JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGY
http://e-journal.uum.edu.my/index.php/jict

Ong, J. H., Ong, P., & Woon, K. L. (2022). Image-based oil palm leaves disease detection using convolutional neural network. *Journal of Information and Communication Technology, 21*(3), 383-410. https://doi.org/10.32890/jict2022.21.4

# Image-based Oil Palm Leaf Disease Detection using Convolutional Neural Network

[1]Jia Heng Ong, *[2]Pauline Ong & [3]Woon Kiow Lee
Faculty of Mechanical and Manufacturing Engineering,
University Tun Hussein Onn Malaysia, Malaysia

ongjiaheng0612@gmail.com
*ongp@uthm.edu.my
wklee@uthm.edu.my
*Corresponding author

Received: 14/3/2022 Revised: 14/4/2022 Accepted: 18/4/2022 Published: 17/7/2022

## ABSTRACT

Over the years, numerous studies have been conducted on the integration of computer vision and machine learning in plant disease detection. However, these conventional machine learning methods often require the contour segmentation of the infected region from the entire leaf region and the manual extraction of different discriminative features before the classification models can be developed. In this study, deep learning models, specifically, the AlexNet convolutional neural network (CNN) and the combination of AlexNet and support vector machine (AlexNet-SVM), which overcome the limitation of handcrafting of feature representation were implemented for oil palm leaf disease identification. The images of healthy and infected leaf samples were collected, resized, and renamed before the model training. These images were directly used to fit the classification models, without the need for segmentation and feature extraction as in the conventional machine learning methods. The optimal architecture of AlexNet CNN and AlexNet-SVM models were then determined and subsequently applied for the oil palm leaf disease identification. Comparative studies showed that the overall performance of the AlexNet CNN model outperformed AlexNet-SVM-based classifier.

**Keywords:** AlexNet, convolutional neural network, leaf disease, oil palm, support vector machine.

## INTRODUCTION

Oil palm (*Elaeis Guineensis*), a West African agricultural plant, was brought into Malaysia by the British in the early 1870s. It is an important economic crop for Malaysia, contributing 2.7 percent to the gross domestic product in 2020. The environmental conditions in Malaysia with adequate sunlight and rain supply throughout the year provide a suitable cultivation environment for oil palms. However, such high ambient humidity promotes the growth of many plant diseases, especially those fungal diseases that spread rapidly under high humidity.

The oil palm diseases can be categorized into three different types based on the region of symptoms, i.e., root, basal stem, and leaf. The detection of oil palm leaf disease is chosen as the main focus of this work instead of root or basal stem because the disease mainly occurs on leaves (Septiarini et al., 2021). Oil palm leaf diseases can be characterized into three classes: leaf rot, leaf spot, and leaf speckle. Similar to basal stem rot disease, leaf disease is infected by fungus type parasites (*Curvularia spp*.), which can widely spread throughout the whole farming area. The formation of brownish or yellowing dots on leaves is the basic symptom of leave diseases, which is visible to the naked eye. Devastating damage to crop health and reduced fruit quantity and quality occur if the disease is left untreated (Sunpapao et al., 2018). It has been pointed out that the area infected by oil palm disease in Malaysia will reach almost 450,000 hectares by 2020 (Roslan & Idris, 2012) and thus, monitoring of oil palm disease needs more work.

Conventionally, the health condition monitoring of oil palm is conducted manually by observation through the naked eye. When disease infection is confirmed, the tree-felling method is applied to delay the disease front. The commonly practiced manual plant disease identification is a laborious process, requiring expertise in plant pathology and continuous monitoring. This visual inspection is time-consuming and thus not feasible for large areas of oil palm plantations (Alaa et al., 2020). Moreover, since the analysis of oil palm disease often requires trained expertise, which is usually lacking in remote areas and small farms, the detection results are prone to human misjudgment (Masazhar & Kamal, 2017). Therefore, the daily crop health monitoring with conventional methods is less efficient for small and medium plantation farms. Moreover, due to workforce shortage, massive import of foreign labor into commercial oil palm plantation sectors causes difficulty for farm holders to provide job-related technical training, not to mention plant disease detection that requires years of experience. Intrinsically, an automated solution for plant disease identification may be helpful for crop health monitoring.

## RELATED WORKS

### Plant Disease Detection using Conventional Machine Learning Techniques

In recent years, numerous studies have been conducted on the integration of computer vision and conventional machine learning techniques in plant disease detection, including the detection of target spots in tomatoes using multilayer neural network (Abdulridha et al., 2020), potato plant leaf disease detection with K-nearest neighbor (KNN), Naïve Bayes (NB) classifier and support vector machine (SVM) (Abdu et al., 2020), identification of wheat yellow rust using SVM (Guo et al., 2020), rice disease detection by radial basis function neural network (Rath & Meher, 2019) and decision tree (Yang et al., 2019), detection of foliar disease in tea plants using clustering method (Yuan et al., 2019), orange fungal decay detection using Fisher's linear discriminant analysis (Ghanei Ghooshkhaneh et al., 2018), detection of brinjal leaf disease using artificial neural network (ANN) (Veni et al., 2017), and identification of wheat yellow rust disease using Gaussian process regression (Ashourloo et al., 2016), to name a few.

Various works have been reported on the detection of oil palm leaf diseases using the integration of computer vision and conventional machine learning techniques as well. Aji et al. (2013) applied the features of RGB values, average brightness, standard deviation, and shape obtained from the leaf images to develop an ANN classifier for the detection of oil palm leaf diseases. The average classification accuracy of 87.75 percent was reported when the number of hidden nodes was assigned as 6. The performances of six different classifiers, namely KNN, NB, C4.5, decision tree, ANN, and SVM, in identifying healthy and infected oil palm leaves were studied by Hamdani et al. (2021). They concluded that applying the principal component analysis to extract the input features for ANN by splitting the histogram of RGB, L*a*b, HIS, and HSV color spaces into eight bins provided the highest accuracy. Masazhar and Kamal (2017) developed a multi-class SVM to classify the oil palm leaf images of Anthracnose and Chimaera diseases. The K-means clustering algorithm was used to separate the region with the disease symptom from the entire image. Subsequently, the texture features extracted from the infected leaf region using the gray level co-occurrence matrix were applied as the input to SVM. The classification accuracies of 97 percent and 95 percent were attained for the Chimaera and Anthracnose diseases, respectively. Septiarini et al. (2021) applied four different classifiers, namely KNN, NB, decision tree, and SVM, to classify the oil palm leaf images into two categories (healthy and unhealthy). The Otsu thresholding method was applied to detect the infected region from the oil palm leaf images, followed by the correction-based feature selection to generate a set of features to be used by the classifiers. The highest classification of 99 percent was reported by the KNN classifier. Despite the good results achieved by classical machine learning techniques, they often require segmentation of plant disease regions and manual extraction of different features and are therefore time-consuming and labor-intensive. Improperly selecting the discriminative features from the data may jeopardize the generalization performance of the machine learning models.

