

Performance Analysis of Efficient Pre-trained Networks based on Transfer Learning for Tomato Leaf Diseases Classification

Sawsan Morkos Gharghory
Computers and Systems Department
Electronics Research Institute
Giza, Egypt

Abstract—Early diagnosis and accurate identification to tomato leaf diseases contribute on controlling the diffusion of infection and guarantee healthy to the plant which in role result in increasing the crop harvest. Nine common types of tomato leaf diseases have a great effect on the quality and quantity of tomato crop yield. The tradition approaches of features extraction and image classification cannot ensure a high accuracy rate for leaf diseases identification. This paper suggests an automatic detection approach for tomato leaf diseases based on the fine tuning and transfer learning to the pre-trained of deep Convolutional Neural Networks. Three pre-trained deep networks based on transfer learning: AlexNet, VGG-16 Net and SqueezeNet are suggested for their performances analysis in tomato leaf diseases classification. The proposed networks are carried out on two different dataset, one of them is a small dataset using only four different diseases while the other is a large dataset of leaves accompanied with symptoms of nine diseases and healthy leaves. The performance of the suggested networks is evaluated in terms of classification accuracy and the elapsed time during their training. The performance of the suggested networks using the small dataset are also compared with that of the-state-of-the-art technique in literature. The experimental results with the small dataset demonstrate that the accuracy of classification of the suggested networks outperform by 8.1% and 15% over the classification accuracy of the technique in literature. On other side when using the large dataset, the proposed pre-trained AlexNet achieves high classification accuracy by 97.4% and the consuming time during its training is lower than those of the other pre-trained networks. Generally, it can be concluded that AlexNet has outstanding performance for diagnosing the tomato leaf diseases in terms of accuracy and execution time compared to the other networks. On contrary, the performance of VGG-16 Net in metric of classification accuracy is the best yet the largest consuming time among other networks.

Keywords—Deep learning; Alex; squeeze; VGG16 networks; tomato leaf diseases diagnosis and classification

I. INTRODUCTION

In the past decades, the plant diseases identification were mostly performed through the optical observation by farmer. Unfortunately, the process of detection and diagnosis of crop diseases by this method were error-prone, expensive and time consuming. In addition, there is no local experience for dealing with any new diseases maybe occur in places that were previously unidentified [1]. Machine learning techniques has

been emerged as an intelligent technique to be used in large scale of this field. They were applied in early stage of plant diseases diagnosis and classification. Li et al. [2] suggested the K-means clustering segmentation method to the grape disease images. The authors in this work proposed a SVM classifier which was designed using thirty one of significant features that were selected to identify both of grape downy mildew disease and grape powdery mildew disease. The classification rates in testing phase were respectively 90% and 93.33%. Athanikar and Badar [3] implemented Neural Network to classify the potato leaf image into category of healthy and diseased. Their results demonstrated that BPNN could effectively detect the spots leaf disease and could particularly categorized the disease type with accuracy 92%.

Recently, the deep learning is getting more interest, particularly the Convolutional Neural Network (CNN). CNN is a type of deep learning structure that was designed for the classification purposes especially for digital image classification and it has been regarded as one of the best approaches for pattern recognition tasks. The Manual extraction of features in the conventional techniques is a tedious task and researches need mostly a lot of time to test and extract the suitable features for the classification. On contrast, CNN can automatically extract features of image by tuning the parameters in both of convolutional and pooling layers. CNN also has another advantage which is the ability of learning from big data set. On contrary, the algorithms of the conventional machine learning usually needs only hundreds samples that are used for training but when larger training sets are used, these algorithms converge slowly or maybe cannot converge. Deep learning method has the robustness and the ability of generalization so that it outperforms in many fields such as: signal processing [4], pedestrian detection [5], face recognition [6], road crack detection [7], and biomedical image analysis [8]. Deep learning techniques have also accomplished impressive outcome in the agriculture field and were benefit for horticultural workers and smallholders including: recognition of weeds [9] selection of fine seeds [10], pest identification [11], fruit counting [12], and research on land cover [13]. The wide spread of deep CNNs in the agriculture field has lead to a big progress especially in plant diseases classification, in which they can find high variance of pathological symptoms in visual appearance. In addition to this, CNN can find the high dissimilarity in intra-class and

even the low similarity between inter-classes that perhaps are noticed only by the botanists [14]. More studies of using CNN in the field of crop disease recognition and identification as a new hot spot research in agricultural field were presented in [15-23]. These studies demonstrated that CNNs have not only reduced the requirements of image preprocessing, but also improved the accuracy of diseases recognition. Lee et al. [24] proposed a CNN approach to identify leaf images and reported an average accuracy of 99.7% on a dataset covering 44 species, but the scale of datasets was very small.

In this paper three pre-trained deep networks based on the transfer learning and fine tuning are suggested for tomato leaf diseases classification. The performance analysis of the three networks in existence of different number of images and with variation to the values of learning parameters are evaluated through the comparison and results analysis. The rest of the paper is organized as follows. Section 2 reviews the related works of the application of transfer learning to CNN in a different fields. Section 3 introduces the structures of the three pre-trained networks: AlexNet, VGG-16Net and SqueezeNet in addition to the used data set of tomato leaf diseases. The experimental results of the fine tuning and the transfer learning-based pre-trained networks are presented in Section 5. Finally Section 6 concludes the analysis to the suggested networks performances and the results in addition to the future work.

II. RELATED WORKS

Deep learning networks can implemented the transfer learning either by using the pre-trained network to extract attributed features that can be applied to a new field or via fine tuning the weights of network through its training with a new data set. The transfer learning has been applied to deep networks in different domains with different applications. In the work [25], the transfer learning was utilized in the biometrics domain in which the joint probabilistic was exploited for face recognition to cope the problem of insufficient images of the wanted identities. Also, among the applications of applying transfer learning-based deep learning in the biometrics are the ear recognition as they were introduced in the works [26-27]. In [28], they also compared AlexNet architecture, the 16-layer VGG model architecture, and the latest SqueezeNet architecture for ear recognition using limited training data. The outcomes of this work showed that the architecture of SqueezeNet trained by using the learned parameters with ImageNet data and its fine tuned through utilizing 1383 of ear images was the best model. In [29], the authors applied transfer learning to the well-known AlexNet Convolution Neural Network (AlexNet CNN) for human recognition based on ear images. The work in [30] compared the performance of the pre-trained CNN AlexNet with the same network but with its fine tuning for the application of Arabic characters recognition. Their results proved that the transfer learning based on fine tuning to AlexNet produced a higher accuracy compared to the same AlexNet model without tuning as a fixed feature extractor. Meanwhile, the plants identification using the 2015 LifeCLEF dataset based on the transfer learning through the fine tuning to the pre-trained deep networks GoogleNet, VGGNet and AlexNet as proposed in [31]. Their output results showed that the most affecting factor on the performance of the transfer learning based-fine tuning

was the number of iterations. Mohanty et al. [17] made a fine tuning to deep learning models that pre-trained on ImageNet to be used in identifying 14 crop species with 26 leaf diseases. The models were tested on the available a public dataset including 54,306 images of healthy and diseased plant leaves collected under controlled conditions. They achieved the best accuracy of 99.35% on a hold-out test dataset. Zhang et al. [32] addressed the detection issue of cherry leaf powdery mildew disease using GoogLeNet which achieved accuracy of 99.6%. Their results also demonstrated that the performance of deep learning model can be boosted by the transfer learning in crop disease identification. In [33], a united convolutional neural networks (CNNs) architecture based on an integrated method is suggested. The proposed United Model is designed to distinguish leaves with the common grape diseases, it achieves an average accuracy of validation 99.17% and accuracy of test 98.57%. Also the work in [34] used the pre-trained models and multiple classifiers for detecting the potato leaf diseases. The logistic regression classifier with VGG19 outperformed the other classifiers by a classification accuracy 97.8% with the test dataset.

