Uluslararası İleri Doğa Bilimleri ve Mühendislik Araştırmaları Dergisi
Sayı 9, S. 298-310, 6, 2025

Araştırma Makalesi

IJANSER

https://as-proceeding.com/index.php/ijanser
ISSN:2980-0811

International Journal of Advanced Natural Sciences and Engineering Researches
Volume 9, pp. 298-310, 6, 2025

Research Article

Determining Categories of Leaf Images Using Transfer Learning of VGGNet

Rıfat Aşlıyan[*,1], Bircan Cemek[2]

[1]*Department of Mathematics, Faculty of Science, Aydın Adnan Menderes University, Türkiye*
[2]*Graduate School of Natural and Applied Sciences, Aydın Adnan Menderes University, Aydın, Türkiye*

[*]rasliyan@adu.edu.tr *Email of the corresponding author*

(Received: 29 June 2025, Accepted: 01 July 2025)

(6th International Conference on Engineering and Applied Natural Sciences ICEANS 2025, June 23-24, 2025)

ATIF/REFERENCE: Aşlıyan, R. & Cemek, B. (2025). Determining Categories of Leaf Images Using Transfer Learning of VGGNet, *International Journal of Advanced Natural Sciences and Engineering Researches*, 9(6), 298-310.

Abstract – Recognizing plant species from their leaves holds great value across many domains, from agriculture and forestry to pharmacology and the assessment of regional biodiversity. This research focused on creating automated systems capable of identifying 12 distinct plant species solely from leaf images. To accomplish this, we utilized the powerful VGGNet16 convolutional neural network. After dividing our leaf images into a train set and a test set, we enriched the training data through an on-the-fly augmentation process. By applying transformations like scaling, translation, and rotation to the images during the network's training epochs, we improved its ability to generalize to future data, thereby enhancing its overall performance. We applied with two approaches as training a standard VGGNet16 network and training a pre-trained VGGNet16 that had already learned features from a massive image database. The effectiveness of these systems, configured with different parameters, was benchmarked on the test dataset using classification metrics as F1-Score , Accuracy, and the ROC curve. Performance of the pre-trained VGGNet16 network was observed to be markedly superior, achieving an F1-score of 98.8% compared to the 93.4% highest score from the standard VGGNet16.

Keywords – Leaf Image Categorization, VGGNet16, Convolutional Neural Networks, Transfer Learning, Augmented Images.

I. INTRODUCTION

In agriculture, using drones for the early detection of diseases from plant leaves enhances productivity by enabling faster intervention. Additionally, farming costs can be reduced. If weeds are detected early, their removal from the field becomes easier. Nutrient deficiencies in plants can be identified from their leaves and can be corrected through appropriate fertilization. Most plant diseases can be identified by examining their leaves, which allows for the application of the correct treatment to cure the disease [1-3].

Through leaf classification studies, identifying plant species in a specific region or forest, monitoring these plants over time, and ensuring their conservation are of great importance [4]. The rapid and accurate identification of plant species is also highly beneficial. The status of rare or endangered plant species can be monitored more easily and effectively. Furthermore, the impact of climate change on plant species can be determined.

Leaf image categorization offers significant contributions to various fields such as agriculture, biodiversity assessment, accurate plant species identification, pharmacology, forestry, and food security.

In many studies, convolutional neural networks have been trained separately using different datasets of leaf images with same backgrounds, and the effects of various parameters on the training process has been examined. A combined dataset, created by merging these individual datasets, has also been used for training and testing purposes. General-purpose leaf classification models were created using convolutional neural networks, and the models ‘results from the categorization have been compared with other works in scientific literature.

Nowadays, higher accuracy rates in plant disease detection can be achieved by applying advanced techniques like artificial neural networks, in addition to traditional methods of artificial intelligence. In this context, the use of artificial intelligence-based systems can be more effective by supporting decision-making processes in agriculture, thereby enhancing the efficiency of farmers and agricultural engineers in disease detection and management. [5].

For this reason, this study will greatly contribute to the proliferation of sustainable practices and smart agriculture technologies. Integrating artificial intelligence with drone technology will aid in mitigating environmental harm and preventing economic losses. The reduction of yield loss in farming is regarded as an important advancement for securing farmers' food supply and increasing their economic sustainability [6].

In this study, leaf species from a leaf dataset were classified using the VGGNet16 convolutional neural network. This dataset consists of twelve different classes. Seventy percent of the dataset was utilized to train the models, and after the models have been built, their performance was evaluated using the test dataset. An image augmentation process was also applied to rise the image size in the dataset. Models have been created from these convolutional networks using different parameters and were then compared with one another. The accuracy and F1-score metrics were chosen for model evaluation [7-11].

The rest of the paper is as the followings. This study’s second section provides a literature review on plant recognition and classification using deep learning models. The third chapter offers fundamental information on convolutional neural networks, their layers, and activation functions. Subsequently, the VGGNet network used in this thesis is defined, and its architecture and operation are explained. Additionally, information is provided about the ROC curve and the leaf dataset used in this thesis. The fourth chapter presents the experimental studies in the form of tables and graphs. The fifth chapter contains the conclusion of the thesis, where the findings from the previous chapter are compared.

