Comparison of Image Texture Based Supervised Learning Classifiers for Strawberry Powdery Mildew Detection

Young K. Chang [1],* , **Md. Sultan Mahmud [1], Jaemyung Shin [1], Tri Nguyen-Quang [1], Gordon W. Price [1] and Balakrishnan Prithiviraj [2]**

[1] Department of Engineering, Faculty of Agriculture, Dalhousie University, Truro, NS B2N 5E3, Canada
[2] Department of Plant, Food, and Environmental Sciences, Faculty of Agriculture, Dalhousie University, Truro, NS B2N 5E3, Canada
* Correspondence: youngchang@dal.ca

Abstract: Strawberry is an important fruit crop in Canada but powdery mildew (PM) results in about 30–70% yield loss. Detection of PM through an image texture-based system is beneficial, as it identifies the symptoms at an earlier stage and reduces labour intensive manual monitoring of crop fields. This paper presents an image texture-based disease detection algorithm using supervised classifiers. Three sites were selected to collect the leaf image data in Great Village, Nova Scotia, Canada. Images were taken under an artificial cloud condition with a Digital Single Lens Reflex (DSLR) camera as red-green-blue (RGB) raw data throughout 2017–2018 summer. Three supervised classifiers, including artificial neural networks (ANN), support vector machine (SVM), and k-nearest neighbors (kNN) were evaluated for disease detection. A total of 40 textural features were extracted using a colour co-occurrence matrix (CCM). The collected feature data were normalized, then used for training and internal, external and cross-validations of developed classifiers. Results of this study revealed that the highest overall classification accuracy was 93.81% using the ANN classifier and lowest overall accuracy was 78.80% using the kNN classifier. Results identified the ANN classifier disease detection having a lower Root Mean Square Error (RMSE) = 0.004 and Mean Absolute Error (MAE) = 0.003 values with 99.99% of accuracy during internal validation and 87.41%, 88.95% and 95.04% of accuracies during external validations with three different fields. Overall results demonstrated that an image texture-based ANN classifier was able to classify PM disease more accurately at early stages of disease development.

Keywords: image processing; colour co-occurrence matrix; artificial neural network; support vector machine; k-nearest neighbors

1. Introduction

Strawberry powdery mildew (*Sphaerotheca macularis*) is a polycyclic disease, caused by a fungal pathogen, which affects petioles, leaves, runners, flowers, and fruits that appear to be specific to strawberry plants [1,2]. Particularly, powdery mildew (PM) disease begins with white powder on the leaf surface, so the front side of the leaf images are needed to detect the disease [3]. Healthy leaves are saw-tooth shape in the edge with dark green colour and have no discoloration or defect on the leaf surface. However, strawberry leaves infected with PM are deformed as the edge faces upward. The edges of the leaves have a purplish hue, and the underside and upside of the leaves appear to have been coated with fine white powder [4]. The disease also impacts the plant's photosynthetic ability influencing fruit quality, growth potential, and productivity [5]. Plant fruit quality and yield are closely

tied to appropriate disease management and control of this disease. The ability to detect disease at early stages when less than 10 white spots begin to appear is essential to be able to apply suitable controls in order to decrease impacts on fruit quality [6]. Plant diseases have unique developmental characteristics and behaviors that can aid in their detection [7]. Although the application of fungicides is the most effective way to control the disease of the crops, it is desirable to minimize the use of fungicides as cost and environmental problems increase for growers. Therefore, understanding the distribution and severity of disease before applying fungicides is very useful information for growers [8]. Development of accurate and rapid techniques to detect plant diseases is of critical importance to the fruit and crop industry.

Image processing-based approaches can be fast and accurate in detecting plant diseases [9]. Digital images form important data units that can be analyzed to generate key pieces of information across a range of applications [10]. Since the early 2000s, imaging techniques such as hyperspectral, multispectral, thermal, and colour imaging have been developed to solve various problems in agriculture. In terms of hyperspectral imaging, Qin et al. [11] extracted the information from the spectrum which was used to determine the health of the crops. The ultra-spectral images of 450–930 nm was utilized to distinguish the canker and other damages on the Ruby red grapefruit with 96% accuracy. Rumpf et al. [12] automatically detected the early disease of the sugar beet leaves with 97% of the classification accuracy by using SVM based on hyperspectral reflectance. In terms of multispectral imaging, Laudien et al. [13] found that red spectra from 630 nm to 690 nm and near infrared spectra from 760 nm to 900 nm are important areas in agricultural applications. They detected and analyzed a fungal sugar beet disease with high resolution multispectral and hyperspectral remote sensing data. In terms of thermal imaging, Chaerle et al. [14] reported that heat around the tobacco mosaic virus (TMV) disease spots on the leaves was monitored before the disease symptoms appeared on the tobacco leaves, and thermal imaging technique was used to visualize that heat in areas infected with TMV. Besides, the red-green-blue (RGB) colour coordinates are cost-effective and the most general colour system. Kutty et al. [15] extracted regions of interest from the RGB colour model and categorized watermelon anthracnose and downy mildew leaf disease by using neural network. Khirade and Patil [16] converted RGB images into Hue, Saturation, and Intensity (HSI) and used them for disease identification. Green pixels were recognized using k-means clustering, and various threshold values were obtained using the Otsu's method. Kim et al. [17] analyzed images collected for grapefruit peel disease recognition and achieved the best classification accuracy of 96.7%.

Various studies have reported success with image processing-based technology, as a plant disease identification mechanism [18,19]. Schor et al. [20] applied image processing techniques as an automated diseases detection tool capable of ensuring timely control of PM and *Tomato spotted wilt virus* (TSWV) diseases. This system also increased crop yield, improved crop quality and reduced the quantity of applied pesticides on bell pepper. A thresholding-based imaging method proposed by Sena Jr et al. [21] aimed to distinguish fall armyworm affected maize plants from healthy ones. Camargo and Smith [22] used colour transformation-based image processing technique to identify the visual symptoms of cotton diseases. Textural feature analysis is also widely used as an image processing approach to extract key plant health information. Spatial variation pixel values are described by image textures [23] to explain regional properties like smoothness, coarseness, and regularity [24]. Colour co-occurrence matrix (CCM) based textural analysis was introduced for plant identification [25] and leaf and stem disease classification [26]. Xie et al. [27] extracted eight features to develop a detection model for leaf early and late blight diseases from tomato leaves with co-occurrence matrix. However, it should be noted that the features themselves are not enough for object identification and need classifiers for further plant disease recognition.

Machine learning techniques, such as artificial neural networks (ANNs), support vector machines (SVMs), k-nearest neighbors (kNNs), and decision trees have been utilized in agricultural research [28] as part of supervised learning. The ANNs, SVMs and kNNs classifiers have classified different plant diseases with higher success rate [29–31]. Wang et al. [32] reported improved control of tomato

diseases by predicting late blight infections using ANNs. Pydipati et al. [33] utilized backpropagation ANNs algorithm and colour co-occurrence textural analysis for citrus disease detection and achieved classification accuracies of over 95% for all classes. The researchers also claimed that an overall 99% mean accuracy was achieved when using hue and saturation textural features. Camargo and Smith [22] detected visual symptoms of cotton crop diseases using SVMs classifier. The measurement of texture used as a useful discriminator for this detection. A method using kNNs to detect nitrogen and potassium deficiencies in tomato crops was proposed by Xu et al. [34]. Conversely, VijayaLakshmi and Mohan [35] suggested some limitations using SVM and kNN for leaf type detection by using colour, shape and texture features. Yano et al. [36] found artificial neural network provide better accuracy compared with kNN and Random Forest (RF) classifier.

The potential to combine image texture-based machine vision and machine learning algorithms for plant disease detection is significant. To date, there has been no research conducted applying machine vision with different machine learning algorithms for PM disease detection in strawberry cropping systems. As the first step of detecting PM disease, an RGB camera was used as an image source, which can be extended to hyper-/multi-spectral camera or thermal camera as in other crop disease detection. Therefore, the main purpose of this research is to compare image texture-based machine vision techniques for PM disease classification using a series of supervised classifiers.