CNN and the three pre-trained deep networks based on the transfer learning and fine tuning are explained in the next section.

III. CONVOLUTIONAL NEURAL NETWORKS

Convolutional Neural Network (CNN) is emerged inspiring from the researches in human brain cortex. It is developed to extract significant features by sequential operations of convolution and pooling [35]. Convolutional layers, pooling layer, activation function layers, dropout layers, and fully-connected layers are the main layers in CNN architecture. **Convolutional layers** carry the outputs of convolution filters or kernels with preceding layer. The main parameters of these filters or kernels are the weights and biases which can be learned in each iteration through optimization function. Purpose of the optimization function is generating kernels that are a good data representation without error. **Pooling layers** are used for the down sampling to lower size of neuron and reduce the performance issue of over-fitting. Max pooling operation is the most type used in pooling layers which captures the maximum value of the pooling window. **Activation function layers** are used to add non-linearity to the network. In the literature, there are a lot of activation functions such as sigmoid, tanh and ReLU that is the most one used [12]. **Dropout layers** are used to overcome the problem of over-fitting by randomly shut down the neurons in the network. **Fully connected layers** are utilized to calculate the scores or probabilities of classes. The classifier inputs are the results of the fully connected layers, the most well-known classifier is the softmax classifier. Since CNN is a supervised learning, the loss between the ground truth data and the network output is calculated and this loss is an input to the optimization algorithm. The most common optimization algorithm is the Stochastic Gradient Descent (SGD) algorithm in which this algorithm updates the weights according to the loss value calculated in each iteration. Both of the loss function and SGD are depicted in equations 1 and 2 as follows:

$$J(\omega) = \frac{1}{n} \sum_{i=1}^n y_i \log(h_{\omega}(X_i)) \quad (1)$$

$$\omega_{t+1} = \mu\omega_t - \alpha\nabla J(\omega_t) \quad (2)$$

where: x is the training sample with number of input data n , y_i is the true label data and $h_\omega(X_i)$ is the predicted label of CNN network in a given current weights ω . Also, μ is the momentum weight for both of current weights ω_t and learning rate α . The most common CNN deep learning architectures such as AlexNet, SequeezeNet, and VGG-16 Net are briefly explained in the next section.

A. AlexNet

The first well known CNN was AlexNet and it was among the early successful architectures of deep learning developed by author in [36] and it consists of several layers of convolution layers, Rectified Linear Unit (ReLU) and may be with batch normalization and Max-pooling in some layers. Each layer has many kernels and each kernel is initialized randomly at the beginning of training and through the optimization function the kernels are learned. The number of kernels at the preceding layer determines the dimension depth of the convolutional layers. Through the convolution, each kernel maps the preceding layer to new space. In this paper, the pre-trained model used in our study consists of five convolutional (conv) layers and three fully connected layers as shown in Fig. 1. The first convolution layer consists of 96 filters, each one with dimension of 11 x 11 x 3 which is the height, width and depth, respectively and it is applied on an input image of size 227 x 227 x 3. Thus, the Rectified Linear Unit (ReLU) from the first convolution layer generates 96 activation map. In the same fashion, the four remaining convolutions layers for performing the convolutional operations are respectively as follows: conv2 includes 256 filters each one with dimension 5x5x48, conv3 includes 384 filters each one with dimension 3x3x256, conv4 contains 384 filters each one with dimension of 3x3x192 and finally conv5 includes 256 filters each of them with dimension 3x3x192. Findings from these layers are activation maps with various neuron which activated in each map. The convolutional layers followed by ReLU, Max-pooling and normalization layers. ReLU is a nonlinear and a non-saturating activation function which is applied to the output of both of the 5 convolution layers and the last two fully connected layers. The function of Max-pooling layers are reducing the dimension of the previous convolution layer output through finding and saving the maximum value in the concerning field. The last two fully connected layers 6 and 7 have 4096 neurons where all of them are linked to each other, while the fully connected layer 8 (fc8) has 1000 output classes as trained with ImageNet data. The objective of Dropout layer is to randomly prevent the number of a network connections for training and this showed its ability to improve the network performance over test phase [37]. The final fc8 layer is followed by the softmax and classifier with 1000 output categories in which the loss function used is the cross entropy.

B. VGG-16Net

VGG-16Net is deeply learning series network and it consists of thirteen convolutional (conv) layers [38], each layer followed by ReLU layer and its architecture is shown in Fig. 2 in which all conv layers are with green color.

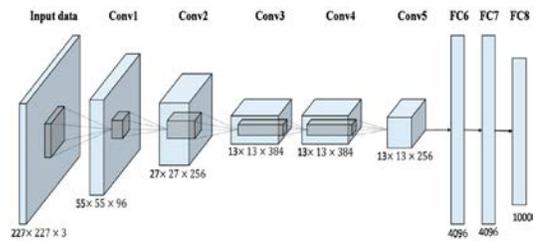


Fig. 1. The AlexNet Structure.

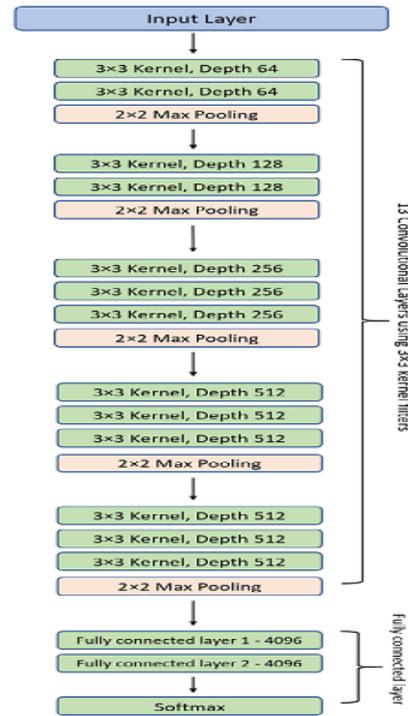


Fig. 2. VGG-16 Net Structure.

