IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 16 January 2023, accepted 1 February 2023, date of publication 13 February 2023, date of current version 23 February 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3244386

RESEARCH ARTICLE

A Lightweight Attention-Based Convolutional Neural Networks for Fresh-Cut Flower Classification

YEQI FEI[1,2], ZHENYE LI[2], TINGTING ZHU[2], AND CHAO NI[2]

[1]School of Intelligent Manufacturing, Nanjing University of Science and Technology ZiJin College, Nanjing 210046, China
[2]College of Mechanical and Electronic Engineering, Nanjing Forestry University, Nanjing, Jiangsu 210037, China

Corresponding authors: Tingting Zhu (tingtingzhu@njfu.edu.cn) and Chao Ni (chaoni@njfu.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 62006120 and Grant 31570714.

ABSTRACT In the process of classifying fresh-cut flowers, the classification accuracy of the algorithm plays a vital role in the control of quality stability, uniformity, and price of fresh-cut flowers, while the classification speed of an algorithm determines the possibility of industrial application. Currently, research on fresh-cut flower classification focuses on the breakthrough of classification accuracy, ignoring the real-time processing speed of the terminal, which seriously affects the use of fresh-cut flower online classification technology. In this study, RGB images and depth information data of 434 rose flowers were collected using a binocular stereo depth camera. Combined with the actual production line classification environment, a set of data argumentation solutions was developed under the condition of limited samples. The architecture was established and optimized based on the ShuffleNet V2 network backbone unit, transfer learning was performed, and an appropriate attention mechanism was invoked to classify flowers of five specifications. The experimental results showed that the proposed network structure had a competitive advantage in terms of parameter quantity, classification speed, and accuracy compared with traditional networks without an attention mechanism and other lightweight networks. The classification accuracy on the 3-channel (RGB channel) flower dataset and the 4-channel (RGB and depth channel) flower datasets were 98.891% and 99.915%, respectively, and the overall prediction classification speed can reach 0.020 seconds per flower. Compared to the fresh-cut flower classification machines currently on the market, the speed of the proposed method has a great advantage. These advantages are of great significance for the design and development of fresh-cut flower classification and grading systems, and the proposed method is instructive for the identification and application of multichannel data in the future.

INDEX TERMS Fresh-cut flower grading, lightweight network, attention mechanism, image augmentation, model parameter optimization.

I. INTRODUCTION

With the improvement of people's quality of life and the continuous expansion of the fresh-cut flower market, fresh-cut flowers must be subdivided according to quality to meet different market needs and enhance the value and market competitiveness of fresh-cut flowers. The classification and grading of different specifications of fresh-cut flowers in terms of type, quality, and length are widely used in the post-harvest processing of flowers. The original classification of fresh-cut flowers is done manually [1], but owing to human subjectivity, visual fatigue, experience differences, and classification efficiency, it is difficult to meet the requirements of standardized production of fresh-cut flowers and cannot guarantee the efficiency, consistency, and stability of grade classification and quality evaluation of fresh-cut flowers. To effectively reduce the labor intensity of workers, manual intervention, and the loss of post-harvest processing of fresh-cut flowers, achieve high efficiency and quality, and meet the demands of automatic processing and classification line of

The associate editor coordinating the review of this manuscript and approving it for publication was Amin Zehtabian.

fresh-cut flowers, it is necessary to use automatic processing and classification systems [2], [3], [4].

In recent years, with the development of artificial intelligence (AI), machine learning and deep learning techniques have been successfully applied to flower species recognition and classification. Earlier studies relied more on manual segmentation and feature selection with traditional machine learning algorithms [5], [6]. Due to the large number of flower categories and the large variability of flowers themselves, the classification accuracy of each category varies depending on the acquisition environment and means, classification criteria, and the type and number of categories, the classification robustness cannot be guaranteed. To obtain a higher and more stable classification accuracy, the classical CNN architecture combined with the transfer learning algorithm has made great progress in flower recognition applications. The common method is to use VGG, Inception, ResNet, and other classic architectures for transfer learning [7], [8], [9], which significantly improves the classification accuracy compared to traditional machine learning algorithms [10], and the classification accuracy can reach 98.5% on Oxford 17 dataset [7], but there is also much room for improvement. Most classical algorithm architectures and their improvements in recent years generally obtain deeper image feature information by adding parallel branches and layers to the network. For example, CNN was proposed as a feature extractor and used in combination with other machine learning classification algorithms to synthesize algorithms with more classification advantages, such as VGG, Alexnet, and DenseNet, used as feature extractors and SVM, RF, etc., used as feature classification tools [11], [12]. The highest classification accuracy of these methods is up to 99.8% [13]. From many studies and experiments, it was found that parallel branch and layer expansion can significantly improve recognition accuracy [14], [15], but it tends to lead to an increase in the number of parameters, which undoubtedly causes the burden of memory usage and hardware computing power, resulting in slow recognition processing speed [16]. In most experiments, the number of parameters and time of identification were not mentioned; therefore, the applicability of these methods to industrial applications is uncertain.