2. Materials and Methods

2.1. Study Area and Experimental Overview

Three strawberry fields were selected in western Nova Scotia, Canada to collect PM infected and healthy plant strawberry image samples. The selected fields were located on two farms in Great Village: the Millen farm site I (Field I: 45.398° N, 63.562° W), Millen farm site II (Field II: 45.404° N, 63.549° W), and the Balamore farm site III (Field III: 45.413° N, 63.567° W). Strawberry leaves were collected throughout two growing seasons of the strawberry, summer and fall, between 10 a.m. and 4 p.m. in 2017–2018. They were randomly selected from fully grown plants which were producing strawberries. Separated leaves (from the plants) were stored in the icebox which had 5–7 °C internal environment and were brought to the lab directly. The images of the leaves were taken around 5–6 p.m. at the same day of leaf collection. Regional climate of Figure 1 is the same as condition of the usual Nova Scotia, with the hottest temperatures of 24 °C in July and August in summer and the coldest of −11.8 °C in January in winter. The average annual precipitation in this area is 779.66 mm [37]. Three strawberry varieties, Albion, Portola and Ruby June, were cultivated in Field I, Field II and Field III, respectively (Figure 1).

Images can be collected under different lighting conditions, which means collection under natural light, and controlled lighting condition. In this experiment, initially two lighting conditions were set and processed for preliminary image acquisition. The first lighting condition was sunny condition and the second lighting condition was cloudy condition which we made the artificial cloud condition (ACC) with the black cloth. According to Steward and Tian [38], they implemented segmentation algorithm with two sets of weed images which were photographed under sunny and cloud conditions. Estimation of weed density was highly related to lighting conditions. The variability occurred more in images that were taken under sunny conditions. The results of the correlation coefficient for images taken under cloud conditions was closer to 1 which means more stable about the variability. This aligned with our preliminary results that showed less accuracy under sunny conditions.

Thus, leaves were harvested from randomly selected plants across each of the fields and individual leaf images were taken under artificial cloud conditions (ACCs) to increase the accuracy.
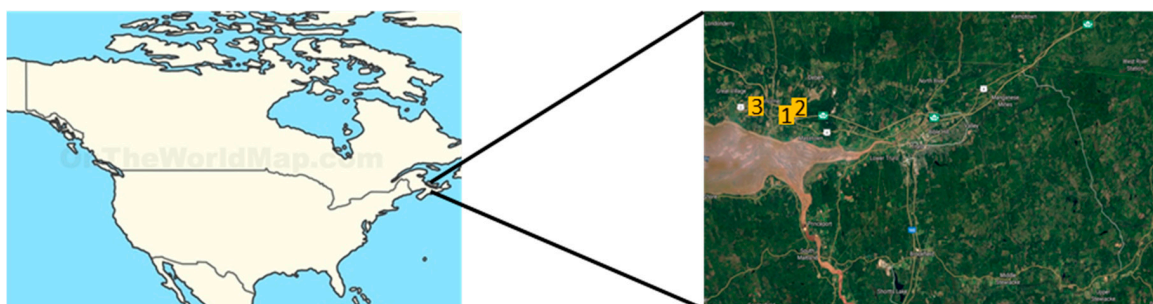
Figure 1. Geographical information about data collection. The yellow mark on the map on the right, 1,2, and 3 matches the information of the sites I, II, and III.

2.2. Performance Evaluation

A total of 450 images were collected from all fields, specifically, 150 leaves from each field (total 3 fields). Each field data consisted of 75 healthy and 75 infected leaves. The 450 images were divided into two sets containing 300 images for training and 150 images for validation in three different classifiers. Internal, external and cross-validations were experimented for performance evaluation (Accuracy (%) = $\frac{\text{correctly classified number}}{\text{Total number}} \times 100$) of this study. Internal validations were conducted with different set of data from same field. Cross-validations were done with 4-fold and 5-fold, splitting the total data from all fields in 4 or 5 different subsets and then 3 or 4 sets of data were utilized for training and one set for validation. K-folds is the method used to maximize the use of available data for model training and testing. This avoids overfitting the predictive model and solves the problem of low data count. The data set was divided into k subsets and a holdout method was repeated k times. At each time, k-1 subsets were utilized for training and k^{th} subsets was utilized for testing. Finally, the average error across all k trials was computed. Thereafter, every data point gets to be in a test set exactly once and gets to be in a training set k-1 times. In terms of external validation, Field I and Field II images were used to train the classifiers and Field III images were used for validation. Accordingly, Field I and Field III were used for training and Field II for validation. Final external validation was conducted with Field II and Field III for training and Field I for validation.

2.3. Image Acquisition

The image acquisition system consisted of two major components: an artificial cloud chamber and a digital single lens reflex (DSLR) camera model: EOS 1300D (Canon Inc., Tokyo, Japan) with a sensor resolution of 5184 × 3456 pixels and a 30 mm focal length lens for taking very detailed images. Individual leaves were collected by separating them from each bundle and the images were taken while the leaf sat on a white paper under the controlled environment, ACC. The images were taken at a height of 30 cm above the leaves and the same height was maintained for all image acquisitions. Exposure time and ISO gain were automatically controlled with an F/8.0 aperture to maintain both same depth of field and same ACCs. The images were processed with a hardware system of Intel® Core™ i5-3320M CPU @ 2.60 gigahertz (GHz) and 4.00 gigabyte (GB) Random Access Memory (RAM) laptop (Lenovo Group Ltd., Morrisville, NC, USA). The images were saved from the camera in the RAW format and were subsequently converted to Windows Bitmap (BMP) format to overcome loss related issues caused by image compressions.

2.4. Image Processing and Data Normalization

A graphic user interface (GUI) program was developed for image pre-processing, textural features extraction, and to save the features in a text file. The first step embraced in textural features extraction was the conversion of Blue, Green, and Red (BGR) channels to National Television System Committee (NTSC) standard for luminance and HSI colour models. Windows Graphics Device Interface (GDI) uses BGR channel space for bitmap representation, so BGR channel was selected when image processing in

order to match with the GDI information. The luminance (Lm) of each pixel constituting the image was converted from 24-bit Bitmap (BMP) into 8-bit brightness image of NTSC standard, and Lm was calculated by the following equation by using BGR [24].

$$Lm = (0.1140\ B + 0.5870\ G + 0.2989\ R) \tag{1}$$

In this study, BGR channels were converted to HSI, which represents the colour similar to how the human eye senses colour, to create three 2-dimensional arrays. The principle of converting BGR to HSI was calculated by Equations (2)–(5). First, θ_h represents the angle in Equation (2) then Hue (H) was calculated and normalized into [0,255]. The original BGR, Luminance, Hue, Saturation, and Intensity based images of healthy and PM infected leaves can be seen Figures 2 and 3.
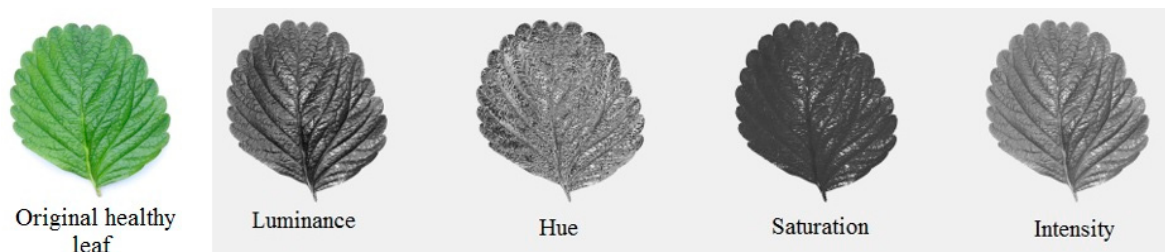

Original healthy leaf   Luminance   Hue   Saturation   Intensity

Figure 2. Original and processed strawberry healthy leaf image used for texture features extraction.