## Plant Disease Detection using Convolutional Neural Network

To overcome the limitations of conventional machine learning techniques, the utilization of deep neural networks, in particular, the convolutional neural network (CNN), has seen a rise in plant disease detection in recent years. The prominent advantage of CNN is that its feature extraction is learned automatically from raw data through a cascade of multiple layers instead of handcrafting the optimum feature representation of data within the domain knowledge. In this regard, Ferentinos et al. (2018) developed a multi-plant species disease detection system with the accuracies of 99.45 percent and 99.53 percent achieved for the AlexNet and VGG models, respectively. Pantazi et al. (2019) applied ANN and deep neural network in paddy leaf disease classification using the Jaya optimization algorithm, with the accuracies of 90 percent and 97 percent attained. Coulibaly et al. (2019) developed a crop disease detection system based on CNN for pearl millet, in which the detection rate of 95 percent was achieved. By utilizing AlexNet and VGG as the pre-trained models, Han and Gao (2019) developed a hyperspectral imaging system for Aflatoxin detection in peanuts based on CNN. The hyperspectral images were used for model training, where a high recognition success rate that exceeded 90 percent was attained. Liu and Wang (2020) exploited Yolo V3 to detect the location and type of tomato diseases and pests, and the results showed a detection accuracy of 92.39 percent, which outperformed Faster R-CNN. Chen et al. (2020) studied plant leaf disease identification of rice and maize using the VGG pre-trained model and Inception module. The proposed method achieved a detection accuracy greater than 91.83 percent. Darwish et al. (2020) combined VGG16 and VGG19 to conduct the detection of plant disease on maize leaves. The results showed that the ensemble model outperformed or was comparable with the optimized VGG16, VGG19, Inception V3, and Xception. Picon et al. (2019) developed a real smartphone application based on CNN to identify three different diseases on wheat images. It achieved an average balanced accuracy of 0.87 for both early and late diseases. Alaa et al. (2020) applied VGG16 CNN to detect the oil palm leaf diseases of leaf spot and blight spot. They concluded that the CNN model could successfully identify the oil palm leaf diseases even if the images were taken at nighttime.

The review of previous studies using the CNN classifier for plant disease detection showed that the classifier gives a high accuracy level for leaf disease detection and classification, but requires a large amount of image dataset for the model training process. However, as compared to the classical machine learning models, the use of CNN does not require image pre-processing such as the contour segmentation of the infected region from the entire leaf region and manual feature extraction. The entire process from feature extraction up to disease classification is integrated into the network architecture.

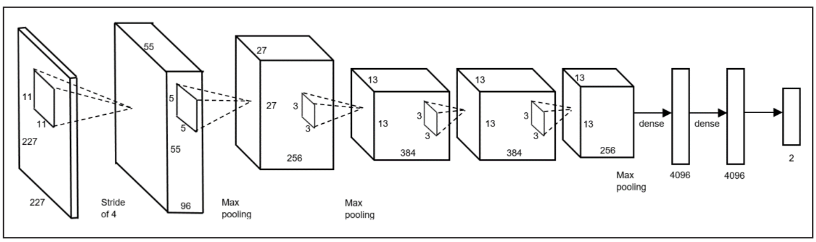
The concept of transfer learning has been widely applied when training a CNN classifier with the implementation of a pre-trained neural network model. The total detection model training time can be reduced significantly through this approach. Therefore, in this study, the AlexNet pre-trained CNN model is utilized for the detection of oil palm leaf disease. Additionally, the use of the AlexNet pre-trained CNN model as the feature extractor and its combination with SVM as the classifier (AlexNet-SVM) in the detection of oil palm leaf disease is examined.

**AlexNet Pre-Trained CNN Model**

Different from the single-hidden-layer ANN, CNN consists of several additional layers that can extract higher-level features from the raw images sequentially. AlexNet (Krizhevsky et al., 2017) is one of the most widely used CNN models. The basic network architecture of the AlexNet CNN model is shown in Figure 1.

**Figure 1**

*The Basic Architecture Network of AlexNet CNN Model*




The AlexNet architecture consists of five convolutional layers. The convolutional layer functions as the image feature extractor and feature map dimension minimizer. The input images with the dimension of 227 x 227 pixels will pass through these convolutional layers, before moving on to the fully connected layers. The pooling layer acts as the dimensional regulator as it resizes the feature map based on the subsequent layer size. Before entering the fully connected layer, the feature map is reduced from 6 x 6 x 256 to 1 x 1 x 4096. The fully connected layer works as the layer activation connector. It links all the layers to be in the same layer. The Softmax activation function is used in the last dense layer. By default, the pre-trained AlexNet CNN model has been designed to classify 1,000 classes of objects. Table

1 summarizes the detailed information of all layers in the AlexNet architecture.

**Table 1**

*Detailed Information of All Layers in the AlexNet CNN Model*

| No. of Layer | Name of Layer | Kernel Size | Function |
|---|---|---|---|
| 1 | Input | 227 x 227 x 3 | Input image with zero normalization |
| 2 | Convolution 1 | 55 x 55 x 96 | First feature map conversion and dimension reduction |
| 3 | ReLU 1 | 55 x 55 x 96 | |
| 4 | Normalization 1 | 55 x 55 x 96 | |
| 5 | Pooling 1 | 27 x 27 x 96 | |
| 6 | Convolution 2 | 27 x 27 x 256 | Second feature map conversion and dimension reduction |
| 7 | ReLU 2 | 27 x 27 x 256 | |
| 8 | Normalization 2 | 27 x 27 x 256 | |
| 9 | Pooling 2 | 13 x 13 x 256 | |
| 10 | Convolutional 3 | 13 x 13 x 384 | Third feature map conversion and dimension reduction |
| 11 | ReLU 3 | 13 x 13 x 384 | |
| 12 | Convolutional 4 | 13 x 13 x 384 | Fourth feature map conversion and dimension reduction |
| 13 | ReLU 4 | 13 x 13 x 384 | |
| 14 | Convolutional 5 | 13 x 13 x 256 | Fifth feature map conversion and dimension reduction |
| 15 | ReLU 5 | 13 x 13 x 256 | |
| 16 | Pooling 5 | 6 x 6 x 256 | |
| 17 | Fully Connected 6 | 1 x 1 x 4096 | First layer connection and modify the layer to 1x1x4096 |
| 18 | ReLU 6 | 1 x 1 x 4096 | |
| 19 | Dropout 6 | 1 x 1 x 4096 | |
| 20 | Fully Connected 7 | 1 x 1 x 4096 | Second layer connection |
| 21 | ReLU 7 | 1 x 1 x 4096 | |
| 22 | Dropout 8 | 1 x 1 x 4096 | Thrid layer connection and reduce the layer size to the intended class number |
| 23 | Fully Connected 8 | 1 x 1 x 1000 | |
| 24 | Softmax | 1 x 1 x 1000 | Feature classification |
| 25 | Output | - | Output classification result |

Transfer learning is a popular approach used in deep learning. The pre-trained CNN model on large annotated image databases is reused to provide domain adaption for a new problem. This allows the new problem to be solved with a relatively small amount of data. The training of the pre-trained CNN model with transfer learning is relatively faster since most layers of the pre-trained model remain unchanged. Only a few layers are retrained for the new problem.

The implementation of the AlexNet pre-trained CNN model with transfer learning in plant disease detection has been widely studied. Singh et al. (2022) applied the AlexNet model to detect the leaf-spot-based and rust-based diseases in maize plants, with an average accuracy of 99.16 percent attained. Rashwan and Elteir (2022) compared the performances of AlexNet with MobileNetV2 in plant leaf disease detection in terms of accuracy and inference performance. The results showed that AlexNet achieved higher detection accuracy than MobileNetV2, but with slower inference performance on a professional class mobile device. Elaraby et al. (2022) used the AlexNet and VGG16 models to diagnose citrus leaf diseases. Significant improvements in terms of accuracy, precision, sensitivity, specificity, and F-score were observed, when both models were trained based on the stochastic gradient descent with momentum (SGDM) optimizer. Elaraby et al. (2022) applied the AlexNet model to detect the plant diseases in corn, cotton, cucumbers, grapes, and wheat. The results revealed that the predictive capability of the AlexNet model was improved when the hyperparameters of AlexNet were optimized using particle swarm optimization. For identifying the diseases in soybean leaves, Jadhav et al. (2021) used the pre-trained AlexNet and GoogleNet CNN models based on a transfer learning approach. The accuracy of AlexNet was 98.75 percent, which outperformed the GoogleNet CNN model with an accuracy of 96.25 percent.

