SUPERVISED MACHINE/DEEP LEARNING TECHNIQUES – A CASE STUDY OF POWDERY MILDEW DETECTION ON THE STRAWBERRY LEAF

by

Jaemyung Shin

Submitted in partial fulfilment of the requirements

for the degree of Master of Science

at

Dalhousie University

Halifax, Nova Scotia

March 2020

DEDICATION

This thesis is dedicated to my loving parents Hyungseob Shin and Aeryung Kang and to my soulmate

Thank you all for helping to give me the life I love today.

Author

Jaemyung Shin

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 3

CHAPTER 4

LIST OF FIGURES

CHAPTER 2

CHAPTER 3

CHAPTER 4

ABSTRACT

This research proposed the algorithm, that can detect powdery mildew and give the highest classification accuracy (CA). Three image processing and two machine learning algorithms (artificial neural network; ANN and support vector machine) were used to find the optimal combination for different image resolutions. Also, data augmentation by using a rotation technique was carried out to simulate the real-world field-situation. The results after data augmentation tended to be underfitting due to the added directional parameter. The results showed the highest CA of 94.34% for the combination of speeded-up robust features and ANN at 908×908 image size. To get better performance, six convolutional neural network algorithms were compared after data augmentation. ResNet-50 as the highest CA, AlexNet as the shortest computation time, and SqueezeNet-MOD2 as the smallest memory would be recommended in the conclusion.

LIST OF ABBREVIATIONS USED

AI - Artificial Intelligence

ANN - Artificial Neural Network

ANOVA – Analysis of Variance

BMP – Bitmap

CA - Classification Accuracy

CNN - Convolutional Neural Network

CPU - Central Processing Unit

CR2 - Canon Raw 2nd

CV - Computer Vision

DL - Deep Learning

DSIFT - Dense Scale-Invariant Feature Transform

DSLR – Digital Single-Lens Reflex

EOS – Electro-Optical System

FN - False Negative

FOV – Field of View

FP - False Positive

FPGA – Field Programmable Gate Array

GLCM - Gray Level Co-occurrence Matrix

GPS - -Global Positioning System

HOG - Histogram of Oriented Gradients

ILSVRC – Imagenet Large Scale Visual Recognition Challenge

IMG – Image

IOU - Intersection Over Union

ML - Machine Learning

MMC - Multiple Mean Comparisons

MSE - Mean Squared Error

MV - Machine Vision

NSERC - Natural Science and Engineering Research Council

NSGS - Nova Scotia Graduate Scholarship

PA - Precision Agriculture

PM - Powdery Mildew

RGB - Red, Green, and Blue

RTK – Real Time Kinematics

RH – Relative Humidity

SA - Spot Application

SIFT - Scale-Invariant Feature Transform

SURF – Speeded-Up Robust Features

SVM - Support Vector Machine

TN - True Negative

TP - True Positive

UAV - Unmanned Aerial Vehicle

UGV - Unmanned Ground Vehicle

ACKNOWLEDGEMENTS

First and foremost, I am extremely thankful to God Almighty and grandparents in heaven for all the blessings on me to complete a master's degree and for any good opportunities that come in my life. Especially, both my grandfather and my father have given me the greatest potential of advanced study and motivation in research as a faculty.

I sincerely would like to express appreciation to my supervisor Dr. Young Ki Chang for his enthusiasm, motivation, and support during the master's studies. His encouragement and professional help throughout the project always motivated me to broaden my perspective and learn new knowledge. He meticulously found the points that I can develop and grow more and gently gave me feedbacks. Also, I feel pleasure in expressing gratitude to my committee members: Drs. Tri Nguyen-Quang, Gordon Price, and Ahmad Al-Mallahi, for their insightful comments and attention to my project throughout the master studies. I could not have imagined doing graduate study without my supervisor, committee members, and Dr. Brandon Heung, who helps me a lot in terms of revision and deep discussion on research. I am indeed grateful to them.

I would also like to thank Dr. Jin Yue who provided me to have the opportunity to work as a Teaching Assistant in a class of Calculus 1 and 2. He always taught me the philosophy of pedagogy and motivated me to become a generous instructor. Also, Drs. Young Ki Chang and Brandon Heung helped and taught me the skills of oral and poster presentations in Communication Skills for Scientists course, which was helpful for me to win the first prize in all three presentation competitions during my master studies.

This work would not have been possible without the support of the Ralph Fraser and Joe Cooper (Balamore Farm Ltd.) and Curtis Millen (Millen Farms Ltd.). They provided the field to collect data and helped me to broaden the perspective of the field situation. Also, thanks to all funding sources that allowed me to continue two years of my master's study: Natural Science and Engineering Research Council (NSERC) of Canada Discovery Grants Program (RGPIN-2017-05815) from my supervisor, Nova Scotia Graduate Scholarship from the Government of Nova Scotia, and In-Course Scholarships (Chartwell Graduate Scholarship and Atlantic Farm Mechanization Show Graduate Scholarship) from Dalhousie University.

Lastly, I take this opportunity to record my deep sense of appreciation for my parents and family members who helped me a lot of unconditionally throughout my life. Without my family support, none of this and where I stand in life would have been possible. Special thanks to my friends: Nishchitha Natesh, Dalhousie Agricultural Association of Graduate Students executive members, and lab members: Prabahar Ravichandran, Sabiha Shahid Antora, Sandoval Ernesto, and Luke Macmillan from Biosystems Automation & Robotic Research lab for their support and encouragement during the graduate study.

CHAPTER 1: INTRODUCTION

From the late 1900s to the early 2000s, changes in agricultural practices led to various outcomes such as increasing productivity, food diversity, and reducing seasonal dependence to help the world produce more food (Kearney, 2010). However, the estimates the world population to increase by more than a billion over the next 15 years and it will reach 11.2 billion in 2100 (United Nations, 2015). Also, per capita, food intake has increased by nearly 400 kcal per person per day worldwide and climatic instability caused by global warming is a factor (Alexandratos et al., 2006; Thompson, 2016). Hence, increasing the production of food should be prepared and a balance between the growing food demand and global agricultural output is required. Plant diseases may pose a serious threat to national food security; hence, the disease management could be a clear solution to meet growing demand. For example, potato blight caused by *Phytophthora infestans* hit Europe in the 1840s and around one million people died of starvation in Ireland (Strange & Scott, 2005). The threat to food security has increased in recent years. For example, banana, a major crop for Africans, has devastated the cultivars of previous varieties due to the new species that cause Panama disease (Ploetz, 1994). Also, wheat rust fungus, which is not resistant to current varieties of wheat, barley, and grains, is spreading from Africa to the Middle East (Thompson, 2016). Pathogens and weeds cause a loss of from 20 to 40% in global agricultural productivity (Teng & Krupa, 1980; Teng & James, 2002; Savary et al., 2012). Destruction of crops due to such devastating diseases could threaten national security (Strange & Scott, 2005).

Powdery mildew (PM) can cause a decrease in strawberry yield by causing a decline in photosynthesis (Scholes et al., 1994; Carisse et al., 2013). Severe PM disease can cover the flowers with mycelium and make them die before they bear fruit. Pollen can also be polluted, reducing pollen retention and creating a poor quality of fruit (Carisse et al., 2013). Strawberries that are grown in polyethylene tunnels are particularly vulnerable to a major fungal disease such as PM because tunnels could make high humidity and warm temperature which is a good condition for infection and the disease can cause up to a 70% yield loss by affecting strawberry production all over the world (Liu, 2017).

Fungicides are the most common and effective way to control crop disease in the field, but over-spraying without the precision of detection is not good because of the environmental concerns and waste of fungicides. The decision to spray fungicides depends on producers observing the distribution and severity of the disease by scouting. This is performed to minimise the amount of fungicide use and determine the appropriate spraying timing of fungicide (Kobayashi et al., 2001). Detecting and finding out the plant disease at an early stage is the key challenge because a bigger problem can arise with the pathogens spreading. Wspanialy and Moussa (2016) indicated that field-scouting of the PM should be investigated, diagnosed, and treated as often as possible at an early stage. Due to the high cost of field-scouting, the inspection of the disease processes on a weekly or bi-weekly basis, and this low frequency of detection cannot meet the demand to find early signs of disease (Wspanialy & Moussa, 2016). Garud and Devi (2017) reported that the existing way of judging disease is normally carried out based on the producer's own experience which would not be accurate. It is also expensive to rely on a large team of experts to monitor the field continuously (Kulkarni & Patil, 2012). Hence, an attempt to integrate

existing knowledge in new ways and apply them to solve real-world problems is needed. Machine vision (MV), which is a technology commonly utilised in imaging-based automated inspection and analysis for applications excels at quantitative measurement under the structured circumstances with speed, accuracy, and repeatability (Steger et al., 2018). MV systems are not only able to recognise the size, shape, colour, and texture of an object, but also provide numerical properties of an imaged object or scene. Among the many steps in MV, machine learning (ML) can be used as one of the developments before applying to the hardware as a role of training algorithms that computers can learn how to make decisions. ML, which is a subset of artificial intelligence (AI), is an algorithm or statistical mode for scientific research to conduct specific work based on the pattern. Deep learning (DL), which is a subset of ML, is a more advanced neural network including the automatic feature extraction steps. AI helps to produce healthier crops by checking crop diseases, defects, and nutrients; hence, AI can increase the overall efficiency in agriculture production and agro-business. Therefore, the key of this research is to find the optimal algorithms in order to detect PM on the strawberry leaves by using MLs (two non-DLs and six DLs) and classify the healthy and infected leaves with high test accuracy (classification accuracy; CA) and short computation time.

CHAPTER 2: LITERATURE REVIEW

2.1. Powdery mildew (PM)

The value and production of strawberries have increased for the past five consecutive years due to the adoption of new varieties and the increase in yields (Statistics Canada, 2017). In Canada, however, strawberry losses range from 9 to 30% of crop yields due to various diseases (Carisse et al., 2013). PM disease is a fungal disease that affects various plants such as wheat, barley, grape, onions, apples, pears, and strawberries (Bushnell & Allen, 1962). Among the various crops, PM commonly happens on the strawberry leaves in Nova Scotia due to a big temperature difference in summer. In the strawberry market, detecting PM disease at the early stage is one of the crucial factors. If producers figure out the existence of PM early, they could prevent it from spreading diseases all over their field. A common fungal pathogen is *Podosphaera macularis* (formerly *Sphaerotheca macularis*) and its symptoms include white clusters of hyphae that are often present on the leaves and can infect fruits themselves in the severe stages (Paulus, 1990) (Figure 2-1).

(a) (b) (c)

Figure 2-1. Symptoms of powdery mildew disease on the strawberry leaves, (a) the individual infected leaf, (b) the bundle infected leaves, and (c) the infected fruits at the severe stage.

Zhang et al. (2012) detected the PM disease through spectroscopic measurement and analysis, which has had a great influence on the production of wheat. They extracted a total of 32 spectral characteristics and suggested that multispectral remote sensing is a cost-effective way to detect PM in wheat. Wspanialy and Moussa (2016) reported that PM is relatively easy to identify with the naked eye and is suitable for automated visual disease detection. When a susceptible host plant is born, the asphalt-coloured spores germinate and cause an infection. PM is passively dispersed by the wind and mass production of spores due to the infection occurs under high humidity and high-temperature conditions. The infected leaves are directed upward, with the underside reddish, and in the worst case, the colour of the edges appears to be burned. The highest germination rate observes at 25℃, relative humidity (RH) 99%, and 1,750 lux light intensity which is a minimal degree and starting from the lowest part of the leaves which are the most adjacent to fruit from the bundle (Jacob et al., 2008). Disease lowers the productivity of infected crops and makes infected strawberries non-marketable, resulting in yield losses of up to 60% in the United States (Nelson et al., 1995).

2.2. Fungicide and spot application (SA)

Fungicide is an important tool for managing PM disease which is a major problem in many crop productions around the world. Unfortunately, these fungicides generally have a particular mode of action, so they are at high risk for resistance and PM fungi have a high potential for resistance development (McGrath, 2001). Esau et al. (2014) studied automated prototype and variable rate in a sprayer for SA of fungicide in a wild blueberry field. Inadequate or excessive use of a fungicide is inefficient in terms of cost and has the potential to increase the risk of environmental pollution. Therefore, using spot-spray

methods with fungicides for diseases in blueberry could maximise farm profitability and minimise environmental pollution (Esau et al., 2014).

According to Statistics Canada (2020a), the number of people who were engaged in Canadian agriculture in 1991 was 390,875, however, in 2016 it decreased to 271,935. In addition, the average of farm operators was 47.5 years of age in 1991, however, it reached 55 years of age in 2016. Therefore, accurate and fast methods should be applied for the detection of PM disease for the preparation of agricultural workforce declining and aging phenomenon in the situation of increasing the world population.

Xue et al. (2012) developed an MV application that allows farm robots to navigate rows with an average error of 1 mm for the distance field of view (FOV) guidance method in the cornfield. They showed that field scouting was used as a justification for robotic research in the field and the developed method was able to be operated in cornfields without damaging the crops. Cubero et al. (2011) showed that MV can detect defects or features that are invisible to the human when grading, quality estimation, and monitoring of fruit processes when they are stored. Blasco et al. (2003) evaluated the efficiency of MV developed for the on-line assessment of fruit quality. The quality of the fruit was assessed through a classification used based on size, colour, and location of the stem. When collecting fruits in batches, they achieved 86% and 93% of repeatability in blemish detection and size estimation, respectively (Blasco et al., 2003).

2.3. Image processing

Images are a versatile tool for both qualitative and quantitative information because they have dual attributes such as visually analyzable entities with shape elements or large

datasets that can be generated (Vidal & Amigo, 2012). Computer vision (CV) is derived from computer science that deals with how computers can gain a high level of understanding through images. MV is the use of the CV in an industrial or practical application that traditionally requires the execution of a specific function or result based on image analysis processed by a vision system (Jain et al., 1995). The components of a basic CV and MV system are typical: imaging devices, camera sensors and lenses, lighting for specific applications, computers, and image processing in software (Jain et al., 1995).

Image processing is an interesting and dynamic part of CV by visually representing the theoretical parts that appeared as writing. Image processing is a way to perform an image modification to extract useful information. This is the processing in which the input is an image and the output can be an image or property associated with that image. Steward and Tian (1999) reported that images taken in sunny areas were more susceptible to pixel noise and more variability at the edges when segmenting. Light and other environmental parameters could be the hamper factors in image processing; hence, these parameters should be generalised through the image processing. The intensity of light, size, and colour of the images could affect the classification accuracy (CA) (Li et al., 2010; Rehman et al., 2019). Patil and Kumar (2011) reported that image processing is helpful when extracting the representative features in order to train and test the algorithms to get the CA in disease detection.

2.3.1. Histogram of oriented gradients (HOG)

HOG has gained momentum in human recognition by condensing huge features into nine bins of the histogram (Dalal & Triggs, 2005). HOG descriptors have been studied to understand the importance of fine-scale gradients, fine orientation binning, relatively

coarse spatial binning, and high-quality local contrast normalisation in overlapping descriptor blocks. HOG feature extraction is usually used with a linear SVM for robust object detection and applied it to human detection. It has considerably better performance than existing methodologies for human detection (Dalal & Triggs, 2005). Zhu et al. (2006) studied a fast and accurate human detection system by combining cascade of rejectors method and HOG and Adaboost to select the representative features and they selected the appropriate set of blocks under the good conditions. In terms of the agriculture field, in order to increase the productivity of soybean crops, soybean disease detection was achieved with an image of a local descriptor and bag of visual words. Pires et al. (2016) got over 98% of CA in the detection of soybean disease and they demonstrated that their approach could be easily applied to other crops.

2.3.2. Speeded-up robust features (SURF)

Bay et al. (2006) suggested a new concept called SURF that generates a point of interest and descriptor that is robust for variation and rotation of images. SURF has a superior performance than scale-invariant feature transform (SIFT), which is the previous method introduced by Lowe (1999) with respect to the blurred images, different hue images, warp images, and red, green, and blue (RGB) images. Lowe (1999) developed SIFT as an object recognition system that uses local image features that are invariant to the deformation. The total time extraction for feature detection and matching in SURF showed three times faster than the times of SIFT (Mistry & Banerjee, 2017). SURF is based on attributes like SIFT and has less complexity than ever before (Bay et al., 2006). Pires et al. (2016) indicated that four local descriptors (SURF, HOG, DSIFT, and SIFT) were used for soybean disease recognition. The reason for using local descriptors is that they are

distinctive and robust to occlusion which usually occurs when detecting features of the leaf in the real field.

2.3.3. Gray level co-occurrence matrix (GLCM)

GLCM, also called the gray level space-dependence matrix is a statistical method of texture inspection that takes into account the spatial relationship of pixels. The way to obtain GLCM is to calculate the frequency at which a pair of pixels are lying in a relationship between a particular value and a specific space. After that, GLCM creates features of the image texture by using the statistical measurements (Sebastian et al., 2012). Yao et al. (2009) extracted texture features to classify rice diseases by using GLCM among colour, shape, and texture features in image processing. GLCM has proved to be an effective method to extract textural features from an image by using statistical variables (Mohanaiah et al., 2013). Using statistical variables such as the co-occurrence matrix gives mathematically valuable information to process a lot of images in an intensive and compact way (Eleyan & Demirel, 2011). Depending on the number of pixels, the statistical variables are classified into first-order, second-order, and higher-order. The GLCM extracts secondary statistical texture features (Albregtsen, 2008). There are six statistical variables: energy, contrast, variance, correlation, entropy, and inverse difference moment (Baraldi & Parmiggiani, 1995). In this study, four statistical variables will be used. Each variable represents the relationship between different pixels in the image, and each variable can be extracted using MATLAB R2019a software (The MathWorks Inc., Natick, MA, U.S.A.).

2.4. Machine learning (ML)

ML is a technique that builds a system that learns, predicts and improves its own performance based on empirical data and algorithms for realizing functions such as human learning ability on a computer (Michie et al., 1994). Iterative aspects are important since models can adapt independently when existing data are exposed to the new data (Russel & Norvig, 2016). Many technologies have automatically identified and developed rules and relationships to simplify the knowledge required from massive and error-prone empirical data. These techniques are assessed for completeness based on their ability to understand the new data that are totally separated from training data (McQueen et al., 1995). There are two types of ML: supervised ML and unsupervised ML. In this study, supervised ML which requires labeled data will be carried out.

2.4.1. Supervised ML (non-deep learning; DL)

Supervised ML is a learning method that predicts values using given data and correct answers. It can be classified into two subgroups: regression and classification. Regression mainly deals with continuous value; however, classification primarily deals with assigning discrete categories rather than sequential quantities. The binary classification which will be mainly covered in this research corresponds to a simple classification. However, the algorithm is important to perform binary classification (Kotsiantis et al., 2007). There are many kinds of supervised ML. SVM is a statistical learning theory developed in the 1990s and is one of the fields of supervised ML. It is a model for pattern recognition and data analysis by being mainly used as a classification and regression analysis (Scholkopf & Smola, 2001). Joachims (1998) used SVM to learn text classifiers and the experimental results demonstrated that the SVM method performs

consistently better in text classification because SVM could do parameter tuning automatically. Ahmed et al. (2012) studied how to effectively classify crops and weeds in digital images using SVM algorithms. They have achieved above 97% accuracy of classification over a set of 224 test images. Artificial neural network (ANN) is inspired by biological neural networks and many researchers are designing ANN to solve various problems related to pattern recognition, prediction, and optimisation (Jain et al., 1996). Huang (2007) used ANN to detect and classify Phalaenopsis seedling diseases and GLCM has been added for texture extraction of lesion sections. The accuracy was 89.6% by successfully detecting and classifying the Phalaenopsis seedling diseases. Jhuria et al. (2013) monitored not only the time of harvest but also the plant during its growth period and detected disease for fruits by using image processing as a tool. Each disease was categorised by colour, texture, and morphology, and 90% accuracy was achieved based on MATLAB software (Jhuria et al., 2013).

