Proceedings of the 6th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/64/20241347

Detection and classification of wilting status in leaf images based on VGG16 with EfficientNet V3 algorithm

Qixiang Li[1], Yiming Ma[2], Ziyang Luo[3], Ying Tian[4,*]

[1]College of Information Technology, Jilin Agricultural University, Jilin, Changchun, 130118, China
[2]College of Plant protection, Jilin Agricultural University, Jilin, Changchun, 130118, China
[3]College of Resources and Environment, Jilin Agricultural University, Jilin, Changchun, 130118, China
[4]College of Information Technology, Jilin Agricultural University, Jilin, Changchun, 130118, China

[*]Corresponding author email: tianying.ai@hotmail.com

Abstract. The aim of this paper is to explore the importance of leaf wilting status detection and classification in agriculture to meet the demand for monitoring and diagnosing plant growth conditions. By comparing the performance of the traditional VGG16 image classification algorithm and the popular EfficientNet V3 algorithm in leaf image wilting status detection and classification, it is found that EfficientNet V3 has faster convergence speed and higher accuracy. As the model training process proceeds, both algorithms show a trend of gradual convergence of Loss and Accuracy and increasing accuracy. The best training results show that VGG16 reaches a minimum loss of 0.288 and a maximum accuracy of 96% at the 19th epoch, while EfficientNet V3 reaches a minimum loss of 0.331 and a maximum accuracy of 97.5% at the 20th epoch. These findings reveal that EfficientNet V3 has a better performance in leaf wilting status detection, which provides a more accurate and efficient means of plant health monitoring for agricultural production and is of great research significance.

Keywords: VGG16, Efficientnet v3, Accuracy, Classification.

1. Introduction

Classification of leaf wilting status detection is an important topic in the field of agriculture, and its research background stems from the need to monitor and diagnose plant growth conditions [1]. With the growing global population and the impact of climate change, the problem of food security is becoming more and more prominent, so it becomes imperative to improve crop yield and quality. And leaf wilting status is often one of the important indicators of plant health [2,3]. By timely and accurately detecting and classifying leaf wilting status, it can help farmers take targeted measures to adjust planting management strategies in a timely manner, so as to improve crop yields, reduce resource wastage, and realise intelligent precision agriculture.

Deep learning image algorithm plays a key role in leaf wilting state detection and classification. Traditional leaf wilting state detection methods mainly rely on manual feature extraction and machine learning algorithms for classification, but this method requires a lot of professional knowledge and experience, and is sensitive to factors such as light, angle, etc., and is easily disturbed [4,5]. In contrast, deep learning algorithms have the advantages of automatic feature extraction, strong adaptability and good generalisation ability, and have achieved great success in the field of image recognition.

In leaf wilting state detection and classification, deep learning image algorithms are usually implemented by convolutional neural networks (CNN) [6]. Firstly, a large number of well-labelled leaf image datasets are used for training so that the network can learn the feature representations of different leaf states. Then, in practical applications, the leaf images to be detected are fed into the trained CNN model, and the probability distributions of each category are calculated through forward propagation and categorised into normal or withered states [7,8].

Deep learning image algorithms have the advantages of high accuracy, automatable processing, adaptability and scalability in the detection and classification of leaf wilting status. Deep learning algorithms are able to learn complex feature representations and train on large-scale datasets, which improves the accuracy of detection and classification. Without manually extracting features, deep learning algorithms can automatically learn the optimal feature representation, simplifying the operation process. The deep learning algorithm is robust to factors such as light and angle, and can still maintain good performance under different environmental conditions. By adjusting the network structure and parameter settings, it can be adapted to the needs of different types of plants, different size scale data sets, etc.

Deep learning image algorithms play an important role in leaf wilting state detection and classification, and with the continuous progress of the technology and the improvement of the dataset, it is believed that it will have a wider application prospect in the field of agriculture. In this paper, we detect and classify the wilting state of leaf images based on the traditional VGG16 image classification algorithm and the now more popular Efficientnet v3 algorithm, and compare the advantages and disadvantages of the two deep learning algorithms in terms of prediction accuracy.

2. Data set sources and data analysis

The dataset selected in this paper is from the Kaggle public dataset, which contains three types of leaf images, namely Early Blight, Healthy and Late Blight, Early Blight contains a total of 1303 leaf images, Healthy contains a total of 816 leaf images, and Late Blight contains a total of 1132 images, and the three types of images are shown in Figure 1.