**AlexNet-SVM Model**

SVM is a classifier that separates the samples into different groups based on the hyperplane concept. The learning of SVM will generate a hyperplane or barrier of separation between two classes based on the nature of the data. In the training process, SVM attempts to search for the optimum hyperplane gradient. The iterative computation increases the margin of its separation plane to the maximum margin. Theoretically, a traditional SVM classifier that utilizes a one-against-one approach can only perform the separation of data into two distinct classes at once, which is not compatible with multi-class classification. However, this can be solved by the multi-class SVMs with the one-against-rest approach, involving the combination of more than one SVM linear classifier. In this study, the AlexNet pre-trained CNN model is also used to extract input image features. The extracted features are then applied as the input of SVM for oil palm leaf disease detection. It has been suggested that the features selected from the

CNN model could be the primary features in most image recognition problems (Roy et al., 2018).

Kawatra et al. (2020) used AlexNet pre-trained CNN with three different configurations for the leaf disease detection of 12 crop species. The first configuration was the original AlexNet model. The second configuration was the combination of AlexNet with the global average pooling layer. The third configuration used the extracted features from AlexNet as the input of SVM. The combination of AlexNet with SVM gave the best results, with a validation accuracy of 99.98 percent attained. To detect the grapevine yellow diseases of grapes, Ampatzidis et al. (2018) applied a linear SVM to classify the features extracted from AlexNet pre-trained CNN. The developed system achieved a classification accuracy of 95.23 percent. For comparison, SVM with the extracted features using local binary patterns and color histogram was considered, in which the attained accuracy was merely 26.7 percent. Muhammad et al. (2021) extracted the features of aloe vera disease and apple disease using AlexNet and VGG19 pre-trained models. Subsequently, these features were used as the input of KNN, SVM, probabilistic neural network, fuzzy logic, and ANN. For both diseases, the best results were obtained when the features of AlexNet were used to develop the SVM classifier. Khan et al. (2020) introduced multiple deep models based on AlexNet and VGG to extract the features of plant diseases and adopted a fusion strategy to combine the extracted features into a single vector. Türkoğlu and Hanbay (2019) considered three different pre-trained CNN models, namely AlexNet, VGG16, and VGG19, for the feature extraction of plant disease detection. The obtained features were then classified by SVM, extreme learning machine, and KNN. The highest accuracy of 95.5 percent was achieved for the combination of AlexNet as feature extractor and SVM as classifier.

## MATERIALS AND METHOD

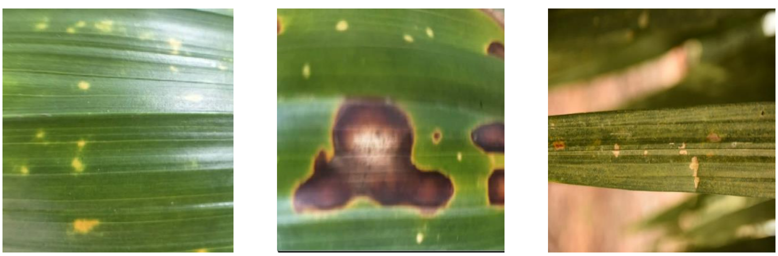
### Data Acquisition

An image acquisition campaign took place in a commercial oil palm farm in Pontian, Johor, Malaysia from September 2019 to February 2020, covering the images of the oil palm disease at different phenological growth stages. The leaf images were captured through a smartphone camera without additional image filters and color compensation functions. The data were taken at 9.00 a.m. (morning), 1.00 p.m. (afternoon), and 5.00 p.m. (afternoon) with the mixture of images captured with and without a macro lens adaptor. Two thousand images of oil palm disease with the combination of healthy (500 images), eye spot (500 images), leaf rot (500 images), and leaf speckle (500 images) disease leaves were acquired. The acquired images were in the size of 4608 x 2592 pixels. Subsequently, the raw images were divided into three different batches: training dataset, testing dataset, and validation dataset. The leaf images were distributed based on the ratio of 70 percent (training), 20 percent (testing), and 10 percent (validation).

Figure 2 presents the example of the acquired oil palm leaf diseases in this study.

**Figure 2**

*The Example of Oil Palm Leaf Diseases (a) Eye Spot; (b) Leaf Rot; and (c) Leaf Speckle*



(a) (b) (c)

## Image Enhancement and Resizing

To reduce the computing unit workload during the training of the CNN classifier, each acquired RGB image in the size of 4608 x 2592 pixels was manually cropped to 227 x 227 pixels. This is also to fit the AlexNet pre-trained CNN model with the input size of 227 x 227 pixels. For each original image, only one selected region showing symptoms of leaf disease was cropped. The histogram equalization function was then applied to compensate for the image contrast for better sharpness. The equalized images were then converted into

L*a*b color space format (Habib et al., 2020) before feeding into the AlexNet pre-trained CNN model.

### Development of AlexNet Pre-Trained CNN Model

A standard AlexNet pre-trained CNN model was applied in this study for the detection of oil palm leaf disease. By default, the pre-trained AlexNet CNN model has been designed to classify 1,000 classes of objects, and thus, the kernel size of the Softmax layer is 1 x 1x 1000 (refer to Table 1). To apply transfer learning, the dimension of the Softmax layer was modified to 1 x 1 x 4 such that the AlexNet pre-trained model could be used for the oil palm leaf disease detection with four different classes (healthy, eye spot, leaf rot, and leaf speckle). The dimensions of other layers before the Softmax layer were fixed as in Table 1. The AlexNet CNN model was optimized in terms of validation accuracy and training time by fine-tuning the hyperparameters of learning rate, maximum epoch, momentum, batch size, and optimizer type.

The performance of the AlexNet CNN model was evaluated based on the evaluation metrics of accuracy, recall, specificity, precision, and F1-score. All simulations were carried out in the MATLAB R2019a environment (MathWorks Inc, Natrick, MA, USA).

## EXPERIMENTAL RESULTS AND DISCUSSION

### Fine-Tuning of AlexNet CNN Classifier

Before the implementation of the AlexNet CNN model in the oil palm leaf disease identification, the optimal network parameters were identified to maximize the prediction accuracy. Five hyperparameters were taken into consideration, namely learning rate, maximum epoch, momentum, batch size, and optimizer.

#### *Learning Rate*

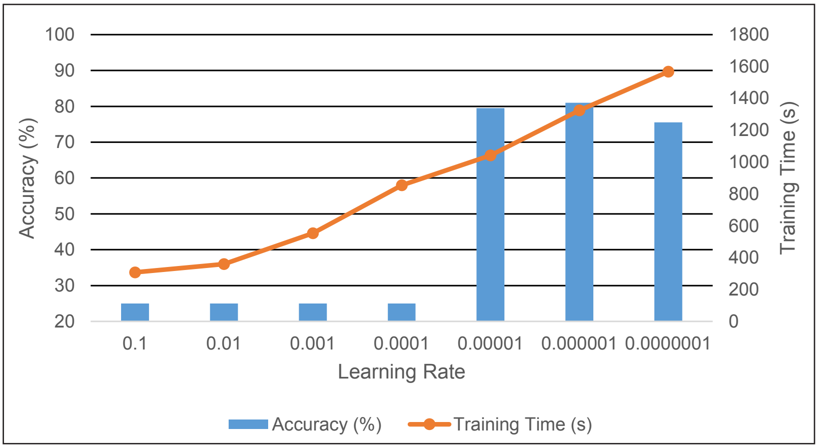
Learning rate is one of the tunable values in the training of a classification model. It is applied to the weights of the network toward the loss gradient across the training process. Generally, the choice of a suitable learning rate is subjective and dependent on the type of network architecture and data. A fast-learning rate will result in high training speed, but the chances of under-fitting are high.