2.4.2. Unsupervised ML (non-DL)

Unsupervised ML is different from supervised ML because there is only a predictor set but no corresponding responses. Unsupervised ML might not be directly applied to a regression or classification problem because it does not know what output data are and retains a pattern in a dataset without referring to known or labeled results. Clustering is the most prominent example in the unsupervised ML. With clustering, the dataset could be separated into groups based on similarity. Unsupervised ML might be useful for later supervised ML implementations. Behmann et al. (2015) studied biological stress detection due to the crop protection and they used k-means clustering and self-organizing maps for clustering. The k-means clustering algorithm is an unsupervised ML based on the

distribution of data points within the feature space to structure the data. Hu et al. (2014) studied how to quantify the damage and aging of the banana surface by using k-means clustering algorithms. The result of the experiments showed that this algorithm was suitable for defect extraction of bananas hand/bundle and showed the potential of this algorithm for segmentation. The k-means clustering can also be used in supervised ML depending on whether the input data are labeled.

2.5. Deep learning (DL)

DL has been renamed by deriving from ANN. The multi-layer perceptron used in ANN has limitations. The first limitation was the vanishing gradient problem in a process of backpropagation by losing data and the second limitation was that the algorithm is difficult to process with the new data which were not shown in the training stage. However, LeCun et al. (2015) solved two problems by introducing pretraining and by developing a dropout technique, respectively. Hence, DL is an ANN that exceeds the limitations of the existing ANN.

2.5.1. AlexNet

AlexNet is an algorithm that won the first prize in the Imagenet large scale visual recognition challenge (ILSVRC) in 2012 on image classification tasks and one of the popular convolutional neural networks (CNNs) , already trained with a million images (Krizhevsky & Sutskever, 2012; LeCun, 2015). AlexNet consisted of the first five convolutional layers, some of them followed by maxpooling layers, and three fully connected layers at the end as a totaling of eight layers (Krizhevsky & Sutskever, 2012). The convolutional layer/filter stands for each network as a form of a slice through image

one by one in a process of the whole architecture and convolutional networks by consisting of multiple filters (i.e., kernels) (O'Shea & Nash, 2015). Compared to the previous CNN architectures used for CV tasks, AlexNet was much larger by carrying 60 million parameters and 650,000 neurons when considering it was developed six years ago. Durmuş et al. (2017) used AlexNet and SqueezeNet to classify using ten different tomato leaf classes (including healthy tomato leaf). They got an accuracy of 95.65% and 94.30%, respectively, on AlexNet and SqueezeNet. The size of the model of SqueezeNet is over 100 times smaller than AlexNet, so they concluded that SqueezeNet is a good candidate for mobile application. Mohanty et al. (2016) used AlexNet and GoogLeNet to detect 14 crop species and 26 diseases and they got the highest accuracy with 99.35% by using GoogLeNet with 80% of the data for training and 20% for the testing combination.

2.5.2. SqueezeNet

Even though many algorithms were developed, large memory requirements in deployment are the shortcoming to the use of such models in real-time applications. SqueezeNet was developed to reach the needs of smaller neural networks with fewer parameters in 2016 while showing still competitive accuracy compared to AlexNet. SqueezeNet has 50 times fewer parameters and is 510 times smaller than AlexNet (Iandola et al., 2016). However, it does not mean that the processing time of SqueezeNet would be faster than AlexNet. Bianco et al. (2018) compared inference time depending on different batch sizes in terms of many deep CNN including AlexNet, SqueezeNet, GoogLeNet, ResNet-50. For the 64 batch sizes used in our study, the sequences of the short computation time were AlexNet, SqueezeNet, GoogLeNet, and ResNet-50. The benefit of SqueezeNet

is small model size not the computational complexity. SqueezeNet has more convolutional layers where the most of computation costs weight on than AlexNet.

2.5.3. GoogLeNet

GoogLeNet, known as Inception, is an algorithm that won the first prize in ILSVRC in 2014 on object detection. The main characteristic is that the network utilisation has been fully improved by increasing the depth and width of the network. It is much deeper and wider architecture than AlexNet by having 22 layers deep CNN (Szegedy et al., 2015); however, the number of parameters is reduced from 60 million (AlexNet) to four million. This algorithm is close to human-level performance with 6.67% (Top-5 error rate). Ramcharan et al. (2017) used CNN architectures to detect cassava disease and concluded that even though the performance of Inception V3 is higher than the original Inception due to the deeper layers (42 layers), the former has a limitation where memory and computational cost is limited. Hence, Inception is recommended to apply for mobile or drone devices (Ramcharan et al., 2017). Brahimi et al. (2017) used AlexNet and GoogLeNet to solve a task of classification with nine types of tomato leaf diseases. They found that the results of GoogLeNet as the highest accuracy of 99.18% are more accurate than those of AlexNet with the highest accuracy of 98.66%.

2.5.4. ResNet-50

ResNet is an algorithm that won the first prize in 2015 on object detection and the main innovative point is this algorithm can successfully train the model with more than 150 layers. This algorithm achieves 3.57% (Top-5 error rate) which beats human-level performance. Among various types of ResNet, ResNet-50 was selected as the last CNN

architecture and it has 50 deep layers and can classify images into 1000 object categories. Fuentes et al. (2017) compared the results by using ResNet-50 and RexNeXt-50, respectively and they got 75.37% from the former and 71.1% from the later. They found out that the networks showed better performance when the networks got deeper.

2.5.5. Overview of DL in agriculture

Modern technology has given us the ability to produce enough food to meet the needs of over seven billion people (Mohanty et al., 2016). CV and object recognition technologies have made tremendous progress over the last few years and the PASCAL VOC Challenge based on the Imagenet dataset and the recent development of the ILSVRC have reduced the error rate of CNNs to 3.57% for the classification of images into 1000 possible categories (Mohanty et al., 2016). ILSVRC is a benchmark that has been running annually since 2010 for object category classification and detection for hundreds of object categories and millions of images (Russakovsky et al., 2015). Mohanty et al. (2016) got 99.35% accuracy in the classification model that distinguishes 26 diseases out of 14 types of crops using 54,306 using the CNN approach. They suggested in-depth learning model training with increasingly larger, publicly available image data present a clear path to diagnosing crop disease with smartphones. Kamilaris and Prenafeta-Boldu (2018) said that DL is the latest technology to process images and analyse data, delivering promising results and great possibilities. In recent years, DL is one of the latest technologies that help in solving some of the food production challenges. Flexible DL due to its highly hierarchical structure and large learning ability can be extended by adding depth and complexity to ML as well as various functions that can express data in a hierarchical manner through various levels of abstraction to convert the data (LeCun & Bengio, 1995; Schmidhuber, 2015). One

great advantage of DL is that it simplifies the steps of image processing. DL does not require feature extraction steps because it can find important features by itself through training (Kamilaris & Prenafeta-Boldu, 2018). The disadvantage of DL is that the training time takes longer, but the testing time is faster than ML (non-DL) methods (Chen et al., 2014). Durmuş et al. (2017) used AlexNet and SqueezeNet to train how to detect various diseases of tomato plant leaves and AlexNet performance was slightly better than SqueezeNet. Brahimi et al. (2017) carried out the CNNs to detect tomato disease with 14,828 datasets including healthy and nine types of diseased leaves. As a result, they got 99.18% accuracy in the detection of tomato diseases. Atole and Park (2018) used an AlexNet to determine the health status of rice plants based on the image of the leaves and achieved 91.23% accuracy with 600 images. Recently, scene classification in remote sensing has been a subject of considerable research; hence, Alhichri et al. (2018) developed for remote sensing by running SqueezeNet with three different image sizes and they used multi-scale CNNs that can accommodate a softmax layer for classification. They suggested a table of CA results for multi-scale classification versus single-scale classification and three different trainable layers.

2.6. Research statement

Many studies to date have been conducted to detect diseases of crops using ML including DL. However, most studies showed a tendency to carry out up to two ML (non-DL) techniques to derive the results by following the best combination, which was suggested from other studies, but only enforcing the best combination suggested might not make the researchers find promising methodologies and the best combination when detecting the diseases of crops. The goal of this research is to suggest the optimised

combination of three image processing techniques and two supervised ML (non-DL), comparing five different image sizes. When evaluating the supervised ML (non-DL), the highest CA would be prioritised; however, analysis of the results from different angles will also be explained in CHAPTER 5. Furthermore, we will compare six DL algorithms with respect to test accuracy/CA and computation time. For plentiful evaluation, the combination that showed the highest CA in supervised ML (non-DL) will be compared to the DL results.

2.7. Objectives

The objectives of this research are:

1. Evaluation of the performance of the two supervised ML algorithms using ANN and SVM, to detect powdery mildew using five image resolutions with a combination of image processing,

2. Analysis of data augmentation effects in supervised ML algorithms using ANN and SVM, considering various leaf orientation in the real field, and

3. Comparison of six CNN algorithms for the detection of powdery mildew disease on strawberry leaves with augmented data.

CHAPTER 3: EFFECTS OF DIRECTIONAL AUGMENTATION WITH SUPERVISED MACHINE LEARNING TECHNIQUES – A CASE STUDY OF POWDERY MILDEW DETECTION ON THE STRAWBERRY LEAF

The study extracts representative features to train a model with supervised machine learning (ML) to detect powdery mildew (*Sphaerotheca macularis f. sp. fragariae*) on the strawberry leaves. Powdery mildew (PM) is a fungal disease that greatly affects the production of strawberry and usually infects under conditions of warming temperatures and high humidity. In this research, we report robust models to detect PM using image processing and ML technologies. Three feature extraction techniques (histogram of oriented gradients; HOG, speeded-up robust features; SURF, and gray level co-occurrence matrix; GLCM) and two supervised ML (artificial neural network; ANN and support vector machine; SVM) were implemented using MATLAB. Images were augmented to 1016 images using a four different angle rotation technique to simulate strawberry leaf bundles in the real field. The classification accuracy (CA) to detect PM was highest at 94.34% with a combination of ANN and SURF with 908×908 image resolution and with SVM and GLCM at 88.98% with 908×908 image resolution. In terms of the extraction time for real-time processing, HOG takes the shortest time to extract features in both ANN and SVM.

3.1. Introduction

Strawberries have been cultivated in North America since 1835 and have become economically important fruit crops. In Canada, strawberries are grown in all provinces and in Nova Scotia, they were the fourth most productive crop (of 13 crops), producing 2,045 t in 2019 (Statistics Canada, 2020b). Various diseases such as bacterial and fungal pathogens have reduced annual strawberry yields by between 9 and 30% (Carisse et al., 2013; Pan et al., 2014). Fungal diseases have the potential to reduce strawberry yields in the field and also infect adjacent healthy strawberries in packaging, thereby reducing their overall quality during transportation and marketing (Kovach et al., 2000).

Among fungal diseases, powdery mildew (PM) in Nova Scotia is enhanced by large temperature fluctuations and high humidity during mid to late summer. Early-onset indications of PM are typically described as when ten white spots are observed on leaves; however, mass production of spores can occur between 4 to7 days after early-onset due to the transport of spores by the wind (Adam & Somerville, 1996). Upon infection, the leaf edge is directed inwards, and white powdery colonies become visible on both the upper surface and underside of the leaves (Jacob et al., 2008).

To manage PM, producers depend on uniformly spraying the fields with fungicides, even though PM exhibits uneven spatial distribution (Bah et al., 2018). However, this method of control is not sustainable since PM fungi have a high potential for developing resistance to fungicides (McGrath, 2001). Moreover, the fungal disease is easily misdiagnosed; hence, incorrect spraying of fungicides may increase the operational costs to farms and lead to negative impacts on the natural environment (Mueller, 2006). Another solution to treat PM is to hire experts, or disease specialists, to manually monitor and

identify diseases, as well as areas of concentration, for targeted application of fungicides (Perini & Susi, 2004). These conventional methods are laborious and time-consuming and can lead to inaccurate predictions over the entire extent of a field (Kobayashi et al., 2001). Due to the limitations of existing methods to control PM on strawberries, a more efficient and accurate approach is required to maintain production and reduce chemical use.

Conventional disease detection methods in agriculture, such as physiological and biological laboratory-based methods are complicated, time-consuming, and invasive (Khaled et al., 2018). However, the rapid development of advanced agricultural technologies provides a unique opportunity to develop non-destructive approaches, using a combination of image processing and cutting-edge, computing technologies (e.g., machine learning; ML) for detecting plant diseases.

Interest in image processing techniques exploded in the 1980s and 1990s, and active applications in remote sensing, technical diagnostics, and autonomous vehicle guidance are currently available (Sonka et al., 2014). Image processing techniques are used to derive a group of image attributes such as shape, colour, and texture that can be put into a detection algorithm using ML. Image attributes such as shape, spatial information (Vidal & Amigo, 2012; Cheng et al., 2014), colour, and morphology (Bhange & Hingoliwala, 2015) may be extracted to detect the severity of diseases using image analysis techniques. Different combinations of feature extraction techniques and supervised ML have been described in the literature (Dalal & Triggs, 2005; Bay et al., 2006; Yadav et al., 2013). Popular techniques for the determination of representative features are histogram of oriented gradients (HOG), speeded-up robust features (SURF), and gray level co-occurrence matrix (GLCM).

The HOG approach was initially invented to recognise humans (Dalal & Triggs, 2005) and the HOG descriptors can be used to isolate fine-scale gradients, bin information, and high-quality local contrast normalisation in overlapping descriptor blocks. In agriculture, HOG has been used to classify healthy and infected leaves with three different types of plant leaves (cabbage, citrus, and sorghum) using colour-based features such as pixels and statistical features (Rahman et al., 2017). They utilised the artificial neural network (ANN), support vector machine (SVM), and random forest ML algorithms after extracting the features by using HOG and have achieved the highest F1-score (>95%) in damaged sorghum data in a combination with SVM. The SURF algorithm generates a point of interest and descriptors, which are robust for variation and rotation of images. The reason for using descriptors is that SURF algorithm is distinctive and robust to occlusion, which is a possible problem when detecting features of the leaf on the spot. Awate et al. (2015) developed a fruit disease detection system for grapes, apples, and pomegranates using SURF and detected healthy and high-quality fruit with a classification accuracy (CA) of 90%. The GLCM is calculated using a statistical method of texture inspection that considers the spatial relationship of pixels and has proved to be an effective method to extract textural features from the images (Mohanaiah et al., 2013). Using a statistical approach, such as the co-occurrence matrix, gives mathematically valuable information when processing many images (Eleyan & Demirel, 2011). The preferred set of statistics (e.g., contrast, correlation, and entropy) has been shown to improve analytical outcomes (Howarth & Rüger, 2004). Gavhale et al. (2014) extracted the features using a GLCM to detect disease of citrus fruits with a CA of 96%.

Within the literature that examines the use of ML techniques for disease detection in agricultural crops, ANN and SVM have been commonly used. For example, Jhuria et al. (2013) used ANN to detect black rot and PM of grape leaves, as well as scab and rot of apples with CA of up to 96%. Similarly, for SVM, Yao et al. (2009) achieved a 97.2% CA when detecting bacterial leaf blight, sheath blight, and blast on rice crops while Pujari et al. (2015) were able to detect symptoms of fungal diseases on various agricultural and horticultural crops. As indicated previously, Rahman et al. (2017) tested both ANN and SVM for detecting diseases on plant leaves.

Despite the increasing literature surrounding the use of image processing and ML technologies for disease detection, there is still limited information on their application for PM detection on strawberry leaves. Hence, the objectives of this research are (1) to extract the representative features from the strawberry leaf images by using three feature extraction techniques (HOG, SURF, and GLCM), (2) to compare and assess two supervised ML (ANN and SVM) among five sizes, and (3) to find the best combination based on the test accuracy and processing time with four directional rotation augmentation to accommodate real field situations.

3.2. Materials and methodology

3.2.1. Experimental method

Strawberry leaves were acquired from fields located in central Nova Scotia, Canada (45°23'57.6"N, 63°33'31.1"W) during the summer of 2018. This time period was chosen as it was an ideal time to find PM infections; the average difference in daily temperature was 18.5°C and humidity generally exceeded 90% in the locations where the leaves were

acquired (The Weather Network, 2020). After collection, the leaves were stored in an icebox and images were taken within 30 minutes of acquiring the leaves. A digital single-lens reflex (DSLR) camera model: EOS 1300D (Canon Inc., Tokyo, Japan), with 3456×5184 pixels (Raw CR2 format), a 46 mm focal length, and 1/15 exposure time, was used to acquire leaf images. Leaf infections were identified with the assistance of disease experts and followed the criteria described in the literature (Adam & Somerville, 1996). The total number of images, including both upper surface and underside of the leaves, consisted of 254 images, which included 134 healthy leaves and 120 infected leaves. It was necessary to take images from both sides of the leaves because PM sometimes appears on the underside of the leaves; furthermore, the leaf edge is directed inwards as the disease progresses.

In computer vision, data augmentation is widely used to improve model accuracy and robustness in data-limited situations (Bloice et al., 2017). Here, the dataset was increased using a data augmentation method by applying an angular rotation technique. To that end, the number of the images (n=254) was increased to make the augmented image dataset (n=1016) by rotating to 90°, 180°, and 270° angles to consider the leaf positions facing various directions in the field (Mohanty et al., 2016; Sladojevic et al., 2016). Figure 3-1 shows the detailed procedure of data augmentation process using four different angles. Figure 3-1 (a) shows the original image acquired, totalling 254 images, Figure 3-1 (b) shows the angle rotation procedure, and Figure 3-1 (c) is the combined dataset after augmentation, totalling 1016 images.

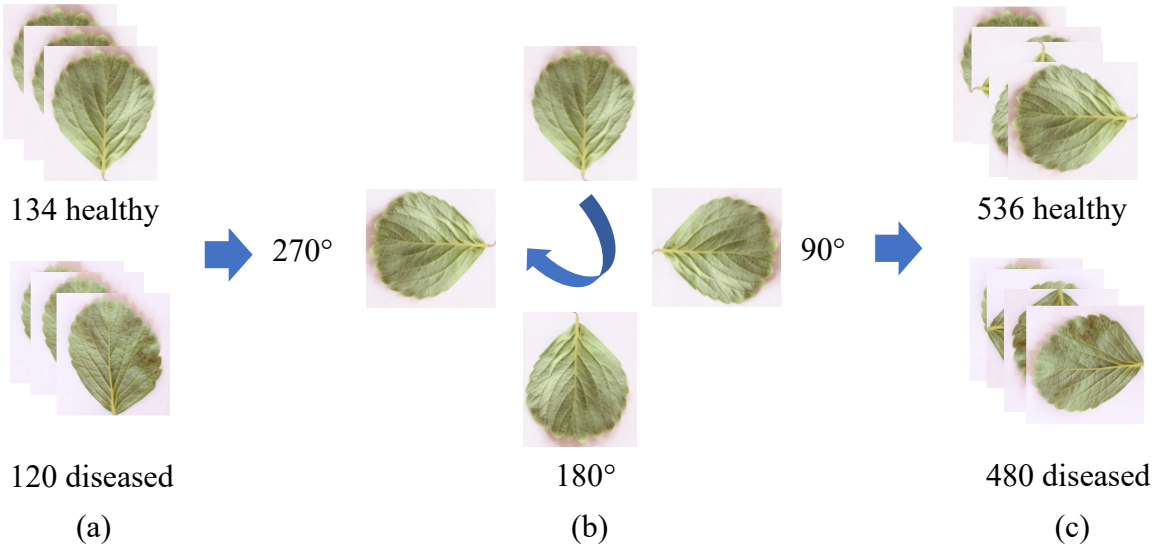

Figure 3-1. The four different angles used to make augmented dataset, (a) the original image dataset (*n*=254), (b) the process of four different angular images, and (c) the final augmented dataset (*n*=1016).

