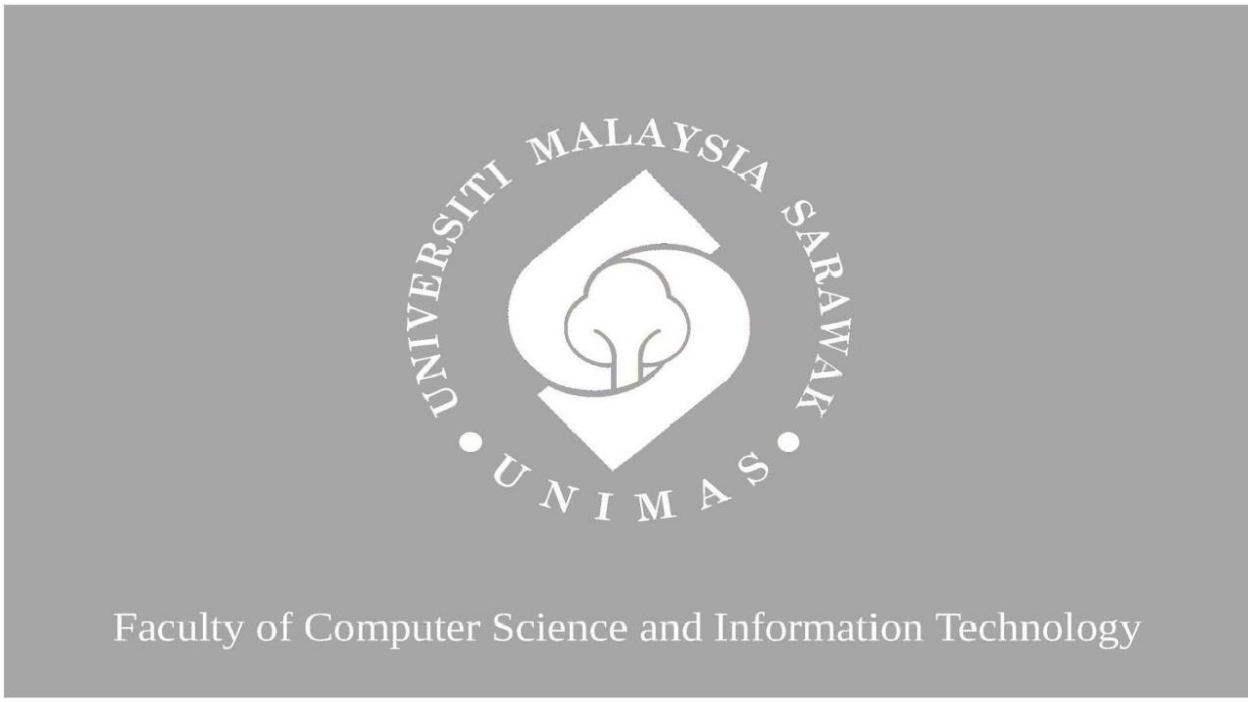

Lettuce Leaf Disease Detection Using Convolutional Neural Network Algorithm

NGU SU HANG

Bachelor of Computer Science with Honours

(Software Engineering)

2023

LETTUCE LEAF DISEASE DETECTION USING CONVOLUTIONAL NEURAL NETWORK ALGORITHM

NGU SU HANG

This project is submitted in partial fulfilment of the
requirements for the degree of
Bachelor of Computer Science and Information Technology
(Software Engineering)

Faculty of Computer Science and Information Technology
UNIVERSITI MALAYSIA SARAWAK
2023

UNIVERSITI MALAYSIA SARAWAK

THESIS STATUS ENDORSEMENT FORM

TITLE LETTUCE LEAF DISEASE DETECTOR USING CONVOLUTIONAL NEURAL NETWORK ALGORITHM

ACADEMIC SESSION: 2019/2020

NGU SU HANG
(CAPITAL LETTERS)

hereby agree that this Thesis* shall be kept at the Centre for Academic Information Services, Universiti Malaysia Sarawak, subject to the following terms and conditions:

1. The Thesis is solely owned by Universiti Malaysia Sarawak
2. The Centre for Academic Information Services is given full rights to produce copies for educational purposes only
3. The Centre for Academic Information Services is given full rights to do digitization in order to develop local content database
4. The Centre for Academic Information Services is given full rights to produce copies of this Thesis as part of its exchange item program between Higher Learning Institutions [or for the purpose of interlibrary loan between HLI]
5. ** Please tick (√)

	CONFIDENTIAL	(Contains classified information bounded by the OFFICIAL SECRETS ACT 1972)
	RESTRICTED	(Contains restricted information as dictated by the body or organization where the research was conducted)
√	UNRESTRICTED	

Validated by

(AUTHOR'S SIGNATURE)

(SUPERVISOR'S SIGNATURE)

Permanent Address

NO.318, LOT2291, LORONG
7/1, JALAN SEMABA, BATU 5,
93250, KUCHING, SARAWAK

Date: 15 JULY 2023

Date: 17 JULY 2023

Note * Thesis refers to PhD, Master, and Bachelor Degree
** For Confidential or Restricted materials, please attach relevant documents from relevant organizations / authorities

DECLARATION OF ORIGINALITY

I hereby declare that this project altogether with all of its contents is none other than that of my work, with consideration of the exception of reference or acknowledgement is not made explicitly in the text, nor has any part had been written for me by another person.

……

(NGU SU HANG)

Faculty of Computer Science and

Information Technology,

Universiti Malaysia Sarawak

30TH JUNE 2023

ACKNOWLEDGEMENT

I would like to thank my FYP supervisor, Professor Dr. Wang Yin Chai for guiding me throughout the final year project. He was there to help me every step of the way, and his motivation is what helped me in the way to complete this project successfully. In addition, I would like to thank my friends and parents who always encouraging me, guiding me and being there to strengthen me with confidence every time I felt weak or discouraged.

ABSTRACT

Plant disease is a major problem towards agriculture, as some of the disease could be infectious, the farmer who are not expert in observing plant disease may lead to the disaster of plant dying. Lettuce is a vegetable that is usually served as salad because of the taste crisp and mild. Although lettuce is a cool season crop, it can be grown in Malaysia by controlling the temperature and the environment. The examples of lettuce disease are Powdery Mildew, Downy Mildew, Bacterial Leaf Spot, and the infection of Mosaic Virus. The diseased lettuce can be healed if it is observed in early stage, but the lesion of disease area in early stage is hard to observe with raw eye. Therefore, this project proposed a lettuce leaf disease detection application using deep learning algorithm which is the convolutional neural network (CNN) to classify whether the image of the lettuce leaf is healthy or diseased. The detection algorithm will be develop based on a modified AlexNet model. The input dataset for the training of model is the images of healthy lettuce, bacterial leaf spot diseased lettuce and powdery mildew diseased lettuce. The images all are undergoing image processing to enrich the image dataset, improve the performance of the model and avoid overfitting problem. Each image will be labelled with the class for the CNN model to classify it. The image dataset will split into three set, training, validation and the testing. The evaluation of the model will be looking at the performance metrics which are precision, recall, F1 score and accuracy. The trained CNN model will then implement using OpenCV for real-time operation and Python language for the programming.

Keywords: Convolutional neural network, image classification, image processing, lettuce leaf disease