Original disease leaf   Luminance   Hue   Saturation   Intensity

Figure 3. Original and processed strawberry powdery mildew affected leaf image used for texture features extraction.

In Equation (6), p (m, n) represents the marginal probability function. Inside of p (m, n), m is the intensity level at a particular pixel, and n is the intensity level of a matching pixel with an offset. Of the offsets from 1 to 5, offset 1 gives the best results at any orientation angle such as 0°, 90°, 180°, and 360° [39]. Shearer and Homes (1990) extracted 10 features based on CCM which contains the luminance and HSI [24] (Table 1).

$$\theta_h = \cos^{-1}\left\{ \frac{\frac{1}{2}[(R-G)+(R-B)]}{\left[(R-G)^2+(R-G)(G-B)\right]^{1/2}} \right\} \tag{2}$$

$$H = \left\{ \begin{array}{l} \frac{\theta_h}{360} * 255 \text{ if } B \leq G \\ \frac{360-\theta_h}{360} * 255 \text{ if } B > G \end{array} \right\} \tag{3}$$

$$S = 255 \times \left\{ 1 - \frac{3}{(R+G+B)}(\text{Min}(R,\ \text{Min}(G,B))) \right\} \tag{4}$$

$$I = \frac{R+G+B}{3} \tag{5}$$

$$p(m,n) = \frac{P(m,n)}{\sum_{m=0}^{N-1}\sum_{n=0}^{N-1}P(m,\ n)} \tag{6}$$

<div align="center">**Table 1.** Textural feature equations (Shearer & Holmes, 1990).</div>

Description	Textural Features	Equation [u]
Contrast	TF1	$\sum_{|m-n|=0}^{N-1}(m-n)^2\sum_{m=0}^{N-1}\sum_{n=0}^{N-1}p(m,n)$
Homogeneity	TF2	$\sum_{m=0}^{N-1}\sum_{n=0}^{N-1}p(m,n)\frac{1}{1+|m-n|}$
Entropy	TF3	$\sum_{m=0}^{N-1}\sum_{n=0}^{N-1}p(m,n)\,In\,p(m,n)$
Dissimilarity	TF4	$\sum_{m=0}^{N-1}\sum_{n=0}^{N-1}p(m,n)|m-n|$
Angular 2nd Moment	TF5	$\sum_{m=0}^{N-1}\sum_{n=0}^{N-1}p(m,n)^2$
Inverse Difference Moment	TF6	$\sum_{m=0}^{N-1}\sum_{n=0}^{N-1}\frac{p(m,n)}{1+(m-n)^2}$
Average	TF7	$\sum_{m=0}^{N-1}m\,p_x(m)$
Sum of Squares	TF8	$\sum_{m=0}^{N-1}(m-\mu)^2 p_x(m)$
Product Moment	TF9	$\sum_{m=0}^{N-1}\sum_{n=0}^{N-1}p(m,n)(m-\mu)(n-\mu)$
Correlation	TF10	$\sum_{m=0}^{N-1}\sum_{n=0}^{N-1}p(m,n)\frac{(m-\mu_m)(n-\mu_n)}{6_m 6_n}$

[u] N is the total number of intensity levels, $p(m,n)$ is the $(m,n)^{th}$ entry in a normalized CCM; and μ is the mean, μ_m is the mean of row, μ_n is the mean of column, 6_m and 6_n are the standard deviation along the mth row and n^{th} column of $p(m,n)$ and $p_x(m)$ was obtained by summation of CCM values in m^{th} row.

A total of 40 features, 10 from each color plane (Luminance, Hue, Saturation, and Intensity), were extracted from one image which were utilized as inputs for classifiers. In order to improve the performance of the classifiers, the input data were normalized [40]. The following equation was used for the normalization of the data.

$$u_i = \frac{(R_i - Min_i)}{(Max_i - Min_i)} \tag{7}$$

where, u_i = Normalized value of input; R_i = Actual value of input; Min_i = Minimum value of input; Max_i = Maximum value of input.

2.5. Classifiers

In this study different types of classifiers were evaluated to determine the most effective ones at classifying strawberry PM disease characteristics using CCM based textural features analysis. In MATLAB software, various classification models from the data were trained by using classification learners. The three classifiers evaluated including the ANN, SVM, and kNN. Firstly, 9 parameters were used for ANN tuning. The epoch size of 15,000 was determined to be enough for the model structure to perform the classification rather than other epochs. The output range of tanh sigmoid was −1 to 1, that is, the transformed version of the logistic sigmoid with an output range of 0 to 1 and it was used as a mathematical function because it tends to be more fit well with neural networks [41]. The proposed settings of the developed model are given in Table 2.

Table 2. Developed model setting of an artificial neural network.

Parameters	Settings
Training pattern	300 out of 450
Optimum Epoch	15,000
Learning rate	0.10
Momentum	0.70
Number of hidden layers	2
Nodes per hidden layer	40
Mathematical Function	Tanh Sigmoid
Internal Validation	50 out of 150 from each field
External validation	150 out of 450

Secondly, 4 different kernels were used as a parameter for SVM: Linear, Quadratic, Cubic, and Fine Gaussian. SVM can train when there is more than one class in the data in the classification learner. Each classifier type differs depending on prediction speed, memory usage, interpretability, and model flexibility. Linear SVM is the most commonly used because of its fast prediction speed and easy interpretability compared to other kernels. Fine Gaussian SVM, on the other hand, is more difficult to interpret than Linear SVM, but offer more model flexibility and finer separation of classes than other kernels.

Lastly, 4 different kernels were used as a parameter for kNN: Fine, Cosine, Coarse, and Cubic. Model flexibility decrease with the number of neighbours setting. kNN with Fine is showing the detailed distinctions between classes and the number of neighbours was set to 1. kNN with Coarse is showing the medium distinctions between classes and the number of neighbours is set to 10. kNN with Cosine uses Cosine distance metric and kNN with Cubic has slower prediction speed than Cosine kNN [42].

2.5.1. ANN Classifier

In this study, textural features of the CCMs were used to provide the input data for training the ANN model. Neural network models were found to incorporate all the textural features in the discrimination scheme. Peltarion Synapse (Peltarion Corp., Stockholm, Sweden) was used for classifying the textural features, as well as images of healthy and PM affected leaves. A back-propagation artificial neural network (BP-ANN) algorithm was applied for training of the proposed network. Total three model structures were examined in this experiment, and one of example is the 40-80-1 ANN network structure which represents forty nodes for the input layer, eighty nodes for two hidden layers, and one node for the output layer for the data analysis (Figure 4). The extracted textural features were selected as inputs for the input layer and corresponding healthy or disease labels were established as an output in the output layer with Synapse Peltarion software. Four different functions, such as the tanh sigmoid, exponential, logistic sigmoid, and linear transfer were used to translate the input signals into output signals ranging from 1 to 2. The predictor model was used Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) to find the best model structure. The formulas for MAE is given by $\frac{1}{n}\sum_{i=1}^{n}|yi - \hat{y}i|$, that is, multiplication between a division the total number of data point and the sum of the absolute value of the residual of the subtraction from actual output value, yi, to the predicted output value, $\hat{y}i$. The formulas for RMSE is given by $\sqrt{\frac{\sum_{i=1}^{n}(yi - \hat{y}i)^2}{n}}$, that is, measurement of the differences between values predicted by a model and the values observed by giving how much error there is between two data sets.
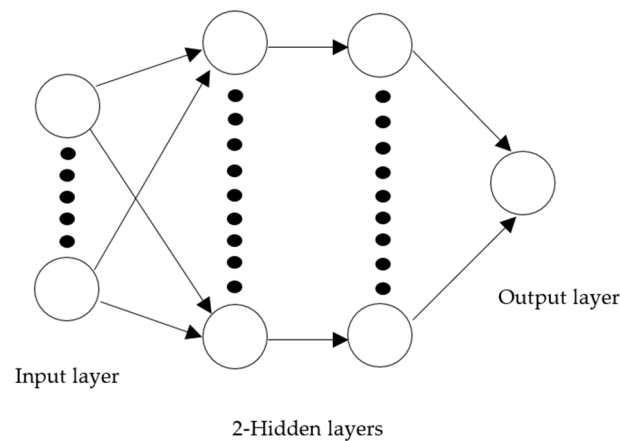
Figure 4. Artificial neural networks (ANN) model structure.