To fine-tune the learning rate, the values of learning rate were varied from 0.1 to 0.0000001 (see Figure 3), while the values of maximum epoch, momentum, batch size, and type of optimizer were fixed as 50, 0.9, 4, and ADAM, respectively. Figure 3 shows the obtained accuracy and computation time during the training process by utilizing different learning rates. As shown in Figure 3, the model accuracy was significantly varied with the changes in learning rate. A smaller learning rate improved the generalization capability of the prediction model, but at the expense of a longer computation time. The results were in accordance with the reported study by Jacobs (1988). The highest accuracy of 81 percent was attained when the learning rate was marked at the value of 0.000001. A further decrease in learning rate did not significantly improve the prediction accuracy but took a longer duration to converge. For this reason, the learning rate of 0.000001 was selected and used for the subsequent analysis.

Learning rate is the scale factor of weight at each step during the training of the classification model. Even though a larger learning rate, such as 0.1, 0.01, and 0.001, shows a faster training process, the obtained training accuracy may not be satisfactory. It can be seen in Figure 3 that the model with a large learning rate experienced under-fitting and failed to obtain higher accuracy, although the training process could be completed in a faster manner than the model with a smaller learning rate. Utilizing a low learning rate may reduce the chances of over-fitting during the model training and lead to a smoother convergence training curve and higher accuracy (Tang et al., 2019). However, too small a learning rate results in slow convergence (Hannan et al., 2020).

**Figure 3**

*The Effect of Using Different Learning Rates to the Obtained Accuracy and Computation Time During the Training Process of the AlexNet CNN Model*



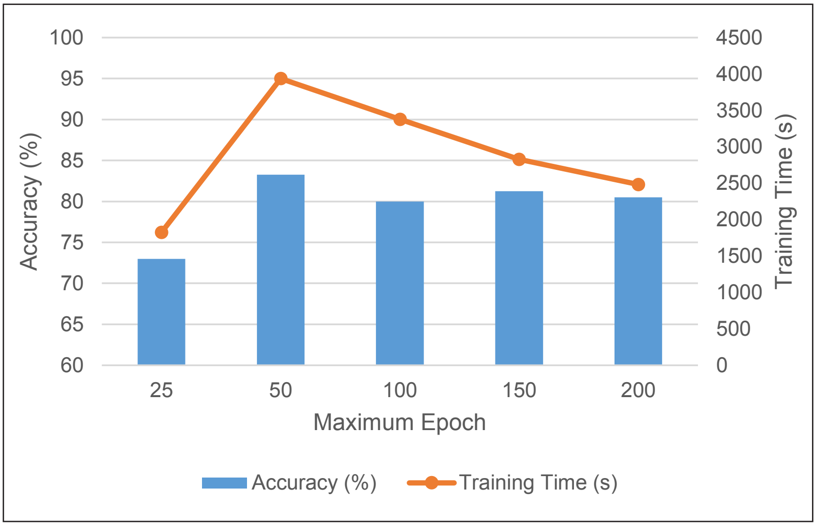

### *Maximum Epoch*

Epoch represents how frequent the entire training datasets pass through the network. It serves as one of the most crucial hyperparameters for the data to be optimally fitted. The concept of optimum value is preferred when training a deep learning model because it minimizes the training time without sacrificing the model's accuracy. The model was tested with four different maximum epoch values: 50, 100, 150, and 200. Other parameters were the learning rate – 0.0001, momentum – 0.9, batch size – 4, and type of optimizer – ADAM.

Figure 4 illustrates the variation of prediction accuracy and total computation time during the training process by considering different maximum epochs. The distinct feature of the prediction model with transfer learning that required fewer epochs is clearly shown in Figure 4. Among all the considered values, the prediction model with 50 maximum epochs achieved the highest accuracy. The declination of the prediction accuracy occurred after this value. The manipulation of this parameter beyond this value did not give significant improvement to the model's validation accuracy. Therefore, the optimum maximum epoch was chosen to be 50.

**Figure 4**

*The Effect of Using Different Maximum Epoch Values to the Obtained Accuracy and Computation Time During the Training Process of the AlexNet CNN Model*




***Momentum***

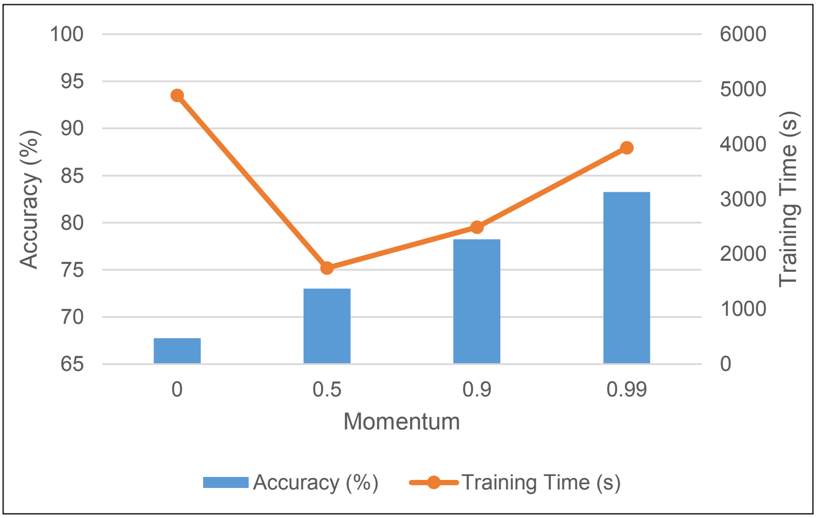
Momentum is the gain parameter applied to the learning algorithm. It increases pressure on a training session for faster training accuracy and speed. A high momentum value allows the model to react faster in the training session. The integration of a momentum parameter into the optimizer of the AlexNet CNN model may move the overall prediction closer to the ideal outcome (Skowron et al., 2020). The determination of the optimum momentum value will be based on the training results with the momentum values of 0, 0.5, 0.9, and 0.99.

Figure 5 shows the relationship between accuracy and total training time against different momentum values used in the ADAM optimizer. As indicated in this figure, an increase in the momentum value was accompanied by increasing accuracy. By varying the momentum from 0 to 0.99, the highest accuracy of 83.25 percent was recorded at the momentum value of 0.99. The momentum applied in the ADAM optimizer allowed the gradient fitted to the best fit situation to accumulate the past gradient as the prediction for the next step.

The setting of high momentum during the training process would improve the accuracy, and the effect on the overall training time was significant. Throughout the model training, the learning curve would keep fluctuating upon completion; the utilization of momentum allowed the learning curve to get out from the local minima of each fluctuation faster and allowed the learning validation updates to respond faster. In this case, prediction accuracy was the priority when selecting a suitable parameter setting and thus, the momentum value of 0.99 would be chosen for further AlexNet modeling.