According to statistics, the post-harvest processing loss of fresh-cut flowers can reach 31.88%, of which the grading loss accounts for 21.74% [17]. Owing to the perishability and vulnerability of fresh-cut flowers, processing efficiency is an important factor in controlling their quality. Traditional deep learning methods have a large number of parameters, long running time, and high requirements for equipment performance and are not feasible in the actual recognition process of classification production lines. Considering the limitations of processing time, quantity of parameters, classification speed, and hardware computing power of the production line, MobileNet series [18], [19], [20], ShuffleNet series [16], [21], SqueezeNet [22], and other lightweight network architectures have attracted increasing attention in the field of object recognition and classification. In this application, transfer learning [23], [24], [25] and fine-tuning [26], [27] were performed on a lightweight network to obtain a more satisfactory recognition accuracy with fewer parameters. However, the above experiments were not tested in an actual production line, and no specific quantification and comparison of the detection time were illustrated, which cannot accurately reflect the processing speed advantage of the lightweight network architecture in practical applications. In the same situation, the lightweight network was more prone to losing some recognition accuracy than the classical CNN network architecture [18].

The attention mechanism can capture more useful features according to the classification goals and requirements without changing the existing network architecture, and the valid and important information are given a higher weight than irrelevant and interference information to improve the classification performance. Qin et al. [28] used a pre-training model based on VGG, adding an attention mechanism before the linear layer, and obtained an 87.6% classification accuracy on the Oxford-102 flower dataset. Zhang et al. [29] fine-tuned the Xception architecture and used the spatial attention and channel attention modules in combination with the residual module, and realized more accurate discrimination of classification categories on the basis of light weight and a low number of parameters. The accuracy was 97.35% for Oxford102. The classification accuracy improved in the experiments.

To meet market demand and improve production efficiency, it was necessary to design a system that could adapt to the production needs of the flower production line. This system would provide a competitive advantage in terms of the number of parameters, classification speed and accuracy compared with traditional networks without an attention mechanism and other lightweight networks.

Aiming at an automatic flower classification line, we designed a set of image data augmentation schemes with Yunnan rose as the acquisition object, proposed an improved algorithm framework based on ShuffleNet, and designed a classification method with a superior attention mechanism to improve the classification speed and accuracy of 3-channel image data and 4-channel image datasets. The main contributions of this study are as follows:

1) A lightweight classification architecture was proposed, which was mainly fine-tuned on the basis of Shufflenet [16], and the original training weights were used for transfer learning, providing the advantageous initial weight and training starting point for the training task. The lightweight detection method determines whether the algorithm can be deployed on common production lines and handheld terminals.

2) In order to improve the classification accuracy, an attention mechanism module ECAnet [30] with superior performance was added to the network architecture.

The rest of this paper is organized as follows. The collection of fresh-cut flower datasets and classification criteria are given in Section II. Experimental methods, working procedures, analysis, and discussion are presented in Section III. The hardware environment of the experiment and its description are presented in Section IV. The conclusions are presented in Section V. The future experimental plans are discussed in Section VI.

II. MATERIALS AND CLASSIFICATION CRITERIA

A. DATA COLLECTION

We collected the data of roses using a depth camera (Intel RealSense D435i) and obtained a total of 434 color images of flowers in jpg format, with a size of 640 × 480×3 (width × height × RGB channels). In addition, 434 pieces of depth information in csv format corresponding to the flowers were obtained, with a size of 640 × 480×1 (width × height × depth). The flower depth information can bring 1 more channel data for later experimental classification. The color images and visualization results of the depth data in the experimental dataset are shown in Figure 1.