The training process was initiated using randomly selected initial weights and biases. The experimental conditions involved a supervised training mechanism providing the network output by labelling. The MSE and RMSE values were used to determine the performance of the model structures. The activation, or non-linear function, was determined by the presence of particular neural features that determined the optimal weights. The neural network performed a non-linear transformation on the input variables (N) to achieve an output (M) by Equation (8).

$$\{M\} = f\ (\{N\}) \tag{8}$$

where, M = Output; f = Non-linear function; and N = Input variables.

Three ANN model structures were developed and tested to discover a satisfactory mathematical function to process image data. Around twenty simulations of the ANN model were conducted to select the three optimal model structures. Although a complex network can be used, it is claimed that a hidden layer is sufficient to represent a continuous nonlinear function because overfitting can occur, as the number of hidden layers increases [43]. All the selected models were run at an epoch size (iterative steps) of 15,000 with learning rate of 0.1. To determine the optimal epoch size, the best selected ANN model was operated at different epoch sizes at an interval of 1000 and the error values of MAE and RMSE were calculated at each interval. According to Madadlou et al. [44], the epoch size has a major influence on error terms. The momentum rule of the developed models was 0.7 to establish a comparison of the processing capabilities of the different mathematical functions. The best model structure and mathematical function were selected based on lower MAE and RMSE values and by comparing the actual and predicted values. Finally, the model developmental process was completed having the acceptable errors (i.e., MAE < 0.003 and RMSE < 0.005) from predicted data set compared with actual data set. The MAE is a measure of absolute difference between actual and predicted observation and the RMSE is the square root of the average of squared differences between prediction and actual observation that measures the average magnitude of the error. After training the model, the performance of the ANN model was tested by employing the internal, external, and cross-validation separately (Figure 5).
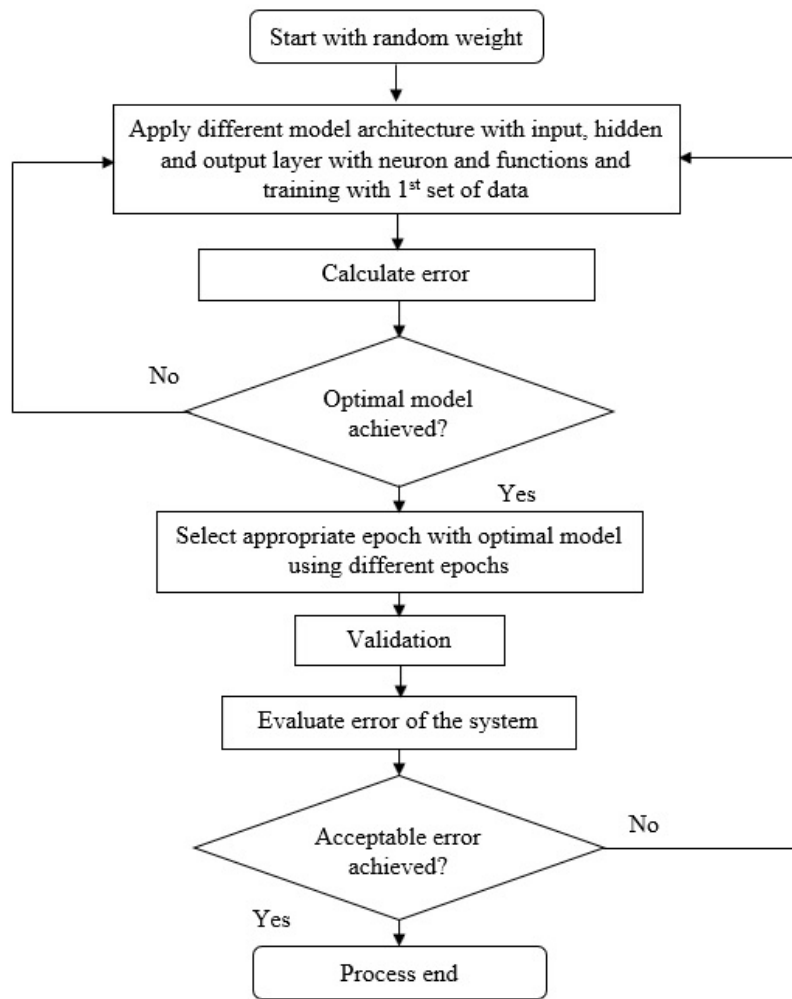
Figure 5. Flowchart of artificial neural network algorithm for model development.

2.5.2. SVM Classifier

An SVM classifier was selected for the experiments based on previously established studies reporting efficient implementation and performance for high dimensional problems and small datasets. MATLAB® Classification Toolbox version 2017a (MathWorks, Natick, MA, USA) was used for SVM classifier evaluation. The general concept of SVM (Figure 6) is to provide a solution of classification problems by plotting the input vectors into a new high-dimensional space through some nonlinear mapping and constructing an optimal separating hyperplane that measures the maximum margin to separate positive and negative classes [45]. The SVM constructs a hyperplane in the space to observe the training input vector in an n-dimensional space for classification that has the largest distance to the closest training data point of any class. A linear and separable sample sets belonging to separate classes are separated by the hyperplane. The generalization ability of SVM model is usually better as the distance between the closest vectors to the hyperplane is maximized. The mathematical form of SVM classifier is as follows:

$$F: U\hat{\ }X \rightarrow Q \tag{9}$$

where, specified data set mapping is made via a map function F from input space into higher dimension feature space Q (dot product space).
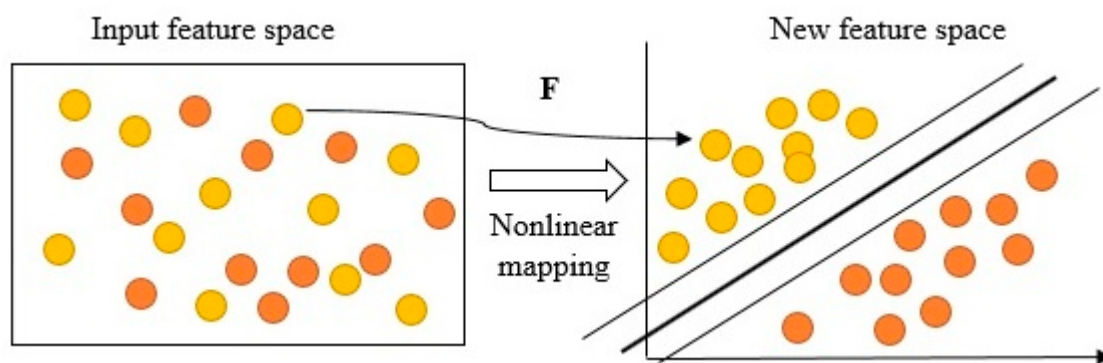
Figure 6. Working principle of support vector machine (SVM) classifier.

A linear learning algorithm was performed in Q which required the evaluation of dot products. If Q is of a higher dimensional value, then the right-hand side of Equation (9) will be very costly to compute [46]. Hence, kernel functions are utilized using the input parameters to compute the dot product in the feature space.

Four different kernels such as linear, quadratic, cubic, and fine Gaussian were utilized for experimentation. Internal, external, and cross-validations were tested for performance evaluation of classifier. Two fields of data were used for training and one field of data was simultaneously utilized for the model validation. The external validation process began with exporting the model after development. The exported model tested the data from separate fields that were not used for training the model. Finally, the accuracy was determined based on correctly classified textural features from disease and healthy strawberry leaves.