**Figure 5**

*The Effect of Using Different Values of Momentum to the Obtained Accuracy and Computation Time During the Training Process of the AlexNet CNN Model*

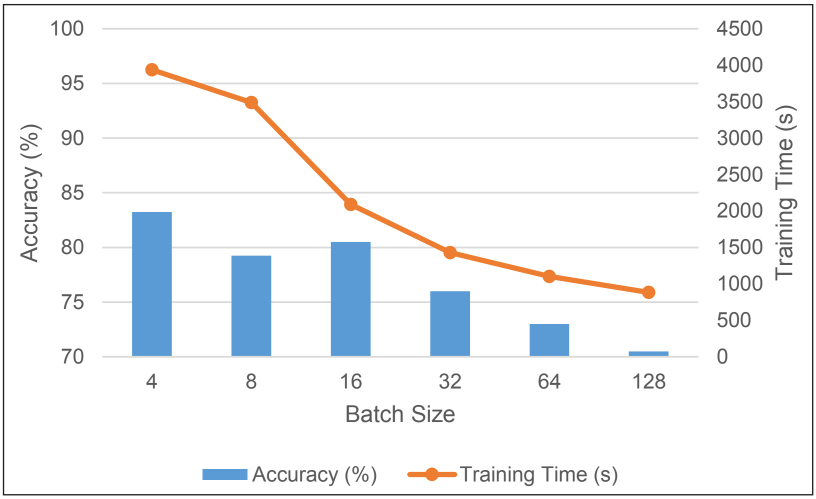



### *Batch Size*

Data batch size represents the number of samples propagating through the network during each cycle of training. It affects the generalization gap difference between the training and testing time and leads to high losses during the training process (Hoffer et al., 2017). The larger the data batch size, the faster the computational speed. However, a large data batch size does not guarantee a promising training performance, particularly when the model training is conducted on a non-parallel computing featured machine (Hoffer et al., 2017). Figure 6 illustrates how the training accuracy and computation time vary when different data batch sizes of 4, 8, 16, 32, 64, and 128 are assigned.

**Figure 6**

*The Effect of Using Different Values of Batch Size to the Obtained Accuracy and Computation Time During the Training Process of the AlexNet CNN Model*




As shown in Figure 6, in general, a decreasing trend in the prediction accuracy could be observed with the increasing batch size. The model accuracy was marked at a peak with the predefined minimum batch size of 4. Furthermore, its total training time of 3877 seconds was below the average of the total time taken by all batch sizes, i.e., 4228.67 seconds. The obtained result was in agreement with the conclusion in Hoffer et al.'s (2017) study. It was found that a small batch size could minimize the generalization gap during the training process, producing relatively higher accuracy results than that of the model with large batch size, with the learning rate, momentum, and gradient clipping fixed for both models (Hoffer et al., 2017).

### *Type of Optimizer*

Optimizer is the main solver for any CNN model training. It allows the training process to complete with the lowest training loss by continuously manipulating the weight and learning rate across the learning cycle. ADAM, SGDM, and root mean squared propagation (RMSProp) are the common optimization algorithms applied to improve the training process through gradient manipulation. The network performance by utilizing these three optimizers is depicted in Figure 7. It is pertinent to note that the low epoch value of 50 was applied.

**Figure 7**

*The Effect of Using Different Types of Optimizers to the Obtained Accuracy and Computation Time During the Training Process of the AlexNet CNN Model (Epoch Value of 50)*

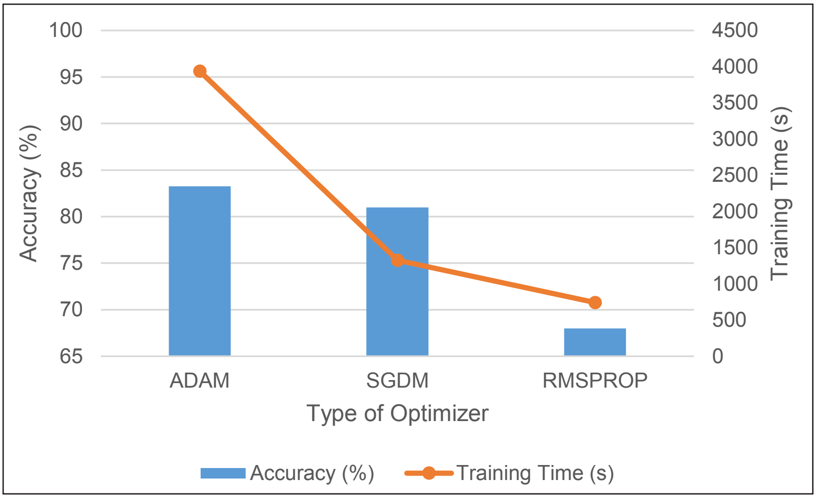
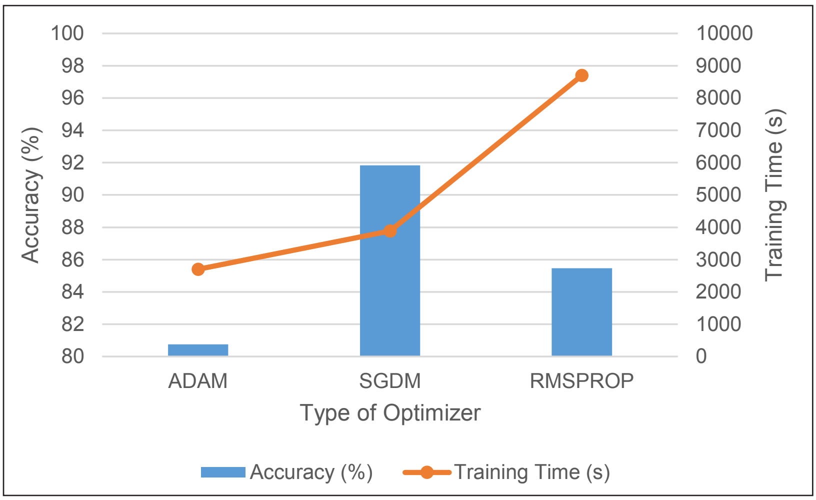



Figure 7 presents that the highest prediction accuracy of 83.25 percent was attained by the ADAM optimizer, while the model accuracy with the SGDM optimizer was 2.7 percent lower. This finding was in agreement with the work reported by Tang et al. (2019). By applying transfer learning with the CIFAR-100 dataset on ResNet-34, the network with the ADAM optimizer gave a relatively higher accuracy than that of the SGDM optimizer at a low epoch value. It was due to the poor generalization performance of SGDM during the training process. However, for the model training iteration approaches to achieve unity, the accuracy of SDGM would be higher than ADAM and RMSProp (Tang et al., 2019). For the ADAM optimizer, the high progressive epoch might lead to the failure of the algorithm to search for the correct path for effective convergence toward the sub-optimal point.

Figure 8 shows the prediction accuracy and total computation time attained by the network with different types of optimizers at the epoch value of 200. With a larger epoch value, SGDM outperformed the ADAM optimizer, with the attained prediction accuracy of 91.84 percent. The result corroborated the findings of Tang et al. (2019) on the effect of a larger epoch value on the improved accuracy of the SGDM optimizer.