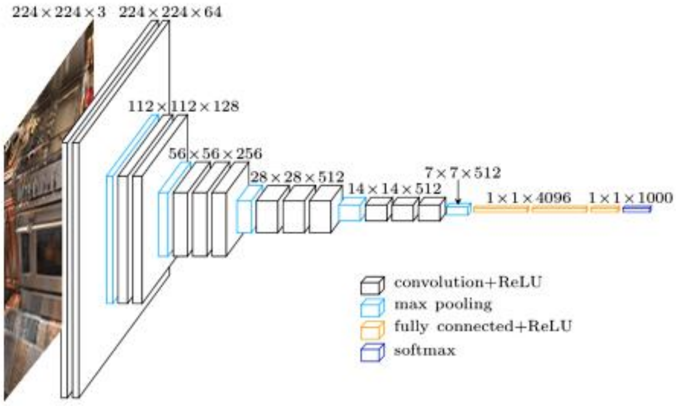
Figure 1. Partial data.
(Photo credit: Original)

3. Method

3.1. VGG16

VGG16 is a classical convolutional neural network architecture proposed by a team of researchers at the University of Oxford. The network achieved excellent results in the ImageNet Large Scale Visual Recognition Challenge in 2014, and is widely used in computer vision tasks such as image classification and target detection. The structure of VGG16 is relatively simple and deep, mainly composed of convolutional layer, pooling layer and fully connected layer. The model structure of VGG16 model is shown in Figure. 2.

Figure 2. VGG16 Structure.



(Photo credit: https://link.springer.com/article/10.1007/s42979-020-0114-9)

Firstly, VGG16 contains 13 convolutional layers and 3 fully connected layers, the overall structure is very deep and all the convolutional layers use convolutional kernels of 3x3 size and a step size of 1. This small size of convolutional kernel and smaller step size helps to increase the nonlinear expressiveness of the network and reduces the number of parameters to decrease the risk of overfitting. In addition, a pooling layer follows between every two convolutional layers, which is used to reduce the spatial dimension of the feature map and extract more abstract features [9].

Second, multiple convolutional kernels with different depths and widths are used in VGG16 to extract features. These convolutional kernels are gradually added to the network, allowing the network to gradually learn abstract feature representations from low to high levels. By stacking such convolutional blocks multiple times, VGG16 can effectively capture complex structural information in images and achieve accurate classification or detection of images [10].

In addition, a maximum pooling layer with smaller size (2x2) and larger step size (2) is used in VGG16 to reduce the feature map size. This helps the network to reduce the computational effort while maintaining an effective range of sensory fields and helps prevent overfitting. By repeatedly stacking such convolutional and pooling blocks, VGG16 can gradually reduce the feature map size and increase the number of channels, ultimately enabling efficient and accurate processing of the input image.

3.2. Eficientnet v3

EfficientNet V3 is an efficient and powerful neural network model proposed by Google, which aims to reduce the computational complexity and the number of parameters while maintaining the accuracy of the model.EfficientNet V3 combines the strengths of the EfficientNet family of models and introduces new features and tricks to further improve the performance. The network architecture of EfficientNet V3's network structure is shown in Figure 3.

DOI: 10.54254/2755-2721/64/20241347

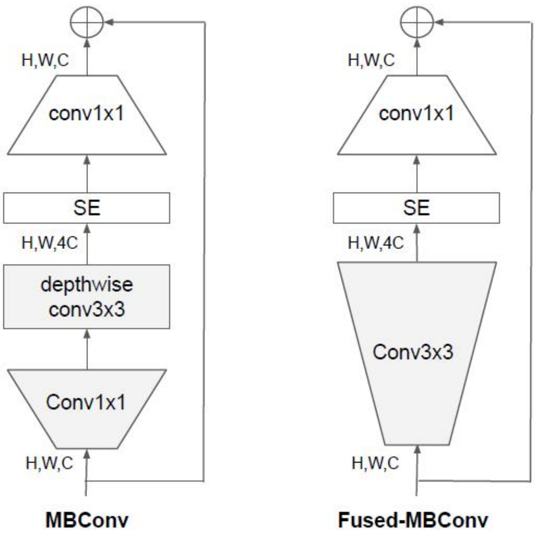


Figure 3. EfficientNet V3's network structure.
(Photo credit: Original)

Firstly, EfficientNet V3 adopts a composite scaling method similar to the EfficientNet family, i.e., constructing networks of different sizes by uniformly scaling the network depth, width and resolution. This composite scaling method can improve the model performance without adding too much computational cost, making EfficientNet V3 more versatile when dealing with different sized datasets or tasks.

Secondly, EfficientNet V3 introduces a novel attention mechanism called "Swish-Excite" to enhance the network's attention to important features. The Swish-Excite mechanism dynamically adjusts the importance of each channel in the feature map by combining the Swish activation function and the Excite gating unit to make the network pay more attention to the information that is useful for the current task and improve model performance.