Pengesanan Penyakit Daun Salad menggunakan Algoritma CNN

ABSTRAK

Penyakit tumbuhan adalah masalah utama terhadap pertanian, kerana sesetengah penyakit boleh berjangkit kepada tumbuhan lain, petani yang tidak pakar dalam memerhati penyakit tumbuhan boleh membawa kepada bencana kematian tumbuhan. Salad adalah sayuran yang biasanya dihidangkan sebagai salad kerana rasanya yang segar dan lembut. Walaupun salad merupakan tanaman musim sejuk, ia boleh ditanam di Malaysia dengan mengawal suhu dan persekitaran. Contoh penyakit salad ialah Cendawan Serbuk, Cendawan Downy, Bintik Daun Bakteria, dan jangkitan Virus Mosaic. Salad yang berpenyakit boleh disembuhkan jika ia diperhatikan pada peringkat awal, tetapi lesi kawasan penyakit pada peringkat awal sukar untuk diperhatikan dengan mata mentah. Oleh itu, projek ini mencadangkan aplikasi untuk pengesanan penyakit daun salad dengan menggunakan algoritma pembelajaran mendalam iaitu rangkaian saraf konvolusi (CNN) untuk mengklasifikasikan sama ada imej daun salad itu keadaan sihat atau berpenyakit. Algoritma CNN akan dibangunkan berdasarkan model AlexNet yang diubahsuai. Set data input untuk latihan model ialah imej salad sihat, salad berpenyakit bintik daun bakteria dan daun salad berpenyakit cendawan serbuk. Semua imej sedang menjalani pemprosesan imej untuk memperkaya set data imej, meningkatkan prestasi model dan mengelakkan masalah overfitting. Setiap imej akan dilabelkan dengan kelas mengikuti untuk model CNN mengelaskannya. Set data imej akan dibahagikan kepada tiga set, iaitu Latihan, pengesahan dan ujian. Penilaian model akan melihat kepada metrik prestasi iaitu ketepatan, ingatan semula, skor F1 dan ketepatan. Model CNN yang terlatih kemudiannya akan melaksanakan menggunakan OpenCV untuk operasi masa nyata dan bahasa Python untuk program tersebut.

Kata Kunci: *Klasifikasi imej, pemprosesan imej, penyakit daun salad, rangkaian saraf konvolusi*

TABLE OF CONTENTS

	Page
DECLARATION OF ORIGINALITY	**i**
ACKNOWLEDGEMENT	**ii**
ABSTRACT	**iii**
ABSTRAK	**iv**
TABLE OF CONTENTS	**v**
LIST OF TABLES	**x**
LIST OF FIGURES	**xi**
CHAPTER 1 INTRODUCTION	**1**
1.1 Introduction	1
1.2 Problem Statement	2
1.3 Project Objective	3
1.4 Methodology	4
1.4.1 Proposed Model	4
1.4.2 Image Acquisition and Pre-processing	5
1.5 Project Scope	5
1.6 Significance of Project	6
1.7 Expected Outcome	6