The first conv1 layer receives an input image with size equal to 224 x 224x3. The input image propagated through a set of conv layers having filters with receptive field of 3x3. Also the net architecture contains five max-pooling layers which are used for down-sampling with stride equals two. Max-pooling layers are implemented over a window of 2x2-pixels and they follow some of the conv layers. In addition, there are three fully connected layers (fc) following the conv layers with channel size equal 4096, 4096, and 1000 respectively. Each neuron in fc layer accepts the input from the activations of the previous neuron layer. The output size of 1000 in final fc layer represents the number of ImageNet categories used in training the global classifier. The final layers in the VGG-16 Net are the soft-max layer and the classification layer. The rectification non-linearity layer (ReLU) equips all the hidden layers [38]. The main advantage of using VGG-16 architecture is that it can be generalized well with any new datasets. From experimental results of VGG-16 Net applications, it was concluded that the features of the previous layers of a pre-trained network usually include information about the edge and color. With other meaning, the later layers hold features more specific to the classes' details. Hence, the earlier layers parameters of VGG-16 network does not need for

the fine-tuning as explained in [39]. Motivation from this, only the fine tuning to the last layers of the network has been proposed in the present literatures [40-41]. VGG-16 was trained by one million images or more, thus it has the capability to categorize the input images into 1000 classes.

C. SqueezeNet

It consists of 68 layers; a squeeze convolution layer comprise nine fire modules that has filters with only of size 1x1 and it feeds an expand layer that has convolution filters with size 1x1 or 3x3 [42]. Also each fire-expand layer feeds a fire-ReLU-expand layer. The final conv10 layer has 1000 output categories and followed by a classification layer. The full architecture of SqueezeNet is shown in Fig. 3 and the structure with two fire modules is shown in Fig. 4.

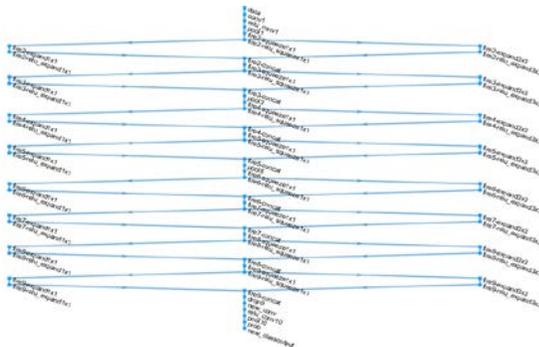


Fig. 3. SqueezeNet Structure with All Layers.

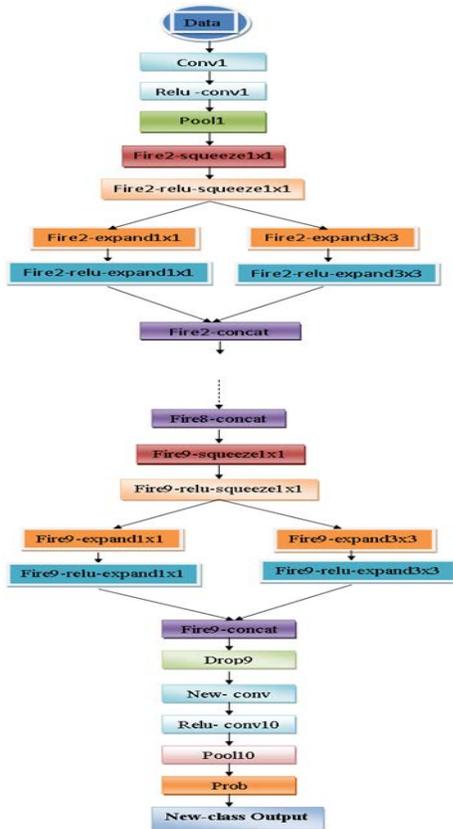


Fig. 4. SqueezeNet Structure with Exposition to the Fire Modules 2 and 9.

IV. THE PROPOSED NETWORKS MODELS AND DATASET

A brief explanation to the architecture of the pre-trained networks based on transfer learning and the fine tuning to the learning parameters in addition to the utilized dataset are introduced in this section.

A. The Pre-trained Networks based on Transfer Learning

In this paper, transfer learning is suggested for the pre-trained deep networks AlexNet, VGG-16Net and SqueezeNet with application to tomato leaf diseases diagnosis. Concerning the transfer learning to the pre-trained AlexNet, the two last fully connected layers have been modified with the desired number of categories under consideration in the field of tomato leaf diseases. Also, with regard to VGG-16 Net, the transfer learning is implemented on it through excluding only the three last layers in VGG-16 Net architecture and retaining the remaining layers of the network structure. The last three layers are substituted by a new layer of fc layer, a softmax layer, and a new classification layer so that its classification output should be suitable to the new classification task. The transfer learning is also applied to the pre-trained SqueezeNet by modifying both of the final convolution layer named conv10 and the classification layer to be convenient with the desired number of classes assigned in our study cases. Also, fine-tuned is suggested to all the pre-trained networks through assigning both factors of the learning rate of weight and learning rate of bias of the fully connected layer to are 10. In case of the fine-tuned AlexNet, the learning rate of both weight and bias are 10 and 5 times the learning rate of the fully connected layer in the global AlexNet. While these learning parameters are respectively 10 and 10 times thos of learning rate of the fully connected layer (fc8) in the global VGG-16Net. Regarding the fine-tuned SqueezeNet, the learning rate of both weight and bias are 10 and 10 times those of learning rate of the final convolutional layer (conv10) of the global SqueezeNet.

B. Data Set

In this paper, the pre-trained networks based on transfer learning and the fine tuning of the learning rate parameters and adjusting the Mini-Batch Size to be at a suitable value are carried out to resolve the diagnosis issue of tomato leaf diseases. Nine different diseases in addition to healthy leaves of tomato crop from Plant Village dataset [43] are utilized in our study. The Nine tomato leaf diseases are Late Blight (LB), Leaf Mold (LM), Septoria leaf Spot (SS), Mosaic Virus (MV), Bacteria Spot (BS), Early Blight (EB), Yellow Leaf Curl Virus (YC), Spider mites Two-spotted Spider mite (SM), Target Spot (TS). The tomato leaf images of Plant Village data set are in three channels of red, green and blue (RGB) with dimension of 256x256 with total numbers equal 18160 images. All the used images are resized to be convenient with the input data size to each network. Images of six samples from each class are shown in Fig. 5. Four diseases and healthy leaves with a few number of images for each class are only used in the first test for evaluating the performance of the suggested networks and the comparison with the work in literature. While all the above mentioned nine leaf diseased and healthy leaves with the aforementioned of total number images are used in the second test for the performance analysis and the evaluation of the suggested networks.