In this study, MATLAB R2019a software (The MathWorks Inc., Natick, MA, U.S.A.) was used to run the models, resize images, and classify the leaves as being healthy or infected. An Intel® Core™ i7-8700 CPU @ 3.20GHz with 48.0 GB RAM and 64-bit Windows 10 operating system was used for data processing. The original image dimension of 3456×5184 pixels was cropped to 3000×3000 pixels, with the leaf in the centre, and resized using bicubic interpolation. The cropped images were downscaled to five different image resolutions: 227×227, 454×454, 681×681, 908×908, and 1135×1135 pixels. Each image resolution was considered as a parameter within the modelling process and the respective CA produced at each resolution was compared. The 227×227 pixels size is the commonly used image dimension when applied to similar studies using that integrate image processing with ML and/or deep learning (DL) (Becherer, 2017). Also, when the size of

the images was reduced by pixel aggregation, antialiasing was carried out to prevent artefacts caused by distortion. Table 3-1 shows all the combinations that were compared.

Table 3-1. Combinations of machine learning, feature extraction, image resolution, and cell size.

ML[a]	Feature extraction	Image resolution	Cell size[b]	Total CA[c] cases
ANN SVM	HOG SURF GLCM	227×227 454×454 681×681 908×908 1135×1135	32×32 64×64 128×128	50

[a] Machine learning
[b] Cell size parameters are only for a HOG feature extraction
[c] Classification accuracy

Figure 3-2 shows the methodological framework of the study. Firstly, the imagery was acquired in the raw CR2 format then converted into 24-bit red, green, and blue (RGB) in the BMP format. Next, HOG, SURF, and GLCM feature extraction were performed by combining an image converted from 24-bit BMP image to an 8-bit grey image and an image inverted from mask to binary was used to conduct image segmentation (Figure 3-3). The outputs from the image processing algorithms provided the feature information used as predictors to help train the ML models. In addition, image processing techniques consisted of several steps, ranging from the importing, conversion, resizing, analysis, and the extraction of output features of the images.

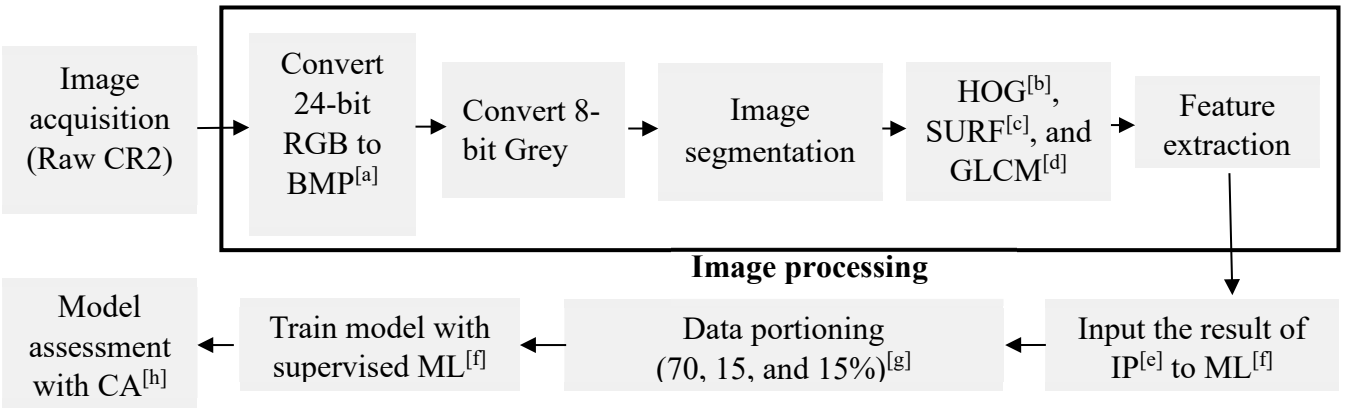

Figure 3-2. Methodological framework of the proposed method.

[a] The basic metabolic panel
[b] Histogram of oriented gradients
[c] Speeded-up robust features
[d] Gray level co-occurrence matrix
[e] Image processing
[f] Machine learning
[g] 70% for training, 15% for validation, and 15% for testing
[h] Classification accuracy

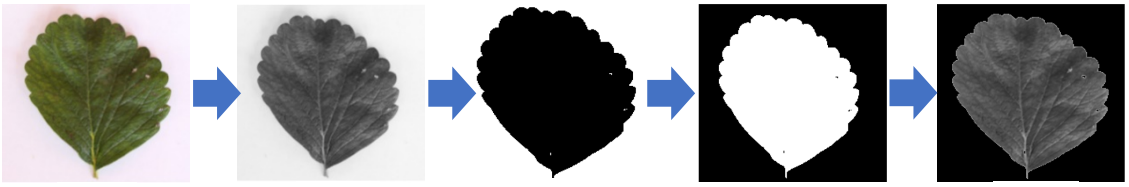

Figure 3-3. Image processing of conversion from RGB to 2-D grayscale image. (a) original strawberry leaf image infected by powdery mildew, (b) grayscale intensity image, (c) image converted into a binary image, (d) image with converted background and leaf pixels, (e) combining (b) and (d) images.

3.2.2. Image processing techniques

3.2.2.1. Histogram of oriented gradients (HOG)

HOG can extract its descriptors at various cell sizes. Here, three different cell sizes with five different image resolutions produced a total of 15 comparisons (Table 3-2). The total number of features corresponded to the number of features (Table 3-2) multiplied with the total number of images (n=1016). The detector window was processed with a grid of blocks when the HOG feature vectors were extracted (Figure 3-4). In Figure 3-4, the white arrows indicate RGB patches, and represent the direction of the gradients, and the length of the arrows indicates the size of the gradients. The advantage of HOG is that it captures the gradient information, which is specific to the shape of the leaf; hence, it is not sensitive to the orientation of the leaf. The extracted features were used to train, validate, and test the ANN and SVM learners. The number of features increased with increasing resolution (i.e., decreasing cell size).

Table 3-2. Number of features based on three feature extraction techniques.

Feature extraction	Image resolution	Cell size	Number of features[a]
HOG	227×227	32×32	1296
		64×64	144
		128×128	9
SURF		-	2,408,448
GLCM		-	4
HOG	454×454	32×32	6084
		64×64	1296
		128×128	144
SURF		-	9,980,928
GLCM		-	4
HOG	681×681	32×32	14,400
		64×64	2916
		128×128	576
SURF		-	22,195,200
GLCM		-	4
HOG	908×908	32×32	26,244
		64×64	6084
		128×128	1296
SURF		-	39,923,712
GLCM		-	4
HOG	1135×1135	32×32	41,616
		64×64	9216
		128×128	1764
SURF		-	61,943,808
GLCM		-	4

[a]The number of features per image.

Figure 3-4. Features extracted with (a) 32×32 cell sizes, (b) 64×64 cell sizes, (c) 128×128 cell sizes. The arrows indicate the direction of the gradient and the length of the arrow indicates the value of the histogram in each direction. The above images were randomly selected from the disease folder.

3.2.2.2. Speeded-up robust features (SURF)

Bag-of-features, which are visual feature descriptors in SURF, was created by assigning healthy leaves a value of 0 and infected leaves a value of 1. The grid method, which is a rectangular subset of the given image, was used when selecting feature point locations and the SURF features (Bay et al., 2008). Table 3-2 shows the number of features that were extracted using SURF at five image resolutions. The optimising the iteration was carried out to maximise within-bag homogeneity and between-bag heterogeneity. The maximum iteration was set as 100 and the optimised number of iterations among 100 was carried out (e.g., 30/100 means 30 times were iterated).

3.2.2.3. Gray level co-occurrence matrix (GLCM)

In GLCM, the image parameters represent the relationship between different pixels within the image using an image that was converted from the original RGB colour space to a grayscale intensity image (Figure 3-3). In this study, four image variables were used: contrast, correlation, energy, and homogeneity. Contrast measures the difference between the highest and the lowest values of a contiguous set of pixels; correlation measures the grey tone linear dependencies of the image; energy measures the textural uniformity that is pixel pair repetitions; and homogeneity measures the larger values for smaller grey tone differences in paired elements (Baraldi & Parmiggiani, 1995; Howarth & Rüger, 2004). In Table 3-2, shows the number of features that were extracted using GLCM at five image resolutions.

3.2.3. Supervised machine learning

3.2.3.1. Artificial neural network (ANN)

Figure 3-5 shows the ANN training process of the extracted features from 908×908 image resolution with 32×32 cell size based on 100 hidden neurons, which is the basic unit of computation in a neural network. The number of hidden neurons was optimised by comparing the accuracies when 10, 20, 50, and 100 neurons were used. Among them, 100 hidden neurons showed the lowest error and was therefore selected as the best model. Input features (26,244) were derived from the 32×32 cell size at 908×908 image resolution (Table 3-2 & Figure 3-5). It was calculated by MATLAB about how many cell sizes are extracted according to the image resolution and a total of nine bins are formulated for each block, and the calculated value is the input value. Outputs consist of a binary classification with values of 0 representing healthy leaves and values of 1 representing infected leaves. In Figure 3-5, *w* and *b* represent vectors of weight and bias, respectively. This network was assigned *w*=100 and *b*=100 in the first layer, and *w*=100 and *b*=1 in the second layer with 100 hidden neurons. The standard network used in this study was a two-layer feedforward network, which is the simplest type of ANN, with a sigmoid transfer function in the hidden layer and a softmax transfer function in the output layer (Figure 3-5). The sigmoid is an activation function and usually attached after the hidden node, while the softmax is attached after the output node (Karlik & Olgac, 2011, Jang et al., 2016). The softmax is mostly used for multi-classification; however, if two classes wanted to be classified then the softmax can be used for binary classification (Liao et al., 2015).

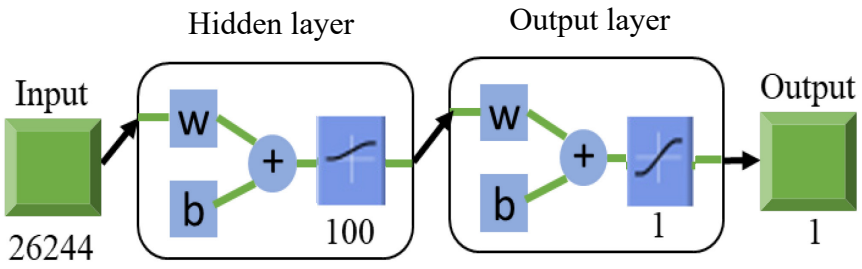

Figure 3-5. The procedure for neural network training.

3.2.3.2. Support vector machine (SVM)

In other applications, SVM learners have been shown to be computationally demanding when attempting to tune the learner's hyperparameters using large datasets (Heung et al., 2016); however, with the limited size of the training data and number of classes, computational limitations were not a concern. This study tested SVM with a linear kernel to delineate the decision boundaries between the two classes (i.e., healthy or infected leaves). The main hyperparameter of SVM is the linear kernel by generating various groups repeatedly and extracting descriptors into *k* clusters.

3.2.4. Training models

3.2.4.1. Training, validation, and testing set

For each model, the full dataset was randomly partitioned into three sets: training (70%), validation (15%), and testing set (15%) where the training data was used to calibrate each model; the validation data was used to optimise the hyperparameters of each model; the testing data was used to calculate the CA (Fira & Goras, 2008; Yadav et al., 2013; Aneece & Thenkabail, 2018). The process of training, validation, and testing the models

was repeated five times and the overall accuracy metrics were reported as the averaged accuracy values of the repetitions. In this study, with a total of 254 images, 178 images were used for training the model and 38 images for validation and 38 images for the testing, respectively. After data augmentation, 712, 152, and 152 images out of a total of 1016 images were used for training the model, validation, and testing, respectively.

3.2.5. Statistical analysis

3.2.5.1. Factorial analysis of HOG

To determine the combination of image resolution and cell size that resulted in the highest CA, the two input factors were set to image resolution and cell size, respectively, and the CA was considered as the response (Eq. 3-1).

The model for factorial analysis is represented as follows:

$$Y_{ijk} = \mu+\alpha_i+\beta_j+ (\alpha\beta)_{ij}+ \varepsilon_{ijk} \quad \text{(Eq. 3-1)}$$

Where;

Y_{ijk} = Classification accuracy (%),

μ = Overall mean,

α_i = Effect of "Image resolution" on response at i_{th} level,

β_j = Effect of "Cell size" on response at j_{th} level,

$\alpha\beta_{ij}$ = Interaction between "Image resolution" and "Cell size",

ε_{ijk} = The error terms (uncontrollable & uncontrolled factors),

i = 1, 2… a; j = 1, 2… b; k = 1, 2… n.

In this study, a = 5 (number of image resolution), b = 3 (number of cell size), and n = 5 (replications).

3.2.5.2. Factorial analysis of SURF and GLCM

To determine which image resolution resulted in the highest CA, the image resolution was considered as an input factor, and the CA was considered as the response (Eq. 3-2).

The model for factorial analysis is represented as follows:

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \quad \text{(Eq. 3-2)}$$

Where;

Y_{ij} = Classification accuracy (%),

μ = Overall mean,

α_i = Effect of "Image resolution" on response at i_{th} level,

ε_{ij} = The error terms (uncontrollable & uncontrolled factors),

i = 1, 2… a; j = 1, 2... n.

In this study, a = 5 (number of image resolution), n = 5 (replications).

Normality tests were carried out to ensure that the CA values followed a normal distribution. The 95% confidence interval was used in this study. An F-value and a p-value were calculated using Minitab 18 (Minitab, LLC, State College, PA, U.S.A). If there is a significant difference in both main effects and interaction effect, the data were analysed using the multiple means comparison (MMC) for only interaction between image

resolution and cell size by using Tukey's test to compare the difference between each pair of means. Tukey's test was performed with the condition that the magnitude of experimental error is low.

3.3. Results and discussion

3.3.1. Comparison of the parameters

Comparisons amongst the five different image resolutions were carried out using the augmented dataset because, in future research, we intend to apply the technique for PM detection in real-time in a strawberry field. Therefore, the augmented dataset, which includes images taken in multiple directions was more realistic and practical for such an experiment. A total of 50 CA values were acquired from the combinations of three feature extraction techniques and two supervised ML (Table 3-1). A CA value was acquired from each test confusion matrix and represents the average from five repetitions. Table 3-3 indicates the elapsed time based on three feature extraction techniques with the augmented dataset (n=1016).

Table 3-3. Elapsed time based on three feature extraction techniques.

Feature extraction	Image resolution	Cell size	Elapsed time (s)[a]
HOG	227×227	32×32	13.78
		64×64	11.47
		128×128	11.84
SURF		-	356.09
GLCM		-	15.94
HOG	454×454	32×32	14.37
		64×64	13.77
		128×128	12.14
SURF		-	1085.94
GLCM		-	40.94
HOG	681×681	32×32	22.24
		64×64	21.41
		128×128	20.22
SURF		-	3871.80
GLCM		-	76.98
HOG	908×908	32×32	26.89
		64×64	26.73
		128×128	25.01
SURF		-	5186.79
GLCM		-	120.04
HOG	1135×1135	32×32	39.28
		64×64	38.09
		128×128	38.00
SURF		-	8867.00
GLCM		-	200.87

[a] Elapsed time for processing of feature extraction with the augmented dataset (n=1016).

When combining HOG and ANN, the results from the two-way ANOVA indicate that image resolution and cell size were both significant factors with $p<.001$ and $p=0.026$, respectively; furthermore, there is also a significant interaction effect for image resolution and cell size with $p<.001$ (Table 3-4 & Figure 3-6). Using HOG and ANN, the highest CA was 79.96% when using an image resolution of 908×908 pixels and a cell size of 32×32. With the exception of an image resolution of 227×227 pixels and a cell size of 128×128, there was no significant difference in CA between image resolution and cell size. When extracting the 128×128 cell size from the 227×227 image resolution, the number of features

was decreased because the number of pixels along the image's width was less than twice the width of the HOG cell size of 128 pixels. The smaller the number of features, the lower the amount of data that is available to train the ANN. Given the general similarity in CA, the most effective approach was determined based on processing time (Table 3-3), where the combination of a 227×227 image resolution with a 64×64 cell size required only 11.47s with the augmented dataset (*n*=1016).

Table 3-4. Analysis of variance of HOG with ANN using the augmented dataset (*n*=1016).

Source	DF[a]	SS[b]	MS[c]	F-value	p-value
Image Resolution	4	526.7	131.67	9.97	<.001
Cell Size	2	102.2	51.12	3.87	0.026
Image Resolution*Cell Size	8	699.3	87.41	6.62	<.001
Error	60	792.2	13.20		
Total	74	2120.4			

[a] Degree of freedom
[b] Sum of squares
[c] Mean square

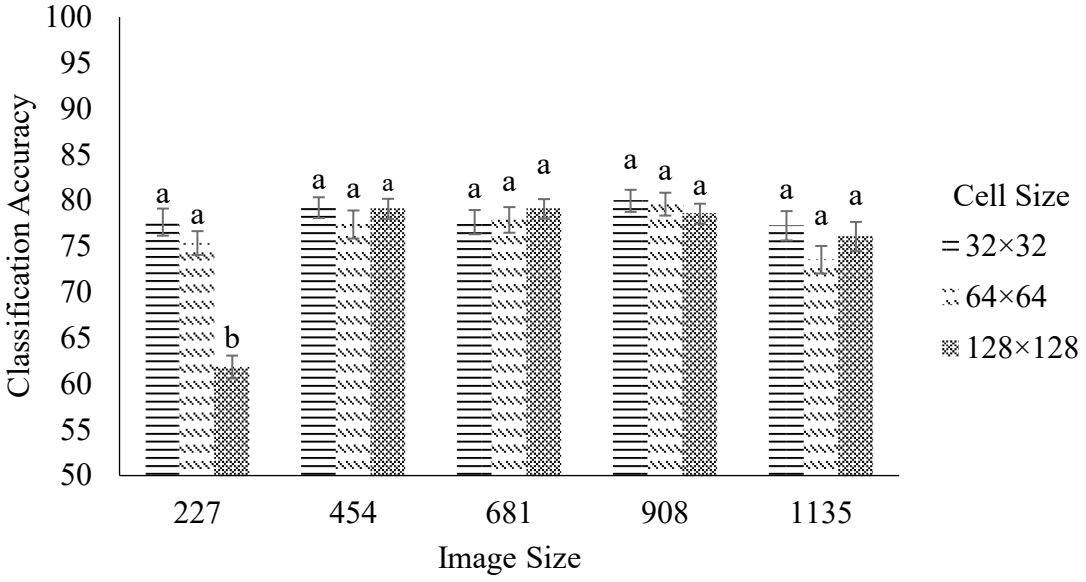

Figure 3-6. Classification accuracy by using HOG with ANN in the augmented dataset (*n*=1016). Actual CA values are presented in section 3.3.2..