CHAPTER 2 LITERATURE REVIEW 7

2.1 Introduction 7

2.2 Lettuce Leaf Disease 8

2.3 Plant Disease Detection 9

2.3.1 Machine Learning Algorithm 9

2.3.2 Deep Learning Algorithm 10

2.4 Learning Method 11

2.5 Algorithms for Plant Disease Detection 12

2.5.1 Identification of Apple Leaf Disease Based on Deep Convolutional Neural Networks 12

2.5.2 Plant Disease Detection and Classification using CNN Model with Optimized Activation Function 14

2.5.3 Leaf Disease Detection Using Support Vector Machine 15

2.5.4 Plant Disease Detection Using k-Nearest Neighbours 17

2.5.5 Plant Disease Detection Using Random Forest 18

2.6 Comparison of Existing Algorithms for Plant Disease Detection 19

2.7 Machine Learning Software Tools 22

2.7.1 TensorFlow 22

2.7.2 PyTorch 23

2.7.3 Google Colaboratory 23

2.8 Comparison of Machine Learning Software Tools 24

2.9 Summary 26

CHAPTER 3 REQUIREMENT ANALYSIS AND DESIGN 28

3.1 Introduction 28

3.2 Image Acquisition and Pre-processing 29

3.2.1 Light Disturbance 29

3.2.2 Direction Disturbance 30

3.2.3 Image Resize 31

3.3 CNN Algorithm Architecture 32

3.3.1 AlexNet Precursor 32

3.3.2 Cascade Inception 33

3.3.3 Convolutional Layer 33

3.3.4 Pooling Layer 34

3.3.5 Inception 35

3.3.6 ReLU Activation Function 36

3.3.7 Fully Connected Layer 36

3.4 Training and Testing Model 37

3.5 Evaluating Model 37

3.6 Development for Prototype of Lettuce Leaf Disease Detection 38

3.6.1 Python 38

3.6.2 Anaconda Navigator 38

3.6.3 OpenCV 39

3.6.4 Visual Studio Code 39

3.6.5 Camera 39

3.7 Conclusion 39

CHAPTER 4 IMPLENTATION 40

4.1 Introduction 40

4.2 Image Collection 40

4.3 Image Pre-processing 42

4.3.1 Light Disturbance 43

4.3.2 Image Rotation 46

4.3.3 Image Mirroring 49

4.4 Implementation of Proposed CNN Model 50

4.5 Build Proposed CNN Model 52

4.6 Train Proposed CNN Model 54

4.7 Implementing the Trained Model in Real Time Method 57

4.8 Conclusion 59

CHAPTER 5 TESTING AND EVALUATION 60

5.1 Introduction 60

5.2 Performance of Trained Model 60

5.3 Evaluate Trained Model 61

5.4 Compare with Modified AlexNet Model 64

5.5 Testing Result of Prototype of Lettuce Leaf Disease Detection Application 65

5.6 Conclusion 69

CHAPTER 6 CONCLUSION AND FUTURE WORKS 70

6.1 Introduction 70

6.2 Summary of Project 70

6.3 Objective Achievement 70

6.4 Work Limitations 72

6.5 Future Works 72

6.6 Conclusion 73

REFERENCE 75

APPENDIX 78

LIST OF TABLES

Table 2.1: Comparison between algorithms 20

Table 2.2: Comparison between machine learning software or tools 25

Table 4.1: Image acquisition for lettuce leaf with healthy condition 41

Table 4.2: Image amount after image pre-processing 49

Table 4.3: Summary of Proposed CNN model 51

Table 4.4: Number of images in each dataset 55

Table 5.1: Performance of trained model in training, validation and testing 61

Table 5.2: Performance of modified AlexNet model on lettuce leaf disease dataset 65

Table 5.3: Testing result of the prototype classifying lettuce leaf 66

Table 6.1: Limitations of this project 72

LIST OF FIGURES

Figure 2.1: Bacterial leaf spot on lettuce (Baysal-Gurel & Miller, 2010). 8

Figure 2.2: Powdery Mildew on Lettuce (Baysal-Gurel& Miller, 2010). 9

Figure 2.3: Plant disease detection by using machine learning (Shruthi et al., 2019). 10

Figure 2.4: Deep learning algorithm (Adedoja et al., 2021). 11

Figure 2.5: Parameters of the convolutional neural network (CNN)-based model (Liu et al., 2018). 14

Figure 2.6: The structure of the convolutional neural network model (Liu et al., 2018). 14

Figure 2.7: The proposed model based on SVM (Das et al., 2020). 16

Figure 2.8: The process of the proposed model (Ramesh et al., 2018). 19

Figure 3.1: Process of proposed methodology 29

Figure 3.2: Image matrix convolution operation with multiple filter matrix (Ajaz et al., 2022) 34

