DETECTION METHOD OF TOMATO LEAF DISEASES BASED ON IMPROVED ATTENTION MECHANISM

/

基于改进注意力机制的番茄叶部病害检测方法研究

Jiapeng QU[1)], Dong XU[1)], Xiaohui HU[2)], Ruihong TAN[1)], Guotian HU*[1)]

[1)]College of Mechanical and Electronic Engineering, Northwest A&F University, Yang ling 712100/ China;

[2)]College of Horticulture, Northwest A&F University, Yang ling 712100/ China;

[3)]Xi'an Aerospace Propulsion Test Technique Institute, Xi an 710000/ China;

Tel: +86 13038595639; E-mail: qujiapengdyx@163.com

DOI: https://doi.org/10.35633/inmateh-70-59

Keywords: *Tomato diseases; MobileNet-V2; Mechanism of attention; Transfer learning*

ABSTRACT

The precise detection and recognition are the premise in accurate prevention and control of tomato diseases. To improve the accuracy of tomato diseases recognition model, nine kinds of sick leaves images including tomato target spot bacteria in Plant Village and healthy leaves images were used. A new attention mechanism module called CBAM-II was created by changing the serial connection between Channel and Spatial attentions of CBAM to parallel connection, and then the results of two modules were added together. CBAM-II had been verified to be effective and universal in the convolutional neural network model. The accuracy of MobileNet-V2 with CBAM-II model was 99.47%，which had increased by 1.13%, 0.93%, 0.7%8 and 1.06 % respectively comparing with MobileNet-V2 model, MobileNet-V2 plus Channel attention module, MobileNet-V2 plus Spatial attention module, and CBAM attention module. Furthermore, the accuracy of AlexNet, Inception-V3 and ResNet50 model has increased 1.73, 0.15 and 0.33 % respectively when the CBAM-II module was added. Results showed that the proposed module CBAM-II created in this experiment is more effective in MobileNet-V2 model for tomato diseases recognition, and could solve interference problems resulted from the serial connection. Additionally, the accuracy of four convolutional neural network models including Mobilenet-V2, AlexNet, Inception-V3 and ResNet50 model had all increased when the CBAM-II module was added, which represented the good universality of CBAM-II module. The results could provide technical support in accurate detection and control of tomato diseases.

摘要

番茄病害准确检测与识别是番茄病害精准化防治的前提。为提高番茄病害识别模型准确率，本文以 Plant Village 中番茄斑点病等 9 类病害叶片及健康叶片图像为研究对象，将 CBAM 的 Channel attention 和 Spatial attention 模块由串行连接变为并行连接并把运算结果相加，提出了一种新的注意力机制模块 CBAM-II，并验证该模块在卷积神经网络模型中的有效性和通用性。在 MobileNet-V2 模型加入 CBAM-II模块后，模型准确率达到了99.47%，准确率较MobileNet-V2 原模型、加入Channel attention 模块、Spatial attention 模块以及CBAM 注意力模块的 MobileNet-V2 模型分别提升了1.13、0.93、0.78、1.06 个百分点；在 AlexNet、Inception-V3 和 ResNet50 加入 CBAM-II模块后，识别准确率较原模型分别提升了 1.73、0.15 和 0.33 个百分点。研究表明，提出的 CBAM-II在番茄病害识别模型 MobileNet-V2 中更有效，且可解决 CBAM 在番茄病害识别过程中串行连接而引起的干扰问题；加入 CBAM-II的 MobileNet-V2、AlexNet 等 4 种卷积神经网络模型较原模型的准确率均有提高，表明通用性好。本文研究可为番茄病害精准检测和防治提供技术支持。

INTRODUCTION

Tomato is one of the most important economical crops due to high nutritional value, high yields, easiness to cultivate and manage, and wide planting in the world (*Zhou et al., 2018*). However, tomato is easy to be infected with diseases during its growth period and most of diseases start from leaves, then spread to other parts of tomato plant, which can cause huge economic losses if it is not properly protected. Therefore,

[1] *Jiapeng, Qu, M.D. Stud. Eng.; Dong, Xu, M.D. Stud. Eng.; Ruihong Tan,Lab Tech.; Guotian Hu, Assoc. Prof. M.D. Eng.* [2]*Xiaohui Hu,Prof. P.H Eng.&Sci.*

rapid and accurate identification of tomato leaf diseases is an important way to prevent diseases spread and ensure quality of tomato.