II. MATERIALS AND METHODS

VGGNet16, which is one of the convolutional neural networks, has been utilized to construct machine learning models in this study. The next subsections mentions the technique and the leaf dataset.

A. *VGGNet Convolutional Neural Network*

VGGNet is a CNN model designed by Simonyan and Zisserman in 2014 and is considered a significant milestone in the deep learning’s field. This model became prominent for its superior performance, particularly in classification of images, and introduced a structure that enabled deep neural networks to learn more complex features. VGGNet gained attention for its success in the ImageNet competition, showcasing the potential of deep learning [12].

One of VGGNet's defining features is its use of exclusively 3x3 convolutional filters to increase its depth. This approach enables the model to learn more complex features while keeping the number of parameters manageable. Consequently, the increased depth of VGGNet allows for better generalization.

The VGGNet architecture is known for its two main versions: VGGNet16 and VGGNet19. VGGNet16 is composed of 16 layers in total, whereas VGGNet19 has 19 layers. Both models operate on 224x224 pixel RGB images. The convolutional layers are followed by consecutive max-pool layers, which decrease the feature maps’ size and lower the computational cost. This structure allows for an increase in the model's depth while simultaneously reducing feature dimensions. The model's learning process increases with Rectified Linear Unit (ReLU) activation function and its accuracy rate has been improved by ReLU [13]. Furthermore, the use of a greater number of filters in subsequent convolutional layers assist to learn more complex features.

In addition, data augmentation techniques applied to large datasets played an important role in the training of VGGNet. Data augmentation increased the variety of the train data, improving the generalization capacity of the models and preventing overfitting. Today, VGGNet is widely used in deep learning research and applications and is also considered a popular base model for transfer learning. Users can get a quick start on their own custom tasks by using VGGNet's pre-trained weights. This enhances VGGNet's flexibility and versatility. Furthermore, VGGNet can be customized for use in various tasks, which ensures its effectiveness across diverse applications [14].

VGGNet's success has inspired the development of many subsequent deep learning architectures. In particular, more complex structures like ResNet are based on the architectural principles of VGGNet. ResNet succeeded in solving the challenges of training deep networks by using skip connections. This, in turn, has enabled the emergence of more complex and effective models in the field of deep learning [15].

Currently, VGGNet is effectively used in numerous fields such as object recognition, face recognition, medical image analysis, and video analysis. Particularly in medical image analysis, VGGNet's deep learning capabilities hold the potential to make a significant contribution to the early diagnosis of diseases and the improvement of treatment processes.

In conclusion, VGGNet is regarded as a key reference point in the deep learning’s field and has become an indispensable tool for both researchers and practitioners. The depth and flexibility offered by VGGNet have constructed a center model in deep learning applications. Future research may contribute to the creation of more efficient and effective deep learning models by further enhancing VGGNet's architecture [16].

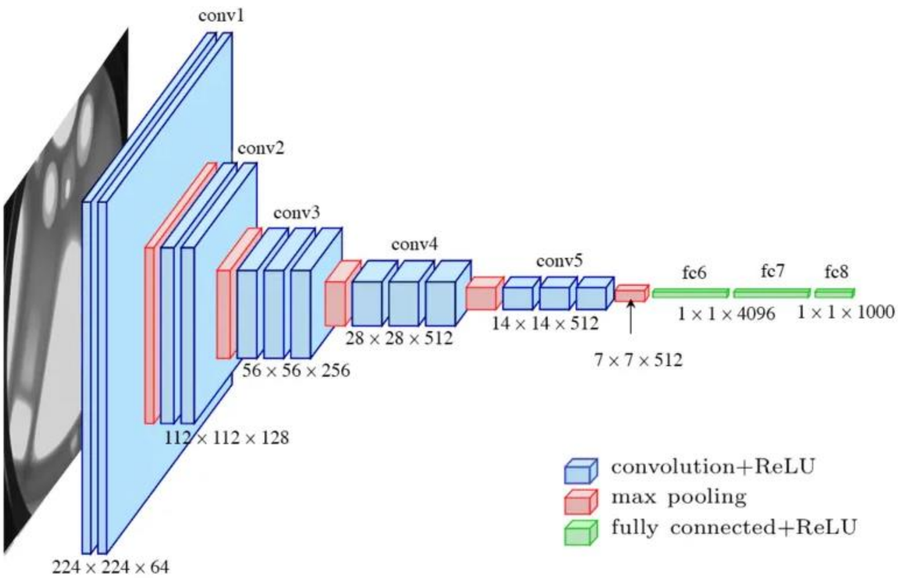

Fig. 1 The structure of VGGNet model

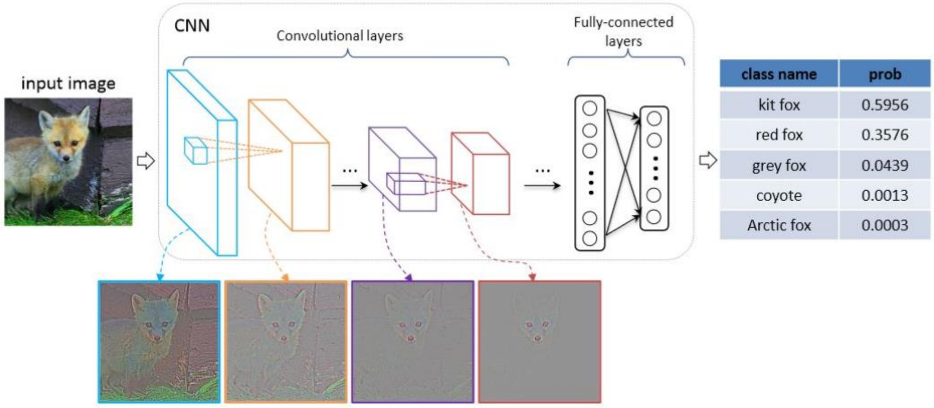

Fig. 2 VGGNet architecture [17]