2.5.3. kNN Classifier

Another supervised learning classifier evaluated in this study was the kNN. The kNN classifier is a non-parametric method that separates a test object according to the class of majority of its kth nearest neighbour in the training set. It applies the Euclidean distance in the multidimensional space as a similarity measurement to separate the test objects. K represents the number of highly data-dependent tuning of neighbours and uses uniform weights meaning which is assigned the value to a query point is computed from a simple majority vote of the nearest neighbours. In other word, the unknown object in the query object is compared to every sample of objects which are previously being used to train the kNN classifier. The distance measurements of kNN classifier were conducted using Euclidean distance with the following equation:

$$E_{Dp}(P,\ Q) = \sqrt{\frac{\sum_{i=1}^{x} (m_i - n_i)^2}{m}} \qquad (10)$$

where P and Q are represented by feature vectors $P = (m_1, m_2, m_{3\ldots\ldots}, m_x)$ and $Q = (n_1, n_2, n_{3\ldots\ldots}, n_x)$ and x is the dimensionality of the feature space. The equation measures the Euclidean distance between two points P and Q.

The performance of kNN varies with different kernel functions. Fine, cosine, coarse and cubic kernels were used for performance evaluation in this study. Like previous classifiers, kNN performance was evaluated by using internal, external, and cross-validations.

3. Results and Discussions

3.1. Classification Using ANN

The experimental outcomes established the most suitable model, i.e., having the lowest RMSE value. A normalized dataset, based on Equation (7), was imported into Synapse Peltarion software and the variables, i.e., input and output, were defined using the software interface. The results from three

network structures evaluated to obtain a suitable ANN network for data processing are listed in Table 3. All the settings of developed models were kept constant, the mathematical functions were changed and MAE and RMSE were recorded. The highest RMSE was derived from the exponential function and compared to the other functions evaluated to classify healthy and PM affected leaves. The exponential function resulted in an infinity error for two model settings, suggesting its non-suitability for data processing. The tanh sigmoid function was able to process data with a reasonably low RMSE (0.004 to 0.019) and MAE (0.003 to 0.012) when compared to the other functions in the networks. Based on low error rates, tanh sigmoid was chosen for further processing. All the networks were operated at an epoch of 15,000.

Table 3. Tested mathematical functions at an epoch size of 15,000 with normalized data.

Model Structure [1]	Tanh Sigmoid		Loigistic Sigmoid		Linear		Exponential	
	MAE [2]	RMSE [3]	MAE [2]	RMSE [3]	MAE [2]	RMSE [3]	MAE [2]	RMSE [3]
1W (40/80) & 1F (80/1) 40 Inputs 1 Output	0.012	0.019	0.034	0.053	0.139	0.179	∞	∞
1W (40/40) 1W (40/40) 1F (40/1) 40 Inputs 1 Output	0.003	0.004	0.030	0.045	0.139	0.179	∞	∞
1W (40/80) 1W (80/80) 1F (80/1) 40 Inputs 1 Output	0.004	0.005	0.026	0.038	0.167	0.216	0.512	0.612

Where, [1] W = weight layer; F = function layer; and (∞) = infinity error. [2] Mean absolute error is the average vertical distance between each point and the identity line. [3] Root mean square error is the standard deviation of the residuals which are a measure of how far from the regression.

After determination of an optimal model, the model was tested and evaluated to select appropriate epoch (iterations) size using different numbers of epoch. In total, five different epoch sizes were evaluated using the optimized model with a tanh sigmoid function (Table 4). The model (1W (40/40) 1W (40/40) 1F (40/1) with an epoch size of 15,000 was selected as the best conditions based on low MAE (0.003) and RMSE (0.004). The optimized model using epochs of 1000, 2000, and 3000 resulted in higher RMSE values suggesting poorer performance than 15,000 epochs (Table 4).

Table 4. Selection of approximate epoch.

Sr. No.	Model Structure	Epoch	Tanh Sigmoid	
			MAE [1]	RMSE [2]
1.		1000	0.064	0.091
2.		2000	0.028	0.039
3.	1W (40/40)1W (40/40) 1F (40/1) 40 Inputs 1 Output	3000	0.024	0.036
4.		5000	0.011	0.015
5.		15,000	0.003	0.004

Where, W = weight layer; F = function layer. [1] Mean absolute error is the average vertical distance between each point and the identity line. [2] Root mean square error is the standard deviation of the residuals which are a measure of how far from the regression line data points are.

Table 5 shows ANN classifier performance based on internal, external and cross-validations with different dataset. The model classified healthy and PM leaves with high accuracy in the case of internal validation. The highest accuracy was observed with internal validation and the lowest accuracy was found with external validation. Results suggested that the overall accuracy which is the averaged value of internal, external, 4-fold cross and 5-fold cross-validations, was 93.81%, using the selected model structure, with an effective epoch size of 15,000.

Table 5. Performance of artificial neural networks (ANN) classifier for powdery mildew disease classification.

Validation	Model Structure	Accuracy (%)
Internal [1]		99.99
External-I [2]		87.41
External-II [3]	1W (40/40) 1W (40/40) 1F (40/1) 40 Inputs 1 Output	88.95
External-III [4]		95.04
4-fold Cross [5]		95.03
5-fold Cross [5]		96.45
Overall accuracy [6]		93.81

[1] Internal is an internal validation; [2] External-I: training with Field I + Field II and validated with Field III; [3] External-II: training with Field I + Field III and validated with Field II; [4] External-III: training with Field II + Field III and validated with Field I. [5] k-fold is splitting the total data from all fields in k different subsets and then k-1 sets of data were used for training and the rest one set for validation; [6] Overall accuracy is an averaged value of internal validation, three external validations, and cross-validations.

3.2. Classification Using SVM

The SVM classifier was implemented to demonstrate the effectiveness of the classification of PM disease leaf from healthy leaves. For this purpose, the feature vectors were applied as the input to an SVM classifier after normalization. Training and testing were evaluated for measuring the performance of the classifier. Experimental results showed that the accuracy of the classifier varied with different kernel functions (Table 6). The highest overall performance was from the linear kernel (96.66%), whereas a 91.14% of accuracy was observed using the fine Gaussian kernel. The linear kernel had an average accuracy > 94% across all validation steps (Table 6). The best kernel was selected by evaluating the accuracy of internal and cross-validation, having the highest overall accuracy of approximately 96.66% (Table 6). Different assessments were applied for calculating the performance of the proposed classifier with linear kernel. The best test performance of SVM classifier is presented in Table 7. The SVM classifier predicts measured values at a high accuracy rate. The highest classification accuracy was 98.67% and the lowest classification accuracy was 81.33% during internal and external-I validations, respectively. The external validations (Table 7) varied with different field data used for validation. The lowest external validation accuracy was found with Field II data and highest accuracy achieved with Field I validation. The cross-validation resulted in an accuracy > 94%, whereas the 5-fold cross and 4-fold cross-validations had accuracies of 96.67% and 94.64%, respectively. The overall average performance result of SVM was about 91.66% including external validation results.

Table 6. Selection of kernel for support vector machine (SVM) classifier.

Kernel [1]	Internal Validation (%)	4-fold Cross Validation [2] (%)	5-fold Cross Validation [2] (%)	Overall Accuracy [3] (%)
Linear	98.67	94.64	96.67	96.66
Quadratic	98.10	92.53	95.02	95.22
Cubic	94.87	89.44	91.17	91.83
Fine Gaussian	94.23	88.26	90.92	91.14

[1] The kernel transforms input data into the required form; [2] k-fold is splitting the total data from all fields in k different subsets and then k-1 sets of data were used for training and the rest one set for validation; [3] Overall accuracy is an averaged value of internal validation, three external validations, and cross-validations.

Table 7. Performance of SVM classifier for powdery mildew disease classification with linear kernel.