Currently, there are two main methods of tomato diseases detection, namely manual detection and deep learning techniques represented by Convolutional Neural Networks (*Sun et al., 2020*). The manual method of leaf diseases detection is dependent on the experience of experts or vegetable farmers, which is not only time-consuming and labor-intensive, but also lacks rigorous scientific evidence (*Wang et al., 2016*). In recent years, the automatic recognition of crop diseases pictures based on image processing technology has gradually achieved intelligent diagnosis of diseases. *Guo et al. (2019)* proposed a convolutional neural network model for multi-scale detection, which can alleviate the problem of sparse image data to a certain extent, and then developed a tomato leaf diseases image recognition system using this model. *Wang et al. (2019)* incorporated the idea of migration learning into AlexNet and studied the recognition of various tomato leaf diseases. This model of AlexNet combined with migration learning could effectively improve recognition of tomato diseases. *Xie and He (2017)* identified tomato disease species by statistically extracting differences in features such as color and shape of different diseases of tomato leaves. But this method relied heavily on the extraction of color features, which caused highly misdiagnosed to some diseases due to their similar disease symptoms. *Wu (2019)* proposed a tomato leaf diseases recognition method based on a deep residual network model, which could solve the problem of degraded performance of over-deep disease recognition network models and achieved 95% accuracy for four common tomato leaf diseases. *Fuentes et al. (2017)* captured pictures using some different resolution cameras based on deep learning methods of Faster R-CNN, VGG and SSD, the experimental results showed that detection system could effectively identify nine kinds of tomato diseases. *Brahimi et al. (2017)* verified the superior performance of deep convolutional networks compared with shallow networks, and confirmed that using training weights to initialize the model parameters could effectively improve the model performance.

Due to the differences in color and shape characteristics of different diseases of tomato leaves, and the similarity in appearance between healthy leaves and those leaves with Septoria leaf spot, or leaf mold, or Mosaic virus, it is easy to confuse them. Therefore, manual feature extraction of diseases makes it is impossible to accurately identify tomato diseases in practice because of the variable shape of tomato disease locations and complex backgrounds. But as a deep learning model, convolutional neural networks through multiple convolutional calculations can learn and express features that are difficult to be extracted by the naked eyes, and can perform various computer vision tasks with high quality (*Luo and Wang., 2021*). Especially, migration learning can solve the problem of insufficient training data and can apply model parameters that trained based on large datasets to new models *(Wang Z. P., et al., 2021*), and speed up the training of models and then improve the accuracy of diseases recognition (*Niu et al., 2022*). However, the recognition accuracy of the convolutional neural network models based on migration learning in tomato leaf diseases recognition task needs to be further improved, and Convolutional Block Attention Module (CBAM) in the tomato leaf diseases recognition models of MobileNet-V2 has interference problems caused by serial connection between spatial attention and channel attention module (*Zhang et al., 2021*).

In this experiment, the nine types of tomato diseases images from the Plant Village dataset were used as the research object. Firstly, comparing the disease recognition effects of four convolutional neural networks based on the full network layer fine-tuning including MobileNet-V2, AlexNet, Inception-V3 and ResNet50, the model that meets the requirements of subsequent experiments was selected. Secondly, a new module named CBAM-II was developed by improving the structure of the conventional convolutional attention module named CBAM, and the effectiveness & generality of CBAM-II were verified which solved the interference problem caused by CBAM.

MATERIALS AND METHODS

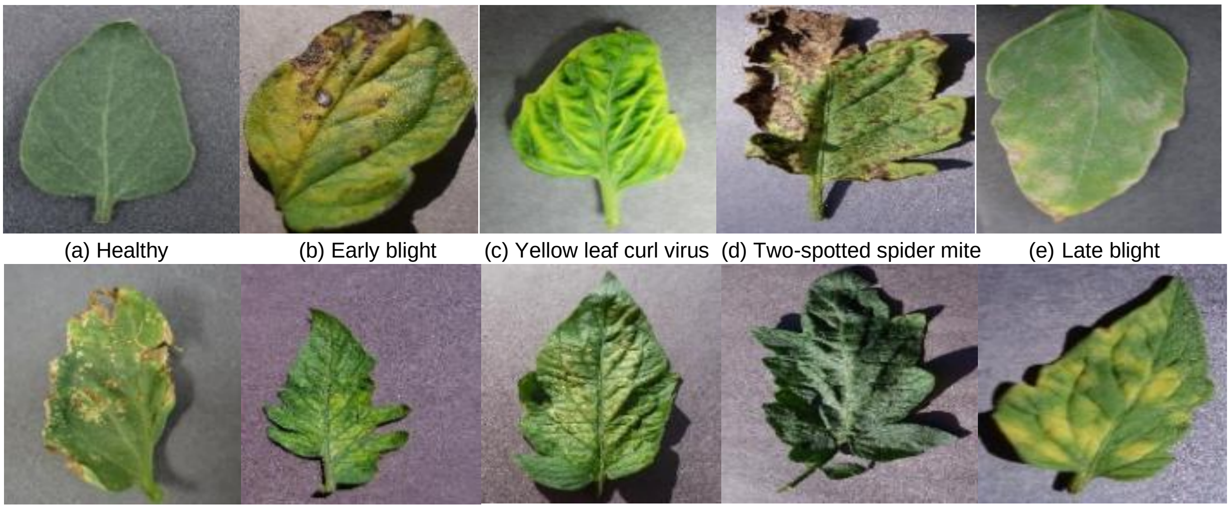
Data Description

The basic samples of tomato leaf disease images used in this paper were obtained from the Plant Village dataset, which contains images of nine types of leaf diseases and one type of healthy leaves. Nine diseases include early blight, yellow leaf curl virus, two-spotted spider mite, late blight, target spot bacteria, mosaic virus, Septoria leaf spot, bacterial spot, leaf mold. The sample set was enhanced and expanded to 17,103 disease images, the details of which are shown in table 1. The sample set was divided into training set, validation set and test set in the ratio of 7:2:1. Figure 1 shows images of healthy tomato leaves and 9 types of diseased leaves.