Figure 3.3: Max pooling process (Ajaz ei al, 2022) 35

Figure 3.4: GoogleNet's Inception (Liu et al., 2018). 36

Figure 4.1: Sample images of healthy lettuce 42

Figure 4.2: Sample images of bacterial leaf spot diseased lettuce 42

Figure 4.3: Sample images of powdery mildew disease lettuce 42

Figure 4.4: Coding for light disturbance image processing 44

Figure 4.5: Original image 45

Figure 4.6: High brightness adjustment 45

Figure 4.7: Low brightness adjustment 45

Figure 4.8: High contrast adjustment 45

Figure 4.9: Low Contrast adjustment 46

Figure 4.10: Sharpen adjustment 46

Figure 4.11: Blurring adjustment 46

Figure 4.12: Coding for rotation disturbance image processing 47

Figure 4.13: Original image 48

Figure 4.14: Rotation in 45 degree 48

Figure 4.15: Rotation in 90 degree 48

Figure 4.16: Rotation in 135 degree 48

Figure 4.17: Rotation in 270 degree 48

Figure 4.18: Coding for mirroring image processing 49

Figure 4.19: Mirror vertical of image 49

Figure 4.20: Mirror horizontal of image 49

Figure 4.21: Coding for convolutional layer 53

Figure 4.22: Coding for Inception layer 53

Figure 4.23: Coding for fully connected layer 54

Figure 4.24: Python code for compiling the model construct 54

Figure 4.25: Load image data from directory 55

Figure 4.26: Separation for training dataset 55

Figure 4.27: Fit model to start training 56

Figure 4.28: Visualize the performance of trained model 57

Figure 4.29: Coding for real time lettuce leaf disease detection 58

Figure 4.30: Screenshot of the prototype 59

Figure 5.1: Coding for model testing 60

Figure 5.2: Accuracy in training 62

Figure 5.3: Accuracy in validation 63

Figure 5.4: Precision versus epoch 63

Figure 5.5: Recall versus epoch 64

Figure 5.6: F1 score versus epoch 64

CHAPTER 1 INTRODUCTION

1.1 Introduction

Plant disease is an abnormal state of a plant when some causal agent continuously disturbed the physiological process and causes the abnormal in structure, growth, function, or activities ("Plant disease",2021). Plant disease is classified as infectious and non-infectious, which infectious disease is disease that is caused by a pathogen organism and able to spread the disease to another plant while non-infectious disease is the disease that is caused by the growing environment ("Plant disease",2021). It becomes a threat towards agriculture since it could destroy the production of the crop.

Lettuce is a vegetable that is usually served as salad because of the taste crisp and mild. Although lettuce is a cool season crop, it can be grown in Malaysia by controlling the temperature and the environment. The varieties of lettuce that can be grown in Malaysia are loose leaf lettuce, Iceberg lettuce, Cos Romaine, Butterhead, Escarole and others. From the statistic of highest hectarage for 10 types of vegetable, and cash crop for the year 2020, lettuce is the tenth vegetable that with 3172.65 hectare of harvest area, 54,583.69 metric ton of production and produce value of RM 141,917,600 (Department of Agriculture Peninsular Malaysia, 2020). The harvest area keeps increasing from year 2017 to year 2020, that is from 2323 hectare to 3173 hectare (Department of Agriculture Peninsular Malaysia, 2020). These data indicate that the harvest area and the production of the lettuce gradually increases, it is believed that in the future it will continually increase to fulfil the demand of the Malaysians.

Growing lettuce is not difficult in Malaysia when the environment is controlled. However, the lettuce still can get diseased even if the growing environment is under control. Disease in crops could cause huge damage to the farmer's effort and their profit. The primary lettuce disease is Powdery Mildew, Downy Mildew, Bacterial Leaf Spot, and the infection of the lettuce Mosaic Virus (Baysal-Gurel & Miller, 2010). Once the lettuce is infected with infectious disease, the farmer must remove it or quarantine it from the farm to prevent it from affecting other lettuce. Farmers have limited knowledge about the plant disease, and confuse the characteristic between the disease, they just relied on their tradition or experiences to detect the disease on the crop (Mahlein, 2016). As the farmer may not be able to always check all the lettuce in the farm and did not have enough knowledge about the lettuce disease, therefore the farmer cannot take the action for the infected lettuce in time, thus causing the increase of lettuce disease infection and profit loss. The proposed solution for the farmer whether the lettuce is diseased and identify the disease of lettuce is using capturing the lettuce and detect the lettuce disease in real-time by using CNN algorithm.

