An Improved MobileNet Model Integrated Spatial and Channel Attention Mechanisms for Tea Disease

Li Zhang, Jiacheng Sun, Minghui Yang
School of Information Engineering, Xinyang Agriculture and Forestry University, China

***Abstract*—Aiming at addressing the challenges of large model parameters, high computational cost, and low accuracy of the traditional tea disease identification model, an improved MobileNet model integrated spatial and channel attention mechanisms (MobileNet-SCA) was proposed for tea disease identification. Firstly, the tea disease identification dataset was augmented through random clipping, rotation transformation, and perspective transformation to simulate diverse image acquisition perspectives and mitigate overfitting effects. Secondly, based on the convolutional neural network (CNN) framework, the Channel Attention (CA) mechanism and Spatial Attention (SA) mechanism were introduced to carry out global average pooling and group normalization operations on input feature maps respectively, and adjust the channel weights using the learned parameters. Then the h-swish activation function was utilized to scale, and the two kinds of attention mechanisms were spliced and mixed to improve the channel and spatial information. In addition, the MobileNetV3 network's structure underwent optimization by adjusting the number of input channels, the size of the convolution kernel, and the number of channels in the residual block. The experimental results showed that the identification accuracy of MobileNet-SCA for tea diseases was 5.39% higher than the original model. This method can balance the identification accuracy and identification time well, and it meets the requirements for accurate and rapid identification of tea diseases.**

Keywords—Tea disease; MobileNetV3; attention mechanism; convolutional neural network component

I. INTRODUCTION

As an important beverage, tea has a long history, and the development of the tea industry has an important impact on the social economy. However, various tea diseases affect the yield and quality of tea, such as tea anthracnose, tea netted blister blight, tea blister blight, and tea algae leaf spot [1]. The impact of tea diseases is not limited to agricultural production but also involves the economic interests of farmers, the sustainable development of the tea industry, and the drinking experience of tea consumers [2]. To ensure the supply of tea and increase the yield, the accurate and efficient identification of tea diseases has become an urgent task.

In the actual tea growing process, most farmers rely on their accumulated agricultural skills and historical experience to identify the tea diseases. However, due to the lack of professional scientific knowledge, manual identification methods often rely on intuition to determine the type of disease, and subjective inference can lead to wrong identification results. Especially in some tea plantations in mountainous areas with steep terrain and far from urban centers, even experienced plant protection experts cannot easily reach the site[3]. In recent years, artificial intelligence has developed rapidly, and technologies such as machine learning, pattern recognition, and computer vision have produced a lot of research results in the fields of biological fermentation [4], intelligent environmental protection [5], and plant phenotype research [6], et al. Among them, in the field of plant phenotype research, artificial neural networks, support vector machines, random forest, and other computational intelligence methods are utilized for image monitoring, identification, and prevention of crop diseases, which is more efficient and faster than traditional crop disease identification methods [7][8]. Chaudhary et al. [9] successfully used the improved random forest algorithm, the attribute evaluator method, and the instance filtering method to accurately identify peanut diseases. Tetila et al. [10] compared different classifiers, including sequential minimum optimization, decision tree, and random forest to improve the performance of soybean leaf disease identification. Ehsan et al. [11] proposed a fuzzy logic identification algorithm to improve the identification efficiency of healthy and diseased strawberry leaves. The extraction of disease features is crucial for achieving high identification accuracy in plant disease identification using classical machine learning techniques such as random forest, decision tree, and support vector machine. However, the identification accuracy of traditional machine-learning methods is hindered by the small differences in color and texture commonly observed in tea diseases.

In recent years, computer vision technology has been applied more and more widely in the field of agriculture, and deep learning technology is ushering in an unprecedented rise, providing a new solution for plant disease identification [12][13]. With the continuous development of computing power, data set scale, and deep learning framework, researchers have made remarkable achievements in using deep learning technology to solve plant disease identification problems [14][15][16]. The introduction of an attention mechanism in the deep learning framework can enhance the model's attention to the disease region, reduce the interference of non-diseased regions and background on the identification results, and promote the development of deep learning in tea disease recognition. Bao et al. [17] added a channel attention mechanism module to the multi-scale feature fusion module to assign network adaptive and optimized weights to each feature mapping channel, enabling the network to select more effective features and facilitate the identification of tea diseases in natural scenes. Xue et al. [18] integrated the Convolutional and Self-attention mechanism module (ACmix) and Convolutional Block attention module (CBAM) into YOLOv5, which enabled the model to pay better attention to tea diseases and improve the accuracy of tea disease recognition. Lin et al. used the self-attention mechanism to enhance the ability to acquire global information on tea diseases and introduced the shuffle attention mechanism to solve the problem that small target tea diseases were difficult to identify, improving the identification accuracy of tea diseases [19]. These studies demonstrate the considerable advantages of incorporating attention mechanisms into deep learning frameworks for tea disease identification. However, how to ensure the accuracy of the disease identification model and realize a more intelligent, lightweight, and efficient model application under a realistic environment such as a lack of data sets, low image quality, and limited computing resources is still an open issue

