JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 907-914
https://publishoa.com
ISSN: 1309-3452

Analytical Study of Deep Learning Architectures for Classification of Plant Diseases

Punitha Kartikeyan[1], Gyanesh Shrivastava[2]

[1]School of Information Technology, MATS University, Raipur, Chhattisgarh, India
[2]School of Information Technology, MATS University, Raipur, Chhattisgarh, India
punitakartikeyan@gmail.com
gyaneshnkshrivastava@gmail.com

Received 2022 April 02**; Revised** 2022 May 20**; Accepted** 2022 June 18**.**

ABSTRACT

Plant disease has a significant impact on crop yield of agriculture resulting in increased economic losses. Plant disease detection is a real hurdle in the agricultural area. Farmers have difficulties is identifying, diagnosing and classifying the actual cause of diseases with the naked eye. Furthermore, identifying the diseases takes time and requires a well-trained person and specialist. In larger farms, disease detection becomes more difficult and challenging task. To address these issues, an automatic computerized plant disease identification and classification system based on Image Processing and Deep Learning techniques could be used to recognize diseases at an earlier stage as well as increase the yield. The Deep Learning technique is more accurate, efficient and reliable than the Machine Learning technique in detecting plant diseases. The goal of this analytical study is to evaluate and compare the accuracy of Deep Learning architectures AlexNet, GoogLeNetand DenseNetwith different optimizers viz SGD, RMS, PROP for tomato plant disease identification and classification.The results showthat the GoogLeNet in combination with ADAM optimizer performed well with an an accuracy of 99.56 % in identification and classification of tomato plant disease. Further, The GoogLeNet architecture trained well with the Adam optimizer achieved the best F1-score of 99.53% as compared to AlexNet and DenseNet with other optimizers. The proposed model GoogLeNet with ADAM optimizer is beneficial for farmers in identifying and classifying tomato diseases as it has high success rate.

Keywords: Deep Learning, Plant Diseases, Classification, AlexNet, GoogLeNet,DenseNet

1. Introduction

Agriculture play a major vital role in food and economic backbone of many countries. Farmers benefit from higher agricultural productivity because they earn more money. Farmers nowadays confront numerous issues such as plant diseases, pest diseases, nutritional deficiencyand other issues that result in less yield in production and significant economic losses[1],[2]. Preventing disease in its early stages can enhance theyield in terms of both quality and quantity. Plant disease detection has traditionally done by visual observation, but this involves longer time, well-trained personneland is difficult to implement on huge farms.

As a result, attention is being focused on the use of artificial intelligence systems to detect plant diseases at an early stage to protect the plant diseases. Plant diseases were previously identified and classified using different Machine Learning (ML) techniques [3]. These strategies, on the other hand, lacked the ability to extract features automatically. Hence, Deep Learning (DL) approaches have recently been employed to identify and classify the diseases with the ability to extract features automatically.

Many prominent Deep Learning architectures, such as LeNet[4],AlexNet[5],VGGNet[6],GoogLeNet[7],ResNet[8]and DenseNet[9]have been used to significantly improve the identification ,detection and accuracy of plant diseases comparable to Machine Learning techniques. They have compared the performance of AlexNet, GoogLeNet, VGG Net, ResNet ,MobileNetand DenseNet approachesbased on performance measures like as Precision, Recall, Sensitivity, Specificity, Harmonic Mean and F1-score.

2. Related Works

Precision agriculture is emphasised in modern technology since it increases productivity. The DL technique is important in plant disease identification because it extracts the relevant features automatically, more resilient and produces accurate results than the ML technique. The DL architectures is used mostly for disease

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 907-914
https://publishoa.com
ISSN: 1309-3452

detection in agriculture are discussed in this review section.

Mohanty et al. [10] used the PlantVillage dataset with 54306 images of plant diseases and healthy to analyse and evaluate two DL architectures, AlexNet and GoogleNet, to identify 14 crop species and 26 diseases. To evaluate the deep learning architecture, two training methods such as transfer learning and training from scratch as well as different dataset types like colour, grayscale oriented and segmented images with different training-test split of datasets like 80–20 %, 60–40 %, 50–50 %, 40–60 %, and 20–80% were employed. They reported thatGoogLeNet architecture with color-oriented dataset as well as split of 80-20% using transfer learning approach achieved 99.35% accuracy and also more effective model.