The statistical model and hypothesis with a combination of HOG and SVM are the same as that of ANN. In Table 3-5, the two-way interaction effect for image resolution and cell size is considered as significantly different. Figure 3-7 shows the two-way interaction effects for image resolution and cell size for HOG with SVM using the augmented dataset. Again the 227×227 image resolution at 128×128 cell size showed the lowest CA among all the treatment groups. The combination of the HOG and SVM in the augmented dataset achieved a maximum CA of 78.36% at 681×681 image resolution with 32×32 cell size. In obtaining the maximum CA, image resolution and cell size did not play a role except for three cases; hence, the next step, that needed to be considered, was processing time. Table 3-3 shows the elapsed time to extract features and the 454×454 image resolution with 64×64 cell size takes the shortest time as 13.77 s with the augmented dataset (*n*=1016).

Table 3-5. Analysis of variance of HOG with SVM using the augmented dataset (*n*=1016).

Source	DF[a]	SS[b]	MS[c]	F-value	p-value
Image Resolution	4	650.10	162.52	35.36	<.001
Cell Size	2	205.40	102.71	22.35	<.001
Image Resolution*Cell Size	8	1014.80	126.85	27.60	<.001
Error	60	275.70	4.59		
Total	74	2146.00			

[a] Degree of freedom
[b] Sum of squares
[c] Mean square

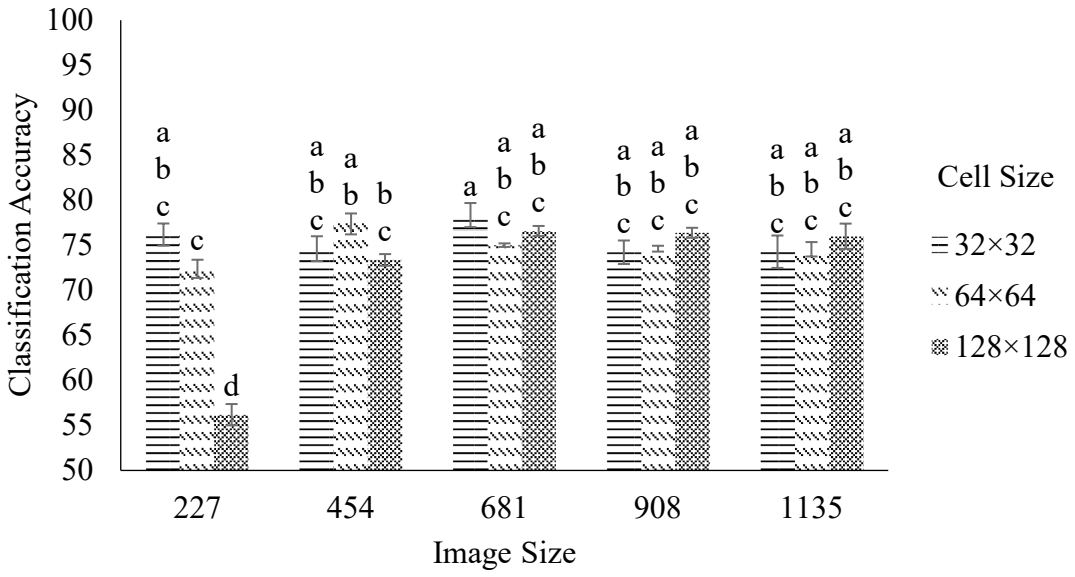

Figure 3-7. Classification accuracy by using HOG with SVM using augmented dataset (*n*=1016). Actual CA values are presented in section 3.3.2..

Overall, HOG feature extraction times were faster than other extraction techniques such as SURF and GLCM (Table 3-3); hence, it was identified as the best feature extraction technique for real-time processing. For applications related to real-time detection of pedestrian and human movement, Xu et al. (2011) and Beiping and Wen (2011) also used a combination of HOG and SVM due to their suitability for processing high-resolution videos, which requires ample computational resources in terms of feature extraction and video processing.

Unlike the HOG algorithm, SURF and GLCM feature extraction techniques do not have a cell size parameter; hence, one-way ANOVA was carried out. The SURF feature extraction technique resulted in different outcomes based on the ML models used (Tables 3-6, 3-7 & Figure 3-8). The combination of SURF and ANN in the augmented dataset achieved a maximum CA of 94.34% although, there was no significant difference with

image resolutions. Table 3 shows the elapsed time to extract features, where the 227×227 image resolution took the least amount of time as 356.09 s for the augmented dataset (*n*=1016). Overall, ANN outperformed SVM when coupled with SURF regardless of image resolution (Figure 3-8). In contrast, when using an image resolution of 227×227, there was a significant decrease in CA when SVM was used. Given the overall similarity in CA when ANN and SURF were used, it was determined that an image resolution of 454×454 was the most computationally efficient approach with a processing time as 1085.94 s with the augmented dataset (*n*=1016) (Table 3-3). It should be noted, however, that the SURF algorithm required a longer feature extraction time when compared to the HOG and the GLCM algorithms.

Table 3-6. Analysis of variance of SURF with ANN using the augmented dataset (*n*=1016).

Source	DF[a]	SS[b]	MS[c]	F-value	p-value
Image Resolution	4	9.67	2.42	0.45	0.768
Error	20	106.29	5.31		
Total	24	115.96			

[a] Degree of freedom
[b] Sum of squares
[c] Mean square

Table 3-7. Analysis of variance of SURF with SVM using the augmented dataset (*n*=1016).

Source	DF[a]	SS[b]	MS[c]	F-value	p-value
Image Resolution	4	233.84	58.46	11.79	<.001
Error	20	99.20	4.96		
Total	24	333.04			

[a] Degree of freedom
[b] Sum of squares
[c] Mean square

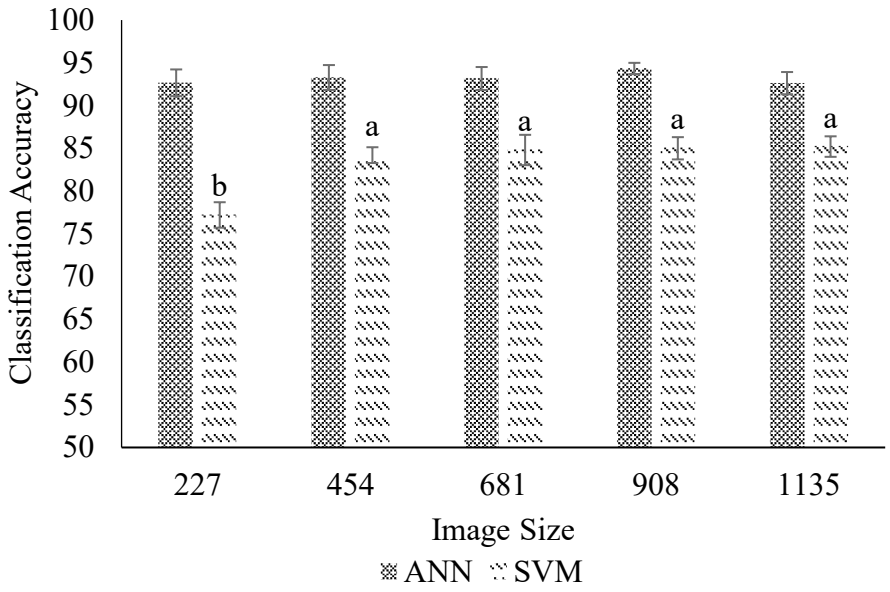

Figure 3-8. Classification accuracy by using SURF and ANN & SURF and SVM using the augmented dataset (*n*=1016). Actual CA values are presented in section 3.3.2..

In a combination of GLCM and ANN, the main effect of image resolution on CA was not significant; however, in a combination of GLCM and SVM, the main effect of image resolution was significant (Tables 3-8 & 3-9). Again, the results showed that ANN consistently outperformed SVM; furthermore, SVM was shown to be more sensitive to image resolution than ANN (Figure 3-9). As a result, the optimal image resolution for coupling GLCM and ANN was 227×227 pixels as it yielded the most efficient processing time as 15.94 s with the augmented dataset (*n*=1016) (Table 3-3). Comparing HOG and GLCM, which showed the fast feature extraction time, GLCM showed higher accuracies; however, the feature extraction time with GLCM increased more than HOG as the image resolution increased. This is similar to the result of Rangayyan et al. (2010) who showed

that it is hard to get reliable GLCM features with too large image resolution due to the excessive number of pixels.

Table 3-8. Analysis of variance of GLCM with ANN using the augmented dataset (*n*=1016).

Source	DF[a]	SS[b]	MS[c]	F-value	p-value
Image Resolution	4	17.41	4.35	1.02	0.422
Error	20	85.59	4.28		
Total	24	103.00			

[a] Degree of freedom
[b] Sum of squares
[c] Mean square

Table 3-9. Analysis of variance of GLCM with SVM using the augmented dataset (*n*=1016).

Source	DF[a]	SS[b]	MS[c]	F-value	p-value
Image Resolution	4	1265.95	316.45	109.93	<.001
Error	20	57.58	2.87		
Total	24	1323.53			

[a] Degree of freedom
[b] Sum of squares
[c] Mean square

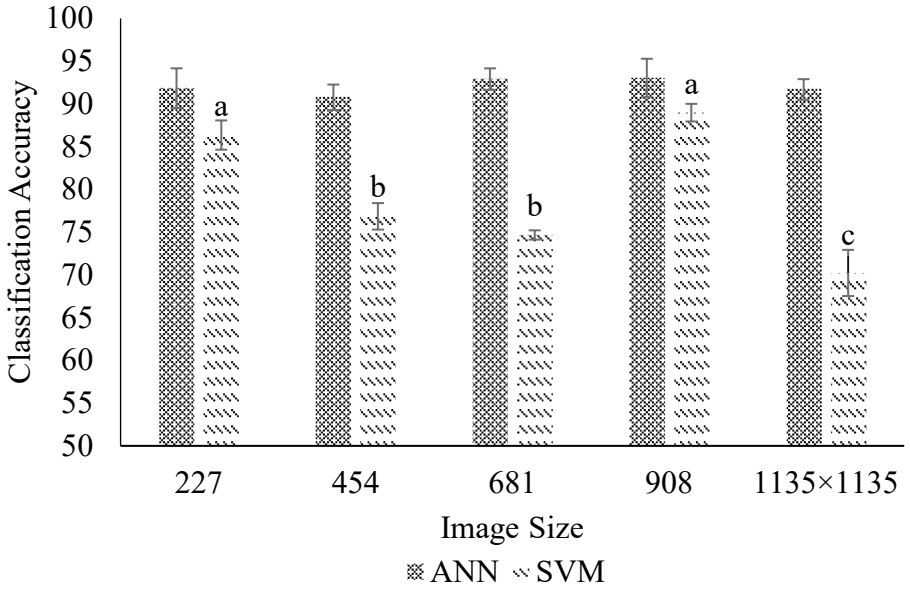

Figure 3-9. Classification accuracy by using GLCM with ANN & GLCM with SVM using the augmented dataset (*n*=1016). Actual CA values are presented in section 3.3.2..

In terms of the HOG feature extraction technique, ANN and SVM were both sensitive to how the image processing techniques were parameterised; for example, CA was influenced by image resolution and HOG cell size. In terms of the SURF and GLCM feature extraction technique, SVM was far more sensitive to how the image processing techniques were parameterised in comparison to ANN. These results were in line with the result of Manevitz and Yousef (2001) as they found that SVM kernels were far more sensitive to model parameters than ANN, which demonstrated greater model stability.

In terms of the CA values from ANN, the results after extraction by using SURF and GLCM were better than those of HOG. The HOG technique depended on the incidence of gradient orientation from the specific portions of the image; hence, the HOG technique

might be affected by the orientation parameter more than SURF and GLCM after data augmentation (Watanabe et al., 2009).

3.3.2. Effects of data augmentation on classification accuracy

Data augmentation is commonly used in image processing and DL to increase the diversity of data available for training models and to prevent overfitting problem (Zhang et al., 2016; Perez & Wang, 2017; Salamon & Bello, 2017). Methods of data augmentation can include horizontal or vertical flipping, random cropping, tilting, and rotation techniques (Fadaee et al., 2017). Data augmentation with supervised ML was not normally found in the literature. Yao et al. (2009) collected 72 samples of each disease image to classify three rice diseases with a CA of 97.2%. Meunkaewjinda et al. (2008) used 497 scab disease samples, 489 rust disease samples, and 492 non-disease samples to derive results to classify types of grape leaf diseases with the CA of 86.03%. ANN was used for detecting an early onset of plant diseases and was used for a variety of plant diseases, where their results yielded a CA of 91% when using ANN after extracting the texture and colour features from the image segmentation process (Kulkarni & Patil, 2012).

Unlike existing literature, this study considered the leaf orientation throughout the disease detection process. Although CA values may be decreased as a result of data augmentation, it more accurately simulates real-world field-conditions whereby leaves have a seemingly random orientation. In general, the addition of noise and uncertainty in the image processing technique will improve the stability and generalisability of the model; furthermore, the conservative estimates of CA are more reflective of the field (Stockwell, 1997, Peterson & Cohoon, 1999). For example, Pauly and Sankar (2016) reported better

results when HOG was applied to a non-augmented dataset-findings that were similar to this study (Tables 3-10 & 3-11).

HOG calculates the occurrence of the gradient direction in the localised part of the image. The HOG and SURF share a lot in common: they form descriptors and shape contexts. The SURF is robust to rotational invariance, blur, and warp transformations (Bay et al., 2006). Since the HOG and the SURF extract orientation-related features, a change in four different angles directional can reduce the CA. That could be the reason why the CA outcomes with the augmented dataset were lower than those of the original data set.

Dandawate and Kokare (2015) discussed the various techniques for detecting and classifying different diseases of the plants. Most methods included feature extraction to detect the diseases as an image processing steps, and then neural network classifiers were selected to classify the diseases as a classification tool. Also, they had proven that the co-occurrence extraction method is helpful to detect many plant diseases based on colour and texture. GLCM quantitatively evaluates texture parameters and representations (Soh & Tsatsoulis, 1999), which resulted in similar CA values when comparing before and after augmentation. Also, when determining co-occurrence factors, it is advisable to use all directions, with tests that compare different directions individually. In accordance with previous studies, CA values with GLCM did not show a significant difference in CA even after data augmentation.

Table 3-10. Classification accuracy based on ANN with the original (*n*=254) and augmented (*n*=1016) datasets.

ML[a]	Feature extraction	Image resolution	Cell size	Original dataset (CA%)[b]	Augmented dataset (CA%)[b]
ANN	HOG	227×227	32×32	88.42±0.48[c]	77.64±1.48
			64×64	84.72±0.80	75.36±1.30
			128×128	85.38±0.94	61.84±1.25
	SURF		-	96.32±0.01	92.66±1.35
	GLCM		-	90.00±0.04	91.84±1.33
	HOG	454×454	32×32	85.24±0.13	79.22±1.13
			64×64	84.20±0.93	77.38±1.53
			128×128	87.38±0.90	79.08±1.10
	SURF		-	93.70±0.99	93.28±1.44
	GLCM		-	90.54±0.79	90.78±1.49
	HOG	681×681	32×32	84.20±0.31	77.64±1.31
			64×64	85.78±0.60	77.88±1.40
			128×128	85.26±0.07	79.08±1.06
	SURF		-	94.74±0.15	93.16±1.39
	GLCM		-	90.52±0.34	92.90±1.26
	HOG	908×908	32×32	84.22±0.20	79.96±1.20
			64×64	84.28±0.84	79.60±1.25
			128×128	81.04±0.09	78.56±1.09
	SURF		-	96.84±0.22	94.34±1.69
	GLCM		-	90.52±0.34	93.04±1.25
	HOG	1135×1135	32×32	80.52±1.77	77.24±1.65
			64×64	85.28±2.00	73.54±1.55
			128×128	85.26±1.43	76.06±1.65
	SURF		-	94.72±1.87	92.62±1.35
	GLCM		-	93.16±1.99	91.70±1.28

[a] Machine learning

[b] Classification accuracy

[c] The standard deviation of five repetitions

Table 3-11. Classification accuracy based on SVM with the original (*n*=254) and augmented (*n*=1016) datasets.

ML[a]	Feature extraction	Image resolution	Cell size	Original dataset (CA%)[b]	Augmented dataset (CA%)[b]
SVM	HOG	227×227	32×32	83.68±1.52[c]	76.19±1.23
			64×64	86.58±1.77	72.39±1.91
			128×128	52.89±0.59	56.19±1.58
	SURF		-	94.09±1.07	77.20±1.49
	GLCM		-	71.58±1.35	86.36±1.91
	HOG	454×454	32×32	86.58±1.85	74.62±1.68
			64×64	82.36±1.03	77.37±1.16
			128×128	83.16±1.88	73.37±1.65
	SURF		-	92.70±1.75	84.20±1.92
	GLCM		-	87.37±1.55	76.85±1.55
	HOG	681×681	32×32	82.89±1.36	78.36±1.73
			64×64	84.73±1.25	75.01±1.20
			128×128	84.99±1.55	76.59±1.56
	SURF		-	87.21±1.53	84.80±1.77
	GLCM		-	85.82±1.06	74.66±0.54
	HOG	908×908	32×32	83.94±1.85	74.23±1.75
			64×64	85.26±1.16	74.62±1.32
			128×128	81.31±1.69	76.39±1.55
	SURF		-	91.30±1.71	84.40±1.71
	GLCM		-	82.22±1.62	88.98±1.03
	HOG	1135×1135	32×32	82.37±0.30	74.29±1.88
			64×64	86.31±1.68	74.56±0.86
			128×128	83.15±1.85	76.01±1.41
	SURF		-	80.09±1.62	85.20±1.28
	GLCM		-	81.57±1.63	70.23±2.77

[a] Machine learning

[b] Classification accuracy

[c] The standard deviation of five repetitions

3.4. Conclusions

In this study, a series of experiments were conducted toward the development of a supervised ML approach that could reduce labour costs for producers who want to detect PM in strawberry fields. This study compared 50 CA values based on three feature extraction techniques and two supervised ML in order to determine the optimal approach for detecting PM on strawberry leaves. In addition, to analyse the effect of data augmentation on the CA, especially in supervised ML, 254 images were rotated with four different angles to make an augmented dataset of 1016 images. The CA values typically decreased after data augmentation because a directional parameter was added through the process of rotating the original image along with four different angles. Considering that leaves in the field are not pointing in one direction, the CA results after data augmentation are likely to be more reflective of real-world conditions. For ANN, a combination with SURF feature extraction showed the highest CA with 94.34% at an image resolution of 908×908 pixels. For SVM, a combination with GLCM feature extraction showed the highest CA with 88.98% at an image resolution of 908×908 pixels. The methods that were used in this study showed the potential to detect PM on strawberry leaves in an accurate and computationally cost-effective way. Future research may explore the use of alternative ML techniques for larger datasets using DL. The findings of this study could facilitate spot-application of fungicide on infected plants; thereby decreasing overall fungicide application and increasing cost-efficiency.