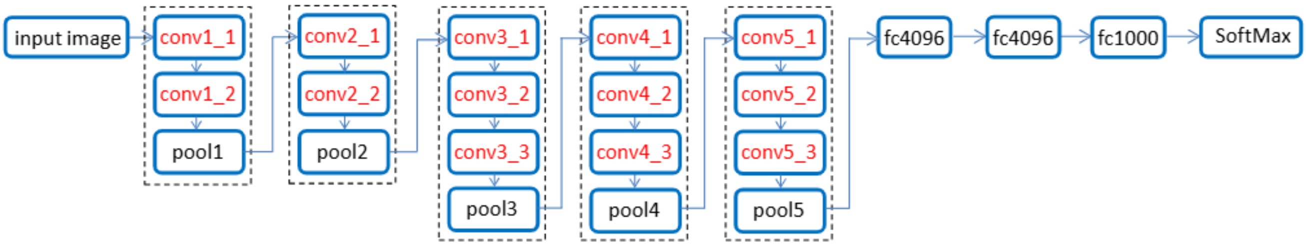

Fig. 3 VGGNet16 layers [17]

B. *Leaf image dataset*

The imbalanced leaf image dataset, obtained from the Mendeley website, consists of 12 classes and a total of 2395 images, as specified in Table 1. From this dataset, 1676 images were allocated for training and 719 for testing. Figure 4 shows some augmented leaf images from the dataset.

Table 1. Leaf image data set.

Leaf Number	Leaf Type	Total Images	Train Images	Test Images
1	Alstonia Scholaris	179	125	54
2	Arjun	220	154	66
3	Bael	118	83	35
4	Basil	148	104	44
5	Chinar	103	72	31
6	Guava	277	194	83
7	Jamun	279	195	84
8	Jatropha	133	93	40
9	Lemon	159	111	48
10	Mango	170	119	51
11	Pomegranate	287	201	86
12	Pongamia Pinnata	322	225	97
	Total	2395	1676	719

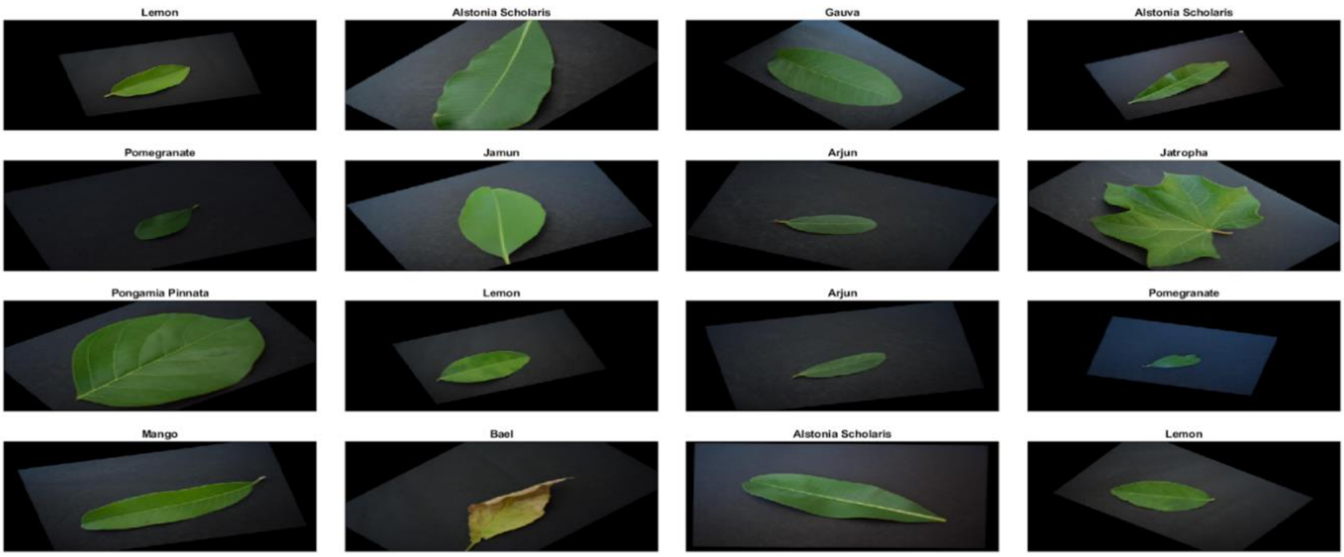

Fig. 4. The augmented images

III. RESULTS AND DISCUSSIONS

In this study, the images from the leaf dataset consisting of 12 different leaf types have been classified using VGGNet16, which is one of the deep learning methods. After the data in the training dataset was trained with this method, models have been constructed. The performance of these developed models on both the training and test datasets has been evaluated using the commonly used classification metrics: Accuracy, Precision, Recall, and F1-Score. In addition, to provide a more comprehensive assessment of the models' performance, their Confusion Matrices and ROC curves were generated and the models have been compared.