Motivated by the above, to realize rapid and accurate identification of tea diseases, an improved MobileNet model integrated spatial and channel attention mechanisms (MobileNet-SCA) was proposed for tea disease identification. The major contributions of the MobileNet-SCA model are as follows.

1) The Spatial Attention (SA) mechanism is suitable for lightweight networks, the model can capture the key features of tea disease more effectively. Meanwhile, Channel Attention (CA) mechanism is added to select the most suitable channel and extract the interested information during image processing.

2) Through data set preprocessing and network fine-tuning, the model can fully cope with the challenge of using small sample data sets, enhance the generalization ability of the model, and improve the identification accuracy of the main tea diseases.

This article is organized as follows. The methodology of the MobileNetV3 model, spatial-attention mechanism, and channel-attention (CA) mechanism are presented in Section II. Section III describes the MobileNet-SCA model in detail. Section IV mainly depicts the collection and processing of the image dataset, and the proposed identification model is compared with other methods. Finally, Section V concludes the article.

II. Methodology

A. MobileNetV3 Model

To solve the problems of high complexity, many parameters and high requirements of application deployment environment of traditional models, lightweight convolutional neural network model application represented by MobileNet came into being. MobileNet series models were introduced by the Google Research team as a mobile-first computer vision model, which uses deep separable convolution to build lightweight models with low computational complexity, including MobileNetV1, MobileNetV2, and MobileNetV3 models [20]. The MobileNetV1 model is mainly composed of deep separable modules. MobileNetV2 model introduces the backward residual and linear bottleneck module, namely the bottleneck residual module. The depth-separable volume module of the MobileNetV1 model and the bottleneck residual module of the MobileNetV2 model are combined in the MobileNetV3 model [21]. Compared with MobileNetV2, some time-consuming layers are redesigned in MobileNetV3, further improving the computational efficiency of the model, and making it more practical for applications on mobile and embedded devices. Meanwhile, the SE attention module and the activation function h-swish(x) are added in MobileNetV3. The learning ability of the network model is enhanced with the SE module by learning channel feature relationships. And h-swish(x) function has strong nonlinear expression ability and progressive saturation characteristic, which show stronger expression ability in the special scene of tea disease, providing the key information in the disease image. Thus, the accuracy and training efficiency of the model is improved [22]. The MobileNetV3 module is shown in Fig. 1.

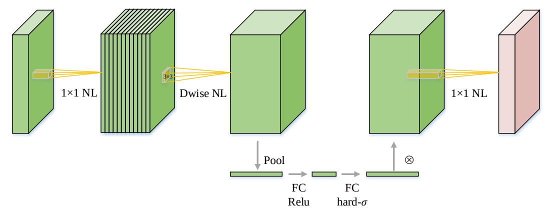

Fig. 1. MobileNetV3 module

NL - Activation function; Dwise - channel by channel convolution; Pool - pooling layer; FC - fully connected layer; Relu - Modified linear unit; hard-σ - hard saturation activation function; V - multiplication of elements

B. Spatial-Attention (SA) Mechanism

The SE attention mechanism of the MobileNetV3 model mainly focuses on internal channel information but does not consider the influence of local regional information on disease images. Due to the local characteristics of tea disease, there are no effective identification features in most leaf regions, and only a few regions with the disease can provide information conducive to the identification of tea disease. In contrast, the spatial attention mechanism performs well in helping the model focus on the local details of the image, focusing on the degree of attention to different regions in the image. Thus, the key features can be captured more effectively, and the complexity of different regions can be adapted, improving the robustness of the model in real scenes [23]. It is crucial for the identification of tea diseases as disease features are often embodied in local areas of the image. By emphasizing these key local features, the characteristics of tea diseases can be captured more accurately, and more important local features can be given greater weight, achieving more accurate and reliable processing of fine-grained identification and disease identification [24]. Combining lightweight applicability, computational efficiency, and local detail attention considerations, the introduction of spatial attention mechanisms enables the model to improve its sensitivity to key features. The spatial attention mechanism module is shown in Fig. 2.