CHAPTER 4: DEEP LEARNING FOR IMAGE-BASED POWDERY MILDEW DISEASE DETECTION ON STRAWBERRY LEAVES

Deep learning (DL) has recently gained popularity in the field of image classification and identification due to its high accuracy and fast learning. In this research, DL was used to detect powdery mildew (PM), a persistent fungal disease in strawberries. High accuracy is required to reduce the amount of unnecessary fungicide use and the number of observations from the specialists. We proposed the most optimised algorithm in accordance with our research goal among AlexNet, SqueezeNet, GoogLeNet, ResNet-50, SqueezeNet-MOD1, and SqueezeNet-MOD2. The initial number of images was 1450 including healthy and infected leaves. Data augmentation was carried out to prevent overfitting and to consider the various shapes and direction of the leaves in the field. A total of eight clockwise rotations (0°; the original data, 45°, 90°, 135°, 180°, 225°, 270°, and 315°) was performed to increase to 11,600 data points. Overall, the six DL algorithms that were used in this research showed on average over 92% classification accuracy (CA). ResNet-50 gave the highest CA of 98.11% in classifying the healthy and infected leaves. However, considering the computation time, we suggested AlexNet which processed testing data in 40.73 seconds. Holistically, SqueezeNet-MOD2 would be recommended for our further research project to build a hardware system to detect the PM on the strawberry leaves.

4.1. Introduction

Fungal diseases are a significant challenge in the strawberry production system of which powdery mildew (PM) is a particularly problematic pathogen. PM can easily infect strawberry plants in regions with warm temperatures and humid climatic conditions. PM sequentially attacks leaves, flowers, and fruits, which can lead to excessive crop loss at the severe stages of infection. Here, white mycelium begins to cover the leaf surface, reducing the amount of photosynthesis and causing water deficiency. Initially, less than 10 white mycelium spots begin to appear with the formation of a white star-shaped dot, and about five days later, white fungal groups begin to cover the surface of the strawberry leaves (Adam & Somerville, 1996). As the infection progresses, infected leaves turn reddish-purple or have small purple spots (Jacob et al., 2008). If the fruit is infected by PM, it causes a bitter taste that reduces the marketability (Liu, 2017).

Due to the significance of PM, effective methods for monitoring and managing are required to control the PM. The primary method used to suppress PM is to apply fungicides as a spray; however, PM is prone to develop a resistance to fungicides (McGrath, 2001). Furthermore, the excessive application of fungicides may lead to environmental degradation such as erosion of beneficial soil components or accumulation of toxic substances in soil (Kalia & Gosal, 2011).

The conventional method for monitoring plant pathogens is to hire disease specialists who can scout a field in order to identify the diseases However, the conventional method is not effective in terms of labour and time, and it is impossible to accurately predict the spread of the disease over the field (Kobayashi et al., 2001). Furthermore, these challenges are compounded by the decreasing workforce within the agricultural sector.

According to Statistics Canada (2020c), the number of people working in the agriculture industry has steadily decreased since 1991 while the average age of workers has gradually increased up to 55 years of age in 2016.

In the 1980s, the concept of precision agriculture (PA) emerged as a major component of the third wave of the modern agricultural revolution (Robert, 2002). PA was initially used for the targeted allocation of fertilizers to suit different soil conditions. Since then, PA has been developed for automatic guidance of agricultural vehicles and tools; autonomous machines and processes; research on farms; and automated management of agricultural production systems (Zhang et al., 2002). Collecting data using sensors mounted on machines (i.e., unmanned ground vehicles (UGV), unmanned aerial vehicles (UAV), satellites, and airplanes) is non-destructive and applicable over large geographical areas. Ground-based and aerial imagery are often critical components in PA (Liaghat & Balasundram, 2010), especially given that image identification and classification techniques.

Machine learning (ML) including deep learning (DL) could be used to detect pathogens, pests, nutrient deficiencies, and other abiotic stressors (Teke et al., 2013). In the case of non-DL in ML, hereinafter referred to as the "non-DL", several feature extraction methods (e.g., histogram of oriented gradients, speeded up robust features, and gray level co-occurrence matrix) would be used to find the best applicable approach, where the best extraction method would then be used as inputs for the learning algorithm. DL is more effective in handling additional complexity and hierarchical structure in the data. The key aspect in DL is that it extracts by itself the best features in the learning procedure (LeCun et al., 2015). In the case of DL, the processing step of DL is more simplified than non-DL

because the optimal feature method is applied to the learning algorithm automatically (Alom et al., 2018). Furthermore, processing time of DL, especially testing time, is faster than non-DL (Chen et al., 2014). Fast processing time is good for algorithms to be applied to the hardware (i.e., field programmable gate array; FPGA or mobile application) and it could solve the needs of producers who need results as soon as possible.

DL can be classified into three types according to the method and purpose of the research. The first type is unsupervised learning that does not require data labels, and this is good to find representative features with little data. As examples of the first type are deep belief network and deep auto encoder (Vincent et al., 2010). The second type is a recurrent neural network which fits well with sequential data processing, such as protein or polymer sequences. The last type is a convolutional neural network (CNN) which has shown the precise results in image recognition and natural language processing (Deng & Platt, 2014). Among three types of DL, CNN has been used as the most advanced computer vision (CV) in various fields since 2014 and is the most popular for detecting disease of the crops due to the high classification accuracy (CA) in image recognition (Szegedy et al., 2016).

Imagenet large scale visual recognition challenge (ILSVRC) is an annual competition to foster the latest algorithm using the annotated Imagenet dataset for the development of CNN (Deng et al., 2009). Outcomes from the ILSVRC are milestone algorithms and techniques in the field of CV and DL. In this study, four state-of-the-art CNNs (AlexNet, SqueezeNet, GoogLeNet, and ResNet-50) were selected, and SqueezeNet-MOD1 and SqueezeNet-MOD2 developed by the author was carried out according to our research goal.

AlexNet, SqueezeNet, GoogLeNet, and ResNet-50 used in this study were found to be used actively in agriculture research. For example, AlexNet (ILSVRC 2012 winner) and SqueezeNet were compared for detecting disease in tomatoes and yielded similar accuracies (Durmuş et al., 2017); however, SqueezeNet would be more appropriate for use with mobile applications due to small memory requirements. GoogLeNet (ILSVRC 2014 winner), also known as Inception, has also been used for disease detection purposes; for example, Ramcharan et al. (2017) compared the Inception and Inception V3 where it was determined that Inception was more effective in detecting diseases on cassava. However, Ramcharan et al. indicated that there were limitations related to memory and computing power. In studies comparing AlexNet with GoogLeNet, the differences were variable; for example, Mohanty et al. (2016) showed that GoogLeNet was more effective than AlexNet in detecting an array of 26 diseases found on 14 different crop species whereas the accuracy rates for the two algorithms were similar when used to detect nine diseases common to tomatoes. Lastly, Fuentes et al. (2017) compared different versions of the ResNet algorithm (ILSVRC 2015 winner) for detecting tomato diseases where it was determined that ResNet-50 outperformed ResNeXt-50 (ILSVRC 2016 winner).

Despite the application and success of different CNNs for detecting crop diseases, there is no existing research for detecting PM on strawberry leaves. Yet, such research is needed in order to inform the targeted application of fungicides thereby mitigating environmental degradation, as well as reduce the labour requirements by producers. Hence, the objectives of this research are (1) to compare the performance between supervised non-DLs and CNNs, (2) to modify the architectures of four CNNs (AlexNet, SqueezeNet, GoogLeNet, and ResNet-50), (3) to develop SqueezeNet-MOD1 and SqueezeNet-MOD2

by modifying the SqueezeNet 1.1 version architectures, that have shown hardware applicability, to increase the performance with respect to the CA and computation time, (4) to evaluate and compare their performance with respect to CA, computational efficiency, and computation time.

4.2. Materials and methodology

4.2.1. Image datasets

Healthy and infected strawberry leaves were collected from Balamore Farm Ltd. (45°24'35.4"N, 63°34'26.3"W) and Millen Farms Ltd. (45°23'57.6"N, 63°33'31.1"W) in the summers of 2018 and 2019. After collecting the leaves, they were brought to the lab as preserved in an icebox within 30 minutes. Images were taken with a digital single-lens reflex (DSLR) camera: EOS 1300D (Canon Inc., Tokyo, Japan), with 3456×5184 pixels (Raw CR2 format). The computational environment was implemented in MATLAB R2019a (The MathWorks Inc., Natick, MA, U.S.A.) using an Intel® Core™ i7-8700 CPU @ 3.20GHz with 48.0 GB RAM and 64-bit Windows 10 operating system. All images were cropped to 908×908 pixels first then resized to 227×227 pixels or 224×224 pixels depending on the requirement of the architectures to ensure that the relative size of the leaves on all images was consistent and to reduce the computational demand. The required input size in AlexNet, SqueezeNet, SqueezeNet-MOD1, and SqueezeNet-MOD2 was 227×227 pixels and in GoogLeNet and ResNet-50 was 224×224 pixels. Images were adjusted to make the relative portions of PM uniform across all images and to reduce the computational burden and improve the efficiency of the results.

The total number of images collected from 2018 to 2019 consisted of 677 healthy and 773 infected leaves for a total of 1450 leaves. Data augmentation in DL is essential for models to be generalised so that it could be applied more easily to real field situations (Taylor & Nitschke, 2017). In addition, Rasti et al. (2019) identified that at least 10,000 observations are required to improve the overall performance of DL algorithms and minimize the problems of overfitting. Hence, data augmentation was performed with the whole dataset to improve the robustness of the architectures and increase the number of observations (Liu et al., 2018). There are two aspects of data augmentation: geometrical transformation (rotating, flipping, cropping, and resizing the images) (Fadaee et al., 2017) and intensity transformation (noise, colour change, brightness enhancement) (Fuentes et al., 2017). Among them, the rotation technique was performed considering the different shape and direction of the leaves in the field. Hence, data augmentation was carried out using clockwise rotation technique along the 0° (the original images), 45°, 90°, 135°, 180°, 225°, 270°, and 315°. Therefore, the original 1450 observations were increased by eight times with 5416 observations representing healthy leaves and 6184 observations representing infected leaves with a total of 11,600 observations.

4.2.2. Comparison of non-DLs and CNNs

In this section, performances with respect to CA in non-DL and CNN techniques will be compared with the original image dataset (n=1450). Shin et al. (2020) compared the classification accuracies between three feature extraction techniques (histogram of oriented gradients; HOG, speeded-up robust features; SURF, and gray level co-occurrence matrix; GLCM) and two supervised non-DLs (artificial neural network; ANN and support vector machines; SVM). According to the results in Shin et al. (2020), the combination

between SURF and ANN, and GLCM and SVM were chosen as the optimal techniques to detect PM on strawberry leaves.

4.2.3. Architectures of CNNs

A CNN has convolutional layers that extract features from input images. The convolutional layers consist of a filter and an activation function and images of the leaves will classify into healthy or infected leaves by using the conventional neural networks based on the extracted features. AlexNet, SqueezeNet, GoogLeNet, and ResNet-50 have all different properties regarding depth, size, a number of parameters, and an image input size (Table 4-1).

Table 4-1. Depth, size, parameters, and image input size per each deep learning architectures.

Network	Depth	Size	Parameters (millions)	Image input size
AlexNet	8	227 MB	61.0	227-by-227
SqueezeNet	18	4.6 MB	1.2	227-by-227
GoogLeNet	22	27 MB	7.0	224-by-224
ResNet-50	50	96 MB	25.6	224-by-224

In this section, four CNNs were introduced in terms of various evaluation metrics to find the best way to detect PM on the strawberry leaves. CNNs consist of a number of hyperparameters that needed to be tuned. Through many optimisation trials, the optimal hyperparameter combination was determined to be as follows: iterations = 5, base learning rate = 0.001, max epochs = 10, mini batch size = 64. This combination of parameters was applied to all CNN algorithms. To find and optimize the best classifier models, randomised hold-back validation was carried out, whereby the dataset was randomly partitioned into

80% (8120 images) for model training and 20% (3480 images) for model testing. The hold-back procedure was repeated five times and the average accuracy metrics were reported.

The learning rate is one of the most important hyperparameters to tune for training CNN and tells the optimizer how much weight to move in the opposite direction from the mini-batch since DL models are primarily trained by stochastic gradient descent optimizers. According to the rate of learning, the benefit of the model would be different. If the learning rate is low, although the optimisation takes a long time by moving forward with the small steps of the minimum loss function, the training process is more stable. On the other hand, if the learning rate is high, the training might not converge but diverge (Sutskever et al., 2013). The number of epochs is also a crucial hyperparameter that defines the times that the learning algorithm will work through the entire training dataset. An epoch can be comprised of one or more batches and mini-batch can be determined according to the size of data. Usually, a batch size of 32 is a good starting point and batches of 62, 128, and 256 can be applied as well (Keskar et al., 2016), and finding the optimised batch size in the given range would be suggested. In this experiment, training and testing division was set as 80% of the whole dataset used for training and 20% for testing. A total of 9280 images out of the 11,600 were used for training the model and the remaining 2320 images were used for testing the model.

4.2.3.1. AlexNet

AlexNet consists of five convolution layers, three fully connected layers, and softmax. In the first convolutional layer, 96 kernels (i.e., filters) of size 11×11×3 can be calculated with a stride of four on 227×227 input images. Among the five convolution layers, the first and second convolution layers are followed by the overlapping max pooling;

however, the remaining convolutional layers are connected directly each other. The last convolutional layers are followed by overlapping max pooling before connecting to the two fully connected layers. The second fully connected layers are linked to the softmax classifiers with two class labels. The architecture of AlexNet in this study was changed from 1×1×1000 into 1×1×2 in the last fully connected layer that classifies them into two (healthy and infected) (Figure 4-1).

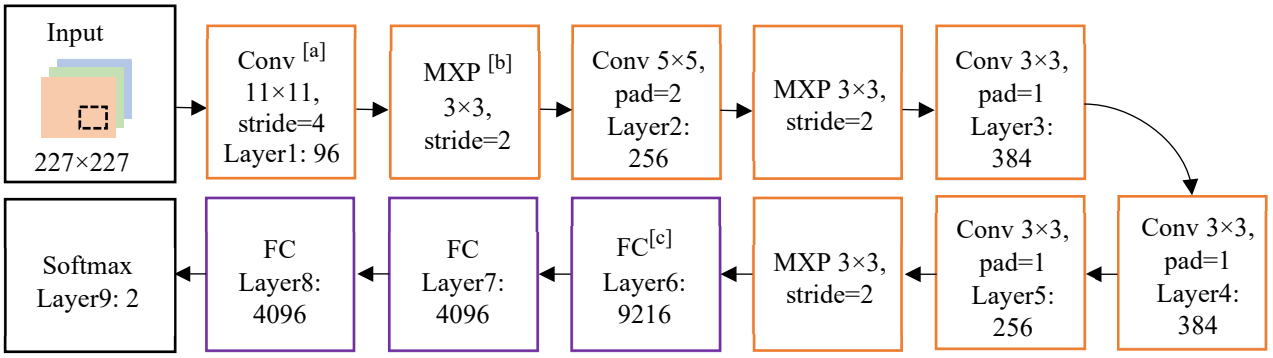

Figure 4-1. Architecture of AlexNet used in this study.

[a] Conv: Convolutional layer
[b] MXP: Max Pooling layer
[c] FC: Fully Connected layer

4.2.3.2. GoogLeNet

GoogLeNet has 22 layers that were designed as a deep network which has more than 2 hidden layers and 12 times fewer parameters than AlexNet with higher accuracy. The major characteristic of GoogLeNet is the first CNN to introduce the Inception module. Inception layers keep a high resolution for small information on the images by covering a bigger area. Hence, the concept of inception runs parallel from 1×1 which is the sophisticated convolutional filter to the 5×5 convolutional filter which is a big size of layer

compared to 1×1 and 3×3 (Figure 4-2). Then, these different sizes of filters can process the better multiple objects scale according to the needs.

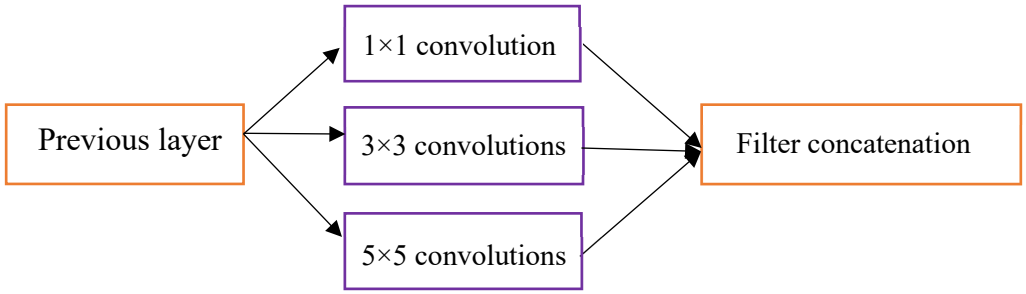

Figure 4-2. Inception module in GoogLeNet.

4.2.3.3. ResNet-50

The original ResNet has 152 layers and ResNet-50 is a smaller version of ResNet-152. ResNet-50 is a CNN that is trained on more than a million images from the Imagenet database and stands for 50 layers deep (He et al., 2016). Each ResNet block is either two layers deep or three layers deep. The input images for the network are 224×224 pixels and each convolution which has three convolution layers and identity block which also has three convolution layers consists of five stages (Figure 4-3).

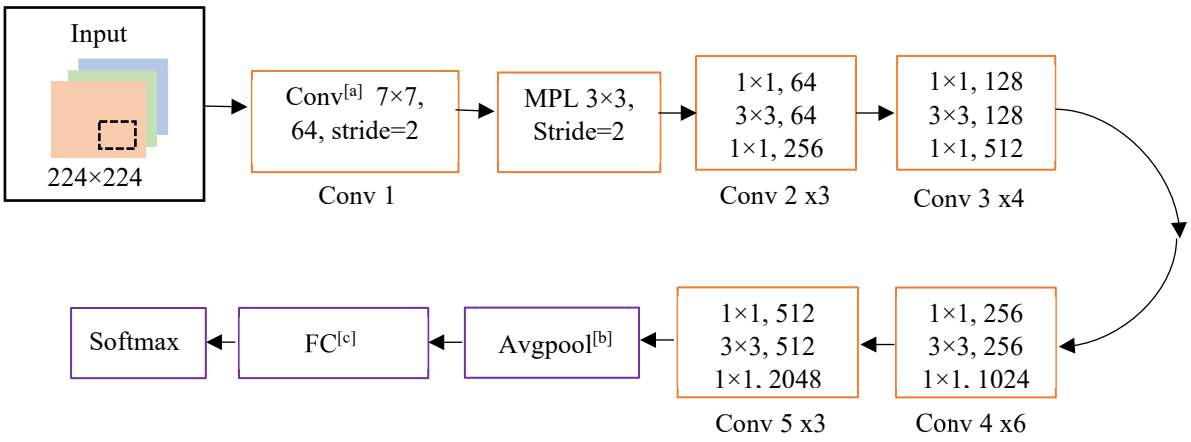

Figure 4-3. ResNet-50 architecture used in this study.

