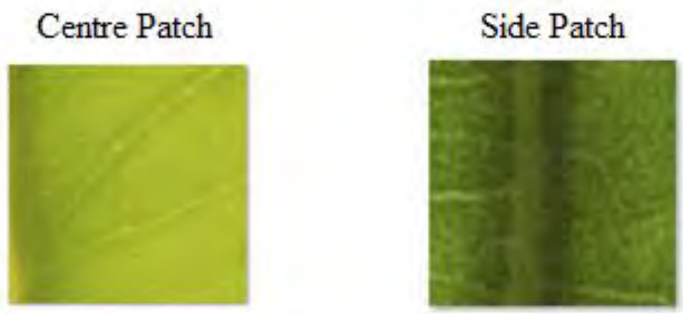


Figure 5.3: Mistakes happened during the patches cropped automatically.

As a result, the combination of all leaf features from the whole leaf images could be the best representative of a specific plant species for feature extraction.

5.8 Vein Morphometric Measurements versus the CNN Approach

Vein morphometric measurements were used to represent conventional computational method for plant species identification and as a benchmark in this study was compared to the deep learning method. The features extracted by the conventional method was evaluated by using ANN classifier with different number of neurons, which ranged between 20 and 100 neurons. The morphometric measurement model with 80 neurons had performed the best result with 99.89% of training accuracy and 66.28% of testing accuracy because the training accuracy tended to decrease after 80 neurons. Besides that, the architecture of the model with higher neuron number was more complicated and time consuming.

The performance of the vein morphometric measurements was not as good as the CNN methods. The vein morphometric measurements obtained a testing accuracy of about 66% only when compared to the CNN methods which achieved more than 90% of testing accuracy. This showed that there was a significant difference between the performance of the vein morphometric measurements and the CNN models.

As mentioned in the previous part, since morphometric measurements is one of the conventional computational approach, it required a series of processes to extract the vein features from the leaf images. This series of processing may cause loss of some important features and information, such as, colour, shape, and texture.

At the same time, the deep learning method can be applied directly on the RGB or coloured leaf images for feature extraction. CNN can extract all the leaf features included in the RGB leaf images. It could extract and condense the features as the extracted features passed from layer to layer. Hence, the last layer of the CNN model would compose of a set of condensed and compacted leaf information from all layers. Generally, the CNN approach encounters all the common features such as shape, vein, colour, etc. during leaf feature extraction. Whereas, in vein morphometric measurement approach, RGB images were converted to grey-scaled images in order to extract the leaf features.

The CNN approach extracts all features in one execution. In contrast, the conventional feature extraction methods extract one type of feature at a time (Beghin et al., 2010; Lin et al., 2008; Kadir et al., 2013). If there are more than one type of features to be extracted, this has to be done separately and manually, which may require a longer time. The combination of all the leaf features (shape, vein, texture and colour) is more appropriate in developing a plant species identification system. In short, the deep learning approach is more beneficial and appropriate than those conventional feature extraction approaches for plant species identification.

5.9 Cross Validation

The proposed D-Leaf approach was tested and examined using the cross-validation method.

k-fold cross-validation approaches was applied to partition the data due to the small number of sample per species (30 samples) used in this study. The 5-fold CV and 10-fold CV obtained 93% of testing accuracy which was comparable to the D-Leaf models with 94.88% accuracy.

5.10 Validation

The D-Leaf approach was further validated with three different leaf datasets which were MalayKew, Flavia and Swedish datasets. The D-Leaf approach was used to extract the leaf features and then classified with the ANN. The D-Leaf approach along with the ANN classifier achieved up to 90% of testing accuracy.

Table 5.4 shows the validation results with the three datasets mentioned above and comparison with the previous studies.

Table 5.4: Validation results and comparison with previous studies.

Dataset	Accuracy	
	By Authors	**D-Leaf Model**
MalayaKew (Lee et al., 2015)	> 97%	90.38%
Flavia (Wu et al., 2007)	90.31%	94.63%
Swedish (Söderkvist, 2001)	--	98.09%

As shown in Table 5.4, the D-Leaf achieved 90.38% with MalayaKew dataset, however, the authors of MalayaKew dataset successfully achieved more than 97% of accuracy in classifying the dataset. Besides that, Wu et al. (2007) obtained 90.31% of accuracy, while D-Leaf performed a higher accuracy at 94.63% by using Flavia dataset. Whereas, D-Leaf performed 98.09% of accuracy in classifying the Swedish dataset. The results of using the D-Leaf approach along with the ANN classifier on the Swedish dataset was up to 98% of testing accuracy. This might be because of the leaf shape of each species which was varied from each other and the Swedish dataset was quite small as compared to the proposed dataset. Small dataset would perform better in classification than those large dataset. However, the D-Leaf approach with the ANN classifier on the MalayaKew dataset produced the least favourable result (90.38% accuracy) which might be due to the black colour background of the provided leaf images. The D-Leaf approach was trained with the leaf images on a white background, thus, it might be less efficient at extracting leaf features from the images on a black background. The Flavia dataset had performed comparably with the proposed dataset since the Flavia dataset was similar to the proposed dataset in terms of the leaf shape of the collected samples.

The results showed that D-Leaf is capable to extract the important leaf features even from the different datasets. It can be confirmed that D-Leaf architecture is practical to be employed to develop an automated plant identification system.

5.11 D-Leaf Plant Species Identification System

A prototype of the graphic user interface (GUI) was designed and developed for automated plant species identification. This can be used to assist the laymen, taxonomists, botanists or other scientists to identify an unknown species by inputting a leaf image. The system requires less time as compared to the traditional methods, which requires searching manually from books to books or from the herbarium collections.