Table 1

Details of the tomato datasets

Types of diseases	Number of images	
	Before augmentation	After augmentation
Healthy	854	1708
Early blight	826	1652
Yellow leaf curl virus	282	1692
Two-spotted spider mite	487	1948
Late blight	621	1863
Target spot bacteria	704	1408
Mosaic virus	270	1620
Septoria leaf spot	352	1760
Bacterial spot	758	1516
Leaf mold	968	1936
Total	6122	17103



(a) Healthy (b) Early blight (c) Yellow leaf curl virus (d) Two-spotted spider mite (e) Late blight

(f) Target spot bacteria (g) Mosaic virus (h) Septoria leaf spot (i) Bacterial spot (j) Leaf mold

Fig. 1 - Image of healthy and 9 diseases tomato leaves

Experimental platform

The computer operating system used in this paper was Windows 64-bit system, and the computer hardware mainly included 2×16 GB of RAM, 1 TB solid state disk (SSD) and 2 TB mechanical hard disk, an Inter i512600KF processor, NVIDIA GeForce3070 (8G) graphics card. The application software environment is CUDA 11.1.0 (Anaconda). Using Python language to program, the model building and training based on Tensorflow 2.4.0 deep learning framework.

Parameter settings

In this paper, we use multiple batches for training. Parameters were set as the Batch size to 32, Epochs to 60, and initial learning rate named Learning rate to 0.001. Optimizer used was Adam optimizer, and the loss function was SoftMax cross entropy loss function, the function is defined as:

$$L = -\sum_{k=1}^{n}\sum_{i=1}^{C} t_{ki} lg y_{ki} \quad (1)$$

n – pixels in a single image;

t_{ki} – probability that pixel point *k* belongs to category *i*;

y_{ki} – probability that a classification network will predict pixel point *k* as category *i*

Data preprocess

Considering that the image size should match the network input size, the image needs to be processed uniformly to 224 x 224 pixels and saved in .jpg format. The images were resized using the function of tf.image.resize_images provided in the Google TensorFlow deep learning framework.

Practically, many factors such as complex growth environments, variable backgrounds, weather changes and capturing angles can affect image quality. Therefore, this paper used python language to expand the sample size. This was done by enhancing or reducing the brightness of the original images by 20% to simulate different weather conditions. In addition, the images were flipped horizontally, flipped vertically and rotated to simulate changes in camera angle.

All images were normalized for converting the pixel values from 0-255 to 0-1, to facilitate the image processing by the convolutional neural network. In this paper, the maximum-minimum normalization pair was used and the equation is:

$$norm = \frac{x_i - min\,(x)}{max(x) - min\,(x)} \tag{2}$$

x_i – pixel point value of the image;

max() – maximum function;

min() – minimum function

MobileNetNet-V2 model

MobileNet-V2 is a lightweight neural network proposed by google team. Its model structure parameters are shown in Table 2 (*Qin., 2021*). MobileNet-V2 introduces Linear Bottleneck based on V1, and also proposes the concept of inverted residual (Zhao and Xu., 2021). MobileNet-V2 was chosen as the preferred model in this paper due to the fact that MobileNet-V2 has higher accuracy and smaller model comparing with MobileNet-V1. As shown in figure 2, the channels of the feature map are expanded by a 1 × 1 point-by-point convolution operation to enrich the number of features and thus improve the accuracy. The 3×3 convolution is then replaced by a 3×3 DW convolution.

The MobileNet-V2 network constructed in this way retained the diversity of feature information and also improved the representational capability of the network *(Zhao et al., 2022*).

$$\longrightarrow \begin{matrix}1\times1\\ \text{PW}\end{matrix} \xrightarrow{\text{ReLu6}} \begin{matrix}3\times3\\ \text{DW}\end{matrix} \xrightarrow{\text{ReLu6}} \begin{matrix}1\times1\\ \text{PW}\end{matrix} \xrightarrow{\text{Linera}}$$

Fig. 2 - Inverted residual processing

Table 2

MobileNet-v2 network structure

Number of layers	Input Size	Operator	Number of repetitions	Stride
1	224×224×3	Conv2d	1	2
2	112×112×32	Bottleneck	1	1
3-4	112×112×16	Bottleneck	2	2
5-7	56×56×24	Bottleneck	3	2
8-11	28×28×32	Bottleneck	4	1
12-14	28×28×64	Bottleneck	3	2
15-17	14×14×96	Bottleneck	3	2
18	7×7×160	Bottleneck	1	1
19	7×7×320	Conv2d	1	1
20	7×7×1280	Avgpool	1	-
21	1×1×1280	Conv2d	1	-