Brahimi et al. [11] have compared the efficacy of AlexNet and GoogLeNet architecture with shallow models of Support Vector Machine (SVM) and Random Forest (RF) for identifying and classifying tomato diseases from plant village dataset having 14828 images such as Early Blight, Late Blight, Target Spot, Bacterial Spot, Septoria Spot, Leaf Mold, Spider Mite, Mosaic Virus and Yellow Leaf curl virus . They observed that AlexNet ,GoogLeNet, SVM and RF achieved 98.66 % , 99.18% , 94.53% and 95.46% accuracies respectively and the GoogLeNet pre trained architecture was the best among them.

Amara et al. [12] suggested a LeNet architecture for banana leaf disease identification and classification of banana speckle and black sigatoka diseases using the plant village dataset with experiment of color and grayscale images. It was discovered that model trained with colour images performed better than grayscale images with a result of accuracy as 98.61%.

The plant village dataset was used by Too et al. [13]and opted transfer learning approach with fine tune parameters for analyzing and examining various DL models viz., ResNet-50, ResNet-101, ResNet-152, DenseNet-121, VGG-16 and Inception V4. They reported that the 30 th epoch, the DenseNet-121 model had the best identification and classification accuracy of 99.75%.

DL GoogLeNet model was pioneered by Barbedo[14] to classify plant leaf diseases using own field images with individual lesion and spots instead of full leaf images. The benefits of this technology were that it could detect the presence of multiple diseases on the same leaf and the disease recognition accuracy improved dramatically. Specific lesions and spots have a 94% overall accuracy results and it was 12% greater than the original image accuracy.

Bi et al [15] used their own field apple plant 334 original images as well as augmented 2004images for diagnosing and classifying the Apple Alternaria Leaf Blotch and Apple Rust using MobileNet, ResNet-152 and Inception V3 models and these modelsachieved accuracy of 73.50%, 75.59%, 77.65% with execution times of 0.22, 0.45, and 0.79 seconds respectively. The approach based on MobileNet was observed to be the most efficient andtook only 0.22 seconds for processing each image. The InceptionV3 model took more than twice as long to identify diseases as the MobileNet model, and ResNet-152 took nearly four times as long as the MobileNet model. They concluded that the MobileNetwas the most efficient and could be adopted and deployed on mobile devices easily.

Ahmad et al. [16] used two data sets, own field dataset and a laboratory dataset, to deploy four Deep Learning models, VGG-16, VGG-19, ResNet, and Inception V3, to identify and classify tomato diseases by using feature extraction and fine-tuning of hyper parameters. In laboratory data set, all of the models responded and performed well. Inception V3 in particular outperformed the other models, with an accuracy of feature extraction and fine tuning of hyper parameter as 93.40% and 99.60% approach on both data sets respectively.

Xu et al. [17] proposed VGG-16 to identify and classify maize leaf disease Leaf Blight, Rust and healthy and achieved an average accuracy as 95.33%.

3. Materials and Methods

The main goal of this research is to categorise leaf disease images using the AlexNet, GoogLeNet and DenseNet models. The key contributions are analysed and examined these three DL models by using fine tuning of hyper parameters and compared the performance metrics of these modes in terms of accuracy, precision, recall and F1-score respectively.

3.1. *Deep Learning Architectures*

In image or object recognition and classification applications, DeepLearning is essential. AlexNet, GoogLeNet, VGGNet, ResNet, and DenseNet are examples of state-of-the-art DL architectures. In the 2012 ImageNet Large Scale Visual Recognition Challenge (ILSVRC), AlexNet was suggested by [5],and it outperformed all other competing models. AlexNet has a similar structure to LeNet, but it

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 907-914
https://publishoa.com
ISSN: 1309-3452

incorporates new approaches such as max pooling, local response normalisation, and ReLUfunciton. Karen and Andrew [6] presented VGGNet network that placed second in the ILSVRC-2014. Szegedy et al. [7] presented GoogLeNet, which integrated various novel ideas such as 1x1 convolutions, smaller feature maps, and increasing the network wider and deeper. The ILSVRC-2015 champion ResNet[8] showed a deeper network that included a novel approach for ashortcut and residual connection. In 2016, [9] came up with a new concept known as dense block with many layers and providing the feature reuse for entire network. In this research, analytical study was carried out to evaluate the performance of the three state-of-the-art DL models AlexNet [5], GoogLeNet [7] and DenseNet [9] with different optimizers for tomato plant disease identification and classification.