The prototype was tested with an input, which took about four seconds to output the top 5 species. It successfully identified the independent samples within the top 5 species even with complex background and plain background. Consequently, it was capable of identifying an unknown sample either in complex or white-plain background within a short time.

5.12 Summary

As discussed above, first, leaf features extracted from FC5 of the DLeaf-ANN model was performed both of the AlexNet models and took a shorter execution time. Second, the performance of ANN classifier was generally better than the other classifiers. Third, extracted features from whole leaf images contained more information and were more efficient in classification compared to leaf patches.

A comparison between CNN-based methods and conventional method (vein morphometric measurement) was made. Conventional method achieved poorer results compare to the CNN-based models due to the loss of some important information during the pre-processing stage. The proposed D-Leaf model was further validated with other leaf datasets in order to validate the performance of the D-Leaf. The results showed that D-leaf model is efficient in identifying plant species with high accuracy. An automated plant species identification system was developed in order to assist botanists, taxonomists, and novices to conveniently identify unknown leaf samples. The automated system is capable of displaying the top five species within four seconds, which is speedy and efficient.

CHAPTER 6: CONCLUSION

6.1 Introduction

This chapter summarises the findings of the plant species identification system with respect to the research objectives, which are: (1) to extract leaf features from the selected tropical plant species using deep learning- based approach; (2) to compare the performance of the extracted features by using deep learning and conventional approaches; (3) to identify the optimum leaf features in deep learning-based plant species identification and (4) to develop an automated plant species identification system using deep learning-based approach. The research constraints and future works in this research are discussed in this chapter. Lastly, the concluding remarks summarise the whole proposed research.

6.2 Research Summary

The aim of this study was applying deep learning-based approach for extracting the most optimal leaf feature set to develop a plant species identification system. A CNN model named D-Leaf was developed and its performances were very promising.

The leaf images used in this study were collected from the University of Malaya. Only 43 plant species with 30 samples per species were collected due to the limitation of cost and time. The leaf images were pre-processed, post-processed and reconstructed based on the needs of different feature extraction methods. The extracted feature set from different feature extraction methods were then tested and evaluated with different classifiers.

The initial experiment carried out in this research was extracting the leaf features with the use of CNN pre-trained AlexNet and fine-tuned AlexNet models. The extracted features were then analysed and evaluated by using different classifiers. It cannot be denied that these models accomplished a good performance in the identification accuracy (AlexNet model – 93.26%, fine-tuned AlexNet model – 95.54%). However, the execution time of these models was longer e and consumed huge computational memory. Thus, a new and simpler version of the CNN model known as D-Leaf, was proposed.

Five different classification approaches were compared and contrasted in terms of the ability to classify and identify the plant species. It was found that the ANN classifier outperformed the other classifiers in the leaf features classification and identification. Thus, the ANN classifier was selected as the classifier for the development of plant species identification system.

Other than that, CNN was seemed to be superior in feature extraction rather than a classifier. It is able to extract an optimal set of leaf features. However, it could not perform well in classifying the leaf features. Hence, CNN was more suitable to be used as a feature extractor.

During the construction of the D-Leaf model, several combinations of parameter setting were tested by trial and error. The quality or revolution of an image was found to be insignificant trait for CNN-based feature extraction models. The images with higher quality and resolution did not aid in improving the performance of plant species identification. Moreover, high resolution images executed in longer time duration and required larger computational memory. Thus, the images with the size of 250 x 250 were found to be the most optimum input in this research.

In addition, three sets of leaf features were extracted from three different fully connected layers of the D-Leaf model. The potential of each feature set was assessed by the classifiers. The extracted feature sets of these three fully connected layers revealed a comparable performance to each other. With regards to that, the features extracted from the FC5 layer produced a slightly better accuracy (94.88%) compared to the FC4 (91.36%) and FC6 (93.84%) layers.

The proposed model – D-Leaf attained a comparable performance with 94.88% of testing accuracy as the fine-tuned AlexNet achieved 95.54% and AlexNet model achieved 93.26%accuracy. Also, the D-Leaf model could be executed in a shorter time which was 7 times faster than AlexNet and fine-tuned AlexNet models. In summary, the D-Leaf model was superior in the execution time as compared to the AlexNex and fine-tuned AlexNet models while the performance of all 3 models were comparable in terms of testing accuracy.

An experiment was conducted to observe the performance of the extracted features from whole leaf and leaf patch images. The whole leaf images carried more information than the leaf patch. This was the reason why the results of classification with the extracted features from the whole leaf images performed better than the leaf patch images. The whole leaf images were more suitable than the leaf patch images in plant species identification.

In addition, a comparison between the D-Leaf model and a model which employed conventional method - vein morphometric measurement for feature extraction was made. Expectedly, the achievement of the vein morphometric measurement model (66.28%) was fairy poor, compared to the D-Leaf model (94.88%).

In the last part of this research, the potential of the D-Leaf approach was further validated and proved by evaluating with cross validation methods and other plant leaf datasets. The results proved that the D-Leaf approach could be employed and applied efficiently and effectively in the development of plant species identification system.

As a summary, the D-Leaf model with extracted leaf features from the FC5 layer and ANN classifier performed the best and is feasible to be employed as a tool for tropical plant species identification.

The analysis and findings of this proposed method have attained our research objectives as listed below:

(i) Pre-trained AlexNet, fine-tuned AlexNet and a proposed CNN model named D-Leaf were used for extracting the leaf features and the results are very promising.
(ii) CNN models outperformed the conventional method – vein morphological measurement.
(iii) The extracted features of Pre-trained AlexNet, fine-tuned AlexNet and D-Leaf models performed a comparable result.
(iv) Whole leaf images achieved better results than leaf centre patches and side patches.
(v) A prototype of the graphic user based on the DLeaf-ANN was developed.