CBAM attention mechanism

CBAM is a lightweight convolutional attention module proposed in 2018, which combined channel and spatial attention mechanism modules to perform attention in both spatial and channel dimensions (Woo et al., 2018). As shown in figure 3, CBAM consisted of two parts of Channel Attention Module (CAM) and Spatial Attention Module (SAM), which performed attention on the channel and space respectively. The overall process of CBAM attention mechanism was that the input feature map passed through the CAM and then the weighted results were sent to the SAM for weighting to obtain the final feature map.

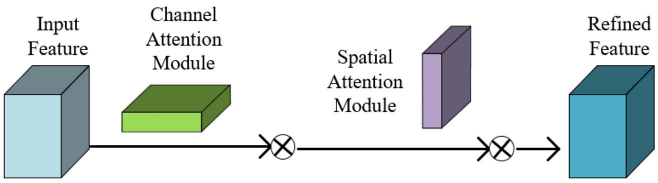


Fig. 3 - CBAM structure

Channel Attention Module

In convolutional neural networks, channel information generally represented information about different features of an image, therefore, the network could pay more attentive to the information useful to the task by selecting channels (*Wang M. H., et al., 2021*). The principle of the channel attention module was to compress the spatial dimension while keeping the channel dimension constant. To achieve channels selection, the global average pooling (Avgpool) and global maximum pooling (Maxpool) information of each feature map was calculated separately using equation (3), and then both global average and maximum information passed through the full connection layer and were added together to obtain channel attention parameters. As shown in figure 4, both Maxpool and Avgpool information shared the same full connection network, and the process was:

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))$$
$$= \sigma(W_1\left(W_0(F^c_{avg})\right) + W_1(W_1(F^c_{max})) \tag{3}$$

σ – sigmoid functions;

F^c_{avg} $and F^c_{max}$ – use maximum pooling and average pooling in the spatial dimension for the feature map *F*;

$W_0 \in R^{C/r \times C}$; $W_1 \in R^{C \times C/r}$

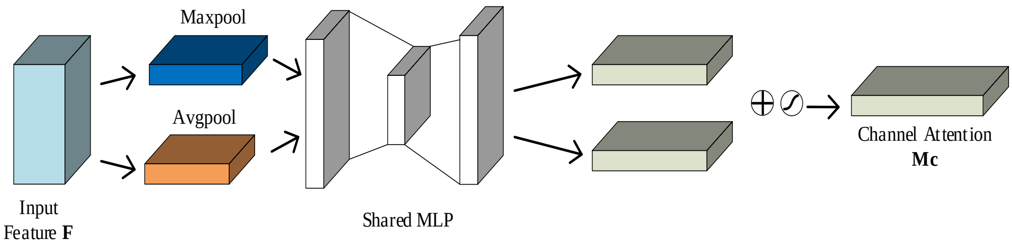


Fig. 4 - Structure of Channel Attention Module

Spatial Attention Module

To improve the model recognition accuracy and robustness, the spatial attention module was used(Fu et al., 2017). The principle of the spatial attention module was to keep the spatial dimension constant and compress the channel dimension which was contrary to the channel attention module. Two feature maps were obtained using equation (4) to global maximum pooling and global average pooling for each coordinate of every channel feature map respectively, and then the spatial attention map was obtained by convolving the feature maps. The structure of spatial attention module is shown in figure 5.

$$M_s(F) = \sigma\left(f^{7\times7}([AvgPool(F); MaxPool(F)])\right)$$
$$= \sigma f^{7\times7}([F^s_{avg}; F^s_{max}])) \tag{4}$$

σ – sigmoid functions;

$f^{7\times7}$ – 7×7 convolution kernel

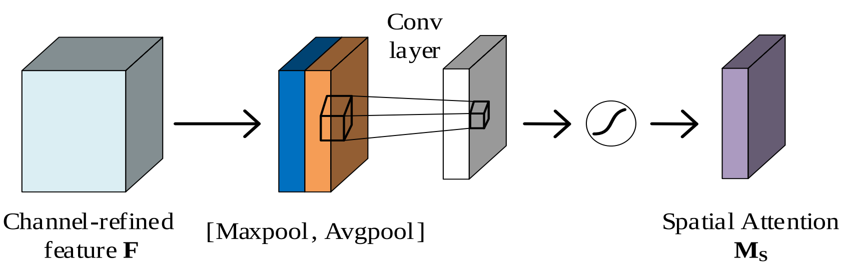


Fig. 5 - Structure of spatial attention module

CBAM improvement

CBAM as a plug-and-play module is itself a serially connection structure, as shown in Figure 3. However, whether the channel attention module or the spatial attention module was in front, the next module weight was generated from the feature maps output by the module in front, which could affect the effectiveness of the latter module in learning features to some extent. Especially in classification tasks, the interference caused by serial linking was highly likely to make the effect of the constructed model unstable, which led to the lower recognition accuracy than expected.