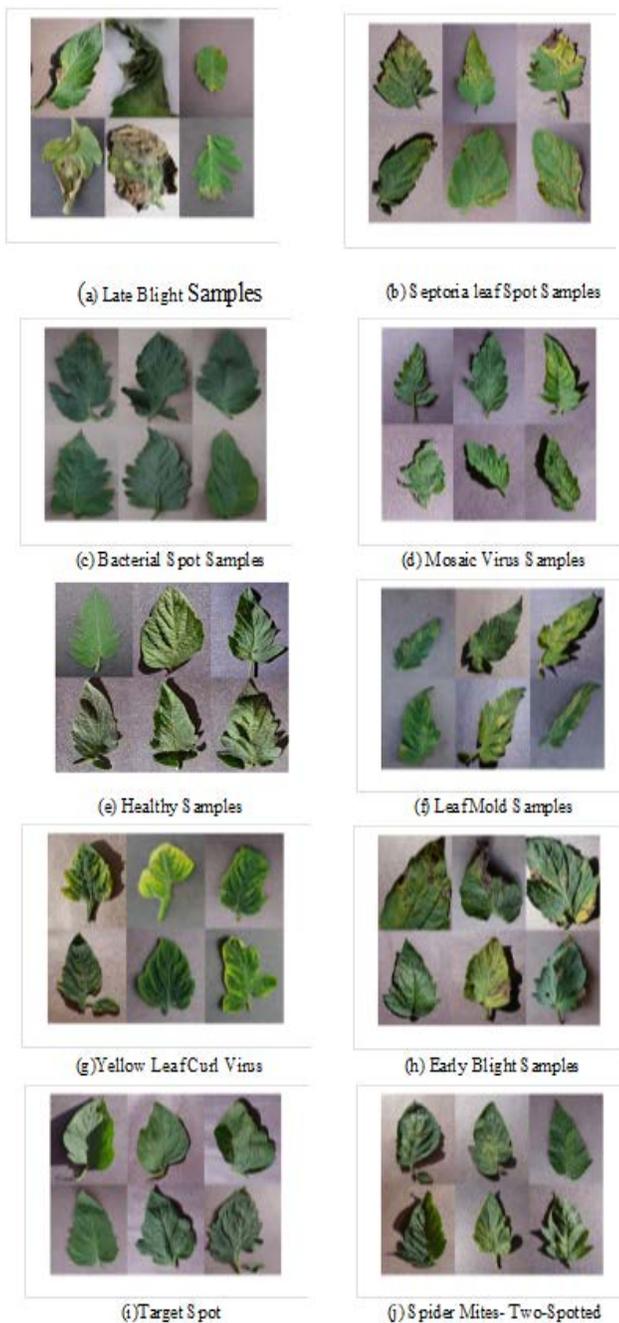


Fig. 5. (a-j) Nine tomato Leaf Diseases and Healthy Tomato Leaf.

V. EXPERIMENTAL RESULTS AND DISCUSSION

The performance of the pre-trained networks based on transfer learning and the fine tuning are verified through their diagnosis results and classification for tomato leaf diseases. Two set of data in which each of them has different numbers of images with different numbers of diseases are used for evaluating the suggested networks performance. In the **first part of our study** to the networks performance evaluation, the suggested fine-tuned networks are carried out on four tomato leaf diseased and healthy leaves. The suggested networks Alex, Squeeze and VGG-16 are trained with the following tomato leaf diseases: Bacteria Spot (BS), Late Blight (LB), Spetoria

Spot (SS), Yellow leaf Curl (YC) and healthy leaves in which 100 images are used from each class. The tomato leaf images in each category are split to 80 images to be used for the training phase and the remaining images are used to the test phase. The learning parameters for training the suggested networks are tuned and selected to be as follows: the Initial Learning Rate is set to be as 0.0001, the maximum number of epochs is chosen to be 15 and finally the Mini-Batch Size parameter to the three networks is tested with the following values 5, 15, 22 and 30 respectively. The performance of the suggested networks in terms of the overall classification accuracy, the accuracy of classification for each category besides the comparison of their performance with that of the state-of-the-art technique are evaluated. The classification accuracy of the suggested networks in test phase and the elapsed time during the training of the three networks using the above mentioned values of Mini-Batch Size are depicted in Tables I and II. Also, the comparison of the three pre-trained networks in terms of the accuracy and the elapsed time during their training using the different Mini-batch size are shown in Fig. 6 and 7. The confusion matrix to the classification outputs of the suggested networks using the Mini-Batch Size values at 30, 15 and 5 are depicted in Tables III, V and VII. The Confusion matrix shows the number of samples of the corresponding detected class for each true class. Also, the classification accuracy of each class, in addition to both of the true and false samples of each class using Alex, Squeeze and VGG-16 networks at different Mini-Batch sizes are depicted in Tables IV, VI and VIII. In addition, the performance of the suggested networks is evaluated through their comparison with the classifier used in literature [15]. The authors in this paper presented a Convolutional Neural Network model and Learning Vector Quantization (LVQ) algorithm based method for tomato leaf disease detection and classification. The dataset used in their work contains 500 RGB images of tomato leaves with four symptoms of diseases in which 20 images from each class are used in the test phase of the classifier. The LVQ classifier had been fed with the output feature vector of convolution part for training the network in which the maximum number of epochs was 300. The classifier performance using the state-of-the-art technique [15] in terms of the confusion matrix, classification accuracy to each category and the average accuracy are depicted in Table IX.

From the results given in tables, the overall accuracy of classification using the three suggested networks ranged from 93% to 99% according to the Min-Batch sizes. The accuracy of classification to recognize the tomato healthy leaves was the best one and reached to 100% with all the suggested networks and at different values of Mini-Batch sizes. The accuracy of classification of all leaf diseases except the Spetoria Spot disease was ranging from 90% to 100% according to the type of used network and the size of Mini-Batch parameter. It was found that the classification accuracy of Spetoria Spot disease using SequeezeNet with Mini-Batch sizes at 30 and 15 was poor compared to that of the other tomato diseases when using the other two suggested networks. The main reason to this low accuracy in diagnosing the Spetoria Spot disease perhaps due to the similarity of its symptom with the other symptoms of diseases and this led to difficult discrimination using squeezeNet at large size of min-batch parameter. The

classification accuracy increased to 100% in diagnosing the Spetoria Spot disease with the three suggested networks when using Mini-Batch size at value 5. Also concerning the elapsed time during training the networks, it was ranging from 8 minutes with AlexNet to almost 160 minutes with VGG-16Net which was the longest time among the other networks. Also from the results given in tables, it was verified that with the small value of Mini-Batch size, the classification accuracy rate of AlexNet is low, while with increasing the Mini-Batch size the accuracy rate of AlexNet is also increasing. On contrast, the elapsed time in training the AlexNet at a small value of Min-Batch size was larger than that time when the Mini-Batch size was large. The accuracy rate of SqueezeNet classification reduces with increasing the value of Min-Batch size, while it increases with reducing the value of Min-Batch size. On the other side, the elapsed time in training the SqueezeNet is inversely proportional to the value of Mini-Batch size. Its training execution time is big when the value of Mini-Batch size is small and vice versa. SqueezeNet takes a smallest execution time during its training when using a small value of Mini-Batch size compared to the two other networks. As VGG-16 Net is network with a deep structure, it generally takes the longest time during its training among other networks. Therefore, the execution process of training VGG-16 Net may be failed as a result of error of out of memory when increasing the size of Mini-Batch parameter to value greater than 30. On other side, at a small Mini-Batch size its classification accuracy rate was larger than its accuracy at a large value of Min-Batch size.