Data Set		Validation	Accurately Classified	Incorrectly Classified	Accuracy (%)
Training	Testing				
300	150	Internal [1]	148	2	98.67
300	150	External-I [2]	128	22	85.33
300	150	External-II [3]	122	28	81.33
300	150	External-III [4]	140	10	93.33
338	112	4-fold cross [5]	106	6	94.64
360	90	5-fold cross [5]	87	3	96.67
Overall accuracy [6]					91.66

[1] Internal is an internal validation; [2] External-I: training with Field I + Field II and validated with Field III; [3] External-II: training with Field I + Field III and validated with Field II; [4] External-III: training with Field II + Field III and validated with Field I; [5] k-fold is splitting the total data from all fields in k different subsets and then k-1 sets of data were used for training and the rest one set for validation; [6] Overall accuracy is an averaged value of internal validation, three external validations, and cross-validations.

3.3. Classification Using kNN

MATLAB software was used for implementing the kNN classifier. After normalizing the textural features, data classification was done between the data of healthy leaves and PM affected leaves using kNN classification method. Table 8 shows classification accuracy using different kernels of kNN model.

Table 8. Selection of kernel for k-nearest neighbors (kNN) classifier.

Kernel [1]	Internal Validation (%)	4-fold Cross Validation [2] (%)	5-fold Cross Validation [2] (%)	Overall Accuracy [3] (%)
Fine	89.33	85.71	87.78	87.61
Cosine	88.64	85.23	87.78	87.22
Coarse	78.70	77.21	77.30	77.74
Cubic	84.23	83.66	84.07	83.99

[1] The kernel transforms input data into the required form; [2] k-fold cross validation is splitting the total data from all fields in k different subsets and then k-1 sets of data were used for training and the rest one set for validation; [3] Overall accuracy is an averaged value of internal validation, three external validations, and cross-validations.

The results from Table 8 show that the highest classification accuracy was found with fine kernel at 89.33% during internal validation. The results also show a high classification accuracy of 87.78% for the two different types of cross-validation with 4 different kernels. The highest classification accuracy of 85.71% for 4-fold cross-validation was achieved using the fine kernel. The best kernel was selected with the highest overall accuracy of 87.61% using fine kernel. Table 9 outlines the performance of kNN classifier in three different validations, e.g., internal, external and cross-validations using fine kernel. As shown in Table 9, the best accuracy achieved was when the kNN classifier worked with internal data (89.33%), and the lowest accuracy was 53.33% which was achieved in external-II validation with Field II data. An overall accuracy was 78.80% which is exposed to be close result in previous study during plant diseases classification. Previous research results have reported classification accuracy of 82.5% with kNN classifier for tomato diseases classification [34].

Table 9. Performance of kNN classifier for powdery mildew disease classification with fine kernel.

Data Set		Validation	Accurately Classified	Incorrectly Classified	Accuracy (%)
Training	Testing				
300	150	Internal [1]	134	16	89.33
300	150	External-I [2]	110	40	73.33
300	150	External-II [3]	80	70	53.33
300	150	External-III [4]	125	25	83.33
338	112	4-fold cross [5]	96	16	85.71
360	90	5-fold cross [5]	79	11	87.78
Overall accuracy [6]					78.8

[1] Internal is an internal validation; [2] External-I: training with Field I + Field II and validated with Field III; [3] External-II: training with Field I + Field III and validated with Field II; [4] External-III: training with Field II + Field III and validated with Field I; [5] k-fold is splitting the total data from all fields in k different subsets and then k-1 sets of data were used for training and the rest one set for validation; [6] Overall accuracy is an averaged value of internal validation, three external validations, and cross-validations.

3.4. Performance Comparison of Three Classifiers

The performance of this study was examined by evaluating the external validations and overall accuracies of each model. The classification methods were tested with the same dataset and the classifier performances are summarized in Table 10. The performance comparisons of classifiers were analyzed based on external validation and overall accuracy. The best overall accuracy score (93.81%) was achieved for the ANN classifier with (1W (40/40) 1W (40/40) 1F (40/1)) model where the external validation results were higher than the other two classifiers. The SVM based predictive model also performed well, with an overall accuracy score over 90%, on the dedicated validation set for same data. However, the kNN based classifier had the lowest accuracy scores in both case of external validation and overall performance. Although ANN and SVM classifiers had an accuracy value over 90%, the SVM classifier is not suggested as the best model due to limitations in training, testing of the data, and external validation accuracy. The speed of data processing was also found to be slower in SVM model development. On the other hand, the ANN classifier with the textural features was suggested as the best classifier in our study. The external validation results were comparatively lower in all the classifiers because the training and testing of classifiers were conducted with different varieties of strawberry. Each variety had some differences associated with leaf texture. Despite some misclassifications between some healthy and PM infected leaves, the proposed use of a machine vision algorithm based on ANN was successfully able to classify strawberry PM diseased leaves with an average accuracy score of 93.81%.

Table 10. Performance evaluation of three classifiers.

Validation	ANN Accuracy (%)	SVM Accuracy (%)	kNN Accuracy (%)
External-I [1]	87.41	85.33	73.33
External-II [2]	88.95	81.33	53.33
External-III [3]	95.04	93.33	83.33
Overall accuracy [4]	93.81	91.66	78.80

[1] External-I: training with Field I + Field II and validated with Field III; [2] External-II: training with Field I + Field III and validated with Field II; [3] External-III: training with Field II + Field III and validated with Field I; [4] Overall accuracy is an averaged value of internal validation, three external validations, and cross-validations.

3.5. Discussions

PM disease identification at earlier stage is very important for strawberry growers to prevent yield loss, to reduce labour costs, and to increase the fruit quality. A computer vision-based disease classification algorithm has been shown to have great potential to provide information for early disease

detection. Wspanialy and Moussa [47] detected PM disease in tomato plants with a detection rate of 85%. Ei-Helly et al. [18] achieved 94% of accuracy for cucumber PM disease detection in leaf.

Since the visual difference of images used in this study was slight, it was important to utilize image colour matrix, i.e., hue, saturation and intensity for textural features extraction. The CCM based textural analysis of 40 features was found to be effective for this specific problem. An ANN based plant disease detection system previously proposed by Kulkarni and Patil [48] with diverse image processing techniques had high recognition rates of up to 91% in pomegranate crop. They suggested that ANNs based classifier detected numerous plant diseases with combination of colour and texture features to recognize Alternaria, Bacterial Blight Disease and Anthracnose diseases. Ramakrishnan [49] found much higher accuracies of disease detection, around 97.41% for Ground nut. Their experiments were done with CCM textural analysis and back propagation ANNs algorithm for detection of leaf disease. Results from our study using the ANN classifier with strawberry leaves and PM appear to support previous assertions as to the accuracy of ANN.

Islam et al. [50] presented an approach that integrates machine learning (i.e., SVM) and image processing techniques to allow diagnosing diseases from potato leaf images. The proposed techniques provide a path toward automated plant disease diagnosis on a large, i.e., field, scale with potential accuracy of detection >95% using SVM classifiers. An image pattern classification was studied by Camargo and Smith [22] for detection of the cotton crop diseases using SVM. CCM, co-occurrence matrix, having five features were extracted on their research and reached a classification accuracy of 90%. Their study suggested that texture-related features and SVM might lead to successful discrimination of plant diseases but a study by VijayaLakshmi and Mohan [35] reported some difficulty using SVM to understand the structure and size and speed limitations both in training and testing the data. Outcomes from our study resulted in similar limitations in speed of training from SVM but detection accuracy was also high.

In contrast, kNN had the lowest accuracy in plant disease classification. Sankaran et al. [30] classified citrus disease in leaves resulting in overall accuracies of 83.3%, 86.8% and 86.8%, for 1st derivative, 2nd derivative, and combined spectral features using kNN, respectively. Similar results were obtained in our study using kNN. The inferior results reported by kNN having the overall accuracy was 78.80% where the external validation results were 73.33%, 53.33% and 83.33%, respectively.

From the analysis, it is revealed that the classification worked successfully when all the textural features are utilized. In this study, we did not perform any method which can reduce the number of textural features. However, reduction method would be required for future real-time application which would require fast calculation. We left the feature reduction for future study.