The findings in (i) accomplished the Objective 1, while, the Objective 2 was attained by the findings in (ii). Other than that, the findings in (iii) and (iv) accomplished the Objective 3. Lastly, the development in (v) accomplished the Objective 4.

6.3 Research Constraints

The sampling progress was heavily dependent on the weather conditions. If it was a rainy day, sampling would not easy to be carried out. Besides that, some of the trees might be infected by certain diseases, and the leaves needed to be excluded from the sampling.

In this research, due to the budget and time constraint, only 43 tropical plant species were included in the dataset.

Deep learning method requires larger computer memory and it is time consuming. Hence, a powerful computer is needed in order to employ deep learning approach in extracting the leaf image features. Besides that, graphics processing unit (GPU) which are efficient and powerful parallel computing is necessary in executing deep learning model. However, a good GPU for deep learning execution is expensive, costing about RM5000.

6.4 Future Works

There are still opportunities for further enhancements of this research. Some future improvements are suggested and described as the following:

First, the number of sample per plant species could be increased by adding more leaf samples for each species. It is suggested to increase the number of leaf sample to 50 samples per species. The increasing in the number of samples of each species could improve the identification accuracy.

Second, the dataset which consisted of 43 tropical plant species could be extended by adding more tropical plant species. It is suggested to include tropical shrub species and herbariums. This could improve and enhance the developed system in identifying more plant species.

Third, since some plant species consists of compound leaves, it is possible to include the compound leaf images into the dataset. The addition of compound leaf images could enhance the system to identify plant species with both single leaf and compound leaves.

Lastly, it is suggested that a web version of the plant species identification system could be developed. This could aid the users in identifying an unknown sample by loading any leaf image to the web-based identification system. It is more convenient that the web-based system could be accessed by the users to identify plant species anywhere and anytime.

6.5 Concluding Remarks

As a summary, the proposed D-Leaf model with ANN classifier provides a computer-based intelligent method to identify the tropical plant species. This model is feasible to assist and aid the botanists, taxonomists and laymen in identifying unknown leaf samples. However, the developed prototype need to be further enhanced and improved by adding more tropical plant species. Although the number of plant species was small, it is hoped that this research will be a stepping stone to encourage more Malaysians to embark in a similar research.

REFERENCES

Aakif, A., & Khan, M. F. (2015). Automatic classification of plants based on their leaves. *Biosystems Engineering, 139*, 66-75.

Adeniyi, D., Wei, Z., & Yongquan, Y. (2016). Automated web usage data mining and recommendation system using K-Nearest Neighbor (KNN) classification method. *Applied Computing and Informatics, 12*(1), 90-108.

Al-Angari, H. M., & Sahakian, A. V. (2012). Automated recognition of obstructive sleep apnea syndrome using support vector machine classifier. *IEEE Transactions on Information Technology in Biomedicine, 16*(3), 463-468.

Amor, N. B., Benferhat, S., & Elouedi, Z. (2004). Naive bayes vs decision trees in intrusion detection systems. In *Proceedings of the 2004 ACM Symposium on Applied Computing* (pp. 420-424). Nicosia, Cyprus: ACM.

Anthimopoulos, M., Christodoulidis, S., Ebner, L., Christe, A., & Mougiakakou, S. (2016). Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. *IEEE Transactions on Medical Imaging, 35*(5), 1207-1216.

Arandjelovic, R., Gronat, P., Torii, A., Pajdla, T., & Sivic, J. (2016). NetVLAD: CNN architecture for weakly supervised place recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 5297-5307). Nevada, Las Vegas.

Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys, 4*, 40-79.

Arun, C. H., Emmanuel, W. S., & Durairaj, D. C. (2013). Texture feature extraction for identification of medicinal plants and comparison of different classifiers. *International Journal of Computer Applications*, *62*(12), 1-9.

Beghin, T., Cope, J. S., Remagnino, P., & Barman, S. (2010). Shape and texture based plant leaf classification. In *International Conference on Advanced Concepts for Intelligent Vision Systems* (pp. 345-353). Berlin, Heidelberg: Springer.

Benard, A., Desmedt, C., Smolina, M., Szternfeld, P., Verdonck, M., Rouas, G., . . . Sotiriou, C. (2014). Infrared imaging in breast cancer: automated tissue component recognition and spectral characterization of breast cancer cells as well as the tumor microenvironment. *Analyst, 139*(5), 1044-1056.

Bhanu, B., & Lee, S. (1994). Image segmentation techniques. In *Genetic Learning for Adaptive Image Segmentation* (pp. 15-24). Boston, MA: Springer.

Bijalwan, V., Kumari, P., Pascual, J., & Semwal, V.B. (2014). Machine learning approach for text and document mining. Retrieved on 31st March 2017, from https://arxiv.org/abs/1406.1580

Bottou, L. (2010). Large-scale machine learning with stochastic gradient descent. In *Proceedings of COMPSTAT'2010* (pp. 177-186). Physica-Verlag HD.

Caglayan, A., Guclu, O., & Can, A. B. (2013). A plant recognition approach using shape and color features in leaf images. In *International Conference on Image Analysis and Processing* (pp. 161-170). Berlin, Heidelberg: Springer.

Camurri, A., Lagerlöf, I., & Volpe, G. (2003). Recognizing emotion from dance movement: Comparison of spectator recognition and automated techniques. *International Journal of Human-Computer Studies, 59*(1-2), 213-225.

Catal, C., Sevim, U., & Diri, B. (2011). Practical development of an Eclipse-based software fault prediction tool using Naive Bayes algorithm. *Expert Systems with Applications, 38*(3), 2347-2353.

Chaki, J., Parekh, R., & Bhattacharya, S. (2015). Plant leaf recognition using texture and shape features with neural classifiers. *Pattern Recognition Letters, 58*, 61-68.