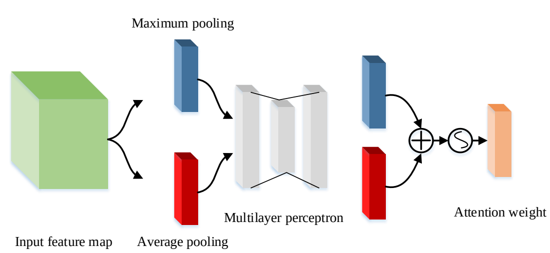

Fig. 2. Spatial attention mechanism module

Spatial attention is the compression of channels. The channel-based maximum pooling and average pooling are performed for input features. Then, the channel dimension is combined and the convolution dimension is reduced for each channel. The sigmoid function is used to generate a spatial attention diagram. The process can be represented by the following formula:

$$\begin{aligned} M_s(F) &= \sigma(f^{7\times7}([AvgPool(F); MaxPool(F)])) \\ &= \sigma(f^{7\times7}([F^s_{avg}; F^s_{\max}])) \end{aligned} \quad (1)$$

where F is the input feature map accepted by the spatial attention module, σ represents the Sigmoid function, conv represents a layer of the convolutional neural network, 7×7 represents the size of the convolutional kernel, *Avgpool(F)* is the average pooling feature, and *MaxPool(F)* is the maximum pooling feature.

C. Channel-Attention (CA) Mechanism

Although MobileNetV3 is a lightweight network, it has some shortcomings in global feature capture. Complex backgrounds may pose challenges to lightweight networks on mobile devices, and the CA channel attention mechanism can help the model better adapt to the relationship between different channels, making it more adaptable and improving its flexibility in complex scenes [25]. By introducing the CA channel attention mechanism, the most important channels for a particular task are emphasized in the model selectively, thereby improving the feature representation ability of the model. The attention to information from different channels is enhanced, selectively highlighting information channels and suppressing irrelevant channels to adaptively recalibrate the feature map. The assigning weights to each channel of the feature map can be obtained, and thus the feature representation can be optimized [26]. The channel attention module consists of a maximum pooling layer, an average pooling layer, and a 3-layer sensing set, as shown in Fig. 3.

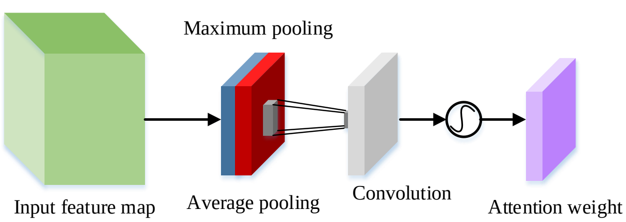

Fig. 3. Channel attention mechanism module

The input feature map is first averaged and globally pooled along the spatial dimension, and the spatial dimension is reduced to a single-channel attention vector. Then it is fed into the 3-layer perceptron, which consists of two fully connected layers. It is projected to the lower dimensional space for dimensionality reduction with the first fully connected layer, and the lower dimensional space is mapped back to the original channel dimension with the second fully connected layer. Thus, the two C×1×1 attention vectors can be generated, and each channel of the feature graph is assigned a weight according to its importance, where C represents the number of channels. Finally, the channel attention diagram is obtained by weighted summation. The calculating channel attention M_C can be written as follows:

$$M_c(F) = \sigma(\mathrm{MLP}(AvgPool(F)) + \mathrm{MLP}(MaxPool(F))) \quad (2)$$

where F is the input feature map accepted by the channel attention module, σ represents the Sigmoid function, MLP represents the 3-layer perceptron, the activation function is ReLU, *Avgpool(F)* represents the average pooling feature, and *MaxPool(F)* represents the maximum pooling feature.

III. MobileNet-SCA Model

While the MobileNetV3 model has significant advantages in lightweight network design, there is still room for improvement, particularly in global feature capture. The SA and CA mechanisms each possess unique advantages, with SA excelling in spatial image analysis, focusing on local details, while CA aids the model in better adapting to inter-channel relationships, thereby enhancing flexibility in complex scenarios. In this study, the MobileNet-SCA model is proposed for tea disease identification. The CA and SA mechanisms are introduced by the sa_layer class. After the sa_layer class is embedded in each bottleneck structure, the input features are divided into two parts in the forward propagation process of the model. Then the CA mechanism and SA mechanism are spliced together and the final output is obtained through a channel mixing operation. The spatial space and channel information are fully paid attention to, which improves the perception of key features, the performance, and the robustness of the model. The architecture of the MobileNet-SCA model is shown in Fig. 4.