[a] Conv: Convolutional layer
[b] Avgpool: Average pool
[c] FC: Fully Connected layer

4.2.3.4. SqueezeNet and SqueezeNet modifications

SqueezeNet can achieve AlexNet level of accuracy with only 0.5 MB which is 510 times smaller than AlexNet (Iandola et al., 2016). According to Iandola et al., it is considered a feasible CNN for deployment to a FPGA or other hardware due to limited memory. Three points are different from the previous big CNN; firstly, nine times of parameters in the convolutional layer are fewer than existing big CNN by replacing 3×3 filter with 1×1 filter. Secondly, the number of channel inputs is decreased to 3×3 filter. Lastly, downsampling late in the network to convolution layers have large activation maps. The key layer introduced in the SqueezeNet is a fire module that consists of a squeeze layer and an expand layer and SqueezeNet constitutes a lot of fire modules and a few pooling layers (Figure 4-4). The role of the squeeze layer shrinks the features down consisting of 1×1 convolutional layer and it can be expanded with a combination of 1×1 and 3×3 convolutional layers. Hence, the feature map would go from small by using squeeze layer to big by using the expand layer. The fire module is made up of these three parameters (e.g., the number of 1×1 convolutional layers used in the squeeze, the number of 1×1 convolutional layers used in the expand, and the number of 3×3 convolutional layers used in the expand) (Figure 4-5). The modified part in SqueezeNet is the activations are changed from 1×1×1000 to 1×1×2. SqueezeNet in this study was developed based on the SqueezeNet 1.1 version which is akin to the SqueezeNet 1.0 version but different in the places of the maxpool. SqueezeNet 1.1 version requires 2.4 times less computation than 1.0 version without negatively impacting accuracy (Bressem et al., 2020).

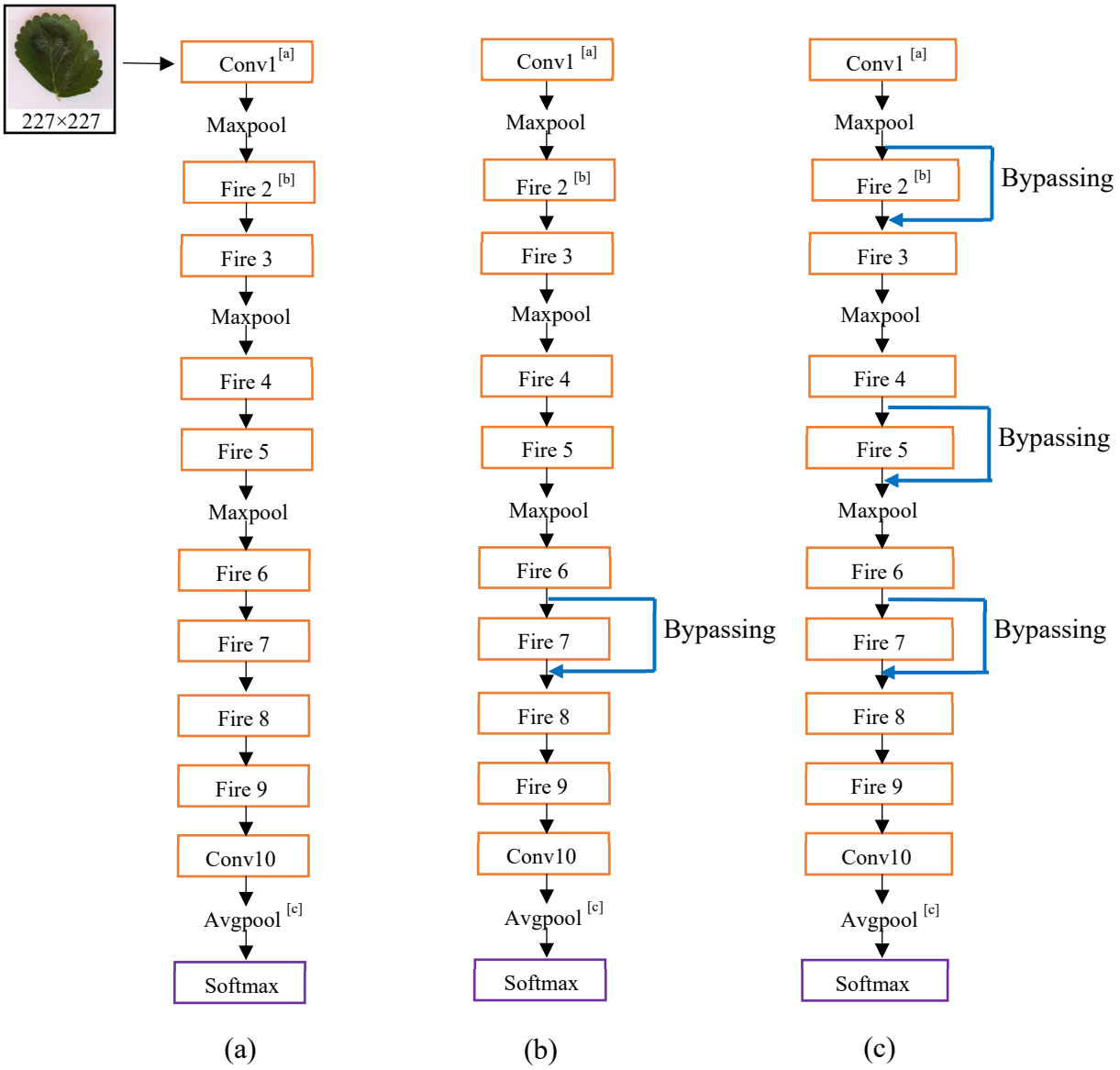

Figure 4-4. The architectures of SqueezeNet and SqueezeNet modifications, (a) SqueezeNet (version 1.1), (b) SqueezeNet-MOD1 with bypassing of Fire 7, (c) SqueezeNet-MOD2 with bypassing of Fire 2, 5, and 7.

[a] Conv: Convolutional layer
[b] Fire: Fire module
[c] Avgpool: Average pool

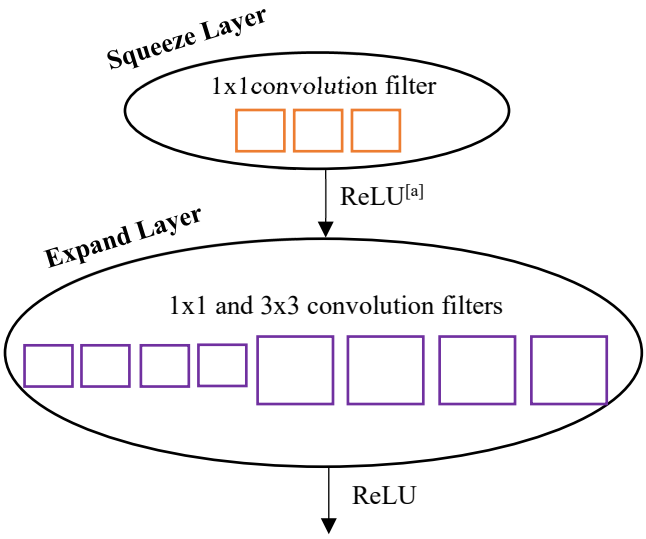

Figure 4-5. Fire module in the SqueezeNet architecture.

[a] ReLu: Rectified Linear Unit which is an activation function

The purpose of modifications in SqueezeNet was to improve the performance with respect to CA and testing time. Besides modifying the end architecture like AlexNet, GoogLeNet, and ResNet-50, SqueezeNet was modified by changing the whole architecture to make this algorithm more applicable to hardware. SqueezeNet-MOD1 was developed by bypassing the seventh fire module in Figure 4-4 (b). SqueezeNet-MOD2 was developed by bypassing the second, fifth, and seventh fire modules in Figure 4-4 (c).

4.2.4. Evaluation of CNNs

Healthy leaves indicate as negative (N and 0) and infected leaves mark as positive (P and 1) (Figure 4-6). FN is classified as Type II error which is related to recall/sensitivity and FP is considered as Type I error which is related to precision. If FN values are high, that means the infected leaves are not identified correctly; on the other hand, if FP values are high, that means the healthy leaves are wrongly identified as infected leaves. By using

a confusion matrix, test accuracy, precision, sensitivity/recall, specificity, and F1-score will be compared. The matrix used to assess the CA values in models included test accuracy, specificity, and the harmonic mean of precision and recall (F1-score). Here, we defined test accuracy/CA as the following:

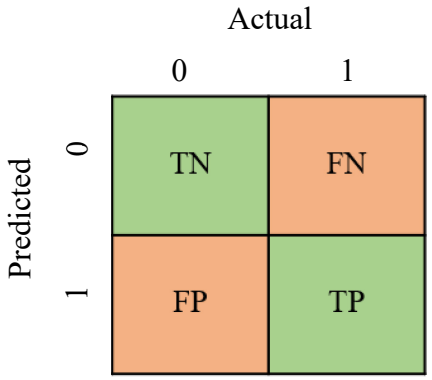

Figure 4-6. Test confusion matrix for evaluation. The rows of the confusion matrix correspond to the predicted class the columns correspond to the actual class. 0: healthy leaves (negative) and 1: infected leaves (positive).

$$\text{Test accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \qquad \text{(Eq. 4-1)}$$

Where TP represents the number of leaves that were correctly classified as being infected and TN represents the number of leaves that were correctly classified as being non-infected. FP, which is considered as a Type I error, reflects the number of leaves that were incorrectly classified as being infected, while FN, which is considered as a Type II error, reflects the number of leaves that were incorrectly classified as being non-infected.

The test accuracy is the most intuitive evaluation metric to check the performance of the model (Eq. 4-1); however, this is useful only in the situation when the values of FP

and FN are balanced. In addition to test accuracy, we also calculated precision which is the fraction of correctly predicted infected leaves among the positively predicted observations (TP+FP) (Eq. 4-2).

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad \text{(Eq. 4-2)}$$

Next, we calculated sensitivity/recall and specificity, which provides a magnitude of the mistakes and a measure of effectiveness of the first treatment respectively. Sensitivity/recall is the fraction of correctly predicted infected leaves among the positively actual observations (TP+FN) (Eq. 4-3). Specificity is the fraction of correctly predicted healthy leaves among the negatively actual observations (TN+FP) (Eq. 4-4). Lastly, we calculated F1-score which considers both FP and FN as a function of precision and sensitivity/recall (Eq. 4-5).

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad \text{(Eq. 4-3)}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad \text{(Eq. 4-4)}$$

$$\text{F1-score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad \text{(Eq. 4-5)}$$

4.2.5. Statistical analysis

4.2.5.1. Methods for factorial analysis

To determine which algorithm provides the highest CA and shortest testing time respectively, each CNN algorithm was considered as an input factor, and the test accuracy and computation time were considered as a response (Eq. 4-6). It should be noted that the statistical analysis of test accuracy/CA and testing time were performed separately. This

statistical experiment was performed in Minitab 18 (Minitab, LLC, PA, State College, PA. U.S.A). The normality test of the data was carried out before doing a one-way ANOVA.

The model for factorial analysis is represented as follows:

$$Yij = \mu + \alpha i + \varepsilon ij \quad \text{(Eq. 4-6)}$$

Where;

Yij = Classification accuracy (%),

μ = Overall mean,

αi = Effect of "Deep learning algorithms" on response at i_{th} level,

εij = The error terms (uncontrollable & uncontrolled factors),

i = 1, 2…, a; j = 1, 2..., n.

In this study, a = 6 (number of algorithms) and n = 5 (replications).

4.2.5.2. Hypothesis

H_o: $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = \alpha_6 = 0$

H_a: at least one of the test accuracies is $\neq 0$

The null hypothesis for test accuracy/CA (or computation time) is there are no significant differences in six CNN algorithms. The alternative hypothesis is at least one test accuracy/CA (or computation time) has a significant difference in six CNN algorithms. If the p-value is less than 0.05, the null hypothesis is rejected and a multiple mean comparison (MMC) was conducted to determine the significant difference among groups. Tukey's test

was selected due to the consideration of the magnitude of experimental error and was used to compare the differences between groups. Since the experiment was carried out in the laboratory, the magnitude of experimental error was low, suggesting that Tukey's test was used as a method for MMC.

4.3. Results and discussion

In this section, six CNN algorithms were compared in terms of the various evaluation metrics and testing time to find the best way to detect PM on the strawberry leaves. The optimised algorithm depended on performances with respect to high test accuracy/CA value, short testing time, and small memory for an application to hardware.

4.3.1. Comparison of non-DLs and CNNs

In ANN, SURF feature extraction was selected as the best combination; In SVM, GLCM feature extraction was selected as the best combination (Shin et al., 2020). With the dataset in this research, both SURF and GLCM feature extraction showed the better performance in terms of ANN. Non-DL techniques showed an underfitting problem due to the added another summer season data in a process of data augmentation and showed an overfitting problem with a small dataset (Shin et al., 2020).

According to the statistical analysis, there was a significant difference ($p<.001$) between test accuracies/CA results from 10 ML techniques (Table 4-2). Among 10 techniques including non-DL and CNN, pre-trained CNN models (AlexNet, SqueezeNet, GoogLeNet, and ResNet-50) showed the highest CA when considering the statistical analysis. Next, SqueezeNet modifications and non-DLs showed good performances in order. CA values are lower than results in Shin et al. (2020) due to an increasing dataset

and adding new parameters as a different season. In conclusion, Table 4-2 shows CNNs are better than non-DLs when various parameters are included in the dataset.

Table 4-2. Performance of non-DLs and CNNs using the original dataset (*n*=1450).

Type	Techniques	Classification accuracy (%)	Grouping
non-DLs	SURF+ANN	85.44±1.12[a]	D
	GLCM+ANN	69.28±0.92	E
	SURF+SVM	55.20±2.28	G
	GLCM+SVM	62.28±2.42	F
CNNs	AlexNet	96.55±0.81	AB
	SqueezeNet	94.62±0.79	ABC
	GoogLeNet	94.42±0.47	ABC
	ResNet-50	98.01±0.47	A
	SqueezeNet-MOD1	93.72±2.75	BC
	SqueezeNet-MOD2	91.01±0.90	C

[a] The standard deviation of five repetitions

4.3.2. Experimental performance

All CA values of six CNN algorithms were over 92%; hence, CNN algorithms proved promising in detecting PM on the strawberry leaves. ResNet-50 had the highest test accuracy/CA (98.11%) and gave the best performance in precision of 98.46% (Table 4-3). Next, SqueezeNet-MOD1 developed by the author had a test accuracy/CA of 96.38% which is slightly higher than CA of GoogLeNet (Table 4-3). ResNet-50 outperformed the other CNN algorithms for various other evaluation metrics, including precision, sensitivity/recall, specificity, and F1-score. High precision means a low FP which means that it is good for saving the costs of fungicides. Regardless of the processing time, if the cost is the most important consideration, then ResNet-50 would be suggested due to the highest precision with CA of 98.46%. In binary classification, recall is expressed as sensitivity/recall and ResNet-50 showed the highest sensitivity/recall value of 97.99%. High sensitivity/recall means that there are a few FN results, which means there are fewer

cases of PM detection mistakes. In the early stages of diagnosis, sensitivity/recall could be considered when trying to find the maximum possible PM leaves. In fact, if PM detection is the most important factor, the values of FN should be reduced. Specificity is encouraged to be considered in the final diagnosis stage or when the effectiveness of treatment is low even after spraying the primary fungicide. Hence, existing in a state of balance with sensitivity/recall and specificity is important. F1-score is useful to understand the results with the uneven class distribution. If the gap between FP and FN values is big, then F1-score would be considered the most. Of the compared CNN techniques, ResNet-50 received the highest F1-score of 98.23% (Table 4-3).

Table 4-3. Performance of six CNN algorithms using the augmented dataset (*n*=11,600).

	AlexNet (%)	SqueezeNet (%)	GoogLeNet (%)	ResNet-50 (%)	SqueezeNet -MOD1 (%)	SqueezeNet -MOD2 (%)
Test accuracy	95.59±0.49 [a]	95.80±0.72	96.36±1.16	**98.11**±0.28	96.38±0.57	92.61±0.84
Precision	96.14±1.93	96.90±1.20	96.94±2.12	**98.46**±0.37	97.00±1.21	92.96±1.21
Sensitivity/ recall	95.64±1.83	95.29±1.64	96.31±2.64	**97.99**±0.42	95.84±1.77	93.48±2.49
Specificity	95.54±2.34	96.49±1.41	96.74±2.61	**98.25**±0.37	96.30±2.30	92.52±1.96
F1-score	95.86±0.44	96.07±0.73	96.56±1.18	**98.23**±0.26	96.29±0.26	94.10±1.18

[a] The standard deviation of five repetitions

* The bold font shows the highest CA among the values

The findings from this study are consistent with other studies that compared a variety of DL algorithms. For example, Wang et al. (2017) compared AlexNet, GoogLeNet, VGG, and ResNet-50 for classifying eight different disease types. Wang et al. achieved the best results by using ResNet-50 to classify eight different pathology classes. Asad and Bais (2019) compared SegNet, UNET, VGG16, and ResNet-50 for predicting weed density in canola fields. Among four CNN algorithms, ResNet-50 showed the best performance with

82.88% mean intersection over union (IOU), 98.69% frequency weighted IOU, and 99.48% accuracy. Cruz et al. (2019) evaluated the performance of six CNN algorithms (e.g., AlexNet, GoogLeNet, Inception V3, ResNet-50, ResNet-101, and SqueezeNet) to detect grapevine yellows disease in a red grapevine and found that ResNet-50 is the most optimal algorithm with CA of 99.18% and training cost. Ashqar and Abu-Naser (2019) reported the potential of DL training by showing CA of 99.84% regarding the tomato leaf disease.

Using ANOVA, there was a significant difference ($p<.001$) between the test accuracies/CA results from the CNN algorithms (Table 4-4); hence, MMC was carried out. Here, it was determined that there was no significant difference between the different algorithms with the exception of ResNet-50. ResNet-50 showed the significantly highest CA of 98.11% and this is aligned with our expectation based on similar results reported by Bianco et al. (2018) (Table 4-5). However, the magnitude of the difference was small. To distinguish performance beyond the CA of the algorithms other than ResNet-50, computation time and memory should be considered. The result analysis of SqueezeNet-MOD1, SqueezeNet-MOD2, GoogLeNet, SqueezeNet, and AlexNet was explained in section 4.3.3..

Table 4-4. Analysis of variance of test accuracy of six CNN algorithms using the augmented dataset (*n*=11,600).

Source	DF[a]	SS[b]	MS[c]	F-value	p-value
Algorithms	5	81.12	16.22	30.30	<.001
Error	24	12.85	0.53		
Total	29	93.98			

[a] Degrees of freedom

[b] Sum of squares

[c] Mean square

Table 4-5. Multiple means comparison for only interaction between algorithms and test accuracy using the augmented dataset (*n*=11,600).