Chang, S.-W., Abdul-Kareem, S., Merican, A. F., & Zain, R. B. (2013). Oral cancer prognosis based on clinicopathologic and genomic markers using a hybrid of feature selection and machine learning methods. *BMC Bioinformatics, 14*(1), 170.

Chen, C., Duan, S., Cai, T., & Liu, B. (2011). Online 24-h solar power forecasting based on weather type classification using artificial neural network. *Solar Energy, 85*(11), 2856-2870.

Chen, H.-L., Huang, C.-C., Yu, X.-G., Xu, X., Sun, X., Wang, G., & Wang, S.-J. (2013). An efficient diagnosis system for detection of Parkinson's disease using fuzzy k-nearest neighbor approach. *Expert Systems with Applications, 40*(1), 263-271.

Choi, I. K., Ahn, H. E., & Yoo, J. (2018). Facial Expression Classification Using Deep Convolutional Neural Network. *Journal of Electrical Engineering & Technology*, *13*(1), 485-492.

Choras, R. S. (2007). Image feature extraction techniques and their applications for CBIR and biometrics systems. *International Journal of Biology and Biomedical Engineering, 1*(1), 6-16.

Cope, J. S., Corney, D., Clark, J. Y., Remagnino, P., & Wilkin, P. (2012). Plant species identification using digital morphometrics: A review. *Expert Systems with Applications*, *39*(8), 7562-7573.

Cope, J. S., Remagnino, P., Barman, S., & Wilkin, P. (2010). Plant texture classification using gabor co-occurrences. In *International Symposium on Visual Computing* (pp. 669-677). Berlin, Heidelberg: Springer.

Cope, J. S., Remagnino, P., Barman, S., & Wilkin, P. (2010). The extraction of venation from leaf images by evolved vein classifiers and ant colony algorithms. In *International Conference on Advanced Concepts for Intelligent Vision Systems* (pp. 135-144). Berlin, Heidelberg: Springer.

Danti, A., Madgi, M., & Anami, B. S. (2012). Mean and range color features based identification of common Indian leafy vegetables. *International Journal of Signal Processing, Image Processing and Pattern Recognition, 5*(3), 151-160.

de Carvalho, M. R., Bockmann, F. A., Amorim, D. S., Brandão, C. R. F., de Vivo, M., de Figueiredo, J. L., ... & Papavero, N. (2007). Taxonomic impediment or impediment to taxonomy? A commentary on systematics and the cybertaxonomic-automation paradigm. *Evolutionary Biology*, *34*(3-4), 140-143.

Díaz-Uriarte, R., & De Andres, S. A. (2006). Gene selection and classification of microarray data using random forest. *BMC Bioinformatics, 7*(1), 3.

Du, J.-X., Huang, D.-S., Wang, X.-F., & Gu, X. (2006). Computer-aided plant species identification (CAPSI) based on leaf shape matching technique. *Transactions of the Institute of Measurement and Control, 28*(3), 275-285

Du, J.-X., Wang, X.-F., & Zhang, G.-J. (2007). Leaf shape based plant species recognition. *Applied mathematics and computation, 185*(2), 883-893.

Ehsanirad, A., & Sharath Kumar, Y. (2010). Leaf recognition for plant classification using GLCM and PCA methods. *Oriental Journal of Computer Science and Technology, 3*(1), 31-36.

Flora Fauna Web - Plant Detail - Cinnamomum iners Reinw. ex Blume. Retrieved on 31st September 2017, from https://florafaunaweb.nparks.gov.sg/Special-Pages/plant-detail.aspx?id=2806.

Friedl, M. A., & Brodley, C. E. (1997). Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment, 61*(3), 399-409.

Gardner, M. W., & Dorling, S. (1998). Artificial neural networks (the multilayer perceptron)—A review of applications in the atmospheric sciences. *Atmospheric Environment, 32*(14-15), 2627-2636.

Gershenson, C. (2003). Artificial neural networks for beginners. In *Cognitive and computing sciences*. East Sussex, U.K.: University of Sussex.

Gibson, E., Robu, M. R., Thompson, S., Edwards, P. E., Schneider, C., Gurusamy, K., ... & Clarkson, M. J. (2017). Deep residual networks for automatic segmentation of laparoscopic videos of the liver. In *Medical Imaging 2017: Image-Guided Procedures, Robotic Interventions, and Modeling* (Vol. 10135, p. 101351M). International Society for Optics and Photonics.

Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1): MIT press Cambridge.

Goswami, G., Bhardwaj, R., Singh, R., & Vatsa, M. (2014). MDLFace: Memorability augmented deep learning for video face recognition. In *Biometrics (IJCB), 2014 IEEE International Joint Conference on* (pp. 1-7). Florida, USA: IEEE.

Grinblat, G. L., Uzal, L. C., Larese, M. G., & Granitto, P. M. (2016). Deep learning for plant identification using vein morphological patterns. *Computers and Electronics in Agriculture, 127*, 418-424.

Guo, H., Liu, B., Cai, D., & Lu, T. (2018). Predicting protein–protein interaction sites using modified support vector machine. *International Journal of Machine Learning and Cybernetics, 9*(3), 393-398.

Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., & Lew, M. S. (2016). Deep learning for visual understanding: A review. *Neurocomputing, 187*, 27-48.

Gupta, S., & Mazumdar, S. G. (2013). Sobel edge detection algorithm. *International journal of computer science and management Research, 2*(2), 1578-1583.

Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine learning, 46*(1-3), 389-422.

Gwo, C. Y., & Wei, C. H. (2013). Plant identification through images: Using feature extraction of key points on leaf contours1. *Applications in plant sciences*, *1*(11), 1200005.