1.2 Problem Statement

Plant disease has always been a problem in food production and agriculture since the day humans learn how to grow their own food (He et al., 2016). It is one of the main problems which cause the food supply shortage. Beside food supply shortage, plant disease also caused an impact to the profit of the farmer especially those small-scale farmers. The existing method for farmers to detect the plant disease is just observing the plant by human eye. The accuracy of this traditional method heavily relies on one's experience, and it is

time consuming as there is not only one plant in the farm (Mahlein, 2016). Lettuce is a type of cool season vegetable that can be grown in Malaysia with temperature control. As the weather in Malaysia is warm and moist, it provides the environment for the disease to grow on the lettuce (University of Florida, n.d.). Once a lettuce is infected with the infectious disease, it will affect other lettuce as well. In 2010, a large scale of lettuce that across several farms had been infected by a fungus disease and caused the huge loss of lettuce production in a city of California (Atallah, 2011). Not every farmer is familiar with the lettuce disease and the disease symptom characteristic on the lettuce, they may not recognize the disease on the lettuce and thus cause the loss in production and profit. Recognizing the disease symptom of the lettuce needs to have a certain level of knowledge and experience. This project will propose a deep learning algorithm that can help the farmers to recognize the diseased lettuce and could let them make decisions more accurately on the diseased lettuce.

1.3 Project Objective

The objectives of this project are:

i. To study convolutional neural network algorithm for lettuce leaf disease detection.
ii. To design a convolutional neural network algorithm for lettuce leaf disease detection to recognize the lettuce leaf disease.
iii. To develop a prototype of the lettuce leaf disease detection application to classify the leaf disease on the lettuce.

1.4 Methodology

The proposed methodology for the lettuce leaf disease detection application is using CNN algorithm. Convolution neural network (CNN) is the mainstream deep learning model used in crop disease detection because it decreases the demand on image processing and increases the accuracy of the detection (Liu et al.., 2018). Besides that, CNN in pattern recognition and classification is one of the best approaches (Liu et al., 2018). The architecture of CNN consists of convolutional, activation, pooling and fully connected layers. Where convolution, activation and pooling layers are responsible for feature extraction and fully connected layer responsible for the classification and generate the output (TM et al., 2018).

1.4.1 Proposed Model

A modified AlexNet model from Liu et al. (2018) works will be used as the CNN model for the lettuce leaf disease detection. It consists of two part, AlexNet Precursor and Cascade Inception. AlexNet Precursor is modified based on AlexNet model, which have 8 layers of it, the first convolutional layer has a smaller kernel size compared to the original AlexNet model because of the lesion of the disease could be small, the noise in the image needs to be filtered. Then is a max-pooling layer which responsible to reduce the not important parameters. Another set of convolutional layer and max-pooling layer followed. After that there are three of convolutional layers and a max-pooling layer. While the Cascade Inception, start with a max-pooling layers, followed by two Inceptions and a max-pooling layer. Inception is a special structure in the GoogleNet, which utilised the intense matrix. Then the sixth convolutional layer is added after the Cascade Inception, here has

replaced the two fully connected layer in original AlexNet model. A fully connected layer with Softmax function is then adjusted to classify two classes of result, which is healthy lettuce and diseased lettuce.

1.4.2 Image Acquisition and Pre-processing

The dataset which contains the images of healthy lettuce, and the diseased lettuce are obtained from https://github.com/chandru11235/Lettuce-plant-leaf-Datasets, https://universe.roboflow.com/ and Google Images. The image dataset consists of 3 classes, healthy lettuce, bacterial leaf spot diseased lettuce and the powdery mildew disease. Each class will contain around 50 images and above to ensure a better trained model, 70% of the images are the training dataset, 20% of images are used as the validation dataset while the other 10% of the images are the testing dataset. Before training the image dataset, images collected will be resized with the same size to ensure a more efficient training. The image in the training dataset is augmented where the image is processed to increase the dataset and to prevent the model overfitting. A few images of each class undergo image processing such as resize image, light disturbance and direction disturbance which include the rotation and mirroring of the image.