Regarding the comparison of our suggested networks with the classifier in literature [15], our pre-trained networks outperformed the classifier in literature. The average rate of the suggested networks accuracy with the same dataset used in literature was ranging from 93% to 99% either with a small or large Mini- Batch size and with maximum number of iterations equals 15 epochs. While, the average accuracy of this work in literature was 86% at maximum number of iterations equals 300 epochs. Hence, the accuracy rate of the suggested networks improved by 8.1% to 15% over the accuracy rate of the classifier introduced in [15].

TABLE I. THE ACCURACY OF CLASSIFICATION USING THE THREE PRE-TRAINED NETWORKS WITH DIFFERENT MINI BATCH SIZE

Pre-Trained Networks	Accuracy of classification in test phase			
	Mini Batch Size	Mini Batch Size	Mini Batch Size	Mini Batch Size
	5	15	22	30
AlexNet	96%	97%	97%	99%
SqueezeNet	98%	94%	94%	93%
VGG-16 Net	99%	97%	97%	96%

TABLE II. THE ELAPSED TIME TAKEN DURING THE TRAINING OF THE THREE PRE-TRAINED NETWORKS WITH DIFFERENT MINI BATCH SIZE

Pre-Trained Networks	Elapsed Time using different Batch size in minutes			
	Mini Batch Size	Mini Batch Size	Mini Batch Size	Mini Batch Size
	5	15	22	30
AlexNet	22 min	11min,29sec	12min, 45sec	8 min, 44sec
SqueezeNet	13min, 57 sec	10min,50sec	10min,50sec	9min, 51 sec
VGG-16 Net	157min	123 min	160min	155 min

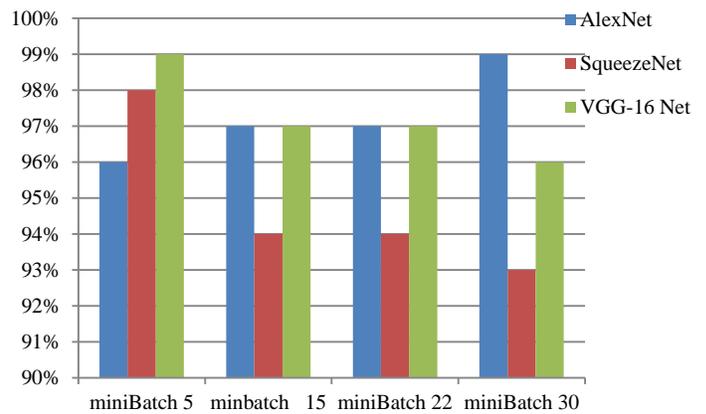


Fig. 6. The Classification Accuracy of the Pre-Trained Networks with Different Sizes of the Mini- Batch Size Parameters.

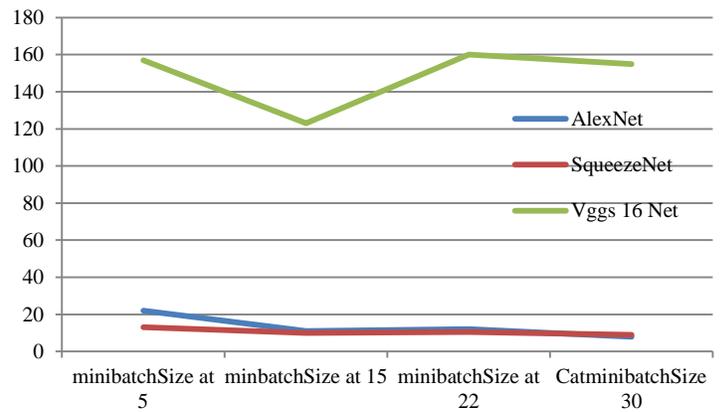


Fig. 7. The Elapsed Time of Training the Pre-Trained Networks with Different Sizes of the Mini- Batch Parameters.

TABLE III. THE CONFUSION MATRIX TO THE CLASSIFICATION OUTPUTS USING THE ALEX, SQUEEZE AND VGG-16 NETWORKS AT MINI-BATCH SIZE=30

Confusion Matrix of the classification with Mini Batch Size=30

AlexNet						SqueezeNet						VGG-16 Net					
True Class	BS	20				18		1	1			18			2		
	H		20				20						20				
	LB			20				19						19	1		
	SS	1			19			1	1	16					20		
	YC									20						19	
		BS	H	LB	SS	YC	BS	H	LB	SS	YC	BS	H	LB	SS	YC	
Predicted Class						Predicted Class						Predicted Class					

TABLE IV. THE ACCURACY OF CLASSIFICATION FOR EACH LABEL LEAF OF TOMATO DISEASES WITH ITS TRUE AND FALSE SAMPLES USING ALEXNET, SQUEEZE NET AND VGG-16 NET AT MINI-BATCH SIZE=30

Leaf Diseases	Accuracy at Batch Size 30								
	AlexNet			SqueezeNet			VGG-16 Net		
	Accuracy	True Sample	False Sample	Accuracy	True Sample	False Sample	Accuracy	True Sample	False Sample
B S	100%	20	0	90%	18	2	90%	18	2
H	100%	20	0	100%	20	0	100%	20	0
L B	100%	20	0	95%	19	1	95%	19	1
S S	95%	19	1	80%	16	4	100%	20	0
Y C	100%	20	0	100%	20	0	95%	19	1

TABLE V. THE CONFUSION MATRIX TO THE CLASSIFICATION OUTPUTS USING THE ALEX, SQUEEZE AND VGG-16 NETWORKS AT MINI-BATCH SIZE=15

Confusion Matrix of the classification with Mini Batch Size=15

AlexNet						SqueezeNet						VGG-16 Net					
True Class	BS	18			1	1	19		1			19			1		
	H		20					20					20				
	LB			20					19	1				19	1		
	SS	1			19		1		2	17					1	19	
	YC										19						20
		BS	H	LB	SS	YC	BS	H	LB	SS	YC	BS	H	LB	SS	YC	
Predicted Class						Predicted Class						Predicted Class					

TABLE VI. THE ACCURACY OF CLASSIFICATION FOR EACH LABEL LEAF OF TOMATO DISEASES WITH ITS TRUE AND FALSE SAMPLES USING THE ALEX, SQUEEZE AND VGG-16 NETWORKS AT MINI-BATCH SIZE=15

Leaf Diseases	Accuracy at Batch Size 15								
	AlexNet			SqueezeNet			VGG-16 Net		
	Accuracy	True Sample	False Sample	Accuracy	True Sample	False Sample	Accuracy	True Sample	False Sample
B S	90%	18	2	95%	19	1	95%	19	1
H	100%	20	0	100%	1	20	100%	20	0
L B	100%	20	0	95%	19	1	95%	19	1
S S	95%	19	1	85%	17	3	95%	19	1
Y C	100%	20	0	95%	19	1	100%	20	0