**Figure 8**

*The Effect of Using Different Types of Optimizers to the Obtained Accuracy and Computation Time During the Training Process of the AlexNet CNN Model (Epoch Value of 200)*




It can be seen that the ADAM and SGDM optimizers worked effectively at the low and high epoch values, respectively. Given these special characteristics of ADAM and SGDM optimizers, the idea of multi-optimizer fine-tuning was applied in this study. The SGDM optimizer was used in the second attempt of model fine-tuning to obtain higher accuracy. Applying multi-optimizers in classification model training had shown to improve the accuracy by about 3 percent after deploying the SGDM optimizer as the second optimizer during the critical liver alteration using CNN models (Arjmand et al., 2020). Attempting to obtain a classification model with optimum performance, the ADAM trained AlexNet model was retrained with the same batch of training datasets using the hyperparameters as shown in Table 2. The implementation of these two levels of fine-tuning resulted in outstanding accuracy. The prediction accuracy was improved from 80.75 percent to 95.5 percent.

**Table 2**

*The Optimized Hyperparameters used for AlexNet CNN Model*

| Hyperparameter | Stage | |
|---|---|---|
| | First Fine-Tuning | Second Fine-Tuning |
| Learning Rate | 0.000001 | 0.000001 |
| Epoch | 50 | 200 |
| Momentum | 0.99 | 0.99 |
| Batch Size | 4 | 4 |
| Optimizer | ADAM | SGDM |

## Fine-Tuning of AlexNet-SVM Classifier

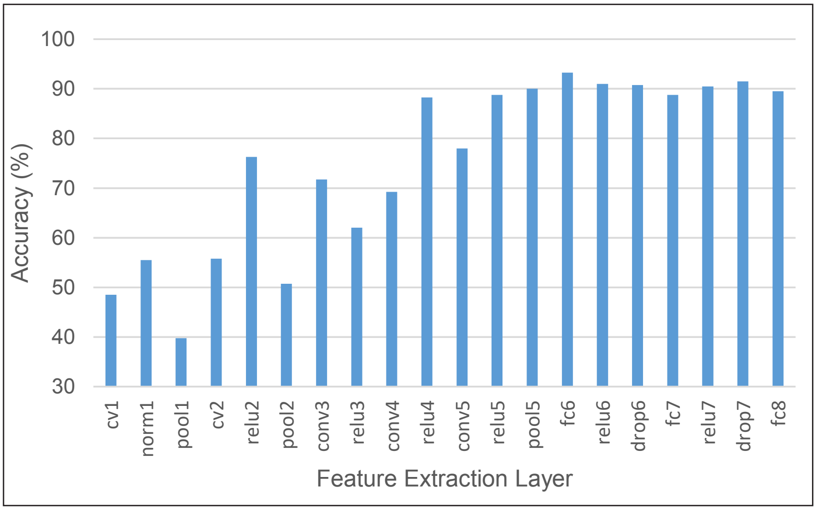
The use of convolutional layer, pooling layer, or the fully connected layer in CNN as the feature extraction has been widely studied. This approach eases the pre-processing stage by removing the segmentation and statistical or gray level co-occurrence matrix feature extraction. The AlexNet architecture consists of several layers, as shown in Table 1. Therefore, selecting the most appropriate layer as the main feature extraction layer for leaf disease classification was first analyzed, followed by the selection of the appropriate optimizer.

### *Selection of the Feature Extraction Layer*

AlexNet consists of 20 image pre-processing layers, where feature maps are extractable from each layer. The identification of a suitable feature map was crucial in this case. This is due to the SVM model requiring the feature map to have the lowest amount of noise and strong disease feature layer as the input data. Figure 9 illustrates the prediction accuracy of the AlexNet-SVM model obtained using the extracted features from different AlexNet feature extraction layers. AlexNet-SVM using the feature map extracted from fully connected layer 6 (fc6) achieved the highest accuracy of 93.3 percent. As the fundamental concept of the feature extraction layer of CNN, an image will be filtered and diminished to form a feature map and coagulated to be smaller and smaller across the layers. The selection of the most appropriate layer allows the SVM classifier to be fitted at a low loss rate and results in higher accuracy. The AlexNet architecture consists of five convolutional layers that serve as the image dimension minimizer. At the first few convolutional layers (cv1 to cv4), the generated image feature maps were not precise enough to represent the leaf disease features, as the image features extracted were simpler and coarser, such as color, shape, and lines that exist in the image data. Therefore, classifier training with these data would result in low classification accuracy. For the deeper layers (cv5 to fc8), disease feature maps were refined and the classifier would be able to learn more complicated and abstract disease features. Therefore, the classification model with these input data would achieve higher validation accuracy.

**Figure 9**

*The Effect of Using Different Feature Extraction Layers to the Obtained Accuracy and Computation Time During the Training Process of the AlexNet-SVM Model*




***Selection of the Optimizer***

The selection of the most appropriate optimizer was performed by evaluating the accuracy of the AlexNet-SVM model using the image features from fc6. Three optimizers were considered here, namely ADAM, SGDM, and RMSprop. The highest classification accuracy of 93.25 percent was achieved when the AlexNet-SVM model was trained with an SGDM optimizer.

## Performance Comparison

A comparison of the AlexNet CNN and AlexNet-SVM models in oil palm leaf disease detection was made. The confusion matrix was applied as the guideline to measure the disease prediction results for different classes. The performance indexes for both models in terms of accuracy, recall, specificity, and precision were evaluated subsequently for comparison. Table 3 summarizes the obtained results of oil palm leaf disease classification using AlexNet CNN and AlexNet-SVM.

**Table 3**

*The Performance Indexes of Oil Palm Leaf Disease Classification Using AlexNet CNN and AlexNet-SVM Models Based on the Confusion Matrix*

| Model | Class | Eye Spot | Healthy | Leaf Rot | Leaf Speckle | Accuracy | Recall | Specificity | Precision | F1-Score |
|---|---|---|---|---|---|---|---|---|---|---|
| AlexNet CNN | Eye Spot | 95 | 1 | 2 | 2 | 0.97 | 0.94 | 0.98 | 0.95 | 0.95 |
| | Healthy | 2 | 96 | 0 | 5 | 0.98 | 0.98 | 0.99 | 0.93 | 0.95 |
| | Leaf Rot | 3 | 0 | 97 | 0 | 0.98 | 0.97 | 0.99 | 0.97 | 0.97 |
| | Leaf Speckle | 1 | 1 | 1 | 97 | 0.98 | 0.93 | 0.98 | 0.97 | 0.95 |
| **Average** | | | | | | **0.98** | **0.96** | **0.99** | **0.96** | **0.96** |
| AlexNet-SVM | Eye Spot | 85 | 3 | 7 | 5 | 0.95 | 0.93 | 0.98 | 0.89 | 0.89 |
| | Healthy | 2 | 95 | 0 | 3 | 0.97 | 0.92 | 0.97 | 0.95 | 0.94 |
| | Leaf Rot | 2 | 1 | 94 | 3 | 0.96 | 0.91 | 0.97 | 0.94 | 0.93 |
| | Leaf Speckle | 2 | 4 | 2 | 92 | 0.95 | 0.89 | 0.96 | 0.92 | 0.91 |
| **Average** | | | | | | **0.96** | **0.91** | **0.97** | **0.92** | **0.92** |

As seen in Table 3, the prediction accuracies for both classifiers were satisfactory, exceeding classification accuracy of 0.90. As a whole, AlexNet with transfer learning with an average classification accuracy of 0.98 outperformed AlexNet-SVM with an average classification accuracy of 0.96. For the AlexNet CNN model, the true positive rates of leaf rot and leaf speckle were the highest (97%) while healthy leaf had the least (93%). For the AlexNet-SVM method classifier, the class with the lowest true positive rate (85%) was eye spot disease, while the healthy class had the highest true positive rate (95%) among the four classes. The oil palm leaf disease detection accuracy was highly dependent on the environmental condition during image acquisition.