In addition, EfficientNet V3 uses a batch normalisation technique called "Ghost Batch Normalization" to reduce memory consumption and computational costs. Ghost Batch Normalisation achieves parameter sharing by applying batch normalization to a smaller number of samples, thus effectively reducing memory consumption and increasing training speed while maintaining model accuracy.

Finally, in terms of network structure, EfficientNet V3 also introduces some novel modules and design ideas, such as Squeeze-and-Excitation blocks and MBConv blocks. These modules and design ideas help to enhance the network's ability for feature representation learning, information transfer and nonlinear transformation, and further optimise the overall network structure.

4. Result

In the experiment of triple classification of leaf wilting status, we will use VGG16 model and EfficientNet V3 model for comparison. The dataset is divided according to 6:4 and divided into training and validation sets. For the VGG16 model, we will use the pre-trained VGG16 as the base network, trained using the Adam optimiser under the cross-entropy loss function, with the learning rate set to 0.001, and combined with data enhancement techniques.

For the EfficientNet V3 model, we will adapt the network structure to fit the triple classification task and pair it with the Swish-Excite attention mechanism and the Ghost Batch Normalisation technique. The computer configuration used for the experiments includes Intel Core i7- 10700K CPU, NVIDIA GeForce RTX 3080 GPU, 32GB DDR4 RAM, 1TB SSD storage and Windows 10 operating system, which ensures the efficiency and stability of the training process. With the above settings, we

DOI: 10.54254/2755-2721/64/20241347

aim to compare the performance of the two models on the leaf wilting state classification task in order to select the most suitable model for the task.

The Loss and Accuracy change curves are output during the training process, and the Loss and Accuracy change curves of the VGG16 model are shown in Figure 4, and the Loss and Accuracy change curves of the EfficientNet V3 model are shown in Figure 5.

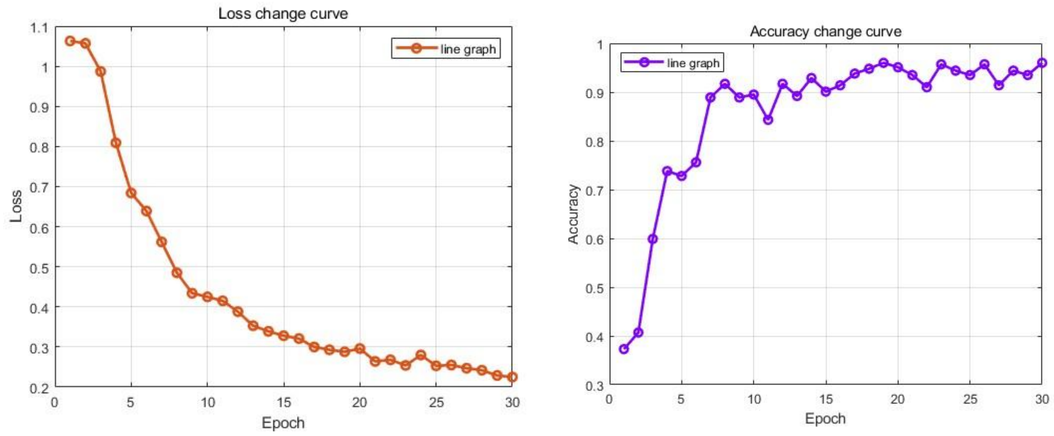


Figure 4. Loss and Accuracy change curves of the VGG16.

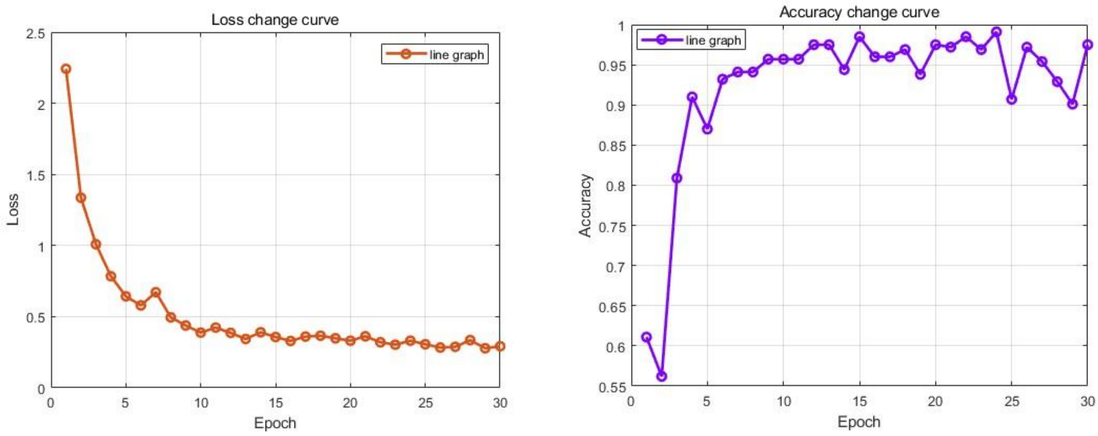


Figure 5. Loss and Accuracy change curves of the EfficientNet V3.