Therefore, a new channel attention module named CBAM-II which based on CBAM was invented aimed to solve above problems in this paper. The key difference between CBAM-II and CBAM was changing the original serial connection structure into a parallel connection and then summation. The parallel connection didn't care about the order of the spatial attention module and channel attention module, and both channel attention module and spatial attention module could learn the features of the original input directly, therefore, parallel connection could avoid interaction between them. The structure of CBAM-II is shown in figure 6. The input feature map *F* was firstly fed into parallel connection of channel attention module and spatial attention to obtain the corresponding weights respectively, and then the two weights were weighted with the original input feature map *F* to obtain F_1 and F_2, which were added together to obtain the final output feature map F_O. The process is shown in Equation (5).

$$\begin{aligned} F_1 &= M_c(F)\otimes F \\ F_2 &= M_s(F)\otimes F \\ F_o &= F_1 + F_2 \end{aligned} \tag{5}$$

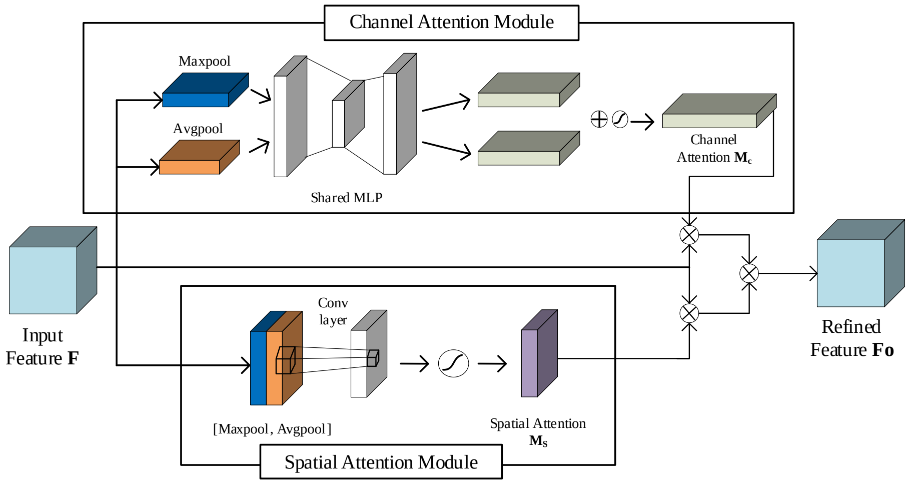


Fig. 6 - Structure Diagram of CBAM - II

Evaluation indicators

Accuracy, Recall and the number of model parameters were used as evaluation indicators in order to objectively evaluate the performance of the model in this paper.

(1) Accuracy: proportion of the number of correctly predicted samples to the total sample, the formula is:

$$Acc = \frac{TP+TN}{TP+TN+FN+FP} \tag{6}$$

TP、FP、TN、FN – number of true positives, false positives, true negatives and false negatives respectively;

(2) Recall: proportion of positive samples among those samples which predicted to be positive samples, the formula is:

$$Recall = \frac{TP}{TP+FN} \tag{7}$$

(3) Number of model parameters: model classification accuracy is an important indicator to evaluate the overall performance of the model, but for the model to be ported to mobile tasks, it also needs to meet the requirements of smaller size and lower memory consumption.

RESULTS AND DISCUSSION

Experimental analysis results of different models

For the tomato leaf diseases recognition task, a MobileNet-V2 network model combined with migration learning was built using the TensorFlow framework. In order to verify the performance of this model, InceptionV3, AlexNet, ResNet50 network models combined with migration learning were used to recognize tomato leaf diseases images. The test accuracy is shown in figure 7, and the detailed indicators are shown in table 3.

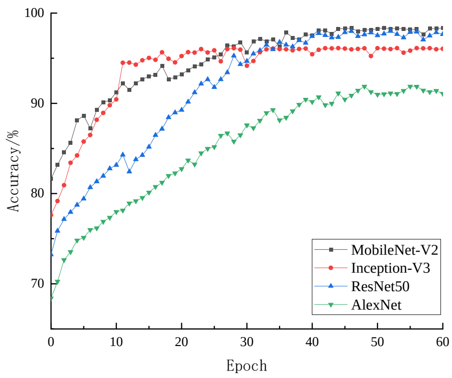


Fig. 7 - Comparison of different model accuracy

Table 3

Results of different model experiments

Model	Accuracy/ (%)	Recall/ (%)	Parameters	Model size /MB
AlexNet	91.39	91.46	60.9×10^6	93.4
Inception-V3	96.12	96.03	24.7×10^6	101.0
ResNet50	98.01	98.07	25.5×10^6	270.06
MobileNet-V2	98.34	98.35	2.34×10^6	27.8