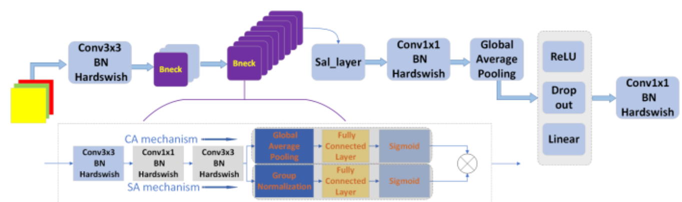

Fig. 4. The architecture of the MobileNet-SCA model

Based on the MobileNetV3 model, a BottleNeck structure is adopted, which includes 3x3 and 5x5 convolution operations to improve the ability to abstract image features. In addition, a shuffle attention layer is introduced between every two adjacent bottlenecks to increase the attention to the spatial information. The CA and SA mechanisms are introduced for each bottleneck structure, and the weights of different channels are adjusted to make the model pay more flexible attention to the most important channel information for a specific task. This helps to improve the feature expression ability and identification accuracy. The SA and CA mechanisms work together on different parts of the model, emphasizing the importance of image space and channel information respectively. These two attention mechanisms performed global average pooling and group normalization operations on the input feature maps, respectively. Then, the h-swish activation function is applied to scale the activations, and the two attention mechanisms are seamlessly combined to enhance both channel and spatial information. Moreover, the architecture of the MobileNetV3 model is optimized by fine-tuning the number of input channels, convolution kernel size, and channel count within the residual blocks. This synergistic effect helps to improve the identification performance of image details, making the model more adaptable to identifying tea disease under complex natural scenes.

IV. Results And Discussion

To verify the effectiveness of the proposed MobileNet-SCA model, experiments are conducted on a computer with the deep learning framework Pytorch, AMD Ryzen5 4600H processor, NVIDIA GeForce GTX1650 12GB, and the running memory was 16GB. The operating system was Windows 11 and Python3.11.5 was used in the integrated development environment Anaconda.

A. Dataset

The tea disease image data used in this study was provided by Professor Jiang Zhaohui's research group at Anhui Agricultural University, which was collected in the natural environment by a single-lens reflex camera and pre-processed by image processing software[27]. After professional and technical personnel screened out 1827 images of tea disease, the original images of the tea anthracnose, tea netted blister blight, tea blister blight, and tea algae leaf spot are shown in Fig. 5(a), (b), (c), and (d).

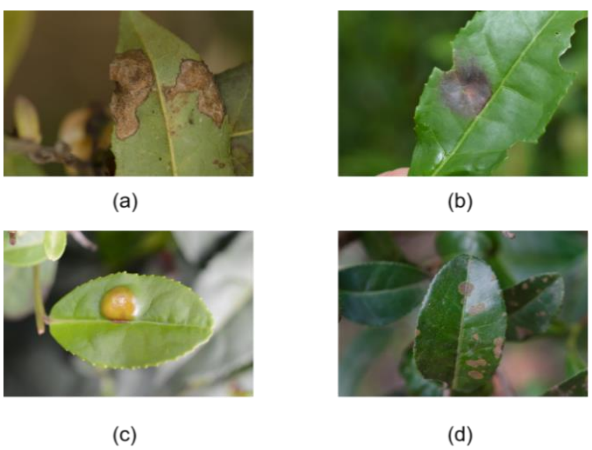

Fig. 5. Original images of tea diseases, (a) tea anthracnose; (b) tea netted blister blight; (c) tea blister blight; (d) tea algae leaf spot

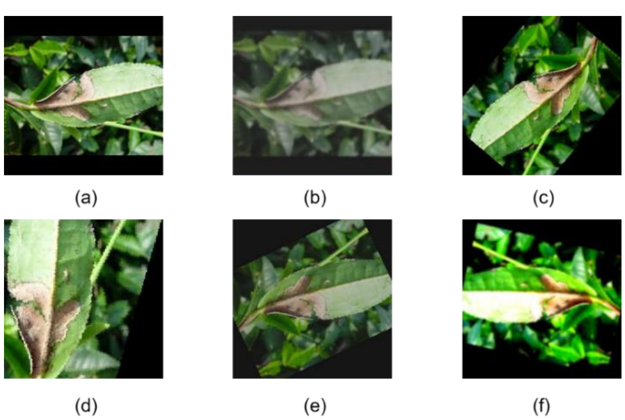

Fig. 6. Examples of image augmentation, (a) Original image of tea anthracnose disease; (b) contrast adjustment, color transformation; (c) rotation, contrast adjustment; (d) rotation scaling, brightness adjustment; (e) rotation, color change; (f) Gaussian blur, brightness adjustment