TABLE VII. THE CONFUSION MATRIX TO THE CLASSIFICATION OUTPUTS USING THE ALEX, SQUEEZE AND VGG-16 NETWORKS AT MINI-BATCH SIZE=5

Confusion Matrix of the classification with MiniBatch Size=5						
AlexNet		SqueezeNet		VGG-16 Net		
True Class	BS	18	1		1	
	H		20			
	LB		1	18		1
	SS				20	
	YC					20
		BS	H	LB	SS	YC
		Predicted Class				

TABLE VIII. THE ACCURACY OF CLASSIFICATION FOR EACH LABEL OF LEAF TOMATO DISEASES WITH ITS TRUE AND FALSE SAMPLES USING THE ALEX, SQUEEZE AND VGG-16 NETWORKS AT MINI-BATCH SIZE=5

Leaf Diseases	Accuracy at Batch Size 5								
	AlexNet			SqueezeNet			VGG-16 Net		
	Accuracy	True Sample	False Sample	Accuracy	True Sample	False Sample	Accuracy	True Sample	False Sample
B S	90%	18	2	90%	18	2	95%	19	1
H	100%	20	0	100%	20	0	100%	20	0
LB	90%	18	2	100%	20	0	100%	20	0
SS	100%	20	0	100%	20	0	100%	20	0
YC	100%	20	0	100%	20	0	100%	20	0

TABLE IX. THE CONFUSION MATRIX AND THE AVERAGE OF CLASSIFICATION ACCURACY OF THE CLASSIFIER USED IN LITERATURE[15]

Leaf diseases	B S	LB	SS	YC	H	Accuracy of each class
BS	18	0	0	2	0	90%
LB	0	17	0	3	0	85%
SS	0	0	16	3	1	80%
YC	0	0	3	17	0	85%
H	0	0	0	2	18	90%
Average Accuracy						86%

In the second part of our study to the networks performance evaluation for tomato leaf diseases classification, the fine-tuned networks are applied on tomato leaves with the aforementioned nine diseases and healthy leaves. Large numbers of the given images for each class are used in this part, the number of images range from 373 images of Mosaic Virus disease to 5357 images of Yellow Curl disease. The dataset is split randomly into 0.8 that is utilized for the training phase and the remaining of them is used for the test phase. In order to adjust the last three layers of the pre-trained Alex and Squeeze networks for the new classification task with 10 categories, both the learning rate of weight and learning rate of bias of the fully connected layer are set to be 10. Also both of AlexNet and SqueezeNet are trained by setting the following parameters values: maximum number of epochs at 15, the

learning rate at 0.0001 and the Mini-Batch size at 30. The confusion matrix to the classification results by the fine-tuned AlexNet is depicted in Table X. Also, the accuracy rate of classification for each class of tomato leaves, the number of true and false samples and the average accuracy of overall classification are depicted in Table XI. The training progress and the loss values during training AlexNet against the number of epochs are shown in Fig. 8. Also, the confusion matrix of classification results by the fine-tuned SqueezeNet is depicted in Table XII. Furthermore, the accuracy rate of classification for each class of tomato leaves, the number of true and false samples and the average accuracy of overall classification are depicted in Table XIII. Fig. 9 shows the training progress and the loss values of training SqueezeNet against the number of epochs.

TABLE X. THE CONFUSION MATRIX OF CLASSIFICATION USING ALEXNET WITH THE LARGE DATASET OF 9 TOMATO LEAF DISEASES AND HEALTHY LEAVES

		Predicted Classes									
		BS	E	H	L	L	M	S	S	T	Y
True Classes	B	42								1	
	S	3									
	E	10	18						1	4	
	B		2		3						
	H		1	31					1	1	
	L			5							
	B		9	1	36	2		1	2	2	
	L				5						
	M	1	1			18	1	1	4	1	1
	V					0					
S					1	7	1				
M						3					
S					1	1	32			6	
S						7					
T	2	2		1	2	1	1	34	2		
S								3			
T	1				2	2	13	3	25	1	
S								9			
Y										107	
C										1	

TABLE XI. ACCURACY OF CLASSIFICATION WITH THE TRUE AND FALSE SAMPLES USING ALEXNET WITH LARGE DATASET OF 9 TOMATO LEAF DISEASES AND HEALTHY LEAVES

Leaf Diseases	Total Samples	Test samples	True samples	False sample	accuracy
B S	2127	425	423	2	99.5%
E B	1000	200	182	18	91%
Healthy	1591	318	315	3	99.1%
L B	1909	382	365	17	95.5%
L M	952	190	180	10	94.7%
M V	373	75	73	2	97.3
S M	1676	335	327	8	97.6%
S S	1771	354	343	11	96.8%
T S	1404	281	259	22	92.2%
Y C	5357	1071	1071	0	100%
Accuracy	97.4%				

TABLE XII. THE CONFUSION MATRIX OF CLASSIFICATION USING THE SQUEEZE NET WITH THE LARGE DATASET OF 9 TOMATO LEAF DISEASES AND HEALTHY LEAVES

		Predicted Classes									
		BS	E	H	L	L	M	S	S	T	Y
True Classes	B	41	3							2	1
	S	7									1
	E		18						1	6	3
	B		1		9						
	H	1		31							1
	L			6							
	B	1	9	1	36	1					4
	L				6						
	M					1	18	2	1		1
	V						5				
S						1	7	1			
M							3				
S						1		33		2	
S							1	1		1	
T	4	5		3	1				33	2	
S									9		
T	5	3		1	1			9	3	25	
S									9		
Y	8							1			
C										106	
										2	

TABLE XIII. THE ACCURACY OF CLASSIFICATION WITH THE TRUE AND FALSE SAMPLES USING THE SQUEEZE NET AND THE LARGE DATA SET OF 9 TOMATO LEAF DISEASES AND HEALTHY LEAVES

Leaf Diseases	Total Samples	Test samples	True samples	False sample	Accuracy %
B S	2127	425	417	8	98
E B	1000	200	181	19	91
H	1591	318	316	2	99
L B	1909	382	366	16	95.8
L M	952	190	185	5	97
M V	373	75	73	2	97
S M	1676	335	331	4	98.8
S S	1771	354	339	15	95.7
T S	1404	281	259	22	92
Y C	5357	1071	1062	9	99
Accuracy	97.2%				

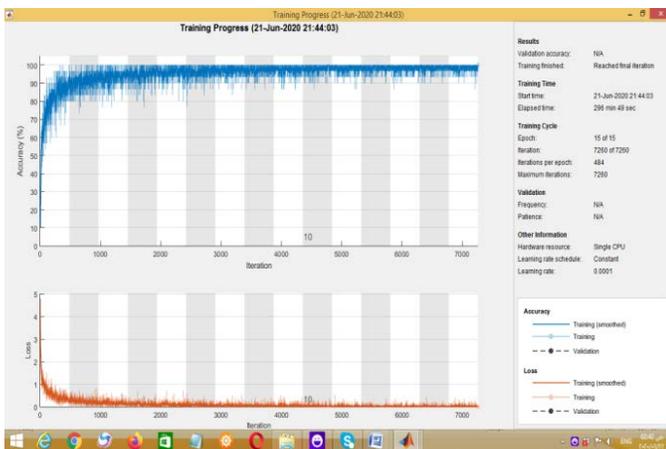


Fig. 8. The Training Progress of the Tuned AlexNet Versus the Number of Epochs.