3.1.1. AlexNet

AlexNet is a network of five convolutional layers and three fully connected layers presented by Alex Krizhevsky et al. [5]. The convolutional layers are followed by normalization and max poolingand the fully connected layer output is feed into softmax layer for classification. This design also used dropout regularisation to reduce overfitting and Rectified Linear Unit (ReLu) non linearity activation function to speed up the training process. It has 60 million parameters and 6,50,000 neurons.

3.1.2. GoogLeNet

Szegedy et al. [7] introduced GoogleNetand it won the ILSVRC in 2014 and it has 22 layers , seven million parameters and nine inception modules. Furthermore, dropout regularisation is used in the fully-connected layer, and ReLU activation is used in all convolutional layers. However, as compared to AlexNet, it has a far smaller amount of network parameters. It employs a number of inception modules, each of which includes pooling, convolutions at various sizes and concatenation operations. The goal of GoogLeNet is to categorize multiclass classification in greater depth and an effective result.

3.1.3. DenseNet

The goal of DenseNet-201[9] is to connect all layers directly in order to ensure maximal features data transfer between the network's middle layers. It has the advantage of having a denser network with fewer parameters. DenseNet is made up of dense blocks, a convolutional layer, and a pooling layer, with each dense block having different convolutions kernels such as 1x1 and 3x3. Simultaneously, this connection approach improves feature and gradient transfer while also making the network easier to learn.

3.2. Datasets

The above three architectures used PlantVillage dataset [10]and which contained 54306 images with 38 different healthy and diseased leaves associated to their 14 plant species. Initially, the image sizes were adjusted to be more appropriate for the corresponding architecture. The dataset was split for training 70%, 20%, and 10% for training, validationand testing process respectively. Tomato plant diseases such as Late Blight (1909), Early Blight(1000), Mosaic Virus (373), Yellow Leaf Curl Virus (5357), Spider Mite (1676), Bacterial Spot(2027), Target Spot (1404), Septoria Leaf Spot (1771) and Leaf Mold (952) and healthy (1591)of18,060 images are taken for this study.

3.3. Effect of Hyper Parameters on Model Efficiency

Hyperparameters are vital and play significant role in DL algorithms because they examine the training and have a major impact on model performance. To improve performance optimization, experts must manually set and adjust different hyperparameter options [18-22]. Two types of hyperparameters are used, one for defining the network structure and other for determining the performance of the network training.

3.3.1 Define the Network Structure

• Kernel Size – the filtersize for edge detection, shape detection, etc.,
• Stride – the movement ratio of the kernel traverses the input image.
• Padding – adding layers of zeros to our input to ensure that the kernel passes over the image boundary or edges.
• Hidden Layer – define that how many hidden layers can be used.
• Activation Functions – it is enabling the model to learn nonlinear prediction boundaries.

3.2.2 Performance of the Network

Hyperparameters such as batchsize, learning rate, momentum, dropoutand optimizer type have the following effects on model efficiency.

3.2.2.1 Batch Size: the batch size defines how many samples or numbers of images are sent across the network.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 907-914
https://publishoa.com
ISSN: 1309-3452

3.2.2.2 Learning Rate: The learning rate determines and regulates the speed at which the model is trained by controlling and adjusting the network's weight modifications. Smaller learning rate can yield accurate results, but it takes much longer time to converge, whereas a high learning rate allows for quick learning, however the weights parameter may not be adequate or optimal. As a result, choosing the right learning rate for the model is important.

3.2.2.3 Momentum: It helps to speed up training and learning rate can aid in bringing the optimization process together and it is required to prevent oscillations in the high-curvature areas of the Stochastic Gradient Descent (SGD) optimizer's loss function.

3.2.2.4 Dropout: Dropout prevents the model from becoming overfit. It is a regularization method that aids in the learning of more powerful differentiating features by the network.