TABLE I. Tea Disease Image Data

Types of disease	Original training set	Extended training set	Training Set	Validation Set	Test Set	Total
tea anthracnose	359	461	749	71	63	883
tea netted blister blight	424	566	906	84	80	1070
tea blister blight	280	332	556	56	53	665
tea algae leaf spot	263	384	595	52	42	689

To better evaluate and optimize the performance of the model, 80% of the disease images were selected as the training data set, 10% as the validation data set, and the rest as the test data set. The images of the training data set were scaled and cropped, and two random operations in rotation, brightness adjustment, contrast adjustment, Gaussian blur, color transformation, and other processing methods were used for data pre-processing of the cropped pictures to generate an extended training set, as shown in Fig. 6(a), (b), (c), (d), (e), and (f). After the low-quality images such as overexposed and blurred images were manually screened, the remaining images were used as the training set by merging the augmented training set and the original training set. Finally, the total data set is expanded to 3307 tea disease images, and the image attribute of the extended data set is adjusted to 256×256 pixels by using the normalization method. The tea disease image data are shown in Table I.

B. Comparison Experiments

In addition to the method proposed in this study, ResNet50, ResNet18, MobileNetV3, MobileNetV2, and MobileNet-SCA were trained. The training was carried out using the augmented tea disease image dataset.

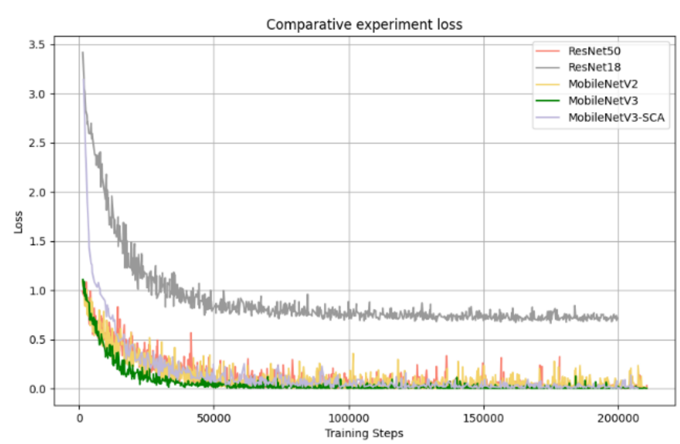

Fig. 7. Loss change curves of 5 kinds of neural networks on the training set

Fig. 7 shows loss change curves of five kinds of neural networks on the training set with 200 iterations. In the initial iteration time of training, four networks converge at similar speeds. With the increase in iteration times, the MobileNetV3 network converges first, followed by MobileNetV2, MobilenetV3-Sal, ResNet50, and ResNet18 networks. The convergence speed of the MobileNetV3-SCA model is relatively faster. After 200 iterations, the accuracy of the MobileNetV3-SCA model is higher than other networks. The accuracy change curves of five neural networks on the validation set are shown in Fig. 8, and the accuracy of all the models reaches more than 80%. Combined with the loss curve, it can be found that the convergence speed of the MobileNet-SCA model can quickly reach the stable convergence result of training, obtaining the highest identification accuracy, which is at least 5% higher than the MobileNetV3 model.