Figure 10(a) shows a misclassified image of eyespot disease, identified as leaf speckle by the classification model. For this sample image, both diseases of leaf speckle and eye spot coexist. The root of failure detection might be due to the complex image background condition, lighting condition, and existence of multiple disease symptoms.

Figure 10(b) exhibits a misclassified image of leaf rot disease, identified as eye spot disease, probably due to the lighting condition.

**Figure 10**

*An Example of Misclassified Leaf Disease Sample Image Due to (a) the Coexistence of Both Diseases and (b) Lighting Condition*

(a) (b)

Besides the average classification accuracy, the AlexNet CNN classification model outperformed the AlexNet-SVM classification model in other performance indexes. In machine learning, the metrics of recall and precision are two extreme indexes in measuring the stability of a model, as it reflects the relevance of the obtained prediction results. The stability of the AlexNet model was higher than that of AlexNet-SVM, where the recall and precision of 0.96 were obtained, as compared to the recall of 0.91 and precision of 0.92 attained by AlexNet-SVM. In most cases, the fraction of recall is always compared to specificity. This is because specificity measures the regularity of a model to generate a true prediction on the testing samples, which is exactly the opposite of recall. The AlexNet CNN model with the specificity of 0.99 tended to score higher in this case as compared to the AlexNet-SVM model with the specificity of 0.92. F1-score reflects the balance between recall and precision of a model. AlexNet excelled in this part as well with a value of 0.96, which was 0.04 higher than the score of the AlexNet-SVM model.

The main classification units of AlexNet CNN and AlexNet-SVM were Softmax and SVM algorithms, respectively. The accuracy and compatibility of these two classifiers are constantly the point of argument in most studies (Li et al., 2019). The main difference between these two classifiers is in the probability computing methods of the algorithms. To be more specific, Softmax is a common linear regression classifier. The prediction of probability is established using the cross-entropy loss function toward all the class labels. On the other hand, SVM is more local biased on prediction, whereby classification is made based on the hinge loss during computation.

The results obtained in this work were in accordance with previous studies on face recognition using AlexNet CNN and AlexNet-SVM (Almabdy & Elrefaei, 2019). It has been reported that AlexNet CNN possessed greater prediction capability than AlexNet-SVM in face recognition, when the dataset of ORL, GTV face, Georgia Tech face, LFW, F_LFW, YouTube face, and FEI faces were taken into account. The AlexNet transfer learning model achieved a higher value of precision, accuracy, recall, and F1-score, within the range of 93 percent to 99 percent as compared to the AlexNet-SVM method with the performance indexes within the range of 90 percent to 99 percent.

## CONCLUSION

In this study, the implementation of deep learning, specifically, AlexNet CNN, into oil palm leaf disease detection was the main focus. Additionally, the AlexNet-SVM model, in which AlexNet CNN was utilized as the feature extractor of SVM, was applied in solving the oil palm leaf disease classification as. The performance of the classification models was evaluated based on accuracy, precision, recall, specificity, and F1-score. The average accuracy, precision, recall, specificity, and F1-score of 0.98, 0.96, 0.99, 0.96, and 0.96, and 0.96, 0.91, 0.97, 0.92, and 0.92 were attained by AlexNet CNN and AlexNet-SVM, respectively. The results showed that the AlexNet-based disease classification model outperformed the AlexNet-SVM model.

## ACKNOWLEDGMENT

The authors would like to express their deepest appreciation to the Ministry of Higher Education Malaysia for funding this project through the Fundamental Research Grant Scheme (FRGS/1/2018/ICT02/UTHM/02/2 Vot K070).

## REFERENCES

Abdu, A. M., Mokji, M. M., & Sheikh, U. U. (2020). Automatic vegetable disease identification approach using individual lesion features. *Computers and Electronics in Agriculture*, *176*, 105660. https://doi.org/https://doi.org/10.1016/j.compag.2020.105660

Abdulridha, J., Ampatzidis, Y., Kakarla, S. C., & Roberts, P. (2020). Detection of target spot and bacterial spot diseases in tomato using UAV-based and benchtop-based hyperspectral imaging techniques [Article]. *Precision Agriculture*, *21*(5), 955–978. https://doi.org/10.1007/s11119-019-09703-4

Aji, A. F., Munajat, Q., Pratama, A. P., Kalamullah, H., Setiyawan, J., & Arymurthy, A. M. (2013). Detection of palm oil leaf disease with image processing and neural network classification on mobile device. *International Journal of Computer Theory and Engineering*, *5*(3), 528.

Alaa, H., Waleed, K., Samir, M., Tarek, M., Sobeah, H., & Salam, M. A. (2020). An intelligent approach for detecting palm trees diseases using image processing and machine learning. *Int. J. Adv. Comput. Sci. Appl.*, *11*(7), 434–441.

Almabdy, S., & Elrefaei, L. (2019). Deep convolutional neural network-based approaches for face recognition. *Applied Sciences*, *9*(20), 4397.

Ampatzidis, Y., Cruz, A., Pierro, R., Materazzi, A., Panattoni, A., De Bellis, L., & Luvisi, A. (2018). Vision-based system for detecting grapevine yellow diseases using artificial intelligence. In *XXX International Horticultural Congress IHC2018: VII Conference on Landscape and Urban Horticulture, IV Conference on 1279* (pp. 225–230).

Arjmand, A., Angelis, C. T., Christou, V., Tzallas, A. T., Tsipouras, M. G., Glavas, E., . . . Giannakeas, N. (2020). Training of deep convolutional neural networks to identify critical liver alterations in histopathology image samples. *Applied Sciences*, *10*(1), 42.

Ashourloo, D., Aghighi, H., Matkan, A. A., Mobasheri, M. R., & Rad, A. M. (2016). An investigation into machine learning regression techniques for the leaf rust disease detection using hyperspectral measurement. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(9), 4344–4351, Article 7533506. https://doi.org/10.1109/JSTARS.2016.2575360

Chen, J., Chen, J., Zhang, D., Sun, Y., & Nanehkaran, Y. A. (2020). Using deep transfer learning for image-based plant disease identification. *Computers and Electronics in Agriculture*, *173*, 105393. https://doi.org/https://doi.org/10.1016/j.compag.2020.105393

Coulibaly, S., Kamsu-Foguem, B., Kamissoko, D., & Traore, D. (2019). Deep neural networks with transfer learning in millet crop images. *Computers in Industry*, *108*, 115–120.

Darwish, A., Ezzat, D., & Hassanien, A. E. (2020). An optimized model based on convolutional neural networks and orthogonal learning particle swarm optimization algorithm for plant diseases diagnosis. *Swarm and Evolutionary Computation*, *52*, Article 100616. https://doi.org/10.1016/j.swevo.2019.100616

Elaraby, A., Hamdy, W., & Alanazi, S. (2022). Classification of citrus diseases using optimization deep learning approach. *Computational Intelligence and Neuroscience*, *2022*, 9153207. https://doi.org/10.1155/2022/9153207

Elaraby, A., Hamdy, W., & Alruwaili, M. (2022). Optimization of deep learning model for plant disease detection using particle swarm optimizer. *Computers, Materials & Continua*, *71*(2), 4019–4031. http://www.techscience.com/cmc/v71n2/45808

Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*, *145*, 311–318.