Hadad, O., Bakalo, R., Ben-Ari, R., Hashoul, S., & Amit, G. (2017). Classification of breast lesions using cross-modal deep learning. In *Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on* (pp. 109-112). Melbourne, Australia: IEEE.

Han, X., Zhong, Y., Cao, L., & Zhang, L. (2017). Pre-Trained AlexNet Architecture with Pyramid Pooling and Supervision for High Spatial Resolution Remote Sensing Image Scene Classification. *Remote Sensing, 9*(8), 848.

Hartley, P., Flamary, R., Jackson, N., Tagore, A., & Metcalf, R. (2017). Support vector machine classification of strong gravitational lenses. *Monthly Notices of the Royal Astronomical Society, 471*(3), 3378-3397.

Hati, S., & Sajeevan, G. (2013). Plant recognition from leaf image through artificial neural network. *International Journal of Computer Applications, 62*(17).

Hefner, J. T., & Ousley, S. D. (2014). Statistical classification methods for estimating ancestry using morphoscopic traits. *Journal of Forensic Sciences, 59*(4), 883-890

Hinton, G. E., & Sejnowski, T. J. (1986). Learning and releaming in boltzmann machines. *Parallel distributed processing: Explorations in the microstructure of cognition, 1*(282-317), 2.

Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A.-r., Jaitly, N., . . . Sainath, T. N. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine, 29*(6), 82-97.

Hrechak, A. K., & McHugh, J. A. (1990). Automated fingerprint recognition using structural matching. *Pattern Recognition, 23*(8), 893-904.

Hsu, C.-W., Chang, C.-C., & Lin, C.-J. (2003). A practical guide to support vector classification. Technical report, Department of Computer Science, National Taiwan University.

Hu, F., Xia, G.-S., Hu, J., & Zhang, L. (2015). Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. *Remote Sensing, 7*(11), 14680-14707.

Hu, W., Huang, Y., Wei, L., Zhang, F., & Li, H. (2015). Deep convolutional neural networks for hyperspectral image classification. *Journal of Sensors, 2015*, 258619-1–258619-12.

Jarasch, A., & Skerra, A. (2017). Aligning, analyzing, and visualizing sequences for antibody engineering: Automated recognition of immunoglobulin variable region features. *Proteins: Structure, Function, and Bioinformatics, 85*(1), 65-71.

Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., ... & Darrell, T. (2014). Caffe: Convolutional architecture for fast feature embedding. In *Proceedings of the 22nd ACM international conference on Multimedia* (pp. 675-678). Florida, USA: ACM.

Jin, J., Fu, K., & Zhang, C. (2014). Traffic sign recognition with hinge loss trained convolutional neural networks. *IEEE Transactions on Intelligent Transportation Systems, 15*(5), 1991-2000.

Kadir, A. (2014). A Model of Plant Identification System Using GLCM, Lacunarity and Shen Features. *Research Journal of Pharmaceutical, Biological and Chemical Sciences*, *5*(2), 1-10.

Kadir, A., Nugroho, L. E., Susanto, A., & Santosa, P. I. (2013). Leaf classification using shape, color, and texture features. *International Journal of Computer Trends and Technology*, *1*(3), 225-230.

Kale, K. V. (2014). Zernike moment feature extraction for handwritten Devanagari (Marathi) compound character recognition. *International Journal of Advanced Research in Artificial Intelligence Elsevier, 3* (1), 68-78.

Khan, J., Wei, J. S., Ringner, M., Saal, L. H., Ladanyi, M., Westermann, F., . . . Peterson, C. (2001). Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. *Nature Medicine, 7*(6), 673.

Kim, C., Son, H., & Kim, C. (2013). Automated construction progress measurement using a 4D building information model and 3D data. *Automation in Construction, 31*, 75-82.

Krenker, A., Bešter, J., & Kos, A. (2011). Introduction to the artificial neural networks. *Artificial Neural Networks: Methodological Advances and Biomedical Applications. InTech*, 1-18.

Krimmel, E., Monfils, A., Marsico, T., Morris, A., Linton, D., & Ruhfel, B. *Taxonomic keys.* Retrieved on 2nd February 2018, from https://collectionseducation.org/identify-specimen/taxonomic-keys/

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).

Kroening, M., Trinklein, D., & Schrock, D. (2007). Plants and their environment. In *Master Garderner*. Retrieved on 15th March 2018, from https://mospace.umsystem.edu/xmlui/bitstream/handle/10355/51765/mg0002-2007.pdf?sequence=1

Kumar, N., Belhumeur, P. N., Biswas, A., Jacobs, D. W., Kress, W. J., Lopez, I. C., & Soares, J. V. (2012). Leafsnap: A computer vision system for automatic plant species identification. In *Computer vision–ECCV 2012* (pp. 502-516). Berlin, Heidelberg: Springer.

Laga, H., Kurtek, S., Srivastava, A., Golzarian, M., & Miklavcic, S. J. (2012). A Riemannian elastic metric for shape-based plant leaf classification. In *Digital Image Computing Techniques and Applications (DICTA), 2012 International Conference on* (pp. 1-7). Fremantle, Western Australia: IEEE.

Lakhani, P., & Sundaram, B. (2017). Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology, 284*(2), 574-582.

Lee, S. H., Chan, C. S., Mayo, S. J., & Remagnino, P. (2017). How deep learning extracts and learns leaf features for plant classification. *Pattern Recognition, 71*, 1-13.

Lee, S. H., Chan, C. S., Wilkin, P., & Remagnino, P. (2015). Deep-plant: Plant identification with convolutional neural networks. In *Image Processing (ICIP), 2015 IEEE International Conference on* (pp. 452-456). Quebec City, Canada: IEEE.

Lee, S. H., Chang, Y. L., Chan, C. S., & Remagnino, P. (2016). Plant Identification System based on a Convolutional Neural Network for the LifeClef 2016 Plant Classification Task. In *CLEF (Working Notes)* (pp. 502-510). Évora, Portugal.