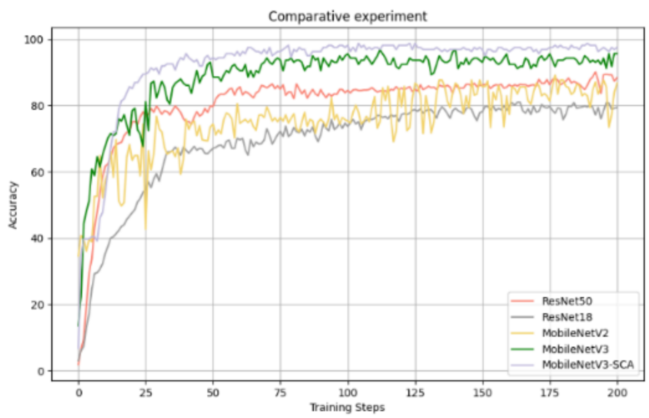

Fig. 8. Accuracy change curves of five neural networks on the validation set

In addition, model size, training time, and accuracy are used as evaluation indicators to evaluate the performance of the proposed identification model. Model size is the number of model parameters and is considered a factor to consider when deploying a model on a mobile device to ensure it fits in a resource-limited environment. The training time is the time it takes for the training to complete, and the testing time is the time it takes for the model to infer a new data set after training. As an important evaluation index to measure the identification task, accuracy refers to the number of successfully identified samples divided by the number of all samples. Suppose that the total number of samples in the dataset is N, the number of samples $P(P \leq N)$ is randomly selected each time for testing, and the number of samples identifying the correct category of the model is $T(T \leq P)$, then the identification accuracy in this identification task is shown as follows:

$$Accuracy = \frac{T}{P} \tag{3}$$

The above formula is the ratio of the number of samples identifying the correct category to the number of samples extracted each time. The higher the accuracy rate, the better the model identification performance.

The identification performance with different models for tea disease is shown in Table II. Compared with lightweight neural networks MobileNetV3, MobileNetV2, and deep neural networks ResNet50 and ResNet18, the MobileNetV3 model can achieve the highest accuracy. The ResNet variant model takes longer training time and requires more computing resources, which is not suitable for tea disease identification. The MobileNet model variant is relatively balanced in terms of model size, training time, and testing time. The model size is moderate, the training time and testing time are relatively shorter, and the identification accuracy is higher.

TABLE II. COMPARISON OF IDENTIFICATION PERFORMANCE WITH DIFFERENT MODELS FOR TEA DISEASE

Model selection	Accuracy	Size(Mb)	Training time completed(h)	Testing time(s)
ResNet50	90.4%	4.3 x10 '	14	1.57
ResNet18	81.32%	2.8 x10 '	11	1.12
MobileNetV2	89.27%	1.7 x10 '	4.6	1.19
MobileNetV3	93%	1.99 x10 '	8	1.21
MobileNetV3-SCA	98.39%	2.3 x10 '	8.7	1.25

C. Ablation Experiments

To show the influence of different modules on the identification performance more intuitively, we present the results of the ablation experiments based on MobileNetV3, shown in Table III.

Compared with the original MobileNetV3 model, the accuracy of MobileNet-SA is improved by 2.31%, which verifies the validity of the SA module of MobileNet-SA. After replacing the original attention mechanism with the CA attention mechanism, the identification accuracy was improved by 2.97%. Compared with the basic lightweight network MobileNetV3, MobileNet-CA achieved better results in tea disease identification. It can also be seen that the accuracy of tea disease identification of the MobileNet-SCA model is 98.39%, which is higher than that of the model with or without any method. However, due to the complexity of the parameters and structure of the MobileNet-SCA model, the training time is longer than the benchmark model MobileNetV3, but better identification performance can be achieved. The time to identify all samples of tea diseases is only about 1.25s, and the parameter number is only slightly higher than that of MobileNetNetV3, which is 2.3 Mb, which is also suitable for the final deployment and identification efficiency on mobile devices.

TABLE III. RESULTS OF ABLATION EXPERIMENTS BASED ON MOBILENETV3

Model selection	Size(Mb)	Accuracy	Testing time(s)
MobileNetV3	1.99 x10 '	93.00%	1.21
MobileNetV3-SA	2.04 x10 '	95.31%	1.16
MobileNetV3-CA	2.16 x10 '	95.97%	1.29
MobileNetV3-SCA	2.3 x10 '	98.39%	1.25

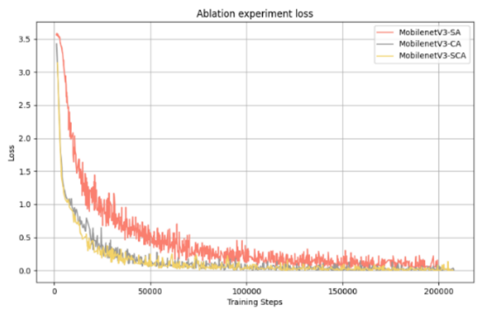

Fig. 9. Loss curves of three MobileNet models combined with different attention mechanisms on the training set