3.2.2.5 Optimizer: By reducing the loss function, the optimizers improve the weight parameters to produce more accurate results. When it comes to training deep learning models, choosing the right optimizer is vital. Several optimizers are used to test the model's performance, including SGD, AdaGrad, Adadelta and ADAM. The tomato plant disease identification and classification can be improved by applying following optimizers and major properties of these optimizers are discussed.

• SGD: One of the most basic optimizers in the deep learning and a constant (static) learning rate for all parameters is required throughout the network training and it has convergence performance rapidly.

• AdaGrad: For each parameter in this system model, this optimizer utilizes a distinct learning rate. It adjusts the learning rate based on the frequency with which each parameter is updated.

• RMSProp: This was developed to minimize the training time and its learning rate.

•Adadelta: Adadelta is intended to speed up the optimization process, for example, by reducing the number of function computations needed to obtain the optima or to enhance the optimization method's capability, for example, by producing a better final result. It is a combination of the AdaGrad and RMSProp methods.

• Adam: It combines the benefits of two enhanced versions of the SGD approach that is Adagrad and RMSProp. It estimates the second moment of gradient average and uses previous gradients for speeding up the learning process. The advantages and disadvantages of various optimizers [23-25] are presented in Table.1

Table 1. Advantages and Disadvantages of Various Optimization Techniques.

Name of Optimizer	Advantages	Disadvantages
SGD	It is easy and efficient for handling huge datasets and not necessary to store the loss function values, hence it uses less memory.	Large number of hyper parameters and iterations are required for SGD. As a result, it is affected by feature scaling. After reaching global minima, it might shoot.
AdaGrad	It is not necessary to manually adjust the learning rate. It works well with data that is sparse	The necessity to estimate the derivative of a function makes it computationally demanding. The rate of learning is constantly reducing, resulting is slow training process.
RMSProp	It is well-suited to stochastic objectives and for min-batch learning.	Manually the learning rate is adjusted.
Adam	Adam is a fast and converges quickly. It overcomes the vanishing learning rate issue that AdaGrad had.	Costly in terms of computation.

3.3.3 Performance Evaluation Metrics

The performance metrics are used to study and examine the performance of the network. In general, it is a standard measure of the trained classifiers performance when compared to new images from the testing set. When compared to the actual class label given to the image, the result of this prediction comes under true positive (TP) or true negative (TN) if correctly classified, and false positive (FP) or false negative (FN) if incorrectly classified. The performance metrics accuracy, precision, recall, sensitivity, specificity, F1-score are used to classify the tasks by using these TP, TN, FP, and FN values. Table 2 lists the performance metrics that were employed in our research to examine and evaluate the performance of classifier with formula and descriptions.

Table 2. Performance Metrics Parameters

Name	Description	Formula
Accuracy	The percentage of predicted correctly observations compared to the total number of all observations.	(TP+TN)/ (TP+FP+TN+FN)
Precision	The percentage of correctly	TP/(TP+FP)

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 907-914
https://publishoa.com
ISSN: 1309-3452

	predicted positive observations to the total predicted positive observations.	
Recall	The percentage of correctly predicted positive observations to all observations in actual class.	TP/(TP+FN)
F1-score	It is the average of precision and recall.	2*Recall*Precision/(Recall+Precision)

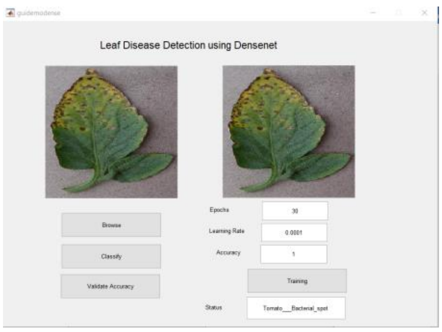
4. Results and Discussion

The analytical study is conducted using DL architectures AlexNet, GoogLeNet and DenseNet as well as their fine tuning of hyper parameters for classification of plant diseases. The performance metrics such as sensitivity, specificity, precision, accuracy, and F1-score are evaluated.

4.1. Analytical Study of Deep Learning Architectures

The objective of this study is not only to identify the diseased as well as healthy leaves but also to improve the performance of the DL architectures.The different hyperparameters mentioned (section 3.3)play significant role in the pre trained DL architecture as they influence the performance of the models. The impact on performance metrics of various optimizers such as SGD, ADAM, RMSPROP have been studied and showed in Figure.1 with respect to plant disease identification and classification.