Leishman, M. R., & Westoby, M. (1992). Classifying plants into groups on the basis of associations of individual traits--evidence from Australian semi-arid woodlands. *Journal of Ecology*, 417-424.

Lévy, D., & Jain, A. (2016). Breast mass classification from mammograms using deep convolutional neural networks. Retrieved on 7th January 2018, from http://arxiv.org/abs/1612.00542

Li, W., Yi, P., Wu, Y., Pan, L., & Li, J. (2014). A new intrusion detection system based on KNN classification algorithm in wireless sensor network. *Journal of Electrical and Computer Engineering, 2014*, 1–8.

Libbrecht, M. W., & Noble, W. S. (2015). Machine learning applications in genetics and genomics. *Nature Reviews Genetics, 16*(6), 321.

Lien, J. J., Kanade, T., Cohn, J. F., & Li, C. C. (1998). Automated facial expression recognition based on FACS action units. In *Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition (FG'98) on* (pp. 390). Nara, Japan: IEEE.

Lin, H., & Peng, H. (2008). Machine recognition for broad-leaved trees based on synthetic features of leaves using probabilistic neural network. In *Computer Science and Software Engineering, 2008 International Conference on* (Vol. 4, pp. 871-877). Wuhan, China: IEEE.

Liu, B., Zhang, Y., He, D., & Li, Y. (2017). Identification of Apple Leaf Diseases Based on Deep Convolutional Neural Networks. *Symmetry, 10*(1), 11.

Maldonado-Bascón, S., Lafuente-Arroyo, S., Gil-Jimenez, P., Gómez-Moreno, H., & López-Ferreras, F. (2007). Road-sign detection and recognition based on support vector machines. *IEEE Transactions on Intelligent Transportation Systems, 8*(2), 264-278.

Malladi, R., & Sethian, J. A. (1996). A unified approach to noise removal, image enhancement, and shape recovery. *IEEE Transactions on Image Processing, 5*(11), 1554-1568.

Mathioulakis, E., Panaras, G., & Belessiotis, V. (2018). Artificial neural networks for the performance prediction of heat pump hot water heaters. *International Journal of Sustainable Energy, 37*(2), 173-192.

Matsugu, M., Mori, K., Mitari, Y., & Kaneda, Y. (2003). Subject independent facial expression recognition with robust face detection using a convolutional neural network. *Neural Networks, 16*(5-6), 555-559.

McCallum, A., & Nigam, K. (1998). A comparison of event models for naive bayes text classification. In *AAAI-98 workshop on learning for text categorization* (Vol. 752, No. 1, pp. 41-48). Madison, Wisconsin.

Medin, D. L., Lynch, E. B., Coley, J. D., & Atran, S. (1997). Categorization and reasoning among tree experts: Do all roads lead to Rome? *Cognitive psychology, 32*(1), 49-96.

Minaee, S., Abdolrashidiy, A., & Wang, Y. (2016). An experimental study of deep convolutional features for iris recognition. In *Signal Processing in Medicine and Biology Symposium (SPMB), 2016 IEEE* (pp. 1-6). Pennsylvania, USA: IEEE.

Moldagulova, A., & Sulaiman, R. B. (2017). Using KNN algorithm for classification of textual documents. In *Information Technology (ICIT), 2017 8th International Conference on* (pp. 665-671). Amman, Jordan: IEEE.

Morse, L. E. (1971). Specimen identification and key construction with time-sharing computers. *Taxon*, 269-282.

Murat, M., Chang, S.-W., Abu, A., Yap, H. J., & Yong, K.-T. (2017). Automated classification of tropical shrub species: A hybrid of leaf shape and machine learning approach. *PeerJ, 5*, e3792.

Nagi, J., Ducatelle, F., Di Caro, G. A., Cireşan, D., Meier, U., Giusti, A., ... & Gambardella, L. M. (2011, November). Max-pooling convolutional neural networks for vision-based hand gesture recognition. In *Signal and Image Processing Applications (ICSIPA), 2011 IEEE International Conference on* (pp. 342-347). Kuala Lumpur, Malaysia: IEEE.

Nakanishi, T., Shio, A., & Ishii, K. I. (1995). Automatic vehicle image extraction based on spatio - temporal image analysis. *Systems and Computers in Japan, 26*(12), 71-82.

Nasir, M. K., & Uddin, M. S. (2013). Hand written bangla numerals recognition for automated postal system. *IOSR Journal of Computer Engineering (IOSR-JCE), 8*(6), 43-48.

Ngiam, J., Khosla, A., Kim, M., Nam, J., Lee, H., & Ng, A. Y. (2011). Multimodal deep learning. In *Proceedings of the 28th international conference on machine learning (ICML-11)* (pp. 689-696). Washington, USA: Hyatt Regency Bellevue.

Nguyen, A., Yosinski, J., & Clune, J. (2015). Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 427-436). Boston, Massachusetts: Hynes Convention Center.

Oskouie, P., Becerik-Gerber, B., & Soibelman, L. (2017). Automated Recognition of Building Façades for Creation of As-Is Mock-Up 3D Models. *Journal of Computing in Civil Engineering, 31*(6), 04017059.

Ou, X., Pan, W., & Xiao, P. (2014). In vivo skin capacitive imaging analysis by using grey level co-occurrence matrix (GLCM). *International journal of pharmaceutics, 460*(1-2), 28-32.

Pal, M. (2005). Random forest classifier for remote sensing classification. *International Journal of Remote Sensing, 26*(1), 217-222.

Pantic, I., Dacic, S., Brkic, P., Lavrnja, I., Jovanovic, T., Pantic, S., & Pekovic, S. (2015). Discriminatory ability of fractal and grey level co-occurrence matrix methods in structural analysis of hippocampus layers. *Journal of Theoretical Biology, 370*, 151-156.