As shown in Table 2, Confusion Matrix contains the counts of TP (True Positive), TN (True Negative), FP (False Positive), and FN (False Negative). To calculate the confusion matrices and evaluation metrics, it is necessary to understand the concepts of TP, TN, FP, and FN. When considering the categorization of a binary dataset, one class is considered "Positive." The other class is then assumed to be "Negative" for the calculations. TP (True Positive) represents the number of samples from the positive class that were correctly categorized. TN (True Negative) indicates the number of samples from the negative class that were correctly categorized. FP (False Positive) is the number of samples, actually belonging to the negative class, that were incorrectly categorized as "Positive." Similarly, FN (False Negative) refers to the number of samples, belonging to the positive actually, that were incorrectly categorized as "Negative." With these concepts, the metrics as F1-Score, Accuracy, Precision and Recall can be computed as shown in Equations 1-4.

Table 2. Confusion matrix

	“Positive” in reality	“Negative” in reality
Predicted “Positive”	TP	FP
Predicted “Negative”	FN	TN

The Accuracy metric indicates the proportion of all samples (for both Positive and Negative class instances) that are classified correctly. Precision indicates what proportion of the instances predicted as Positive were, in fact, correctly identified as Positive. But, Recall expresses what proportion of the actual Positive class instances were correctly predicted as Positive. The F1-Score is the harmonic average of the Precision and Recall rates. In some cases, either Precision or Recall can be high while the other is low. In such situations, it is more appropriate to perform the evaluation by taking the mean of the Precision and Recall scores.

$$\text{Precision} = TP/(TP + FP) \tag{1}$$

$$\text{Recall} = TP/(TP + FN) \tag{2}$$

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \tag{3}$$

$$\text{F1-Score} = (2 \times \text{Precision} \times \text{Recall})/(\text{Precision} + \text{Recall}) \tag{4}$$

The training and testing of the models have been performed in the MATLAB 2024b environment. These applications were coded on a laptop computer featuring a 12th Generation Intel Core i7-12650H 2.30 GHz CPU, 40 GB of main memory, and a 64-bit Windows 11 operating system. The use of a GPU is particularly important for running deep learning algorithms, as it enables very fast training. The training processes were made significantly faster by using an NVIDIA GeForce RTX 4050 GPU.

As shown in Figure 5 and Table 3, the performance values of the pre-trained VGGNet16 network are presented according to different learning rates. For learning rates between 0.1 and 0.01, the F1-score performance rates were low. When the learning rate was set to 0.0001, the highest performance among the learning rates was achieved, yielding an F1-score of 98.4%.

Table 3. F1-Score success rates of pre-trained VGGNet16 according to learning rates.

Learning Rate	**Macro Average F1-Score**
0.1	0.134
0.01	0.134
0.001	0.955
0.0001	**0.984**
0.00001	0.974
0.000001	0.895