Fig .1. Tomato Plant Disease Detection and Classification



The performance of three pretrained network architectures: AlexNet, GoogLeNet, and DenseNet, as well as with three optimizers SGD, ADAM, RMSPROPfor tomato plant disease detection and classification areanalyzed and the impact of hyperparameters on model efficacy is presented in the Table 3 and also in Figures 2 , 3&4. It can be seen from the table and figures that all the optimizes studied under AlexNet, GoogLeNet and DenseNet showed improved performance in terms of accuracy in detecting tomato plant diseases. The SGD optimizer under AlexNet had an accuracy of 96.30 %, which increased to 98.39% and 97.35% when trained underGoogLeNet and DenseNet respectively. The ADAM and RMSPROP also showed increase in accuracy when trained under GoogLeNet and DenseNet . However, the performance of these three optimizers were improved and gave higher accuracy as 98.39%, 99.56% and 98.57 in GoogleNet platform.

The AlexNet pretrained architecture model attained accuracy as 96.30% when the model was trained using the SGD optimizer which was compared with other two optimizers. The percentage increase in the accuracy is 1.41 in Adam optimizer, whereas, it is 0.52% increase in RMSProp optimizer. The Adam optimizer enhanced its accuracy to 97.66% under this pretrained model. Under the GoogLeNet pretrained model, it attained 98.39% accuracy when it was trained with the SGD optimizer. This optimizer was compared with others and the percentage increase was 1.19% was achieved under Adam optimizer. While it is recorded 0.18% increase in RMSProp. The accuracy is significantly improved in Adam optimizer.Anaccuracy of 97.35% is achieved in DenseNet model with SGD Optimizer. The percentage increase in the accuracy of other optimizers ADAM, RMSProp are recorded to be 1.59% and 1.25% respectively when compared with SGD. The highest improved accuracy is obtained in Adam optimizer. Further, the percent increase in accuracy by Adam optimizer ranged from 0.65 to 1.95 % among these three architectures. Overall, the performance of Adam optimizer is found superior and gave the highest accuracy of 99.56% under GoogLeNet architecture. The ADAM optimizer under GoogLNet gave the highest F1 score of 99.53.

Table 3. Efficiency of different optimizers on DL performance

Optimizer	Sensitivity/Recall	Specificity	PreciSion	Accuracy	F1-Score
		egatnecreP			
		AlexNet			
SGD	96.35	96.25	96.29	96.30	96.30
ADAM	98.44	96.88	96.92	97.66	97.67
RMS PROP	96.88	96.72	96.76	96.80	96.79
		GoogLeNet			
SGD	98.12	98.93	98.29	98.39	98.20
ADAM	99.50	99.50	98.60	99.56	99.53
RMS PROP	98.00	98.70	98.00	98.57	98.76

		DenseNet (201)			
SGD	97.35	99.69	97.34	97.35	97.34
ADAM	98.37	99.87	98.73	98.92	98.53
RMS PROP	98.36	97.95	98.28	98.59	98.11