Figure 7 and table 3 show that the model of combined migration learning had high accuracy from the start of training and converged quickly because they already had parameters trained on the large ImageNet dataset and did not need to update them. The ResNet50 model gradually increased in accuracy as the number of iterations increased, and the curve converged when epoch reached 30, with an accuracy of 98.01% and a model size of 270.06 MB. The MobileNet-V2 model accuracy was 98.34%, which showed MobileNet-V2 achieved better results in tomato leaf disease identification than the other three methods due to its network structure and migration learning. The number of MobileNet-V2 parameters was only 2.34×10^6 and the model size was 27.8 MB, which were better than the other three models. The results showed that MobileNet-V2 was more accurate than the other three models. Therefore, the MobileNet-V2 model is an ideal model for subsequent experiments.

Comparison of the attention modules embedded in MobileNet-V2

Considering that the future model needs to be deployed on mobile terminal and the large size model is not suitable for embedding, the best performing MobileNet-V2 model was selected to implant the CBAM-II attention module for comparison experiments to verify the recognition effect of CBAM-II. In the above experiments in which migration learning was employed, the parameter weights trained on ImageNet were used. If the attention module was added to each inverted residual module of MobileNet-V2, it would corrupt the pre-trained model parameters. Therefore, to ensure the integrity of the pre-trained model parameters, comparison experiments were conducted after adding channel attention, spatial attention, CBAM and CBAM-II to the last convolutional layer of the MobileNet-V2 model respectively to verify the recognition effect of CBAM-II on the lightweight model. The experimental results are shown in figure 8, and the indicators are shown in table 4.

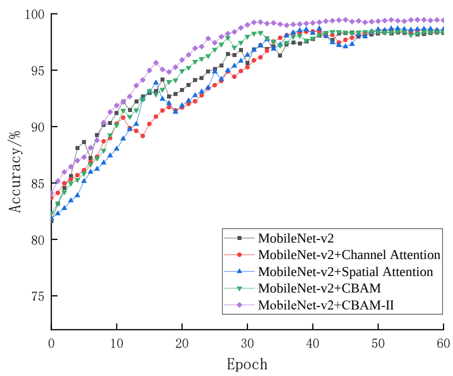


Fig. 8 - MobileNet-V2 Embedded Attention Module Accuracy Comparison

Table 4

Results of different attention module experiments in MobileNet-V2

Model	Accuracy /（%）	Recall /（%）	Parameters	Model size /MB
MobileNet-V2	98.34	98.35	2.34×10^6	27.8
MobileNet-V2+channel attention	98.54	98.54	3.95×10^6	27.8
MobileNet-V2+spatial attention	98.69	98.57	2.33×10^6	27.4
MobileNet-V2+CBAM	98.41	98.44	3.95×10^6	27.7
MobileNet-V2+CBAM-II	99.47	99.49	3.95×10^6	27.9

Compared with other types of attention modules, the accuracy of MobileNet-V2 with CBAM-II model improved significantly with an improvement of 1.13, 0.93, 0.78 and 1.06 percentage points over the original MobileNet-V2 model, the MobileNet-V2 model with the channel attention module, the spatial attention module and the CBAM attention module respectively, which indicated MobileNet-V2+CBAM-II had best results in the tomato leaf diseases identification task. The accuracy of adding CBAM was lower than that of adding the channel attention module or spatial attention module alone, which further indicated that serial connection could cause interfere between the channel attention and spatial attention. The results showed that CBAM-II could effectively improve the recognition accuracy of the MobileNet-V2 model. In addition, the size of the model after adding CBAM-II was only 27.9 MB, which was almost unchanged from that of the MobileNet-V2 model embedded in CBAM, it could provide the conditions for subsequent implantation of the model into mobile terminal.

Experimental comparison of different neural networks with CBAM-II

To further verify the generality of CBAM-II, CBAM-II was added to the last convolutional layer of Inception-V3, AlexNet and ResNet50 respectively, and the experimental results are shown in figure 9 and table 5.

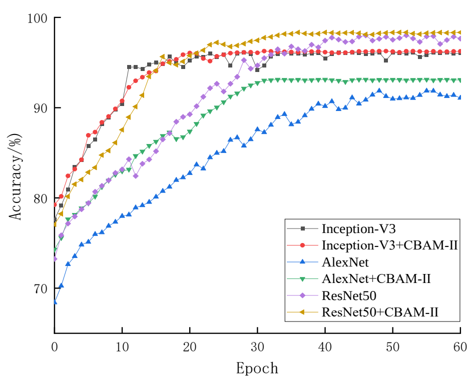


Fig. 9 - Comparison of results for different networks with CBAM-II

Table 5

Results of adding CBAM-II to different neural networks

Model	Accuracy / (%)	Recall / (%)	Parameters	Model size /MB
AlexNet+CBAM-II	93.12	93.17	12.4×10^7	191.2
Inception-V3+CBAM-II	96.27	96.24	51.1×10^7	204.4
ResNet50+CBAM-II	98.34	98.34	54.7×10^7	542.2

As table 3 and table 5 showed, both the number of parameters and model size doubled after CBAM-II module was added to AlexNet, Inception-V3 and ResNet50 respectively, which was not ideal for the task of model deployment on mobile terminal. But the model recognition accuracy of the three convolutional neural networks added to CBAM-II were more accurate than the original networks by 1.73, 0.15 and 0.33 percentage points respectively, and the training process was smoother, which indicated that CBAM-II had some generality in improving the model accuracy.