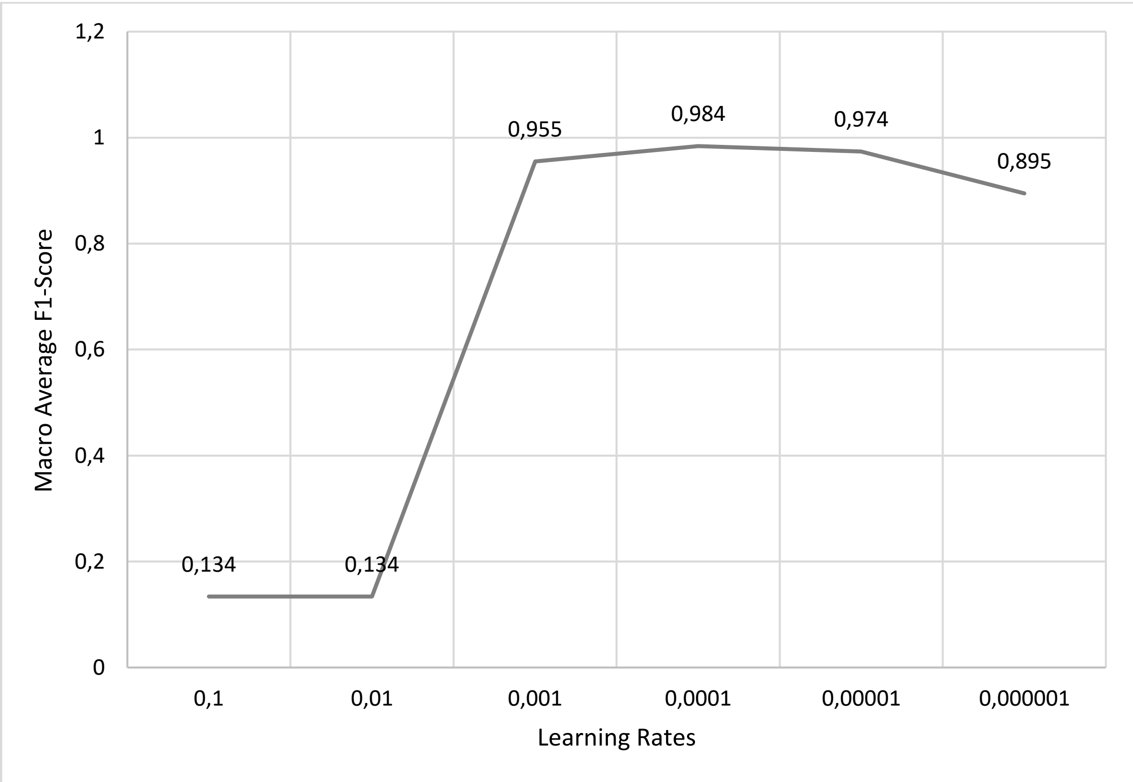

Fig. 5. The graph of learning rate and F1-score for pre-trained VGGNet16

Figure 6 and Table 4 show the F1-score performance results of the pre-trained VGGNet16 network according to the learning rules used. Performance increases across the learning rules in the order of SGDM, Adam, and RMSPROP. According to the F1-score, the highest success rate with 98.8% has been achieved with SGDM.

Table 4. F1-Score success rates of pre-trained VGGNet16 according to learning rules

Learning Rules	Macro Average F1-Score
SGDM	**0.988**
ADAM	0.955
RMSPROP	0.91

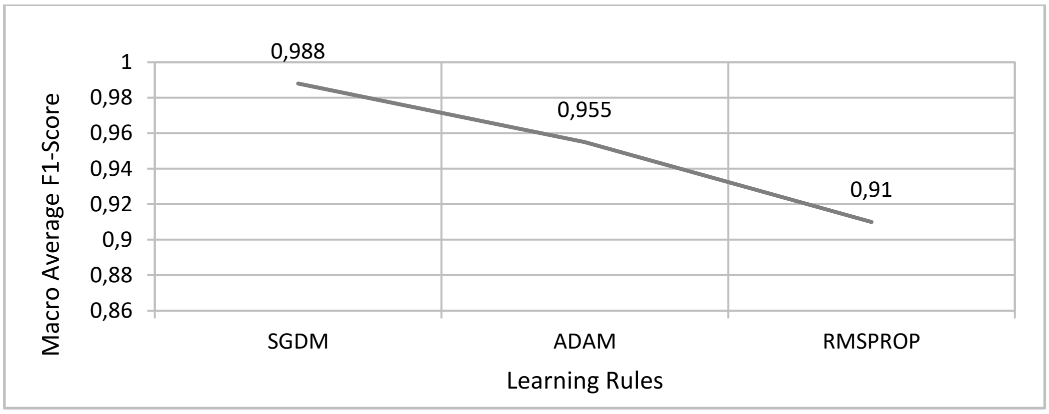

Fig. 6. The graph of learning rule and F1-score for pre-trained VGGNet16

When we examine the results in Figure 7 and Table 5, we see that as the number of epochs for the pre-trained VGGNet network increases, the performance rates sometimes increase and at other times decrease. According to these results, although there is very little difference among the performance rates for 15, 25, and 30 epochs, the best performance based on the number of epochs has been achieved at 20 epochs, with F1-score of 97.4%.

Table 5. F1-Score success rates of pre-trained VGGNet16 according to epoch number

Epoch Number	Macro Average F1-Score
5	0.919
10	0.884
15	0.955
20	**0.974**
25	0.956
30	0.956

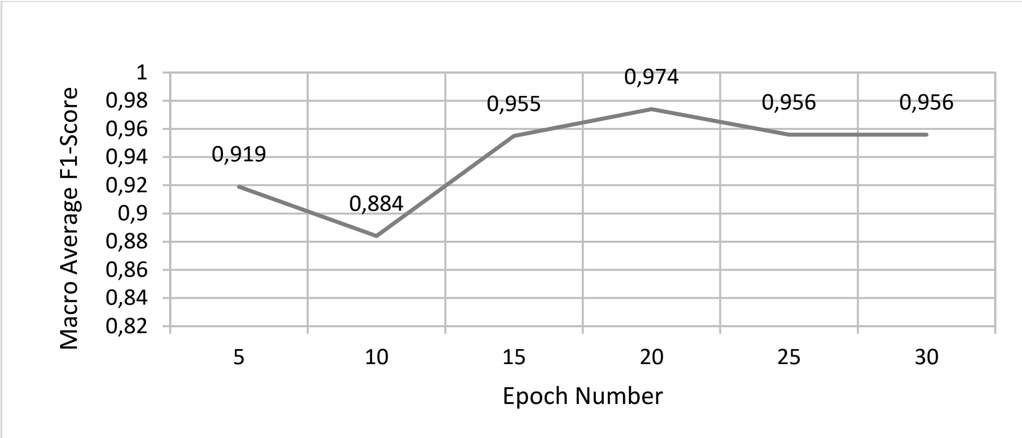

Fig. 7. The graph of epoch number and F1-score for pre-trained VGGNet16

As shown in Figure 8 and Table 6, the F1-score performance rates generally increase as the batch size values increase. However, the performance decreases slightly when the batch size is 128. The best performance, F1-score of 98.4%, has been achieved with a batch size of 64.