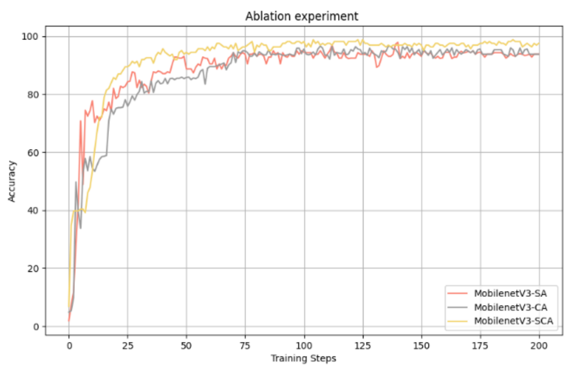

Fig. 10. Accuracy curves of three MobileNet models combined with different attention mechanisms on the training set

In addition, AdamW is adopted as the optimization algorithm in the training process of the MobileNet-SCA model, which can effectively control the complexity of the model by attenuating the weights, thus improving the generalization ability of the model. The loss curves and accuracy curves of three MobileNet models combined with different attention mechanisms on the training set are shown in Fig. 9 and Fig. 10. It can be seen that the MobileNet-SCA model can converge quickly and has better training identification accuracy.

V. CONCLUSION

Tea disease image data have different attributes such as shape, color, and size, the training categories are unbalanced and the amount of training data is insufficient. It is difficult to learn a general and reliable identification model with traditional machine learning methods. The MobileNet-SCA model was proposed for tea disease identification integrated spatial and channel attention mechanisms, which make the modified network pay full attention to spatial space and channel information. The structure of the MobileNetV3 network was fine-tuned with the number of input channels, the size of the convolution kernel, and the number of channels of the residual block. The perception ability of key features and the characteristics of fast convergence and strong portability can be obtained. Compared with other MobileNetV3 models, the identification accuracy of MobileNet-SCA is increased by at least 2.42% with a small model structure, which reduces the need for storage capacity and computing power of the device. This approach significantly contributes to the precise identification of tea leaf disease, laying a solid foundation for rapid, accurate detection and effective disease prevention.

In addition, due to the few tea disease types and images, in our future study, the knowledge learned by the MobileNet-SCA model on other image data sets will be transferred to the task of tea disease recognition with the help of transfer learning, to realize the efficient recognition of tea diseases and provide scientific guidance for the prevention and control of tea diseases

ACKNOWLEDGMENT

This work was supported by the Fund for Henan Scientific and Technological Research Project (222102210300).

We are thankful to Professor Zhaohui Jiang of Anhui Agricultural University for providing the tea disease data.

REFERENCES

[1] Y. K. Wang, R. J. Xu, D. Bai, and H. F. Lin, “Integrated learning-based pest and disease detection method for tea leaves,” Forests, vol. 14, no. 5, p. 1012, 2023.

[2] J. Chen, Q. Liu, and L. W. Gao, “Visual tea leaf disease recognition using a convolutional neural network model,” Symmetry (Basel)., vol. 11, no. 3, p. 343, 2019.

[3] G. S. Hu, H. Y. Wu, Y. Zhang, and M. Z. Wan, “A low shot learning method for tea leaf’s disease identification,” Comput. Electron. Agric., vol. 163, p. 104852, 2019.

[4] X. F. Yuan, L. Li, and Y. L. Wang, “Nonlinear dynamic soft sensor modeling with supervised long short-term memory network,” IEEE Trans. Ind. Informatics, vol. 16, no. 5, pp. 3168–3176, May 2020.

[5] J. F. Qiao, Z. Q. Hu, and W. J. Li, “Soft measurement modeling based on chaos theory for biochemical oxygen demand (BOD),” Water, vol. 8, no. 12, p. 581, 2016.

[6] A. K. Singh, B. Ganapathysubramanian, S. Sarkar, and A. Singh, “Deep learning for plant stress phenotyping: trends and future perspectives,” Trends Plant Sci., vol. 23, no. 10, pp. 883–898, 2018.

[7] M. K. R. Gavhale and U. Gawande, “An overview of the research on plant leaves disease detection using image processing techniques,” J. Comput. Eng., vol. 16, no. 1, pp. 10–16, 2014.

[8] T. Domingues, T. Brandão, and J. C. Ferreira, “Machine learning for detection and prediction of crop diseases and pests: A comprehensive survey,” Agriculture, vol. 12, no. 9, p. 1350, 2022.

[9] A. Chaudhary, S. Kolhe, and R. Kamal, “An improved random forest classifier for multi-class classification,” Inf. Process. Agric., vol. 3, no. 4, pp. 215–222, 2016.