Park, Y.-S., Céréghino, R., Compin, A., & Lek, S. (2003). Applications of artificial neural networks for patterning and predicting aquatic insect species richness in running waters. *Ecological modelling, 160*(3), 265-280.

Patil, T. R., & Sherekar, S. (2013). Performance analysis of Naive Bayes and J48 classification algorithm for data classification. *International Journal of Computer Science and Applications, 6*(2), 256-261.

Pawara, P., Okafor, E., Surinta, O., Schomaker, L., & Wiering, M. (2017). Comparing Local Descriptors and Bags of Visual Words to Deep Convolutional Neural Networks for Plant Recognition. In *ICPRAM* (pp. 479-486).Peng, P., Zhao, X., Pan, X., & Ye, W. (2018). Gas Classification Using Deep Convolutional Neural Networks. *Sensors, 18*(1), 157.

Pornpanomchai, C., Rimdusit, S., Tanasap, P., & Chaiyod, C. (2011). Thai herb leaf image recognition system (THLIRS). *Kasetsart Journal-Natural Science*, *45*(3), 347 – 351.

Prasvita, D. S., & Herdiyeni, Y. (2013). Medleaf: mobile application for medicinal plant identification based on leaf image. *International Journal on Advanced Science, Engineering and Information Technology, 3*(2), 103-106.

Randler, C. (2008). Teaching Species Identification—A Prerequisite for Learning Biodiversity and Understanding Ecology. *Eurasia Journal of Mathematics, Science & Technology Education*, *4*(3).

Razavian, A. S., Azizpour, H., Sullivan, J., & Carlsson, S. (2014). CNN features off-the-shelf: an astounding baseline for recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (pp. 806-813). Columbus, Ohio: Greater Columbus Convention Center.

Refaeilzadeh, P., Tang, L., & Liu, H. (2009). Cross Validation, Encyclopedia of Database Systems (EDBS). *Arizona State University, Springer*, 6.

Rizwan, M., & Anderson, D. V. (2014). Using k-Nearest Neighbor and speaker ranking for phoneme prediction. In *Machine Learning and Applications (ICMLA), 2014 13th International Conference on* (pp. 383-387). Detroit, MI USA: IEEE.

Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J. P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing, 67*, 93-104.

Sariyanidi, E., Gunes, H., Gökmen, M., & Cavallaro, A. (2013). Local Zernike Moment Representation for Facial Affect Recognition.In *BMVC*, (2 pp. 3). Bristol, England: University of Bristol.

Sarraf, S., & Tofighi, G. (2016). Classification of alzheimer's disease using fmri data and deep learning convolutional neural networks. Retrieved on 21st Jun 2017, from https://arxiv.org/abs/1603.08631..

Sawaf, H., Zaplo, J., & Ney, H. (2001). Statistical classification methods for Arabic news articles. *Natural Language Processing in ACL2001, Toulouse, France*.

Sermanet, P., Chintala, S., & LeCun, Y. (2012). Convolutional neural networks applied to house numbers digit classification. In *Pattern Recognition (ICPR), 2012 21st International Conference on* (pp. 3288-3291). Tsukuba Science City, Japan: IEEE.

Sharma, S., & Gupta, C. (2015). Recognition of plant species based on leaf images using multilayer feed forward neural network. *International Journal of Innovative Research in Advanced Engineering, 2*(6), 104-110.

Shirvan, R. A., & Tahami, E. (2011). Voice analysis for detecting Parkinson's disease using genetic algorithm and KNN classification method. *Proceeding of 18th Iranian conference on biomedical engineering*,(pp. 278-283). Tehran, Iran.

Silva, P. F., Marcal, A. R., & da Silva, R. M. A. (2013, June). Evaluation of features for leaf discrimination. In *International Conference Image Analysis and Recognition* (pp. 197-204). Berlin, Heidelberg: Springer.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. In *ICLR*. Banff, Canada

Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., & Stefanovic, D. (2016). Deep neural networks based recognition of plant diseases by leaf image classification. *Computational intelligence and neuroscience, 2016*, 3289801.

Söderkvist, O. (2001). *Computer vision classification of leaves from swedish trees* (Master's thesis). Retrieved on 21st July 2017, from http://citeseerx.ist.psu.edu

Sommer, C., & Gerlich, D. W. (2013). Machine learning in cell biology–teaching computers to recognize phenotypes. *J Cell Sci, 126*(24), 5529-5539.

Sulc, M., & Matas, J. (2014). Texture-based leaf identification. In *European Conference on Computer Vision*(pp. 185-200). Cham: Springer.

Sünderhauf, N., McCool, C., Upcroft, B., & Perez, T. (2014). Fine-Grained Plant Classification Using Convolutional Neural Networks for Feature Extraction. In *CLEF (Working Notes)* (pp. 756-762). Sheffield, UK: University of Sheffield.

Svetnik, V., Liaw, A., Tong, C., Culberson, J. C., Sheridan, R. P., & Feuston, B. P. (2003). Random forest: a classification and regression tool for compound classification and QSAR modeling. *Journal of Chemical Information and Computer Sciences, 43*(6), 1947-1958.

Swain, S., & Sarangi, S. S. (2013). Study of Various Classification Algorithms using Data Mining. *International Journal of Advanced Research in, 2*(2), 110-114.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1-9). Boston, Massachusetts: Hynes Convention Center.

Tan, H. (1970). Butterflies of Singapore. Retrieved on 12th March 2018, from http://butterflycircle.blogspot.my/2015/05/larval-food-plant-for-butterflies-wild.html

Tan, M. S., Tan, J. W., Chang, S.-W., Yap, H. J., Kareem, S. A., & Zain, R. B. (2016). A genetic programming approach to oral cancer prognosis. *PeerJ, 4*, e2482.