Table 6. F1-Score success rates of pre-trained VGGNet16 according to batch size

Batch Size	Macro Average F1-Score
8	0.958
16	0.966
32	0.968
64	**0.984**
128	0.973

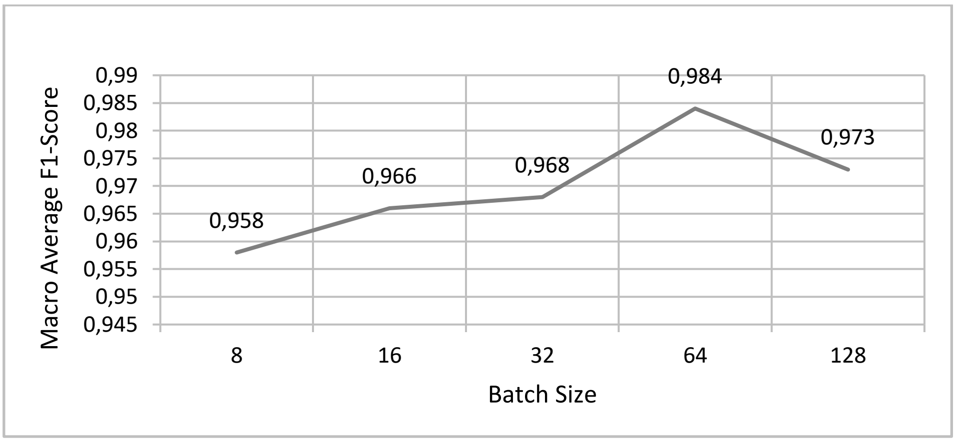

Fig. 8. The graph of batch size and F1-score for pre-trained VGGNet16

Figure 9 displays the confusion matrix obtained from the test results of the VGGNet16 network on the leaf images. According to this confusion matrix, the 3rd, 4th, 5th, 8th, 9th, 10th, and 11th leaf classes were classified with 100% accuracy. The other classes were also classified with very high success.

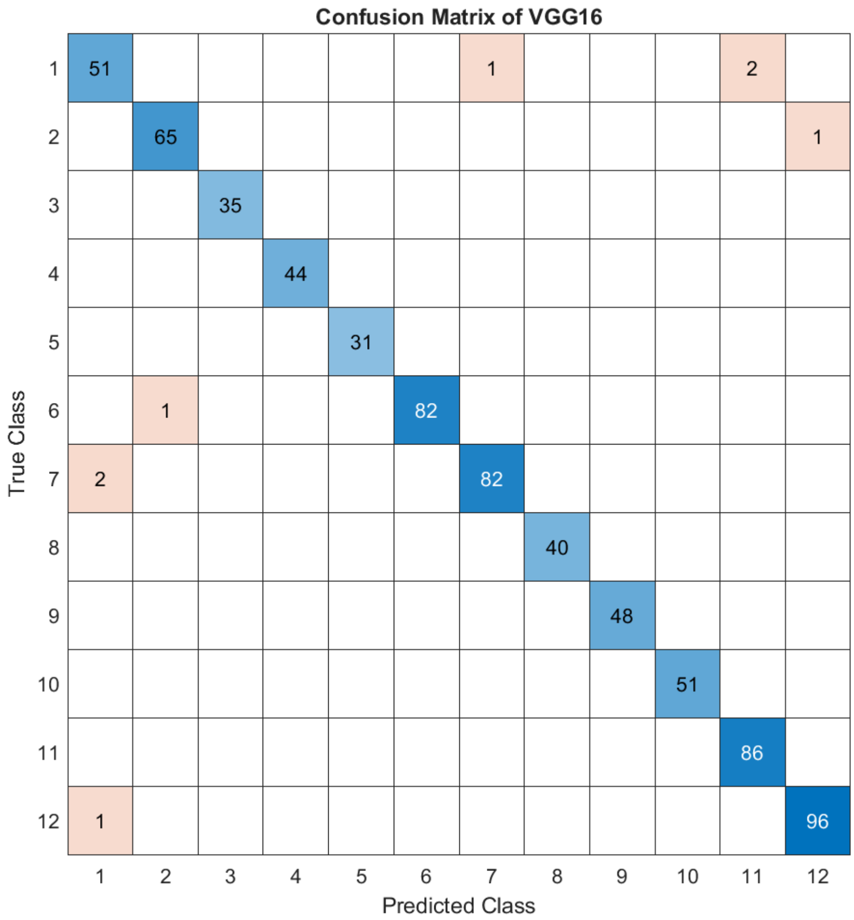

Fig. 9. The confusion matrix for VGGNet16 when the learning rule is SGDM, the learning rate is 0.0001, the batch size is 64, and the maximum number of epochs is 20

Figure 10 shows the ROC curve for the pre-trained VGGNet16 network. It displays the values of the true positive rate and the false positive rate in the dataset. As was also seen in the confusion matrix, the Bael, Basil, Chinar, Gauva, Jatropha, Lemon, and Mango leaf images were all classified correctly. This is because the AUC (Area Under the Curve) values for these classes are 1.0, which signifies a perfect, 100% successful classification.