[10] E. Castelao Tetila, B. Brandoli Machado, N. A. D. S. Belete, D. A. Guimaraes, and H. Pistori, “Identification of soybean foliar diseases using unmanned aerial vehicle images,” IEEE Geosci. Remote Sens. Lett., vol. 14, no. 12, pp. 2190–2194, 2017.

[11] E. Kiani and T. Mamedov, “Identification of plant disease infection using soft-computing: Application to modern botany,” in Procedia Computer Science, 2017, vol. 120, pp. 893–900.

[12] G. S. Liu, J. L. Peng, and A. A. A. El-Latif, “SK-MobileNet: A Lightweight Adaptive Network Based on Complex Deep Transfer Learning for Plant Disease Recognition,” Arab. J. Sci. Eng., vol. 48, no. 2, pp. 1661–1675, 2023.

[13] X. H. Zhang, Y. Qiao, F. F. Meng, C. G. Fan, and M. M. Zhang, “Identification of maize leaf diseases using improved deep convolutional neural networks,” IEEE Access, vol. 6, pp. 30370–30377, 2018.

[14] K. P. Ferentinos, “Deep learning models for plant disease detection and diagnosis,” Comput. Electron. Agric., vol. 145, pp. 311–318, 2018.

[15] Y. Lu, S. J. Yi, N. Y. Zeng, Y. R. Liu, and Y. Zhang, “Identification of rice diseases using deep convolutional neural networks,” Neurocomputing, vol. 267, pp. 378–384, 2017.

[16] J. He, T. Liu, L. J. Li, Y. H. Hu, and G. X. Zhou, “MFaster R-CNN for Maize Leaf Diseases Detection Based on Machine Vision,” Arab. J. Sci. Eng., vol. 48, no. 2, pp. 1437–1449, 2023.

[17] W. X. Bao, T. Fan, G. S. Hu, D. Liang, and H. D. Li, “Detection and identification of tea leaf diseases based on AX-RetinaNet,” Sci. Rep., vol. 21, no. 1, p. 2183, 2022.

[18] Z. Y. Xue, R. J. Xu, D. Bai, and H. F. Lin, “YOLO-tea: A tea disease detection model improved by YOLOv5,” forests, vol. 14, no. 2, p. 415, 2023.

[19] J. Lin, D. Bai, R. J. Xu, and H. F. Lin, “TSBA-YOLO: An improved tea diseases detection model based on attention mechanisms and feature fusion,” Forests, vol. 14, no. 3, p. 619, 2023.

[20] E. Elfatimi, R. Eryigit, and L. Elfatimi, “Beans leaf diseases classification using mobilenet models,” IEEE Access, vol. 10, pp. 9471–9482, 2022.

[21] L. Q. Jia et al., “MobileNet-CA-YOLO: An improved YOLOv7 based on the MobileNetV3 and attention mechanism for Rice pests and diseases detection,” Agriculture, vol. 13, no. 7, p. 1285, 2023.

[22] X. R. Liang, J. F. Liang, T. Yin, and X. Y. Tang, “A lightweight method for face expression recognition based on improved MobileNetV3,” IET Image Process., vol. 17, no. 8, pp. 2375–2384, 2023.

[23] C. F. Chen, D. H. Gong, H. Wang, Z. F. Li, and K. Y. K. Wong, “Learning spatial attention for face super-resolution,” IEEE Trans. Image Process., vol. 30, pp. 1219–1231, 2021.

[24] Z. Q. Lv, W. D. Wang, Z. Q. Xu, K. H. Zhang, Y. H. Fan, and Y. Song, “Fine-grained object detection method using attention mechanism and its application in coal-gangue detection,” Appl. Soft Comput., vol. 113, p. 107891, 2021.

[25] R. Y. Chen, H. X. Qi, Y. Liang, and M. C. Yang, “Identification of plant leaf diseases by deep learning based on channel attention and channel pruning,” Front. Plant Sci., vol. 13, p. 1023515, 2022.

[26] X. T. Yang, Y. C. Luo, M. Li, Z. K. Yang, C. H. Sun, and W. Y. Li, “Recognizing pests in field-based images by combining spatial and channel attention mechanism,” IEEE Access, vol. 9, pp. 162448–162458, 2021.

[27] Y. Y. Sun, Z. H. Jiang, L. P. Zhang, W. Dong, and Y. Rao, “SLIC_SVM based leaf diseases saliency map extraction of tea plant,” Comput. Electron. Agric., vol. 157, pp. 102–109, 2019.