Kurtulmus et al. [51] stated that an irregular amount of the illumination outdoors is a major hindrance to crop detection with machine vision. Aggelopoulou et al. [52] shot images between 11 a.m. and 1 p.m. to reduce the effects of sunlight and removed distorted images due to light from the collected images. Furthermore, a black cloth is placed behind the tree, so the white of the flower contrasts with the black background and it results in reducing the influence of the light. Therefore, the controlled environmental condition of the ACCs presented in this paper is the important factors for future real-time application. Our results also revealed that analysis using the ANN classifier provide higher accuracy values than SVM and kNN. The poorest result was reported using kNN in diseases classification and there were some difficulties acquiring with SVM classifiers. The outcome of this study indicated that analysis of images texture by using a CCM technique and an ANN approach had high precision for classifying the PM affected leaves in strawberry crops.

4. Conclusions

In this study, colour co-occurrence matrix textural analysis and supervised learning classifiers were implemented for classification of healthy and PM disease leaves. The CCM based textural analysis was executed for extracting image features. Forty features were extracted, which were utilized for classification after normalization. Three supervised classifiers, i.e., ANN, SVM, and kNN were

evaluated and the best result was generated using the ANN classifier while the kNN had the lowest overall accuracy. The SVM had high accuracy of disease detection but some limitations associated with speed of training and testing of data were found. Results suggested that the ANN is capable to model non-linear relationships and performed better classification. The study experimented on ACCs, artificial cloud conditions, environment, which reduced different environmental factors like light and leaf density during image acquisition.

The smaller image size used as an input in the machine learning process, the more information could be vanished because the value of the pixel is summarizing. However, the smaller size of the image, the less processing time is required to train and evaluate the model [53]. Also, different image processing techniques besides colour co-occurrence matrix, will need to be examined to find optimum supervised machine learning technique. Histogram of Oriented Gradients (HOG), Scale Invariant Feature Transform (SIFT), and Speeded up Robust Features (SURF) are possible follow up research areas for the detection of PM. Therefore, it is necessary to study the variation of classification accuracy and processing time according to various image sizes and different image processing techniques. As an area for following study, inclusion of other factors would need to occur in model development for utilization in real-time PM disease classification under field conditions.

Also, the development of an unmanned ground vehicle (UGV) to collect the image data automatically will be an essential tool. The camera will be set up in the proper place on the UGV to take a row of crops in the field and a custom image acquisition program should be developed to save images or videos in the computer. In that case, segmentation techniques and adjusting the height due to the resolution could be needed to consider overlapping leaves. Furthermore, deep learning would be more effective than traditional machine learning approaches because it performs automatic feature extraction. Fast processing of deep learning can be effectively used in robots for real-time decision making, which also needs to be explored as a following study. This paper would motivate more researcher to experiment with machine learning and deep learning to solve agricultural problems involving classification or prediction.

References

1. Spencer, D.M. (Ed.) Powdery mildew of strawberries. In *The Powdery Mildews*; Academic Press: New York, NY, USA, 1978; pp. 355–358.
2. Maas, J.L. (Ed.) *Compendium of Strawberry Diseases*; APS Press: St. Paul, MN, USA, 1998; Volume 98.
3. Kanto, T.; Miyoshi, A.; Ogawa, T.; Maekawa, K.; Aino, M. Suppressive effect of potassium silicate on powdery mildew of strawberry in hydroponics. *J. Gen. Plant Pathol.* **2004**, *70*, 207–211. [CrossRef]
4. Nelson, M.D.; Gubler, W.D.; Shaw, D.V. Inheritance of powdery mildew resistance in greenhouse-grown versus field-grown California strawberry progenies. *Phytopathology* **1995**, *85*, 421–424. [CrossRef]
5. Amsalem, L.; Freeman, S.; Rav-David, D.; Nitzani, Y.; Sztejnberg, A.; Pertot, I.; Elad, Y. Effect of climatic factors on powdery mildew caused by Sphaerotheca macularis f. sp. fragariae on strawberry. *Eur. J. Plant Pathol.* **2006**, *114*, 283–292. [CrossRef]
6. Adam, L.; Somerville, S.C. Genetic characterization of five powdery mildew disease resistance loci in Arabidopsis thaliana. *Plant J.* **1996**, *9*, 341–356. [CrossRef] [PubMed]

7.   Meunkaewjinda, A.; Kumsawat, P.; Attakitmongcol, K.; Srikaew, A. Grape leaf disease detection from color imagery using hybrid intelligent system. In Proceedings of the 5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, ECTI-CON 2008, Krabi, Thailand, 14–17 May 2008; Volume 1, pp. 513–516.

8.   Kobayashi, T.; Kanda, E.; Kitada, K.; Ishiguro, K.; Torigoe, Y. Detection of rice panicle blast with multispectral radiometer and the potential of using airborne multispectral scanners. *Phytopathology* **2001**, *91*, 316–323. [CrossRef] [PubMed]

9.   Al-Hiary, H.; Bani-Ahmad, S.; Reyalat, M.; Braik, M.; ALRahamneh, Z. Fast and accurate detection and classification of plant diseases. *Mach. Learn.* **2011**, *17*, 31–38. [CrossRef]

10.   Rastogi, R.; Chadda, V.K. *Applications of Image Processing in Biology and Agriculture J.K. Sainis, Molecular Biology and Agriculture Division*; BARC Newsletter: Mumbai, India, 1989.

11.   Qin, J.; Burks, T.F.; Ritenour, M.A.; Bonn, W.G. Detection of citrus canker using hyperspectral reflectance imaging with spectral information divergence. *J. Food Eng.* **2009**, *93*, 183–191. [CrossRef]

12.   Rumpf, T.; Mahlein, A.K.; Steiner, U.; Oerke, E.C.; Dehne, H.W.; Plümer, L. Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance. *Comput. Electron. Agric.* **2010**, *74*, 91–99. [CrossRef]

13.   Laudien, R.; Bareth, G.; Doluschitz, R. Comparison of remote sensing based analysis of crop diseases by using high resolution multispectral and hyperspectral data–case study: Rhizoctonia solani in sugar beet. *Geoinformatics* **2004**, *7*, 670–676.

14.   Chaerle, L.; Van Caeneghem, W.; Messens, E.; Lambers, H.; Van Montagu, M.; Van Der Straeten, D. Presymptomatic visualization of plant–virus interactions by thermography. *Nat. Biotechnol.* **1999**, *17*, 813. [CrossRef]

15.   Kutty, S.B.; Abdullah, N.E.; Hashim, H.; Kusim, A.S.; Yaakub, T.N.T.; Yunus, P.N.A.M.; Rahman, M.F.A. Classification of watermelon leaf diseases using neural network analysis. In Proceedings of the 2013 Business Engineering and Industrial Applications Colloquium (BEIAC), Langkawi, Malaysia, 7–9 April 2013; pp. 459–464.

16.   Khirade, S.D.; Patil, A.B. Plant disease detection using image processing. In Proceedings of the 2015 International Conference on Computing Communication Control and Automation, Pune, India, 26–27 February 2015; pp. 768–771.

17.   Kim, D.G.; Burks, T.F.; Qin, J.; Bulanon, D.M. Classification of grapefruit peel diseases using color texture feature analysis. *Int. J. Agric. Biol. Eng.* **2009**, *2*, 41–50.

18.   Ei-Helly, M.; Rafea, A.; Ei-Gamal, S.; Whab, R.A.E. Integrating diagnostic expert system with image processing via loosely coupled technique. *Cent. Lab. Agric. Expert Syst.* **2004**. Available online: https://pdfs.semanticscholar.org/52c8/ecd947726e11997d9ec76de76b6d36752e9e.pdf (accessed on 3 September 2019).

19.   Weizheng, S.; Yachun, W.; Zhanliang, C.; Hongda, W. Grading method of leaf spot disease based on image processing. In Proceedings of the 2008 International Conference on Computer Science and Software Engineering, Wuhan, China, 12–14 December 2008; Volume 6, pp. 491–494.