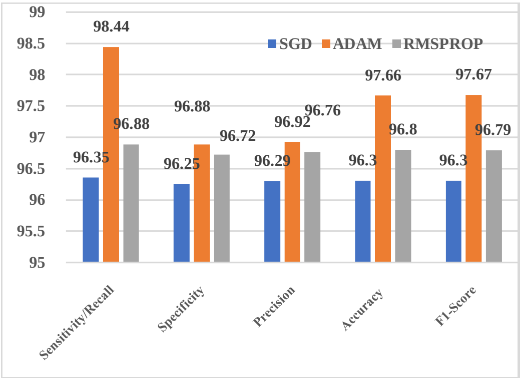


Fig .2. Comparison of SGD, ADAM, RMSPROP optimizer in AlexNet

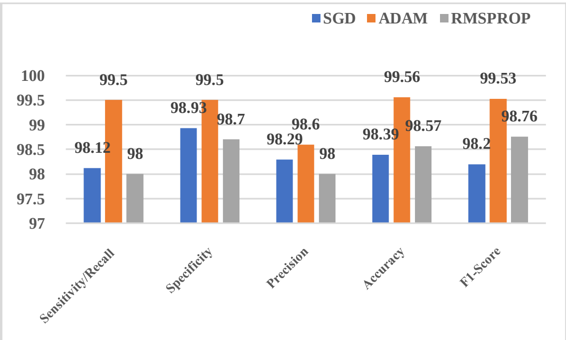


Fig.3.Comparison of SGD, ADAM, RMSPROP optimizer in GoogLeNet

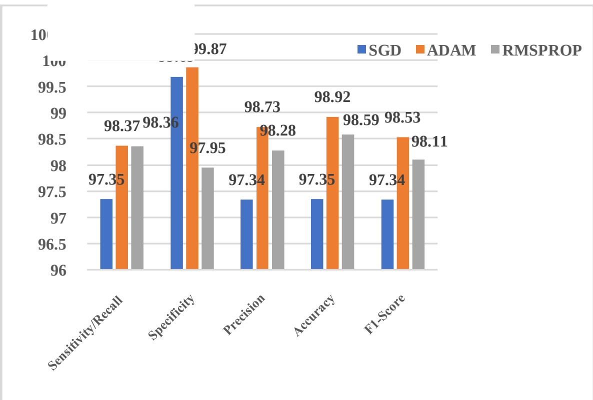


Fig. 4. Comparison of SGD, ADAM, RMSPROP optimizer in DenseNet

This analytical study illustrated the impact of hyper parameter optimizers, the best hyper parameters is selected for optimal DL architecture for identification and classification of tomato plant diseases.

5. Conclusion

Deep Learning pre-trained architectures with fine tuning hyper parameters are used to diagnose and classifying the tomato plant diseases and minimise their severity as well. In this research, analytical study of different well known state-of-the-art architectures AlexNet, GoogLeNet, and DenseNetare experimented with various optimizer SGD, ADAM, RMSProp for their performance in terms of Sensitivity, Specificity, Precision, Accuracy and F1-score. The GoogLeNet trained with Adam optimizer gave the best performance to an accuracy of 99.56% followed by RMSProp optimizer, with a 98.57% accuracy in identification and classification of maize leaf diseases. The proposed study indicates that hyperparameter optimizationshelped to improve the performance of the pretrained DL models. Future study will focus on improving theefficacy of the method for identifying plant disease by data augmentation, evaluating other optimization algorithms and hybrid algorithms, among other things.

References

[1]. R. N. Strange, P. R. Scott, “Plant disease: a threat to global food security,” Annual review of phytopathology, 43(2005), 83-116

[2]. Raghuvanshi, A., Singh, U., Sajja, G., Pallathadka, H., Asenso, E., & Kamal, M. et al. (2022). Intrusion Detection Using Machine Learning for Risk Mitigation in IoT-Enabled Smart Irrigation in Smart Farming. Journal Of Food Quality, 2022, 1-8. doi: 10.1155/2022/3955514

[3]. L. Chen, Y. Yuan, Agricultural disease image dataset for disease identification based on machine learning, in: J. Li, X. Meng, Y. Zhang, W. Cui, Z. Du (Eds.), Big Scientific Data Management. BigSDM, Lecture Notes in Computer Science, Springer, Cham, 2019, p. 11473, https://doi.org/10.1007/978- 3-030-28061-1_26.

[4]. Hemamalini, V., Rajarajeswari, S., Nachiyappan, S., Sambath, M., Devi, T., Singh, B., & Raghuvanshi, A. (2022). Food Quality Inspection and Grading Using Efficient Image Segmentation and Machine Learning-Based System. Journal Of Food Quality, 2022, 1-6. doi: 10.1155/2022/5262294

[5]. Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. In Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1,