Source	Mean	Grouping
ResNet-50	98.11±0.28[a]	A
SqueezeNet-MOD1	96.38±0.57	B
GoogLeNet	96.36±1.16	B
SqueezeNet	95.80±0.72	B
AlexNet	95.59±0.49	B
SqueezeNet-MOD2	92.61±0.84	C

[a] The standard deviation of five repetitions

Means that do not share a letter are significantly different

4.3.3. Computation time

Computation time is a crucial that needs to be considered in order to move real-time processing forward. Even though the popularity of DL and computational power is increasing, the amount of time required to train the algorithms still needs to be a consideration (Justus et al., 2018). The training and testing time in Table 4-6 were calculated as the training time per epoch and the number of epochs that should be performed to reach the desired level of accuracy. Training time is the time taken to train the algorithms with 9280 images, which were 80% of the total dataset, and the testing time of 2330 images, which were 20% of the total dataset.

In 4.3.2., ResNet-50 was recommended giving the best test accuracy/CA; however, in terms of computation time, it took an incredibly long time compared to other algorithms (Table 4-6). Therefore, when considering computation time, AlexNet performed fastest to detect PM on strawberry leaves. The training time of 80% data with AlexNet was 3451.88 seconds and the testing time of 20% data with AlexNet was 40.73 seconds (Table 4-6). Both training time and testing time in AlexNet were excelled as the fastest computation

time among six DL algorithms. As shown in Table 4-6, the order of CNN algorithms with the time taken is the accordance with the results in Bianco et al. (2018).

Shafiee et al. (2017) indicated that SqueezeNet is good for deployment on mobile hardware applications (Table 4-1). The results of the processing time with ResNet-50 is consistent with our expectation that it took a long time due to the model complexity. Due to the many layers and parameters, DL might require much longer training time compared to non-DL; however, the testing time is shorter than that of non-DL (Busseti et al., 2012).

Holistically, considering the test accuracy/CA, computation time, and memory for a hardware implementation, SqueezeNet-MOD1 and SqueezeNet-MOD2 would be recommended. Both have the smallest memory among other CNN algorithms presented in this study. Here, SqueezeNet-MOD1 was presented for high test accuracy/CA and SqueezeNet-MOD2 for short testing time with small memory respectively.

Table 4-6. Computation time of six CNN algorithms using the augmented dataset (n=11,600).

Deep learning algorithms	Training time (s)[a]	Testing time (s)[b]
AlexNet	**3451.88**±40.22	**40.73**±0.33
SqueezeNet	3555.40±261.15	73.40±3.81
GoogLeNet	9365.25±41.61	87.13±2.38
ResNet-50	21,738.29±420.16	178.20±8.51
SqueezeNet-MOD1	5141.80±27.31	68.55±3.99
SqueezeNet-MOD2	**3832.87**±95.46	**45.70**±0.85

[a] Testing time (seconds) for processing with 9280 data

[b] Testing time (seconds) for processing with 2320 data

***** The bold font shows the short computation time among the values which has no significant difference.

Using ANOVA, there was a significant difference ($p<.001$) between the testing time from CNN algorithms (Table 4-7); hence MMC was carried out. As a result of MMC,

there was a significant difference in terms of testing time among six CNN algorithms. In that case, AlexNet and SqueezeNet-MOD2 show the shortest computation time (Table 4-8). The recommended CNN algorithm varies depending on whether the purpose of the study is to use a small memory with higher test accuracy/CA or small memory with shorter testing time.

Table 4-7. Analysis of variance of computation time of six CNN algorithms using the augmented dataset (*n*=11,600).

Source	DF[a]	SS[b]	MS[c]	F-value	p-value
Algorithms	5	62771	12554.2	127.64	<.001
Error	24	2361	98.4		
Total	29	65131			

[a] Degrees of freedom

[b] Sum of squares

[c] Mean square

Table 4-8. Multiple means comparison for only interaction between algorithms and testing time using the augmented dataset (*n*=11,600).

Source	Mean	Grouping
ResNet-50	178.20±8.51[a]	A
GoogLeNet	87.13±2.38	B
SqueezeNet	73.40±3.81	B
SqueezeNet-MOD1	68.55±3.99	B
SqueezeNet-MOD2	45.70±0.85	C
AlexNet	40.73±0.33	C

[a] The standard deviation of five repetitions

Means that do not share a letter are significantly different

4.4. Conclusions

This study began with a comparison of non-DLs and CNNs with the original dataset (*n*=1450). As we expected, CNNs showed the better performance than non-DLs when the dataset increased and parameters were added. With CNN's superior justification, we

decided to focus on CNN algorithms. The best CNN algorithm was suggested in accordance with our research to develop a hardware system to detect PM and to spray fungicides on infected areas of strawberry leaves. The 11,600 data were constructed using eight different angles of rotation technique based on the original number of 1450 leaves. ResNet-50 showed the highest CA of 98.11% and as a result of statistical analysis, the test accuracy in the rest of the algorithms was not significantly different and was comparable to ResNet-50. When considering the high accuracy with small memory, SqueezeNet-MOD2 would be recommended. In terms of testing time, ResNet-50 required the most time with 178.20 seconds and AlexNet took the least time with 40.73 seconds. The testing time showed a significant difference among CNN algorithms. Hence, AlexNet and SqueezeNet-MOD2 would be recommended in order to process in a short time; however, when considering the small memory, the latter would be suggested.

The experimental results showed that the CNN technique is a promising tool and an easily deployable strategy for detecting PM on the strawberry leaves. Furthermore, this research could be further extended to develop a fully automated hardware (i.e., FPGA or mobile application) in order to help millions of producers who are struggling with PM disease. Future work will be on connecting these CNN algorithms with the hardware and developing the disease management platforms, that are easy for producers to use by providing accurate and fast results.

CHAPTER 5: CONCLUSIONS AND FUTURE RECOMMENDATIONS

5.1. Conclusions

The overall goal of this study was to detect powdery mildew (PM) on strawberry leaves with the highest classification accuracy (CA) and shortest computation time. In order to reach the final goal, three objectives were presented step by step. The first objective was to detect PM by using image processing and supervised machine learning (ML). The combination of three image processing techniques (histogram of oriented gradients; HOG, speeded up robust features; SURF, and gray level co-occurrence matrix; GLCM) and two supervised ML (artificial neural network; ANN and support vector machine; SVM) was tested and compared in different image resolutions (227×227, 454×454, 681×681, 908×908, and 1135×1135) and cell sizes (32×32, 64×64, and 128×128; only for HOG). As a result of the first objective, the combination of SURF and ANN showed the highest CA of 94.34% with 908×908 image resolution and the combination of GLCM and SVM showed the highest CA of 88.98% with 908×908 image resolution when considering only highest CA. Also, there was a significant difference between image resolutions that were seen only in a combination of HOG and ANN, HOG and SVM, SURF and SVM, GLCM and ANN. Therefore, an additional result analysis was performed to reduce computational complexity by selecting the smallest image resolution among the combinations that do not have a significant difference in CA. The combination of SURF and ANN showed the highest CA of 92.66% with 227×227 image resolution and the combination of GLCM and SVM showed the highest CA of 86.36% with 227×227 image resolution.

The second objective was to evaluate the effects of data augmentation in supervised ML by rotating the images by using 90°, 180°, and 270°. The 254 images were increased to 1016 by applying an angular rotation technique. The number of data is important in training the image data due to the overfitting problem and data normalisation. To simulate strawberry leaf bundles in the real field, data augmentation with various directions was necessary. As a result of the second objective, the effects on data augmentation were challenging in accordance with the tendency not to show improvement after a certain point even if data increases in traditional ML unlike deep learning (DL).

For a plentiful results analysis, a comparison between supervised non-DLs and CNNs was carried out. CA values were compared by using an original dataset (n=1450) in CHAPTER 4 in 10 ML techniques including non-DLs and CNNs. Non-DLs were decided on the combinations that were chosen as the optimized feature extraction techniques and supervised ML in CHAPTER 3. For the uniformity of image resolution, it will be compared with the combination considering statistical analysis and computational complexity, not the combination just showing the highest CA. The combination of SURF and ANN showed the CA of 85.44% with 227×227 image resolution and the combination of GLCM and ANN showed the CA of 69.28% with 227×227 image resolution. Compared to non-DLs, CNNs results showed much higher CA. As we learned from section 3.3.2., the performance of non-DLs tended to decrease as the dataset increased. Results were overfitting with a small dataset and results were underfitting with a larger dataset. In conclusion, CNN techniques were much superior and robust in added parameters than non-DLs with large datasets.

The third objective was the detection of PM on the strawberry leaves by using six CNN algorithms based on an augmented dataset(n=11,600) to increase the CA and to reduce the testing time. Image dataset was acquired during the summer in 2018 and 2019, and DL normally required a huge dataset of at least 10,000; hence, the total number of images was increased by eight different angles of rotation technique. The four famous algorithms (AlexNet, SqueezeNet, GoogLeNet, and ResNet-50) with a low error rate were used in this study. Also, SqueezeNet-MOD1 and SqueezeNet-MOD2 were developed by the author to improve the performance compared to SqueezeNet 1.1 version. The CA was acquired as the average value from the five repetitions with data that were selected randomly each time. Among six CNN algorithms, ResNet-50 showed the best performance in terms of all evaluation metrics (test accuracy, precision, sensitivity, specificity, and F1-score). After ResNet-50, the performance was good in the order of SqueezeNet-MOD1, GoogLeNet, SqueezeNet, AlexNet, and SqueezeNet-MOD2. As a result of the statistical analysis, except for ResNet-50, there was no significant difference in SqueezeNet-MOD1, GoogLeNet, SqueezeNet, and AlexNet. The next thing to consider was the computation time. ResNet-50 took the longest time and AlexNet took the shortest time. There was no significant difference in testing time between GoogLeNet, SqueezeNet, and SqueezeNet-MOD1. Also, there was no significant difference in testing time between SqueezeNet-MOD2 and AlexNet. When considering applying algorithms to hardware, SqueezeNet-MOD2 would be recommended due to shorter testing time and smaller memory than AlexNet.

5.2. Future recommendations

The overall study was designed to reduce the burden on producers of strawberries. As a starting point for this big project, we developed advanced algorithms to detect PM and classify healthy and infected leaves using image processing, non-DL and CNN techniques. Future work will be to combine the algorithm (SqueezeNet-MOD2) that was suggested in conclusions with hardware (i.e., FPGA or mobile applications) and implement it in the actual field. To conduct the spot application, real-time kinematics-global positioning system (RTK-GPS) would be applied to carry out mapping PM points on the field. Furthermore, a connecting nozzle with this system could be a great help to control the PM on the strawberry leaves.

In addition to PM detection, the development of the algorithms in detection on flowers which will turn into strawberries and crown rot would be recommended. The degree of flowering can be used as an indication to predict strawberry yields for the season and to prepare for how much labour producers need. Therefore, the development of algorithms that detect flowering is recommended. The strawberry farms in Nova Scotia also have a large export business to the United States of daughter strawberry stems, which grows from mother strawberry. The red colour of the daughter strawberry stem stands for malnutrition or burning due to the black colour of the plastic mulch. It is anticipated that more effective and efficient farm operations would be possible if the recommended algorithms are integrated into hardware.

REFERENCES

Adam, L., & Somerville, S. C. (1996). Genetic characterization of five powdery mildew disease resistance loci in Arabidopsis thaliana. *The Plant Journal, 9(3),* 341-356.

Ahmed, F., Al-Mamun, H. A., Bari, A. H., Hossain, E., & Kwan, P. (2012). Classification of crops and weeds from digital images: A support vector machine approach. *Crop Protection*, 40, 98-104.

Albregtsen, F. (2008). Statistical texture measures computed from gray level coocurrence matrices. *Image processing laboratory, department of informatics, university of oslo, 5(5).*

Alexandratos, N., Bruinsma, J., Bödeker, G., Schmidhuber, J., Broca, S., Shetty, P., & Ottaviani, M. G. (2006). World agriculture: towards 2030/2050. Interim report. *Rome: Global Perspective Studies Unit, Food and Agriculture Organization of the United Nations (FAO).*

Alhichri, H., Alajlan, N., Bazi, Y., & Rabczuk, T. (2018, May). Multi-Scale Convolutional Neural Network for Remote Sensing Scene Classification. In *2018 IEEE International Conference on Electro/Information Technology (EIT)* (pp. 1-5). IEEE.

Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., ... & Asari, V. K. (2018). The history began from alexnet: A comprehensive survey on deep learning approaches. *arXiv preprint arXiv:1803.01164.*

Aneece, I., & Thenkabail, P. (2018). Accuracies Achieved in Classifying Five Leading World Crop Types and their Growth Stages Using Optimal Earth Observing-1 Hyperion Hyperspectral Narrowbands on Google Earth Engine. *Remote Sensing*, 10(12), 2027.

Asad, M. H., & Bais, A. (2019). Weed Detection in Canola Fields Using Maximum Likelihood Classification and Deep Convolutional Neural Network. *Information Processing in Agriculture.*

Ashqar, B. A., & Abu-Naser, S. S. (2019). Image-Based Tomato Leaves Diseases Detection Using Deep Learning.

Atole, R. R., & Park, D. (2018). A multiclass deep convolutional neural network classifier for detection of common rice plant anomalies. *INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS*, 9(1), 67-70.

Awate, A., Deshmankar, D., Amrutkar, G., Bagul, U., & Sonavane, S. (2015, October). Fruit disease detection using color, texture analysis and ANN. In *2015 International Conference on Green Computing and Internet of Things (ICGCIOT)* (pp. 970-975). IEEE.

Bah, M. D., Hafiane, A., & Canals, R. (2018). Deep learning with unsupervised data labeling for weeds detection on uav images. *arXiv preprint arXiv:1805.12395.*

Baraldi, A., & Parmiggiani, F. (1995). An Investigation of the Textural Characteristics Associated with Gray Level Cooccurrence Matrix Statistical Parameters. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, 33(2).

Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008). Speeded-up robust features (SURF). *Computer vision and image understanding*, *110*(3), 346-359.

Bay, H., Tuytelaars, T., & Van Gool, L. (2006, May). Surf: Speeded up robust features. In *European conference on computer vision* (pp. 404-417). Springer, Berlin, Heidelberg.

Becherer, N. C. (2017). *Transfer Learning in Convolutional Neural Networks for Fine-Grained Image Classification* (No. AFIT-ENG-MS-17-M-005). AIR FORCE INSTITUTE OF TECHNOLOGY WRIGHT-PATTERSON AFB OH WRIGHT-PATTERSON AFB United States.

Behmann, J., Mahlein, A. K., Rumpf, T., Römer, C., & Plümer, L. (2015). A review of advanced machine learning methods for the detection of biotic stress in precision crop protection. *Precision Agriculture*, 16(3), 239-260.

Beiping, H., & Wen, Z. (2011). Fast Human Detection Using Motion Detection and Histogram of Oriented Gradients. *JCP*, 6(8), 1597-1604.

Bhange, M., & Hingoliwala, H. A. (2015). Smart farming: Pomegranate disease detection using image processing. *Procedia Computer Science, 58,* 280-288.

Bianco, S., Cadene, R., Celona, L., & Napoletano, P. (2018). Benchmark analysis of representative deep neural network architectures. *IEEE Access, 6*, 64270-64277.

Blasco, J., Aleixos, N., & Moltó, E. (2003). Machine vision system for automatic quality grading of fruit. *Biosystems engineering*, 85(4), 415-423.

Bloice, M. D., Stocker, C., & Holzinger, A. (2017). Augmentor: an image augmentation library for machine learning. *arXiv preprint arXiv:1708.04680.*

Brahimi, M., Boukhalfa, K., & Moussaoui, A. (2017). Deep learning for tomato diseases: classification and symptoms visualization. *Applied Artificial Intelligence, 31(4),* 299-315.

Bressem, K. K., Adams, L., Erxleben, C., Hamm, B., Niehues, S., & Vahldiek, J. (2020). Comparing Different Deep Learning Architectures for Classification of Chest Radiographs. arXiv preprint arXiv:2002.08991.

Bushnell, W. R., & Allen, P. J. (1962). Induction of disease symptoms in barley by powdery mildew. *Plant physiology, 37(1),* 50.

Busseti, E., Osband, I., & Wong, S. (2012). Deep learning for time series modeling. *Technical report, Stanford University,* 1-5.

Carisse, O., Morissette-Thomas, V., & Van der Heyden, H. (2013). Lagged association between powdery mildew leaf severity, airborne inoculum, weather, and crop losses in strawberry. *Phytopathology*, 103(8), 811-821.

Chen, Y., Lin, Z., Zhao, X., Wang, G., & Gu, Y. (2014). Deep learning-based classification of hyperspectral data. *IEEE Journal of Selected topics in applied earth observations and remote sensing, 7(6),* 2094-2107.

Cheng, Y., Zhao, X., Huang, K., & Tan, T. (2014, August). Semi-supervised learning for rgb-d object recognition. In *2014 22nd International Conference on Pattern Recognition* (pp. 2377-2382). IEEE.

Cruz, A., Ampatzidis, Y., Pierro, R., Materazzi, A., Panattoni, A., De Bellis, L., & Luvisi, A. (2019). Detection of grapevine yellows symptoms in Vitis vinifera L. with artificial intelligence. *Computers and electronics in agriculture*, 157, 63-76.

Cubero, S., Aleixos, N., Moltó, E., Gómez-Sanchis, J., & Blasco, J. (2011). Advances in machine vision applications for automatic inspection and quality evaluation of fruits and vegetables. *Food and bioprocess technology, 4(4),* 487-504.

Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. In *International Conference on Computer Vision & Pattern Recognition (CVPR'05)* (Vol. 1, pp. 886-893). IEEE Computer Society.

Dandawate, Y., & Kokare, R. (2015, August). An automated approach for classification of plant diseases towards development of futuristic Decision Support System in Indian perspective. In *2015 International conference on advances in computing, communications and informatics (ICACCI)* (pp. 794-799). IEEE.

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition* (pp. 248-255). Ieee.

Deng, L., & Platt, J. C. (2014). Ensemble deep learning for speech recognition. In Fifteenth Annual Conference of the International Speech Communication Association.

Durmuş, H., Güneş, E. O., & Kırcı, M. (2017, August). Disease detection on the leaves of the tomato plants by using deep learning. In *2017 6th International Conference on Agro-Geoinformatics* (pp. 1-5). IEEE.

Eleyan, A., & Demirel, H. (2011). Co-occurrence matrix and its statistical features as a new approach for face recognition. *Turkish Journal of Electrical Engineering & Computer Sciences*, 19(1), 97-107.

Esau, T. J., Zaman, Q. U., Chang, Y. K., Schumann, A. W., Percival, D. C., & Farooque, A. A. (2014). Spot-application of fungicide for wild blueberry using an automated prototype variable rate sprayer. *Precision agriculture*, *15(2),* 147-161.

Fadaee, M., Bisazza, A., & Monz, C. (2017). Data augmentation for low-resource neural machine translation. *arXiv preprint arXiv:1705.00440.*

Fira, C. M., & Goras, L. (2008). An ECG signals compression method and its validation using NNs. *IEEE Transactions on Biomedical Engineering, 55(4),* 1319-1326.

Fuentes, A., Yoon, S., Kim, S., & Park, D. (2017). A robust deep-learning-based detector for real-time tomato plant diseases and pest recognition. *Sensors*, 17(9), 2022.

Garud, P. S., & Devi, R. (2017). Detection of Diseases on Plant Leaf with the Help of Image Processing. *International Journal of Engineering Technology Science and Research (IJETSR),* 912-918.