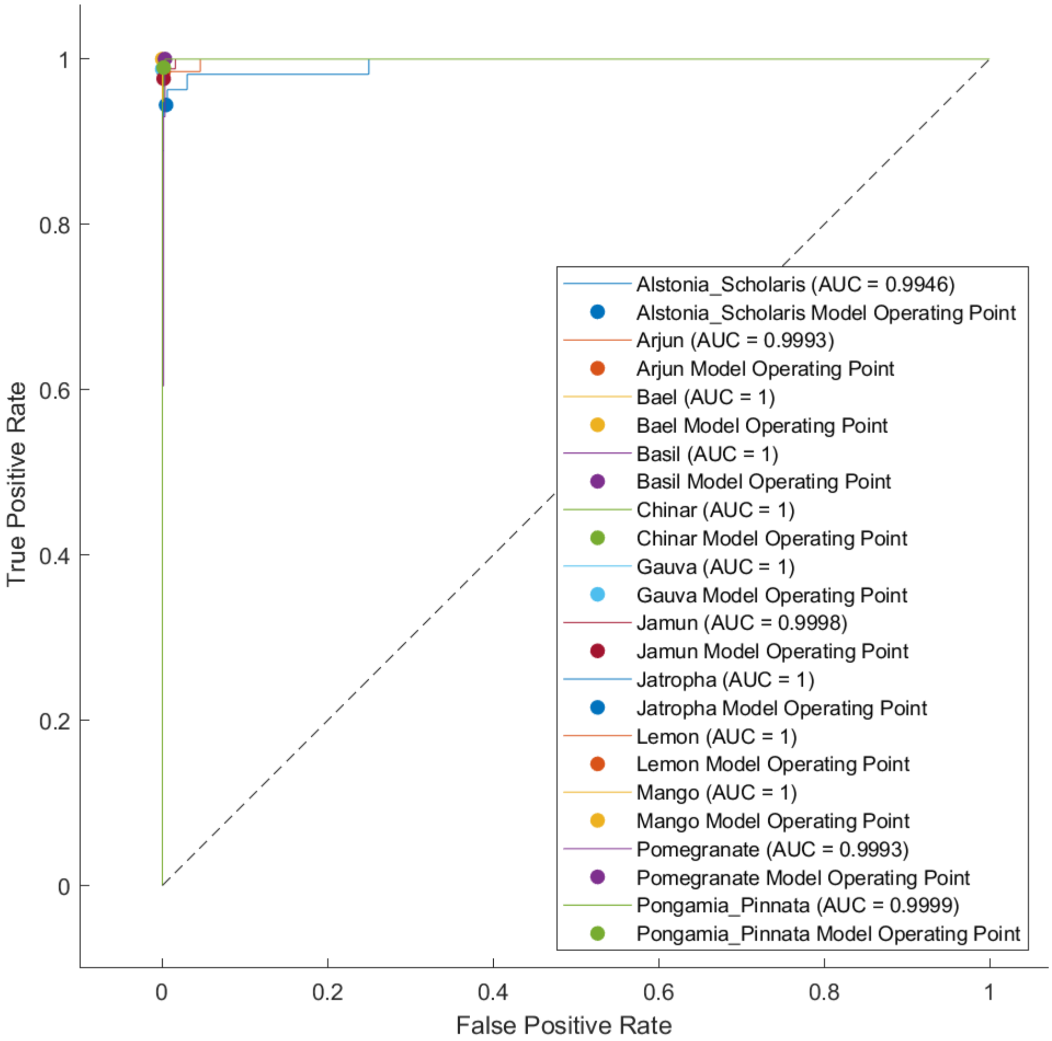

Fig. 10. ROC curve for VGGNet16 when the learning rule is SGDM, the learning rate is 0.0001, the batch size is 64, and the maximum number of epochs is 20

Figure 11 shows the training graph for the VGGNet16 network, based on the accuracy rate at each iteration. Examining this graph, it can be said that a high accuracy rate was achieved in a short amount of time.