Table 1. Model evaluation parameter.

Model	Best epoch	Loss	Accuracy
VGG16	19	0.288	0.96
EfficientNet V3	20	0.331	0.975

From the Loss and Accuracy change curves of VGG16 and EfficientNet V3, it can be seen that both of them tend to converge in the process of model training, and the convergence speed of EfficientNet V3 is faster. At the same time, the classification accuracy of both of them on the leaf wilting state is increasing, in general, EfficientNet V3 is better than the accuracy of VGG16, the best round of VGG16 training results is the 19th epoch, and the best round of EfficientNet V3 training results is the 20th epoch, and the VGG16 loss is the lowest 0.288, and the EfficientNet V3 loss is the lowest 0.288, and the EfficientNet V3 loss is the lowest 0.288, and the EfficientNet V3 loss is the highest 0.288. and EfficientNet V3 loss is as low as 0.331, the highest accuracy of VGG16 is 96% and the highest accuracy of EfficientNet V3 is 97.5%.

DOI: 10.54254/2755-2721/64/20241347

5. Conclusion

This study is dedicated to detecting and classifying the wilting state of leaf images using the traditional VGG16 image classification algorithm and today's popular EfficientNet V3 algorithm. By comparing the performance of the two deep learning algorithms in terms of prediction accuracy, it is found that they both show a convergence trend during model training, while EfficientNet V3 exhibits a faster convergence rate. Both of them keep improving in the classification accuracy of leaf wilting state, and in general, EfficientNet V3 outperforms VGG16, in which VGG16 achieves the best training result in the 19th epoch, with the lowest Loss of 0.288 and the highest accuracy of 96%, while EfficientNet V3 achieves the best training result in the 20th epoch, with the lowest Loss of 0.288 and the highest accuracy of 96%. training results with a minimum Loss of 0.331 and a maximum accuracy of 97.5%. By comparing the performance of traditional VGG16 image classification algorithm and EfficientNet V3 algorithm on leaf wilting state detection and classification, this study provides a more accurate and efficient method for plant growth monitoring and diagnosis in the field of agriculture, which can help to improve crop yield and quality.

References

[1] Dang, Bo, et al. "Enhancing Kitchen Independence: Deep Learning-Based Object Detection for Visually Impaired Assistance." Academic Journal of Science and Technology 9.2 (2024): 180- 184.

[2] Li, Shaojie, et al. "Utilizing the LightGBM Algorithm for Operator User Credit Assessment Research." arXiv preprint arXiv:2403. 14483 (2024).

[3] Zang, Hengyi, et al. "Evaluating the social impact of ai in manufacturing: A methodological framework for ethical production." Academic Journal of Sociology and Management 2. 1 (2024): 21-25.

[4] Ma, Danqing et al. "Implementation of computer vision technology based on artificial intelligence for medical image analysis." International Journal of Computer Science and Information Technology (2023): n. pag.

[5] Ma, Danqing, Bo, Dang, Shaojie, Li, Hengyi, Zang, Xinqi, Dong. "Implementation of computer vision technology based on artificial intelligence for medical image analysis". International Journal of Computer Science and Information Technology 1. 1(2023): 69–76.

[6] Ao Xiang, Zongqing Qi, Han Wang, Qin Yang, Danqing Ma. "A Multimodal Fusion Network For Student Emotion Recognition Based on Transformer and Tensor Product." (2024).

[7] Emel S,Tuncay S .A performance comparison of YOLOv8 models for traffic sign detection in the Robotaxi-full scale autonomous vehicle competition[J].Multimedia Tools and Applications,2023,83(8):25005-25035.

[8] Mudawi A N,Qureshi M A,Abdelhaq M, et al.Vehicle Detection and Classification via YOLOv8 and Deep Belief Network over Aerial Image Sequences[J].Sustainability,2023, 15(19):

[9] Karakuş S,Kaya M,Tuncer A S.Real-Time Detection and Identification of Suspects in Forensic Imagery Using Advanced YOLOv8 Object Recognition Models[J].Traitement du Signal,2023,40(5):

[10] HyoYeon J,CheolHong H .Obscuration Threshold Database Construction of Smoke Detectors for Various Combustibles.[J].Sensors (Basel, Switzerland),2020,20(21):6272-6272.