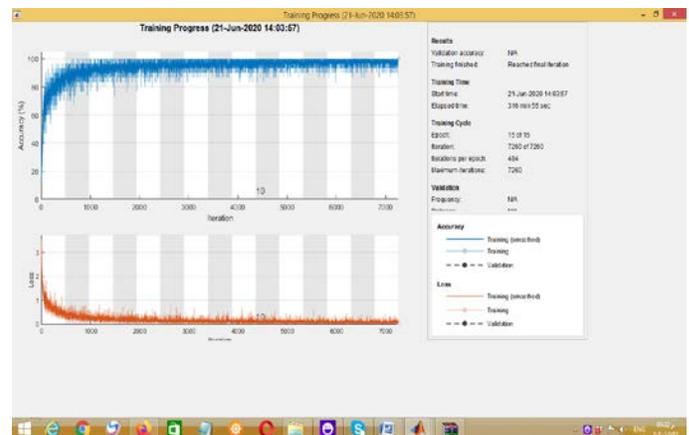


Fig. 9. The Training Progress of the Tuned SqueezeNet Versus the Number of Epochs.

From the results given in tables, the accuracy of AlexNet of tomato leaf diseases classification in test phase was 97.4% while the elapsed time taken in its training was 296 minutes and 50sec. On the other side from the results given in tables with SqueezeNet, the accuracy of tomato leaf diseases classification in test phase was 97.2% and the elapsed time taken in training the SqueezeNet was almost 316 minutes and 50sec (5hour and 17minutes). The classification accuracy of tomato leaves with symptoms of Early Blight either using AlexNet or SqueezeNet was the lowest among the diagnosis accuracy of other tomato diseases. The reason may be back to that the Early Blight disease appears first on old and mature leaves near the base at stem end of fruits with a spot of ring shape and this made a difficulty in its discrimination. AlexNet achieves high accuracy of diagnosing YC disease that reaches to 100% and was larger than SqueezeNet that achieves accuracy reaching to 99%. The accuracy of classification in diagnosing Target Spot disease with both of the fine-tuned Alex and Squeeze networks has an acceptable and was almost 92.2% due to its similar symptoms with the Spider Mites symptoms. Also, because the similarity of symptoms of both Spider mites and target spots, there are false numbers of them with both of their attributed categories. On the other side, AlexNet and SqueezeNet prove their ability in diagnosing the other diseases categories with high classification accuracy reached up to 99%. Also due to the deep learning structure of VGG-16 Net comparing to the structure of other networks, it was verified that this network was costly computationally. It was found that the time needed for training VGG-16 Net reached to 69 hours in which it was the largest among other networks. Therefore, only the results of AlexNet and SqueezeNet were enough for exposition.

VI. CONCLUSION

In this paper, the classification of tomato leaf diseases utilizing the images from Plant Village dataset was performed by the suggested pre-trained deep networks AlexNet, SqueezeNet and VGG-16 Net. The main challenge in tomato diseases diagnosis and classification in our study was that the symptoms of tomato leaf diseases are very similar to each other which results in some leaves may be embedded and classified into wrong classes. The accuracy of classification of AlexNet, SqueezeNet and VGG-16Net using 500 images of tomato leaves as assigned to **the first part** of the work and with Mini-batch size at 30 were 99%, 93% and 96% respectively. Whereas, the classification accuracy of AlexNet, SqueezeNet and VGG-16 Net using the same number of images and with Mini-batch size at 5 were 96%, 98% and 99%, respectively. Furthermore, the performance of the three fine-tuned networks for tomato leaf diseases diagnosis is evaluated through the comparison with that of the-state-of-the-art technique. The accuracy rate of our pre-trained networks increased by 8.1% to 15% over the value of accuracy rate of the classifier introduced in literature. The execution time of training AlexNet using small dataset and with Mini-Batch size at 30 was the shortest among training times for other networks. Also, the performance of AlexNet in terms of both classification accuracy and elapsed time using dataset of 18160 images as assigned to **the second part** of the work was efficient network and outperformed over to other networks. It achieves

classification accuracy of 97.4% with elapsed time in training of almost 296 minutes. On contrary, VGG-16 Net has large execution time during its training either using small Mini-Batch size or large Mini-Batch size compared to that of other networks.

In the future work, Internet of Things and mobile applications are suggested with the deep learning CNN to identify and classify the plant diseases type.

REFERENCES

- [1] Srdjan S, Marko A, Andras A, Dubravko C, Darko S. "Deep neural networks based recognition of plant diseases by leaf image classification". *Comput Intell Neurosci* 2016;2016:1–11
- [2] Li G, Ma Z, Wang H. Image recognition of grape downy mildew and grape powdery mildew based on support vector machine. In: *Proc international conference on computer and computing technologies in agriculture*. Beijing, China. pp. 151–162, 2012
- [3] Athanikar G, Badar P. Potato leaf diseases detection and classification system. *Int J Comp Sci Mob Comput* 2016;5 (2):76–88.
- [4] Xie D, Zhang L, Bai L. Deep learning in visual computing and signal processing. *Appl Comput Intell Soft Comput* 2017;2017:1–13.
- [5] Ouyang W, Wang X. Joint deep learning for pedestrian detection. In: *Proc. IEEE international conference on computer vision*. Sydney, Australia. p. 2056–63.
- [6] Sun Y, Wang X, Tang X. Hybrid deep learning for face verification. *IEEE Trans Pattern Anal. Mach Intell* 2016;38 (10):1997–2009.
- [7] Zhang L, Yang F, Zhang Y, Zhu Y. Road crack detection using deep convolutional neural network. In: *Proc. IEEE international conference on image processing*. Arizona, USA. p. 3708–12, 2016.
- [8] Zhou Z, Shin J, Zhang L, Gurudu S, Liang J. Fine-Tuning convolutional neural networks for biomedical image analysis: actively and incrementally. In: *Proc. IEEE conference on computer vision and pattern recognition*. Honolulu, Hawaii. p. 7340–9.
- [9] Lottes P, Behley J, Milioto A, Cyrill Stachniss C. Fully convolutional networks with sequential information for robust crop and weed detection in precision farming. *IEEE Robot Autom Lett* 2018;3(4):2870–7.
- [10] Wang X, Cai C. Weed seeds classification based on PCANet deep learning baseline. In: *Proc IEEE Asia-Pacific signal and information processing association annual summit and conference*. Hong Kong, China. p. 408–15.
- [11] Cheng X, Zhang Y, Chen Y, Wu Y, Yue Y. Pest identification via deep residual learning in complex background. *Comput Electron Agric* 2017;141:351–6.
- [12] Rahneemofar M, Sheppard C. Deep count: fruit counting based on deep simulated learning. *Sensors* 2017;17(4):905.
- [13] Zhu H, Liu Q, Qi Y, Huang X, Jiang F, Zhang S. Plant identification based on very deep convolutional neural networks. *Multimed Tools Appl* 2018;77(22):29779–97.
- [14] Kussul N, Lavreniuk M, Skakun S, Shelestov A. Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geosci Remote S* 2017;14(5):778–82..
- [15] Melike Sardogan, Adem Tuncer and Yunus Ozen, "Plant Leaf Diseases Detection and Classification Based on CNN with LVQ Algorithm" pp. 382-385, 3rd International Conference on Computer Science and Engineering, 2018 IEEE.
- [16] Mohanty S, Hughes D, Salathe´ M. "Using deep learning for image-based plant disease detection". *Front Plant Sci* 2016;7:1419.
- [17] Mohanty, S.P.; Hughes, D.; Salathe, M. "Inference of Plant Diseases from Leaf Images through Deep Learning", arXiv. 2016. Available online: <https://www.semanticscholar.org/paper/Inference-of-Plant-Diseases-from-Leaf-Images-throu-Mohanty-Hughes/62163ff3cb2fbbf5361e340f042b6c288d3b8e6a> (accessed on 28 December 2017).
- [18] Sladojevic, S.; Arsenovic, M.; Anderla, A.; Culibrk, D.; Stefanovic, D. Deep neural networks based recognition of plant diseases by leaf image