Tan, S., Cheng, X., Wang, Y., & Xu, H. (2009). Adapting naive bayes to domain adaptation for sentiment analysis. In *European Conference on Information Retrieval* (pp. 337-349). Berlin, Heidelberg: Springer.

Tarter, R. E. (1990). Evaluation and treatment of adolescent substance abuse: A decision tree method. *The AmericanJournal of Drug and Alcohol Abuse, 16*(1-2), 1-46.

To-Anun, C., Hidayat, I., & Meeboon, J. (2011). Genus Cercospora in Thailand: Taxonomy and phylogeny (with a dichotomous key to species). *Plant Pathology & Quarantine, 1*(1), 11-87.

Tooke, T. R., Coops, N. C., Goodwin, N. R., & Voogt, J. A. (2009). Extracting urban vegetation characteristics using spectral mixture analysis and decision tree classifications. *Remote Sensing of Environment, 113*(2), 398-407.

Tyrl, R. J. (2010). Critic's Choice Essay:" Being a Method proposed for the ready finding... to what Sort any Plant belongeth". *Oklahoma Native Plant Record, 10*(1).

Tzotsos, A., & Argialas, D. (2008). Support vector machine classification for object-based image analysis. In *Object-Based Image Analysis* (pp. 663-677). Berlin, Heidelberg: Springer.

University of Malaya. In *Google Maps*. Retrieved on 10th February 2018, from https://www.google.com/maps/place/University+of+Malaya/@3.1200852,101.6534117,540m/data=!3m1!1e3!4m5!3m4!1s0x31cdb47024217187:0x1e85ebc65d47d641!8m2!3d3.1201068!4d101.6545492

Vincent, O. R., & Folorunso, O. (2009). A descriptive algorithm for sobel image edge detection. In *Proceedings of Informing Science & IT Education Conference (InSITE)* (pp. 97-107). California: Informing Science Institute.

Wang, B., Gao, Y., Sun, C., Blumenstein, M., & La Salle, J. (2017). Can Walking and Measuring along Chord Bunches Better Describe Leaf Shapes? In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 6119-6128). Honolulu, Hawai.

Wang, D., Khosla, A., Gargeya, R., Irshad, H., & Beck, A. H. (2016). Deep learning for identifying metastatic breast cancer. Retrieved on 7th May 2017, from https://arxiv.org/abs/1606.05718v1.

Wang, X. F., Du, J. X., & Zhang, G. J. (2005. Recognition of leaf images based on shape features using a hypersphere classifier. In *International Conference on Intelligent Computing* (pp. 87-96). Berlin, Heidelberg: Springer.

Wang, Z. (2015). The applications of deep learning on traffic identification. Retrieved on 19th August 2017, from https://www.blackhat.com

Wang, Z., Sun, X., Zhang, Y., Ying, Z., & Ma, Y. (2016). Leaf recognition based on PCNN. *Neural Computing and Applications, 27*(4), 899-908.

Willis, K. J. (2017). State of the world's plants 2017. *Report. Royal Botanic Gardens, Kew*. Retrieved on 26th September 2017, from https://stateoftheworldsplants.org

Wilson, R. L., & Sharda, R. (1994). Bankruptcy prediction using neural networks. *Decision Support Systems, 11*(5), 545-557.

Wu, S. G., Bao, F. S., Xu, E. Y., Wang, Y. X., Chang, Y. F., & Xiang, Q. L. (2007). A leaf recognition algorithm for plant classification using probabilistic neural network. In *Signal Processing and Information Technology, 2007 IEEE International Symposium on* (pp. 11-16). Cairo, Egypt: IEEE.

Wu, Z., Wang, X., Jiang, Y. G., Ye, H., & Xue, X. (2015). Modeling spatial-temporal clues in a hybrid deep learning framework for video classification. In *Proceedings of the 23rd ACM international conference on Multimedia* (pp. 461-470). Brisbane, Australia: ACM.

Yadav, A. R., Dewal, M. L., Anand, R. S., & Gupta, S. (2013). Classification of hardwood species using ANN classifier. In *Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), 2013 Fourth National Conference on* (pp. 1-5). Rajasthan, India: IEEE.

Zaitoun, N. M., & Aqel, M. J. (2015). Survey on image segmentation techniques. *Procedia Computer Science, 65*, 797-806.

Zeiler, M. D., & Fergus, R. (2013). Stochastic pooling for regularization of deep convolutional neural networks. Retrieved on 8th Jun 2017, from https://arxiv.org/abs/1301.3557

Zeng, S., Zhang, B., & Du, Y. (2017). Joint distances by sparse representation and locality-constrained dictionary learning for robust leaf recognition. *Computers and Electronics in Agriculture, 14*2, 563-571.

Zhong, Z., Jin, L., & Xie, Z. (2015,). High performance offline handwritten chinese character recognition using googlenet and directional feature maps. In *Document Analysis and Recognition (ICDAR), 2015 13th International Conference on* (pp. 846-850). Nancy, France: IEEE.

LIST OF PUBLICATIONS AND PAPERS PRESENTED

[1] **Tan, J. W.**, Chang, S.-W., Abdul-Kareem, S., Yap, H. J., & Yong, K.-T. (2018). Deep Learning for Plant Species Classification using Leaf Vein Morphometric, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*. doi:10.1109/TCBB.2018.2848653 (ISI Publication-Q1)

[2] **Tan, J. W.**, Chang, S.-W., Abdul-Kareem. (2017). The Application of Leaf Venation, Feature Selection and Classification Techniques in Plant Species Identification, Biological Sciences Graduate Congress (BSGC) 2017. (Poster Presentation)