Comparison with other model results

Currently, there are many research methods for detecting tomato leaf diseases using image processing techniques. To further make an objective evaluation of the experimental results, the results of the research algorithms in this paper were compared with those existing studies of the same type, as shown in table 6.

Table 6

Comparison of recognition accuracy of different methods

Model	Datasets	Number of diseases	Accuracy / (%)	Parameters	Model size /MB
Model 1 (Zhuang., 2021)	Plant Village	10	98.60	2.28×10^7	-
Model 2 (Fang and Shi., 2020)	Plant Village	8	98.58	-	19.0
Model 3 (Durmus et al., 2017)	Plant Village	10	97.22	-	-
Model 4 (Wang S. H., 2021)	Plant Village	5	96.36	-	-
This model	Plant Village	10	99.47	3.95×10^6	27.9

The dataset used both in this paper and in literatures (table 6) were from Plant Village datasets. Zhuang (2021) used model 1 with a larger number of parameters； Fang and Shi. (2020) used model 2 with high accuracy and a model size of 19 MB, which could be used for tomato leaf diseases detection tasks; Durmus et al. (2017) used model 3 with an accuracy of 97.22%, but the model was over-fitted; Wang S. H. (2021) was able to achieve a high recognition rate using model 4, but styles of tomato leaf diseases were small and the generalizability needs further study. The model used in this paper achieved 99.47% recognition accuracy, which was 0.87%, 0.89%, 2.25%, and 3.11% higher than the models used in the above literatures, respectively. Furthermore, our model had the characteristics of MobileNet-V2 network with deep separation and convolution, the model size was only 27.9 MB, and it also had the advantages of migration learning to suppress the overfitting of small sample data and reduce the training time. The effectiveness of our model was fully proved through comparative analysis.

CONCLUSIONS

To solve the problem of low accuracy of traditional tomato leaf diseases identification, this paper conducted a study using the Plant Village tomato leaf diseases dataset as the research object. Firstly, four convolutional neural networks such as MobileNet-V2 combined with migration learning were selected for experimental comparison, and then MobileNetNet-V2 was selected as the subsequent research model after several validations of model performances. Secondly, a CBAM-II attention module with parallel connection structure is proposed to solve the problem of mutual interference between spatial attention and channel attention modules caused by the serial connection of traditional CBAM. Finally, to verify the effectiveness of CBAM-II, channel attention, spatial attention, CBAM and CBAM-II were added into MobileNet-V2 model respectively for comparison of recognition effects; to verify the generality of CBAM-II, CBAM-II modules were added into AlexNet, Inception-V3 and ResNet50 for comparison with the original model in terms of accuracy and other evaluation indicators. The results showed that the MobileNet-V2 model with the CBAM-II module could improve the recognition accuracy by 0.93, 0.78 and 1.06 percentage points, respectively, compared with the model with other attention modules. Three models with the addition of CBAM-II including AlexNet, Inception-V3, and ResNet50 could improve the accuracy by 1.73%, 0.15% and 0.33% respectively, compared with the three original models. This indicates that CBAM-II has better recognition effect in tomato leaf diseases identification task and is generalized among the four convolutional neural network models. Further research needs to collect different tomato leaf diseases images to validate the effectiveness of the CBAM-II module and to further optimize the model structure to reduce the model size and provide technical support for the development of portable instruments.

ACKNOWLEDGEMENT

The work in this paper was supported by China Agriculture Research System (CARS-23-D06) and Scientific & Technological Innovative Research Team of Shaanxi Province (2021TD-34).

REFERENCES

[1] Brahimi M., Boukhalfa K., Moussaoui A., (2018). Deep learning for tomato diseases: classification and symptoms visualization. *Applied Artificial Intelligence*, 31, 299-315.

[2] Durmus H., Gunes E. O., Kirci M., (2017). Disease detection on the leaves of the tomato plants by using deep learning. *International Conference on Agro-Geoinformatics*.

[3] Fu J., Zheng H., Mei T., (2017), Look closer to see better: Recurrent attention convolutional neural network for fine-grained image recognition. Proceedings - *30th IEEE Conference on Computer Vision and Pattern Recognition*, 4476-4484.

[4] Fuentes A., Yoon S., Kim S. C., Park D. S., (2017), A Robust Deep-Learning-Based Detector for Real-Time Tomato Plant Diseases and Pests Recognition. *Sensors*, 17.