Gavhale, K. R., Gawande, U., & Hajari, K. O. (2014, April). Unhealthy region of citrus leaf detection using image processing techniques. In *International Conference for Convergence for Technology-2014* (pp. 1-6). IEEE.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

Heung, B., Ho, H. C., Zhang, J., Knudby, A., Bulmer, C. E., & Schmidt, M. G. (2016). An overview and comparison of machine-learning techniques for classification purposes in digital soil mapping. *Geoderma*, 265, 62-77.

Howarth, P., & Rüger, S. (2004, July). Evaluation of texture features for content-based image retrieval. In *International conference on image and video retrieval* (pp. 326-334). Springer, Berlin, Heidelberg.

Hu, M. H., Dong, Q. L., Liu, B. L., & Malakar, P. K. (2014). The potential of double K-means clustering for banana image segmentation. *Journal of Food Process Engineering, 37(1),* 10-18.

Huang, K. Y. (2007). Application of artificial neural network for detecting Phalaenopsis seedling diseases using color and texture features. *Computers and Electronics in agriculture, 57(1),* 3-11.

Iandola, F. N., Han, S., Moskewicz, M. W., Ashraf, K., Dally, W. J., & Keutzer, K. (2016). SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size. *arXiv preprint arXiv:1602.07360.*

Jacob, D., David, D. R., Sztjenberg, A., & Elad, Y. (2008). Conditions for development of powdery mildew of tomato caused by Oidium neolycopersici. *Phytopathology, 98(3),* 270-281.

Jain, A. K., Mao, J., & Mohiuddin, K. M. (1996). Artificial neural networks: A tutorial. *Computer*, *29*(3), 31-44.

Jain, R., Kasturi, R., & Schunck, B. G. (1995). *Machine vision* (Vol. 5, pp. 309-364). New York: McGraw-Hill.

Jang, E., Gu, S., & Poole, B. (2016). Categorical reparameterization with gumbel-softmax. *arXiv preprint arXiv:1611.01144.*

Jhuria, M., Kumar, A., & Borse, R. (2013, December). Image processing for smart farming: Detection of disease and fruit grading. In *2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013)* (pp. 521-526). IEEE.

Joachims, T. (1998, April). Text categorization with support vector machines: Learning with many relevant features. In *European conference on machine learning* (pp. 137-142). Springer, Berlin, Heidelberg.

Justus, D., Brennan, J., Bonner, S., & McGough, A. S. (2018, December). Predicting the Computational Cost of Deep Learning Models. In 2018 *IEEE International Conference on Big Data (Big Data)* (pp. 3873-3882). IEEE.

Kalia, A., & Gosal, S. K. (2011). Effect of pesticide application on soil microorganisms. Archives of Agronomy and Soil Science, 57(6), 569-596.

Kamilaris, A., & Prenafeta-Boldu, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture, 147*, 70-90.

Karlik, B., & Olgac, A. V. (2011). Performance analysis of various activation functions in generalized MLP architectures of neural networks. *International Journal of Artificial Intelligence and Expert Systems, 1(4),* 111-122.

Kearney, J. (2010). Food consumption trends and drivers. Philosophical transactions of the royal society B: *biological sciences*, 365(1554), 2793-2807.

Keskar, N. S., Mudigere, D., Nocedal, J., Smelyanskiy, M., & Tang, P. T. P. (2016). On large-batch training for deep learning: Generalization gap and sharp minima. *arXiv preprint arXiv:1609.04836.*

Khaled, A. Y., Abd Aziz, S., Bejo, S. K., Nawi, N. M., Seman, I. A., & Onwude, D. I.

(2018). Early detection of diseases in plant tissue using spectroscopy–applications and limitations. *Applied Spectroscopy Reviews, 53(1),* 36-64.

Kobayashi, T., Kanda, E., Kitada, K., Ishiguro, K., & Torigoe, Y. (2001). Detection of rice panicle blast with multispectral radiometer and the potential of using airborne multispectral scanners. *Phytopathology*, 91(3), 316-323.

Kotsiantis, S. B., Zaharakis, I., & Pintelas, P. (2007). Supervised machine learning: A review of classification techniques. Emerging artificial intelligence applications in computer engineering, 160, 3-24.

Kovach, J., Petzoldt, R., & Harman, G. E. (2000). Use of honeybees and bumble bees to disseminate Trichoderma harzianum 1295-22 to strawberries for Botrytis control. *Biological Control, 18(3)*, 235-242.

Krizhevsky, A., & Sutskever, I. (2012). H. Geoffrey E.,"Alex Net,". *Adv. Neural Inf. Process. Syst*, *25*, 1-9.

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

LeCun, Y. (2015). LeNet-5, convolutional neural networks. URL: http://yann. lecun. com/exdb/lenet, 20, 5.

LeCun, Y., & Bengio, Y. (1995). Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks, 3361(10),* 1995.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436.

Li, D., Yang, W., & Wang, S. (2010). Classification of foreign fibers in cotton lint using machine vision 842 and multi-class support vector machine. *Computers and electronics in agriculture,* 74(2), 274-279.

Liaghat, S., & Balasundram, S. K. (2010). A review: The role of remote sensing in precision agriculture. *American journal of agricultural and biological sciences*, 5(1), 50-55.

Liao, B., Xu, J., Lv, J., & Zhou, S. (2015). An image retrieval method for binary images based on DBN and softmax classifier. IETE Technical Review, 32(4), 294-303.

Liu, B. (2017). *Sustainable strawberry production and management including control of strawberry powdery mildew* (Doctoral dissertation, University of Hertfordshire).

Liu, B., Zhang, Y., He, D., & Li, Y. (2018). Identification of apple leaf diseases based on deep convolutional neural networks. *Symmetry, 10(1),* 11.

Lowe, D. G. (1999, September). Object recognition from local scale-invariant features. In *Proceedings of the seventh IEEE international conference on computer vision* (Vol. 2, pp. 1150-1157). Ieee.

Manevitz, L. M., & Yousef, M. (2001). One-class SVMs for document classification. *Journal of machine learning research*, 2(Dec), 139-154.

Mayuri, K. P., & Priya, C. H. (2018). Role of image processing and machine learning techniques in disease recognition, diagnosis and yield prediction of crops: a review. *International Journal of Advanced Research in Computer Science*, *9*(2).

McGrath, M. T. (2001). Fungicide resistance in cucurbit powdery mildew: experiences and challenges. *Plant disease*, 85(3), 236-245.

McQueen, R. J., Garner, S. R., Nevill-Manning, C. G., & Witten, I. H. (1995). Applying machine learning to agricultural data. *Computers and electronics in agriculture*, 12(4), 275-293.

Meunkaewjinda, A., Kumsawat, P., Attakitmongcol, K., & Srikaew, A. (2008, May). Grape leaf disease detection from color imagery using hybrid intelligent system. In *2008 5th international conference on electrical engineering/electronics, computer, telecommunications and information technology* (Vol. 1, pp. 513-516). IEEE.

Michie, D., Spiegelhalter, D. J., & Taylor, C. C. (1994). Machine learning. *Neural and Statistical Classification*, 13(1994), 1-298.

Mistry, D., & Banerjee, A. (2017). Comparison of feature detection and matching approaches: SIFT and SURF. *GRD Journals-Global Research and Development Journal for Engineering*, 2(4), 7-13.

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M. (2013). Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*.

Mohanaiah, P., Sathyanarayana, P., & GuruKumar, L. (2013). Image texture feature extraction using GLCM approach. *International journal of scientific and research publications*, 3(5), 1.

Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in plant science*, 7, 1419.

Mueller, D. S. (2006). Fungicides: Terminology.

Nelson, M. D., Gubler, W. D., & Shaw, D. V. (1995). Inheritance of powdery mildew resistance in greenhouse-grown versus field-grown California strawberry progenies. *Phytopathology,* 85(4), 421-424.

O'Shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458.*

Pan, L., Zhang, W., Zhu, N., Mao, S., & Tu, K. (2014). Early detection and classification of pathogenic fungal disease in post-harvest strawberry fruit by electronic nose and gas chromatography–mass spectrometry. *Food research international, 62*, 162-168.

Patil, J. K., & Kumar, R. (2011). Advances in image processing for detection of plant diseases. *Journal of Advanced Bioinformatics Applications and Research*, 2(2), 135-141.

Paulus, A. O. (1990). Fungal diseases of strawberry. *HortScience*, 25(8), 885-889.

Pauly, L., & Sankar, D. (2016). Non Intrusive Eye Blink Detection from Low Resolution Images Using HOG-SVM Classifier. *International Journal of Image, Graphics & Signal Processing, 8(10).*

Perez, L., & Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning. *arXiv preprint arXiv:1712.04621.*

Perini, A., & Susi, A. (2004). Developing a decision support system for integrated production in agriculture. *Environmental Modelling & Software, 19(9),* 821-829.

Peterson, A. T., & Cohoon, K. P. (1999). Sensitivity of distributional prediction algorithms to geographic data completeness. *Ecological modelling, 117(1),* 159-164.

Pires, R. D. L., Gonçalves, D. N., Oruê, J. P. M., Kanashiro, W. E. S., Rodrigues Jr, J. F., Machado, B. B., & Gonçalves, W. N. (2016). Local descriptors for soybean disease recognition. *Computers and Electronics in Agriculture,* 125, 48-55.

Ploetz, R. C. (1994). Panama disease: return of the first banana menace. *International journal of pest management, 40(4)*, 326-336.

Pujari, J. D., Yakkundimath, R., & Byadgi, A. S. (2015). Image processing based detection of fungal diseases in plants. *Procedia Computer Science, 46,* 1802-1808.

Rahman, H., Ch, N. J., Manzoor, S., Najeeb, F., Siddique, M. Y., & Khan, R. A. (2017). A comparative analysis of machine learning approaches for plant disease identification. *Advancements in Life Sciences, 4(4),* 120-126.

Ramcharan, A., Baranowski, K., McCloskey, P., Ahmed, B., Legg, J., & Hughes, D. P. (2017). Deep learning for image-based cassava disease detection. *Frontiers in plant science, 8*, 1852.

Rangayyan, R. M., Nguyen, T. M., Ayres, F. J., & Nandi, A. K. (2010). Effect of pixel resolution on texture features of breast masses in mammograms. *Journal of digital imaging, 23(5),* 547-553.

Rasti, P., Ahmad, A., Samiei, S., Belin, E., & Rousseau, D. (2019). Supervised Image Classification by Scattering Transform with Application to Weed Detection in Culture Crops of High Density. *Remote Sensing, 11(3),* 249.

Rehman, T. U., Mahmud, M. S., Chang, Y. K., Jin, J., & Shin, J. (2019). Current and future applications of statistical machine learning algorithms for agricultural machine vision systems. Computers and Electronics in Agriculture*, 156*, 585-605.

Revathi, P., Revathi, R., & Hemalatha, M. (2011). Comparative study of knowledge in Crop diseases using Machine Learning Techniques. *Inter-national Journal of Computer Science and Information Technologies (IJCSIT)*, *2*(5), 2180-2182.

Robert, P. C. (2002). Precision agriculture: a challenge for crop nutrition management. In Progress in Plant Nutrition: Plenary Lectures of the XIV International Plant Nutrition Colloquium (pp. 143-149). *Springer,* Dordrecht.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... & Berg, A. C. (2015). Imagenet large scale visual recognition challenge. *International journal of computer vision, 115(3)*, 211-252.

Russell, S. J., & Norvig, P. (2016). *Artificial intelligence: a modern approach.* Malaysia; Pearson Education Limited.

Salamon, J., & Bello, J. P. (2017). Deep convolutional neural networks and data augmentation for environmental sound classification. *IEEE Signal Processing Letters, 24(3),* 279-283.

Savary, S., Ficke, A., Aubertot, J. N., & Hollier, C. (2012). Crop losses due to diseases and their implications for global food production losses and food security.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61, 85-117.

Scholes, J. D., Lee, P. J., Horton, P., & Lewis, D. H. (1994). Invertase: understanding changes in the photosynthetic and carbohydrate metabolism of barley leaves infected with powdery mildew. *New Phytologist,* 126(2), 213-222.

Scholkopf, B., & Smola, A. J. (2001). *Learning with kernels: support vector machines, regularization, optimization, and beyond.* MIT press.

Sebastian, V., Unnikrishnan, A., & Balakrishnan, K. (2012). Gray level co-occurrence matrices: generalisation and some new features. *arXiv preprint arXiv:1205.4831.*

Shafiee, M. J., Li, F., Chwyl, B., & Wong, A. (2017). SquishedNets: Squishing SqueezeNet further for edge device scenarios via deep evolutionary synthesis. *arXiv preprint arXiv:1711.07459.*

Shin, J., Chang, Y. K., Heung, B., Nguyen-Quang, T., Price, G. W., & Al-Mallahi, A. (2020). Effect of directional augmentation using supervised machine learning technologies: A case study of strawberry powdery mildew detection. *Biosystems Engineering*, 194, 49-60.

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.*

Soh, L. K., & Tsatsoulis, C. (1999). Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices*. IEEE Transactions on geoscience and remote sensing,*

37(2), 780-795.

Sonka, M., Hlavac, V., & Boyle, R. (2014). *Image processing, analysis, and machine vision.* Cengage Learning.

Statistics Canada. (2017). Statistical Overview of the Canadian Fruit Industry.

Statistics Canada. (2020a). Table 32-10-0169-01 Number of farm operators by sex, age and paid non-farm work, historical data.

Statistics Canada. (2020b). Area, production, and farm gate value of marketed fruits. Retrieved from Statistics Canada website: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3210036401&pickMembers%5B0%5D=1.6&pickMembers%5B1%5D=2.1&pickMembers%5B2%5D=4.4

Statistics Canada. (2020c). Table 001-0009 8 Area, Production and Farm Gate Value of Fresh and Processed Fruits.

Steger, C., Ulrich, M., & Wiedemann, C. (2018). *Machine vision algorithms and applications*. John Wiley & Sons.

Steward, B. L., & Tian, L. F. (1999). Machine-vision weed density estimation for real-time, outdoor lighting conditions. *Transactions of the ASAE*, 42(6), 1897.

Stockwell, D. R. (1997). Generic predictive systems: an empirical evaluation using the Learning Base System (LBS). *Expert Systems with Applications, 12(3),* 301-310.

Strange, R. N., & Scott, P. R. (2005). Plant disease: a threat to global food security. *Annual review of phytopathology*, *43*. 83-116.

Sutskever, I., Martens, J., Dahl, G., & Hinton, G. (2013, February). On the importance of initialization and momentum in deep learning. In *International conference on machine learning* (pp. 1139-1147).

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).

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818-2826).

Taylor, L., & Nitschke, G. (2017). Improving deep learning using generic data augmentation. *arXiv preprint arXiv:1708.06020.*

Teke, M., Deveci, H. S., Haliloğlu, O., Gürbüz, S. Z., & Sakarya, U. (2013, June). A short survey of hyperspectral remote sensing applications in agriculture. In *2013 6th International Conference on Recent Advances in Space Technologies (RAST)* (pp. 171-176). IEEE.

Teng, P. S., & James, W. C. (2002). Disease and yield loss assessment. *Plant Pathologist's Pocketbook 3rd Edition,* 35.

Teng, P. S., & Krupa, S. V. (1980). Crop loss assessment.

The Weather Network. (2020). Retrieved from https://www.theweathernetwork.com/ca/weather/nova-scotia/debert

Thompson, S. (2016, September 21). Four threats to global food security and what we can do about them. Retrieved from http://theconversation.com/four-threats-to-global-food-security-and-what-we-can-do-about-them-65676

United Nations. (2015). World population prospects: The 2015 revision. United Nations Econ Soc Aff, 33(2), 1-66.

Vidal, M., & Amigo, J. M. (2012). Pre-processing of hyperspectral images. Essential steps before image analysis. *Chemometrics and Intelligent Laboratory Systems*, *117*, 138-148.

Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., & Manzagol, P. A. (2010). Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *Journal of machine learning research*, *11*(Dec), 3371-3408.

Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., & Summers, R. M. (2017). Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2097-2106).

Watanabe, T., Ito, S., & Yokoi, K. (2009, January). Co-occurrence histograms of oriented gradients for pedestrian detection. In *Pacific-Rim Symposium on Image and Video Technology* (pp. 37-47). Springer, Berlin, Heidelberg.

Wspanialy, P., & Moussa, M. (2016). Early powdery mildew detection system for application in greenhouse automation. *Computers and Electronics in Agriculture*, 127, 487-494.

Xu, G., Wu, X., Liu, L., & Wu, Z. (2011, June). Real-time pedestrian detection based on edge factor and Histogram of Oriented Gradient. In *2011 IEEE International Conference on Information and Automation* (pp. 384-389). IEEE.

Xue, J., Zhang, L., & Grift, T. E. (2012). Variable field-of-view machine vision-based row guidance of an agricultural robot. *Computers and Electronics in Agriculture*, 84, 85-91.

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

Yao, Q., Guan, Z., Zhou, Y., Tang, J., Hu, Y., & Yang, B. (2009, May). Application of support vector machine for detecting rice diseases using shape and color texture features. In *2009 international conference on engineering computation* (pp. 79-83). IEEE.

Zhang, C., Bengio, S., Hardt, M., Recht, B., & Vinyals, O. (2016). Understanding deep learning requires rethinking generalization. *arXiv preprint arXiv:1611.03530.*

Zhang, J. C., Pu, R. L., Wang, J. H., Huang, W. J., Yuan, L., & Luo, J. H. (2012). Detecting powdery mildew of winter wheat using leaf level hyperspectral measurements. *Computers and Electronics in Agriculture*, 85, 13-23.

Zhang, N., Wang, M., & Wang, N. (2002). Precision agriculture—a worldwide overview. *Computers and electronics in agriculture*, 36(2-3), 113-132.

Zhu, Q., Yeh, M. C., Cheng, K. T., & Avidan, S. (2006). Fast human detection using a cascade of histograms of oriented gradients. In *2006 IEEE Computer Society*

Conference on Computer Vision and Pattern Recognition (CVPR'06) (Vol. 2, pp. 1491-1498). IEEE.

APPENDIX

March 25, 2020

Journal Article / Ticket number 200325-013021
Biosystems Engineering
ELSEVIER PUBLISHERS

Dear Sir/Madam,

I am preparing my M. Sc thesis for submission to the Faculty of Graduate Studies at Dalhousie University, Halifax, Nova Scotia, Canada. I am asking your permission to include a manuscript version of the following paper as a chapter in the thesis:

Effect of directional augmentation using supervised machine learning technologies – A case study of strawberry powdery mildew detection

Your sincerely,

Jaemyung Shin
Dalhousie University

ELSEVIER

Dear Jaemyung Shin,

Thank you for your query.

Please note that, as one of the authors of this article, you retain the right to reuse it in your thesis/dissertation. You do not require formal permission to do so. You are permitted to post this Elsevier article online if it is embedded within your thesis. You are also permitted to post your Author Accepted Manuscript online.

However posting of the final published article is prohibited.

"As per our Sharing Policy, authors are permitted to post the Accepted version of their article on their institutional repository – as long as it is for ***internal institutional use only.***

It can only be shared publicly on that site once the journal-specific embargo period has lapsed. For a list of embargo periods please see: Embargo List.

You are not permitted to post the Published Journal Article (PJA) on the repository."

Please feel free to contact me if you have any queries.

Regards,
Prem

Permissions Helpdesk
ELSEVIER |Operations