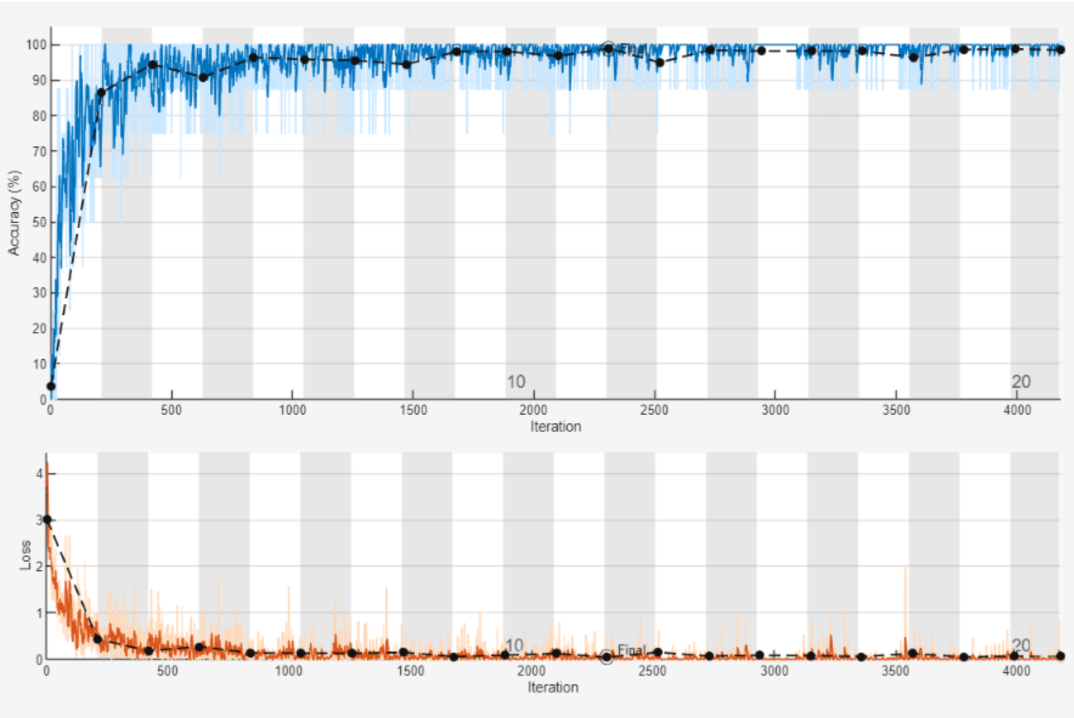

Fig. 11. Training progress graph for VGGNet16 when the learning rule is SGDM, the learning rate is 0.0001, the batch size is 64, and the maximum number of epochs is 20

IV. CONCLUSION

In this study, classification models have been developed using the VGGNet16 convolutional neural network on a dataset composed of 12 leaf image classes. This dataset has been split into a 70% training set and a 30% test set. After the convolutional neural network has been trained with the training data, the models' performance has been measured using the test set. For each model, the training and test sets have been randomly generated five times, and the performance results were calculated accordingly. The success of the models has been evaluated with the average results of these five models. The models were built using the VGGNet16 network with various training parameters. When using the SGDM learning rule, learning rate of 0.0001, batch size of 64, and 20 epochs, the VGGNet16 network achieved an F1-score of 98.8%.

For future work, our plan is to create more successful models using other CNNs as EfficientNet, DarkNet, and MobileNet, as well as through transfer learning approaches.

REFERENCES

[1] K. He, X. Zhang, S. Ren, and J. Sun, Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 770–778, 2016.

[2] D. Han, Q. Liu, and W. Fan, A new image classification method using CNN transfer learning and web data augmentation. *Expert Systems with Applications*, 95, 43-56, 2018.

[3] K. Hanbay, L. Sivrikaya, and D. Hanbay, Classification of Apricot Diseases by using Deep Convolution Neural Network. *BEU Journal of Science*, 9(1), 2020.

[4] N. Kumar, P. Belhumeur, A. Biswas, D. Jacobs, W.J. Kress, I. Lopez, and J. Soares, LNCS 7573 - Leafsnap: A Computer Vision System for Automatic Plant Species Identification, 2012.

[5] G. Altan, Performance Evaluation of Capsule Networks for Classification of Plant Leaf Diseases. *International Journal of Applied Mathematics Electronics and Computers*, 8(3), 57-63, 2020.

[6] Ai A. Kadir, L. Edi Nugroho, A. Susanto, and P. Insap Santosa, Leaf Classification Using Shape, Color, and Texture Features, *International Journal of Computer Trends and Technology*, 2011.

[7] Vg1 L.A. Gatys, A.S. Ecker, and M. Bethge, Image style transfer using convolutional neural networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2414–2423, 2016.

[8] Vg2 O. Ronneberger, P. Fischer, and T. Brox, U-Net: Convolutional networks for biomedical image segmentation. *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 234–241, 2015.

[9] A. Canziani, A. Paszke, and E. Culurciello, An analysis of deep neural network models for practical applications, 2017. arXiv. https://arxiv.org/abs/1605.07678.

[10] S. Han, J. Pool, J. Tran, and W.J. Dally, Learning both weights and connections for efficient neural networks. Advances in Neural Information Processing Systems (NIPS), 28.

[11] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, Plant disease detection using deep learning, 24th Telecommunications Forum (TELFOR), pp. 1-4, 2016.

[12] K. Simonyan and A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, *Proceedings of the International Conference on Learning Representations (ICLR)*. http://arxiv.org/abs/1409.1556, 2014.

[13] V. Nair and G. Hinton, Rectified Linear Units Improve Restricted Boltzmann Machines, 2010.

[14] S. Lee, C. Chan, S. Mayo and P. Remagnino, How deep learning extracts and learns leaf features for plant classification, *Pattern Recognition*, 71, 1-13, 2017.

[15] S. Metlek, Disease Detection From Cassava Leaf Images With Deep Learning Methods In Web Environment. *International Journal of 3D Printing Technologies and Digital Industry*, 5(3), 2021.

[16] E. Prasetyo, N. Suciati, C. Fatichah, Multi-level residual network VGGNet for fish species classification, *Journal of King Saud University Computer and Information Sciences*, 34(8), pp. 5286-5295, 2022.

[17] W. Yu, K. Yang, Y. Bai, T. Xiao, H. Yao, and Y. Rui, "Visualizing and comparing AlexNet and VGG using deconvolutional layers", *In Proceedings of the 33 rd International Conference on Machine Learning*, 3, pp. 43-76, 2016.