20.   Schor, N.; Berman, S.; Dombrovsky, A.; Elad, Y.; Ignat, T.; Bechar, A. Development of a robotic detection system for greenhouse pepper plant diseases. *Precis. Agric.* **2017**, *18*, 394–409. [CrossRef]

21.   Sena, D.G., Jr.; Pinto, F.A.C.; Queiroz, D.M.; Viana, P.A. Fall armyworm damaged maize plant identification using digital images. *Biosyst. Eng.* **2003**, *85*, 449–454. [CrossRef]

22.   Camargo, A.; Smith, J.S. Image pattern classification for the identification of disease causing agents in plants. *Comput. Electron. Agric.* **2009**, *66*, 121–125. [CrossRef]

23.   Chen, C.H.; Pau, L.F.; Wang, P.S.P. *Handbook of Pattern Recognition and Computer Vision*; World Scientific: Singapore, 1998; pp. 207–248.

24.   Gonzalez, R.C.; Wood, R.E. *Digital Image Processing*, 4th ed.; Pearson, Inc.: New York, NY, USA, 2018.

25.   Shearer, S.A.; Holmes, R.G. Plant identification using color co-occurrence matrices. *Trans. ASAE* **1990**, *33*, 1237–1244. [CrossRef]

26.   Al Bashish, D.; Braik, M.; Bani-Ahmad, S. A framework for detection and classification of plant leaf and stem diseases. In Proceedings of the 2010 International Conference on Signal and Image Processing (ICSIP), Dallas, TX, USA, 14–16 December 2010; pp. 113–118.

27.   Xie, C.; Shao, Y.; Li, X.; He, Y. Detection of early blight and late blight diseases on tomato leaves using hyperspectral imaging. *Sci. Rep.* **2015**, *5*, 16564. [CrossRef]

28. Mucherino, A.; Papajorgji, P.; Pardalos, P.M. A survey of data mining techniques applied to agriculture. *Oper. Res.* **2009**, *9*, 121–140. [CrossRef]

29. Al Bashish, D.; Braik, M.; Bani-Ahmad, S. Detection and classification of leaf diseases using K-means-based segmentation and Neural-networks-based classification. *Inf. Technol. J.* **2011**, *10*, 267–275. [CrossRef]

30. Sankaran, S.; Mishra, A.; Maja, J.M.; Ehsani, R. Visible-near infrared spectroscopy for detection of Huanglongbing in citrus orchards. *Comput. Electron. Agric.* **2011**, *77*, 127–134. [CrossRef]

31. Omrani, E.; Khoshnevisan, B.; Shamshirband, S.; Saboohi, H.; Anuar, N.B.; Nasir, M.H.N.M. Potential of radial basis function-based support vector regression for apple disease detection. *Measurement* **2014**, *55*, 512–519. [CrossRef]

32. Wang, X.; Zhang, M.; Zhu, J.; Geng, S. Spectral prediction of Phytophthora infestans infection on tomatoes using artificial neural network (ANN). *Int. J. Remote Sens.* **2008**, *29*, 1693–1706. [CrossRef]

33. Pydipati, R.; Burks, T.F.; Lee, W.S. Statistical and neural network classifiers for citrus disease detection using machine vision. *Trans. ASAE* **2005**, *48*, 2007–2014. [CrossRef]

34. Xu, G.; Zhang, F.; Shah, S.G.; Ye, Y.; Mao, H. Use of leaf color images to identify nitrogen and potassium deficient tomatoes. *Pattern Recognit. Lett.* **2011**, *32*, 1584–1590. [CrossRef]

35. VijayaLakshmi, B.; Mohan, V. Kernel-based PSO and FRVM: An automatic plant leaf type detection using texture, shape, and color features. *Comput. Electron. Agric.* **2016**, *125*, 99–112. [CrossRef]

36. Yano, I.H.; Santiago, W.E.; Alves, J.R.; Toledo, L.; Mota, M.; Teruel, B. Choosing classifier for weed identification in sugarcane fields through images taken by uav. *Bulg. J. Agric. Sci.* **2017**, *23*, 491–497.

37. Debert, Nova Scotia Canada Yearly/Monthly/Daily Climate Data (n.d.). Available online: https://eldoradoweather.com/canada/climate2/Debert.html (accessed on 3 September 2019).

38. Steward, B.L.; Tian, L.F. Machine-vision weed density estimation for real-time, outdoor lighting conditions. *Trans. ASAE* **1999**, *42*, 1897. [CrossRef]

39. Chang, Y.K.; Zaman, Q.U.; Schumann, A.W.; Percival, D.C.; Esau, T.J.; Ayalew, G. Development of color co-occurrence matrix based machine vision algorithms for wild blueberry fields. *Appl. Eng. Agric.* **2012**, *28*, 315–323. [CrossRef]

40. Lee, S.; Ryu, J.H.; Won, J.S.; Park, H.J. Determination and application of the weights for landslide susceptibility mapping using an artificial neural network. *Eng. Geol.* **2004**, *71*, 289–302. [CrossRef]

41. Paul, A.K.; Das, D.; Kamal, M.M. Bangla speech recognition system using LPC and ANN. In Proceedings of the 2009 Seventh International Conference on Advances in Pattern Recognition, Kolkata, India, 4–6 February 2009; pp. 171–174.

42. Choose Classifier Options (n.d.). Available online: https://www.mathworks.com/help/stats/choose-a-classifier.html#bunt0n0-1 (accessed on 3 September 2019).

43. Torrecilla, J.S.; Otero, L.; Sanz, P.D. A neural network approach for thermal/pressure food processing. *J. Food Eng.* **2004**, *62*, 89–95. Available online: http://dx.doi.org/10.1016/S0260-8774(03)00174-2 (accessed on 3 September 2019). [CrossRef]

44. Madadlou, A.; Emam-Djomeh, Z.; Mousavi, M.E.; Ehsani, M.; Javanmard, M.; Sheehan, D. Response surface optimization of an artificial neural network for predicting the size of re-assembled casein micelles. *Comput. Electron. Agric.* **2009**, *68*, 216–221. [CrossRef]

45. Cristianini, N.; Shawe-Taylor, J. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*; Cambridge University Press: Cambridgeshire, UK, 2000.

46. Schölkopf, B.; Smola, A.J. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*; MIT Press: Cambridge, MA, USA, 2002.

47. Wspanialy, P.; Moussa, M. Early powdery mildew detection system for application in greenhouse automation. *Comput. Electron. Agric.* **2016**, *127*, 487–494. [CrossRef]

48. Kulkarni, A.H.; Patil, R.A. Applying image processing technique to detect plant diseases. *Int. J. Mod. Eng. Res.* **2012**, *2*, 3661–3664.

49. Ramakrishnan, M. Groundnut leaf disease detection and classification by using back probagation algorithm. In Proceedings of the 2015 International Conference on Communications and Signal Processing (ICCSP), Chengdu, China, 10–11 October 2015; pp. 0964–0968.

50. Islam, M.; Dinh, A.; Wahid, K.; Bhowmik, P. Detection of potato diseases using image segmentation and multiclass support vector machine. In Proceedings of the 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE), Windsor, ON, USA, 30 April–3 May 2017; pp. 1–4.

51. Kurtulmus, F.; Lee, W.S.; Vardar, A. Green citrus detection using 'eigenfruit', color and circular Gabor texture features under natural outdoor conditions. *Comput. Electron. Agric.* **2011**, *78*, 140–149. [CrossRef]

52. Aggelopoulou, A.D.; Bochtis, D.; Fountas, S.; Swain, K.C.; Gemtos, T.A.; Nanos, G.D. Yield prediction in apple orchards based on image processing. *Precis. Agric.* **2011**, *12*, 448–456. [CrossRef]

53. Rehman, T.; Zaman, Q.; Chang, Y.K.; Schumann, A.; Corscadden, K.; Esau, T. Optimizing the parameters influencing performance and weed (Goldenrod) identification accuracy of color co-occurrence matrices. *Biosyst. Eng.* **2018**, *170*, 85–95. [CrossRef]