- classification. *Comput. Intell. Neurosci.* **2016**, 2016. [CrossRef] [PubMed]
- [19] Hanson, A.M.J.; Joy, A.; Francis, J. "Plant leaf disease detection using deep learning and convolutional neural network", *Int. J. Eng. Sci. Comput.* **2017**, 7, 5324–5328.
- [20] Lu, Y.; Yi, S.J.; Zeng, N.Y.; Liu, Y.; Zhang, Y. "Identification of rice diseases using deep convolutional neural networks". *Neurocomputing* **2017**, 267, 378–384. [CrossRef]
- [21] Tan, W.X.; Zhao, C.J.; Wu, H.R. "CNN intelligent early warning for apple skin lesion image acquired by infrared video sensors. *High Technol. Lett.* **2016**, 22, 67–74.
- [22] Kawasaki, Y.; Uga, H.; Kagiwada, S.; Iyatomi, H. "Basic study of automated diagnosis of viral plant disease using convolutional neural networks. In *Proceedings of the 12th International Symposium on Visual Computing, Las Vegas, NV, USA, 12–14 December 2015*; pp. 638–645.
- [23] Fuentes, A.; Yoon, S.; Kim, S.C.; Park, D.S. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors* **2017**, 17, 2022. [CrossRef] [PubMed].
- [24] Lee SH, Chan CS, Wilkin P, Remagnino P. Deep-plant: plant identification with convolutional neural networks. In: *Proc IEEE international conference on image processing. Quebec, Canada.* p. 452–6.
- [25] N. Nuo, W. Tay, J. Botzheim, and N. Kubota, "Joint probabilistic approach for real-time face recognition with transfer learning," *Rob. Auton. Syst.*, vol. 75, pp. 409–421, 2016.
- [26] Ž. Emeršič, D. Štepec, V. Štruc, and P. Peer, "Training convolutional neural networks with limited training data for ear recognition in the wild," *arXiv preprint arXiv:1711.09952*, 2017.
- [27] Y. Zhang, and Z. Mu, (2017). "Ear detection under uncontrolled conditions with Multiple Scale Faster Region-Based Convolutional Neural Networks," *Symmetry*, vol. 9, no. 4, pp. 1-19, 2017.
- [28] F. N. Iandola, M. W. Moskewicz, K. Ashraf, S. Han, W. J. Dally, and K. Keutzer, "Squeezenet: Alexnet-level accuracy with 50x fewer parameters and < 0.5 MB model size. *arXiv preprint arXiv:1602.07360*, 2016.
- [29] Ali Abd Almisreb, Nursuriati Jamil and N. Md Din, "Utilizing AlexNet Deep Transfer Learning for Ear Recognition". 2018 Fourth International Conference on Information Retrieval and Knowledge Management, March, 2018 IEEE, pp. 8-12.
- [30] C. Boufenar, A. Kerboua, and M. Batouche, "Investigation on deep learning for off-line handwritten Arabic character recognition," *Cogn. Syst. Res.*, 2017. high-level semantics," *Pattern Recognit. Lett.*, vol. 72, pp. 82–90, 2016.
- [31] M. M. Ghazi, B. Yanikoglu, and E. Aptoula, "Plant identification using deep neural networks via optimization of transfer learning parameters," *Neurocomputing*, vol. 235, pp. 228–235, August 2017.
- [32] Zhang K, Zhang L, Wu Q. Identification of cherry leaf disease infected by *Podosphaera Pannosa* via convolutional neural network. *Int J Agric Environ Inf Syst* 2019;10(2):98–110.
- [33] Miaomiao Ji, Lei Zhang, Qiufeng Wu, "Automatic grape leaf diseases identification via UnitedModel based on multiple convolutional neural networks, *INFORMATION PROCESSING IN AGRICULTURE*, October 2019, pp. 1-9, journal homepage: www.elsevier.com/locate/inpa.
- [34] Divyansh Tiwari Mritunjay AshishNitish Gangwar, "Potato Leaf Diseases Detection Using Deep Learning", *Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2020) IEEE Xplore Part Number:CFP20K74-ART; ISBN: 978-1-7281-4876-2*.
- [35] Y. LeCun and Y. Bengio, *Convolutional networks for images, speech, and time series*, in *The Handbook of Brain Theory and Neural Networks*, Vol. 3361 (MIT Press, 1995), pp. 255–258.
- [36] A. Krizhevsky, I. Sutskever and G. H. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems*, 2012.
- [37] Srivastava, N., Hinton, G., Krizhevsky, A, Sutskever, I., Salakhutdinov, R. (2015) "Dropout: a simple way to prevent neural network from overfitting". *Journal of Machine Learning Research*, 15, 1929-1958.
- [38] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv Prepr. arXiv1409.1556*, 2014.
- [39] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?," in *Advances in neural information processing systems*, 2014, pp. 3320–3328.
- [40] Hughes, D.P., Salathe, M. (2015) "An Open Access Repository of Images on Plant Health to Enable the Development of Mobile Disease Diagnostics", *arXiv:1511.08060*.
- [41] P. K. Sonawane and S. Shelke, "Handwritten Devanagari Character Classification using Deep Learning," in *2018 International Conference on Information, Communication, Engineering and Technology (ICICET)*, 2018, pp. 1–4.
- [42] Forrest N. Iandola et al. "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <1MB model size". In: *CoRR abs/1602.07360*, (2016), url: <http://arxiv.org/abs/1602.07360>.
- [43] S. Lu, Z. Lu, and Y.-D. Zhang, "Pathological brain detection based on AlexNet and transfer learning," *J. Comput. Sci.*, vol. 30, pp. 41–47, 2019.