[5] Fang C. C., Shi F. H., (2020), Image recognition of tomato diseases based on improved deep residual network (基于改进深度残差网络的番茄病害图像识别). C*omputer application*, 40(S1), 203-208.

[6] Guo X. Q., Fan T. J., Shu X., (2019), Image recognition of tomato leaf diseases based on improved Multi-Scale AlexNet (基于改进 Multi-Scale AlexNet 的番茄叶部病害图像识别). *Transactions of the CSAE*, 35(13), 162-169.

[7] Luo R. M., Wang W. X., (2021), A review on the identification of plant diseases and insect pests based on convolutional neural network (基于卷积神经网络的植物病虫害识别研究综述). *Automation and Information Engineering,* 42(05): 1-10.

[8] Niu X. D., Gao B. P., Nan X. Y., SHI Y. F., (2022), Tomato leaf disease detection based on improved DenseNet convolutional neural network (基于改进 DenseNet 卷积神经网络的番茄叶片病害检测). *Jiangsu Agricultural Journal*, 38(01), 129-134.

[9] Qin J. Q., (2022), A MobileNet-based method for identification of crop leaf diseases (基于 Mobilenet 的农作物叶片病害识别方法). *Information and Computers* (Theory Edition), 33(18), 181-184.

[10] Sun H., Li S., Li M. Z., Liu H. J., Qiao L., (2020), Research Progress in Agricultural Information Imaging Perception and Deep Learning Applications (农业信息成像感知与深度学习应用研究进展). *Transactions of the Chinese Society for Agricultural Machinery,* 51(5), 1-17.

[11] Woo S., Park J. C., Lee J. Y., Kweon I. S., (2018), CBAM: Convolutional block attention module. arXiv.

[12] Wang M. H., Wu Z. X., Zhou Z. G., (2021), Fine-grained Identification research of crop pests and disease based on improved CBAM via attention (基于注意力改进 CBAM 的农作物病虫害细粒度识别研究). *Journal of Agricultural Machinery*, 52(04), 239-247.

[13] Wang S. H., (2021), *Research on classification of tomato leaf diseases based on convolutional neural network* (基于卷积神经网络的番茄叶片病害分类研究). Jilin Agricultural University.

[14] Wang X. Y., Wen H. J., Li X. X., Fu Z. T., Lu X. J., (2016), Analysis of research progress in detection and early warning technology of major agricultural diseases (农业主要病害检测与预警技术研究进展分析). *Transactions of the Chinese Society for Agricultural Machinery*, 47(9), 266-277.

[15] Wang Y. L., Zhang H. L., Liu Q. F., Zhang Y. S., (2019), Image classification of tomato leaf disease based on transfer learning (基于迁移学习的番茄叶片病害图像分类). *Journal of China Agricultural University*, 24(06), 124-130.

[16] Wang Z. P., Zhang M., Dong W., Zhu J. B., Kong J. J., Qian R., (2021), Research on multi-model rice disease identification method based on transfer learning (基于迁移学习的多模型水稻病害识别方法研究). *Anhui Agricultural Sciences*, 49(20), 236-242.
[17] Wu H. R., (2019), Tomato leaf disease recognition method based on deep residual network (基于深度残差网络的番茄叶片病害识别方法). *Wisdom agriculture*, 1(04), 42-49.
[18] Xie C. Q., He Y., (2017), Identification of early blight disease degree of tomato leaf using spectral reflectance (利用光谱反射特性对番茄叶片早疫病害程度的识别). *China Sciencepaper*, 12(06), 671-675.
[19] Zhang N., Wu H. R., Han X., Miu Y. S., (2021), Tomato disease recognition method based on multi-scale and attention mechanism (基于多尺度和注意力机制的番茄病害识别方法). *Acta Agriculturae Zhejiangensis*, 33(07), 1329-1338.
[20] Zhao S., Liu Z. L., Zheng A. L., Gao Y., (2022), Real-time classification and detection of SSD garbage based on MobileNetV2 and IFPN (基于 MobileNetV2 和 IFPN 改进的 SSD 垃圾实时分类检测方法). *Computer application*, 42(S1), 106-111.
[21] Zhao Y., Xu J., (2021), Face recognition system based on MobileNetV2 and Raspberry Pi (基于 MobileNetV2 与树莓派的人脸识别系统). *Computer system applications*, 30(08), 67-72.
[22] Zhou R., Zhao T. M., Zhao L. P., Wang Y. L., Song L. X., Yu W. G., (2018), High gloss tomato breeding and fruit surface glossiness (高光泽番茄育种及果实表面光泽度). *Jiangsu Journal of Agriculture*, 34(6), 1437-1440.
[23] Zhuang Z. F., (2021), *Research on tomato leaf disease identification based on convolutional neural network* (基于卷积神经网络的番茄叶片病害识别研究). Anhui Agricultural University.