NIPS'12, pages 1097–1105, USA, 2012. Curran Associates Inc. https://doi.org/10.1145/3065386
[6]. KarenSimonyan, Andrew Zisserman,. Very deep convolutional networks for large-scale image recognition. ariXiv,CoRR, abs/1409.1556, 2014.
[7]. C. Szegedy, Wei Liu, Yangqing Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich,. Going deeper with convolutions. In 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 1–9, June 2015. https://doi.org/10.1109/CVPR.2015.7298594
[8]. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 770–778, June 2016. https://doi.org/10.1109/cvpr.2016.90
[9]. G. Huang, Z. Liu, L. V.D. Maaten, K. Q. Weinberger,. Densely connected convolutional networks. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 2261–2269, July 2017. https://doi.org/10.1109/CVPR.2017.243
[10]. S.P.Mohanty,D.P. Hughes; M.Salathé, Using deep learning for image-based plant disease detection. Front. Plant Sci. 7(2016), 1419. [CrossRef] https://doi.org/10.1016/j.compaq.2020.105393
[11]. M. Brahimi, K.Boukhalfa, A.Moussaou,. Deep learning for tomato diseases: classification and symptoms visualization. Applied Artificial Intelligence 31(2017) 299–315, Taylor-Francis Group, https://doi.org/10.1080/08839514.2017.1315516
[12]. J. Amara, B.Bouaziz, A.Algergawy, A deep learning-based approach for banana leaf diseases classification. In: Lect Notes in Informatics (LNI), Proc - Ser Gesellschaft Fur Informatik, Bonn 266 (2017) 79–88.
[13]. Edna Chebet Too, Li Yujiana, Sam Njukia, A. Liu Yingchun, comparative study of fine-tuning deep learning models for plant disease identification, Edna Chebet Tooa, Li Yujiana , Sam Njukia , Liu Yingchun, Computers and Electrnoics in Agriculture,Elsevier, 161(2018) 272–9. https://doi.org/10.1016/J.COMPAG.2018.03.032
[14].J.G.A Barbedo, Plant disease identification from individual lesions and spots using deep learning. Biosystem Engineering 180 (2019) 96–107. https://doi.org/10.1016/j.biosystemseng.2019.02.002.
[15]. C. Bi, J. Wang, Y. Duan, B. Fu, J-R. Kang, Y. Shi, ''MobileNet based apple leaf diseases identification,'' Mobile Networks and Applications, Springer, 10, pp. 1–9, Aug. 2020, https://doi.org/10.1007/s11036-020-01640-1
[16]. I. Ahmad, M. Hamid, S. Yousaf, S. T. Shah, M. O. Ahmad, ''Optimizing pretrained convolutional neural networks for tomato leaf disease detection,'' Hindawi Complexity, vol. 2020, pp. 1–6, 2020, https://doi.org/10.1155/2020/8812019
[17]. J.-H. Xu, M.-Y. Shao, Y.-C. Wang, W.-T. Han, ''Recognition of corn leaf spot and rust based on transfer learning with convolutional neural network,'' Trans. Chin. Soc. Agriculture. 51(2020) 230–236
[18]. E.T, Chebet, Li Y, Sam N, Y,Liu,. A comparative study of finetuning deep learning models for plant disease identification. Computer Electron Agric 161(2019) 272–279. https://doi.org/10.1016/J.COMPAG.2018.03.032
[19]. M.M. Ghazi,B..Yanikoglu, E Aptoula,. Plant identification using deep neural networks via optimization of transfer learning parameters. Neurocomputing 235(2017) 228–235. [CrossRef]
[20]. N,Japkowicz, M,Shah,. Evaluating Learning Algorithms: A Classification Perspective; Cambridge University Press: Cambridge, UK, 2011.
[21]. E.C. Too, L.Yujian,S.Njuki, L. Yingchun,. A comparative study of fine-tuning deep learning models for plant disease identification. Comput. Electron. Agric. 161(2018) 272–279. [CrossRef] https://doi.org/10.1016/J.COMPAG.2018.03.032
[22]. S. Ruder,. An overview of gradient descent optimization algorithms. arXiv 2016, arXiv:1609.04747. 37. van Laarhoven, T. L2 regularization versus batch and weight normalization. arXiv 2017, arXiv:1706.05350.
[23]. M.Mehdipour Ghazi, B,Yanikoglu, E,.Aptoula, Plant identification using deep neural networks via optimization of transfer learning parameters. Neurocomputing 235(2017) 228–235. [CrossRef] https://doi.org/10.1016/j.neucom.2017.01.018

[24]. P.Li,. Optimization Algorithms for Deep Learning. 2017.Available online: http://lipiji.com/docs/li2017optdl.pdf

[25]. D.P.KingmaBa J. Adam, A method for stochastic optimization. arXivpreprint arXiv:1412.6980 (2014)25-33