Ghanei Ghooshkhaneh, N., Golzarian, M. R., & Mamarabadi, M. (2018). Detection and classification of citrus green mold caused by Penicillium digitatum using multispectral imaging. *Journal of the Science of Food and Agriculture*, *98*(9), 3542–3550. https://doi.org/10.1002/jsfa.8865

Guo, A., Huang, W., Ye, H., Dong, Y., Ma, H., Ren, Y., & Ruan, C. (2020). Identification of wheat yellow rust using spectral and texture features of hyperspectral images. *Remote Sensing*, *12*(9), Article 1419. https://doi.org/10.3390/RS12091419

Habib, M. T., Majumder, A., Jakaria, A., Akter, M., Uddin, M. S., & Ahmed, F. (2020). Machine vision based papaya disease recognition. *Journal of King Saud University-Computer and Information Sciences*, *32*(3), 300–309.

Hamdani, H., Septiarini, A., Sunyoto, A., Suyanto, S., & Utaminingrum, F. (2021). Detection of oil palm leaf disease based on color histogram and supervised classifier. *Optik*, *245*, 167753. https://doi.org/https://doi.org/10.1016/j.ijleo.2021.167753

Han, Z., & Gao, J. (2019). Pixel-level aflatoxin detecting based on deep learning and hyperspectral imaging. *Computers and Electronics in Agriculture*, *164*, 104888.

Hannan, M. A., How, D. N., Lipu, M. H., Ker, P. J., Dong, Z. Y., Mansur, M., & Blaabjerg, F. (2020). SOC estimation of li-ion batteries with learning rate-optimized deep fully convolutional network. *IEEE Transactions on Power Electronics*, *36*(7), 7349–7353.

Hoffer, E., Hubara, I., & Soudry, D. (2017). Train longer, generalize better: Closing the generalization gap in large batch training of neural networks. *Advances in Neural Information Processing Systems*, *30*.

Jacobs, R. A. (1988). Increased rates of convergence through learning rate adaptation. *Neural Networks*, *1*(4), 295–307.

Jadhav, S. B., Udupi, V. R., & Patil, S. B. (2021). Identification of plant diseases using convolutional neural networks. *International Journal of Information Technology*, *13*(6), 2461–2470.

Kawatra, M., Agarwal, S., & Kapur, R. (2020). Leaf disease detection using neural network hybrid models. In *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)* (pp. 225–230).

Khan, M. A., Akram, T., Sharif, M., & Saba, T. (2020). Fruits diseases classification: Exploiting a hierarchical framework for deep features fusion and selection. *Multimedia Tools and Applications*, *79*(35), 25763–25783. https://doi.org/10.1007/s11042-020-09244-3

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90.

Li, X., Tang, Z., Chen, W., & Wang, L. (2019). Multimodal and multi-model deep fusion for fine classification of regional complex landscape areas using ZiYuan-3 imagery. *Remote Sensing*, *11*(22), 2716.

Liu, J., & Wang, X. (2020). Tomato diseases and pests detection based on improved Yolo V3 convolutional neural network [Original Research]. *Frontiers in Plant Science*, *11*(898). https://doi.org/10.3389/fpls.2020.00898

Masazhar, A. N. I., & Kamal, M. M. (2017). Digital image processing technique for palm oil leaf disease detection using multiclass SVM classifier. In *2017 IEEE 4th International Conference on Smart Instrumentation, Measurement and Application (ICSIMA)* (pp. 1–6).

Muhammad, N., Rubab, Bibi, N., Song, O.-Y., Khan, M. A., & Khan, S. A. (2021). Severity recognition of aloe vera diseases using AI in tensor flow domain. *CMC-Computers Materials & Continua*, *66*(2), 2199–2216.

Pantazi, X. E., Moshou, D., & Tamouridou, A. A. (2019). Automated leaf disease detection in different crop species through image features analysis and one class classifiers. *Computers and Electronics in Agriculture*, *156*, 96–104.

Picon, A., Alvarez-Gila, A., Seitz, M., Ortiz-Barredo, A., Echazarra, J., & Johannes, A. (2019). Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild. *Computers and Electronics in Agriculture*, *161*, 280–290. https://doi.org/https://doi.org/10.1016/j.compag.2018.04.002

Rashwan, S. A., & Elteir, M. K. (2022). Plant leaf disease detection using deep learning on mobile devices. *International Journal of Computational Vision and Robotics*, *12*(2), 156–176.

Rath, A. K., & Meher, J. K. (2019). Disease detection in infected plant leaf by computational method. *Archives of Phytopathology and Plant Protection*, *52*(19–20), 1348–1358. https://doi.org/10.1080/03235408.2019.1708546

Roslan, A., & Idris, A. (2012). Economic impact of ganoderma incidence on Malaysian oil palm plantation–a case study in Johor. *Oil Palm Industry Economic Journal*, *12*(1), 24–30.

Roy, S. S., Ahmed, M., & Akhand, M. A. H. (2018). Noisy image classification using hybrid deep learning methods. *Journal of Information and Communication Technology*, *17*(2), 233–269.

Septiarini, A., Hamdani, H., Hardianti, T., Winarno, E., Suyanto, S., & Irwansyah, E. (2021). Pixel quantification and color feature extraction on leaf images for oil palm disease identification. In

*2021 7th International Conference on Electrical, Electronics and Information Engineering (ICEEIE)* (pp. 1–5).

Singh, R. K., Tiwari, A., & Gupta, R. K. (2022). Deep transfer modeling for classification of Maize Plant Leaf Disease. *Multimedia Tools and Applications*, *81*(5), 6051–6067. https://doi.org/10.1007/s11042-021-11763-6

Skowron, M., Orlowska-Kowalska, T., Wolkiewicz, M., & Kowalski, C. T. (2020). Convolutional neural network-based stator current data-driven incipient stator fault diagnosis of inverter-fed induction motor. *Energies*, *13*(6), 1475.

Sunpapao, A., Chairin, T., & Ito, S.-i. (2018). The biocontrol by streptomyces and trichoderma of leaf spot disease caused by Curvularia oryzae in oil palm seedlings. *Biological Control*, *123*, 36–42.

Tang, M., Huang, Z., Yuan, Y., Wang, C., & Peng, Y. (2019). A bounded scheduling method for adaptive gradient methods. *Applied Sciences*, *9*(17), 3569.

Türkoğlu, M., & Hanbay, D. (2019). Plant disease and pest detection using deep learning-based features. *Turkish Journal of Electrical Engineering & Computer Sciences*, *27*(3), 1636–1651.

Veni, S., Vishnu Priya, P. M., Aishwarya Mala, G. M., Kayartaya, A., & Anusha, R. (2017). Computer aided system for detection and classification of brinjal leaf diseases using thermal and visible light images. *Journal of Theoretical and Applied Information Technology*, *95*(19), 5224–5236. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85031745861&partnerID=40&md5=e3825478384efd8ea10593d779bf0b75

Yang, N., Qian, Y., El-Mesery, H. S., Zhang, R., Wang, A., & Tang, J. (2019). Rapid detection of rice disease using microscopy image identification based on the synergistic judgment of texture and shape features and decision tree–confusion matrix method. *Journal of the Science of Food and Agriculture*, *99*(14), 6589–6600. https://doi.org/10.1002/jsfa.9943

Yuan, L., Yan, P., Han, W., Huang, Y., Wang, B., Zhang, J., . . . Bao, Z. (2019). Detection of anthracnose in tea plants based on hyperspectral imaging. *Computers and Electronics in Agriculture*, *167*, 105039. https://doi.org/https://doi.org/10.1016/j.compag.2019.105039