1.5 Project Scope

The CNN algorithm for lettuce leaf disease detection shall be able to detect the health condition of the lettuce leaf in 3 classes which are healthy, bacterial leaf spot disease and powdery mildew disease. There are various types of lettuce that can be grown in Malaysia, this algorithm will focus on leaf disease detection on lettuce species which are in green colour. The common leaf diseases found on lettuce are bacterial leaf spot, Mosaic Virus,

Downy mildew and powdery mildew, the application will focus on detecting bacterial leaf spot and powdery mildew on lettuce leaf. The application shall detect the leaf disease on the lettuce and show the label “Disease Name” on the screen to let farmers know which disease the lettuce is infected, and “Healthy” for the healthy lettuce. The environment for the detection should be controlled.

1.6 Significance of Project

The significance of this project is the farmer can observe the health condition of the lettuce and recognize the disease of lettuce by using trained CNN algorithm. The prototype of the lettuce leaf disease detection should be operated in real time. The farmer shall be notified when a disease is recognized on a lettuce. The recognition of the disease helps farmers to make decisions with the sick lettuce by knowing which lettuce is diseased. The loss of production and profit can be reduced if the farmer takes the action early to the infected lettuce. The accuracy of the detection model should be high to reduce the error detection.

1.7 Expected Outcome

The expected lettuce leaf disease detection application for lettuce shall work in real time and capture the image of the lettuce. The application should be able to recognize the lettuce is healthy or diseased and labelled it in real time.

CHAPTER 2 LITERATURE REVIEW

2.1 Introduction

This chapter review about the algorithm that used in plant disease detection. Both machine learning and deep learning algorithm will be reviewed to find out the best method for plant disease detection.

Plant disease detection has been worked by many techniques whether using machine learning or deep learning. The examples of the machine learning algorithm for plant disease detection are Support Vector Machine, k-Nearest Neighbours, Genetic algorithm, Decision Tree and Random Forest. Machine learning algorithm required a lot of effort on image preprocessing to ensure the smooth process in image segmentation and feature extraction, while deep learning algorithm required minimal preprocessing to avoid overfitting (Marzougui et al., 2020). Convolution Neural Network (CNN) is one of the representative deep learning algorithms as it has well performance in pattern recognition and classification (Liu et al.., 2018). As the lettuce leaf disease is bacterial leaf spot, the lesions are in brown, dark brown or black colour, the algorithm need to be able to classify the lesion on the leaf, in feature extraction, the colour and shape will be extracted from the pixels of the image and compare or learning them. There are many techniques for extraction of the colour and shape, technique for extracting out the lesion will be reviewed. The machine learning software tool will be review as it is an important tool when we need to train for the model.

The prototype of the lettuce leaf disease should be work in real time, when the camera is operation and capture any lesion on the lettuce leaf, a label that surround the

lesion with label “Diseased” should be appear, at the same time, the accuracy and precision need to be high.

2.2 Lettuce Leaf Disease

The lettuce leaf disease for detection is bacterial leaf spot and powdery mildew. For bacterial leaf spot, it is hard to detect in the early stage as it can be as small as 0.25 inch in diameter (University of California Agriculture and Natural Resources, n.d.). As shown in Figure 2.1, the lettuce which infected the bacterial leaf spot show a lot of brown spots, these lesions will then dry up and turn into black or dark brown with a papery surface (University of California Agriculture and Natural Resources, n.d.). The condition for lettuce to infect bacterial leaf spots is the same as downy mildew, which is a wet and cool condition.

Figure 2.1: Bacterial leaf spot on lettuce (Baysal-Gurel & Miller, 2010).

Powdery mildew is a fungal disease which favors the mild, dry condition, and mostly occurred in mature lettuce (University of California Agriculture and Natural Resources, n.d.). The symptom of powdery mildew is both the upper and lower side of the lettuce leaf found gray-white or powdery growth on it which is shown in Figure 2.2.