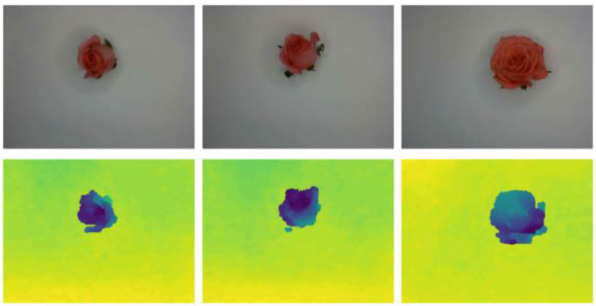

FIGURE 1. Examples of the flower dataset. The first row is the color image data, and the second row shows the depth images generated from the visualization of depth information.

B. CLASSIFICATION CRITERIA FOR ROSE

The classification of fresh-cut flowers is mainly carried out according to the type of flowers, length of flower stalks, and overall quality of flowers. In this experiment, we used Yunnan ''Fire shadow'' roses as the object, and sorted the roses according to their overall quality and appearance, such as flower shape (bud openness and uniformity), maturity (sepal calyx and petal petals opening degree), and defects. According to the domestic trade industry standard of the People's Republic of China (SB/T 11098.2-2014), a flower with sepals separated but opened less than three pieces was Grade1. The flower with petals opened 3-5 pieces was classified Grade2. The flowers with petals opened more than five pieces, but less than 50% were Grade3. The flower with petals opened by 50% was Grade4. The flowers with petals opened at more than 50% were Grade5. According to the classification criteria, the 434 flowers collected were classified and labeled manually. An example of the classification of the roses is shown in Figure 2.

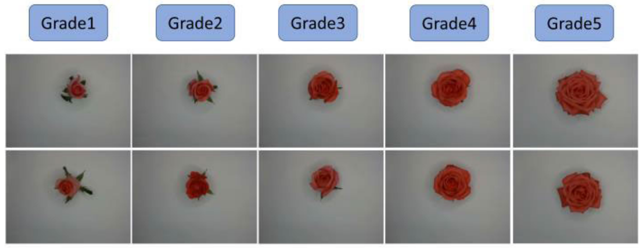

FIGURE 2. Representative pictures of rose flowers in the five grades.

III. METHOD

The overall process of flower data preprocessing and classification proposed in this paper is shown in Figure 3, which mainly includes two parts: the first part was data preprocessing: (1) Edge detection of flower image data, searching for the smallest rectangular box to segment the flower area, and extracting effective target objects. (2) The cropped flower dataset was combined with the actual classification environment for data augmentation and the image size was adjusted to a uniform size. After flower image segmentation, data augmentation, and image resizing, the 3C flower dataset composed of the RGB 3-channel flower dataset was finally obtained, and the 4-channel flower dataset was synthesized using RGB 3-channel data and one-channel depth data, which was called the 4C flower dataset. These two datasets were fed into the recognition and classification frameworks. The second part was the setting of the classification system. The ShuffleNet lightweight unit, several feature deep extraction modules, and attention mechanism modules designed by us were used to construct the training framework. The effective attention mechanism and hyperparameters were set, and the training weights of the ImageNet dataset on ShuffleNet were used for transfer learning. When the feature extraction was completed, our proposed model, named Opti-SA (Optimized ShuffleNet with Attention), was used to classify the flowers on 3-channel dataset and 4-channel dataset respectively. The overall framework of flower data preprocessing and classification in the experiment is shown in Figure 3.

A. DATA PREPROCESSING

1) EDGE DETECTION AND IMAGE SEGMENTATION

In the process of image acquisition, there will be a lot of useless information added, and it will consume a lot of time and computing resources if the information is directly input into the architecture for training and classification. In the experiment, the target object was processed via filtering, denoising, edge detection, and image segmentation. Figure 4 shows the segmentation and extraction processes used to obtain accurate flower-feature data. The experiment was divided into two parts: Experiment (a) operated on color images of flowers and Experiment (b) operated on visualized images of flower depth information.

In experiment (a), Step 1 selected a Gaussian filter to perform linear smoothing and denoising on the original image, and displayed in the six-channel color space of R, G, B, H, S, and V. Experimental results showed that the feature information extracted by the S-channel was clearer and had

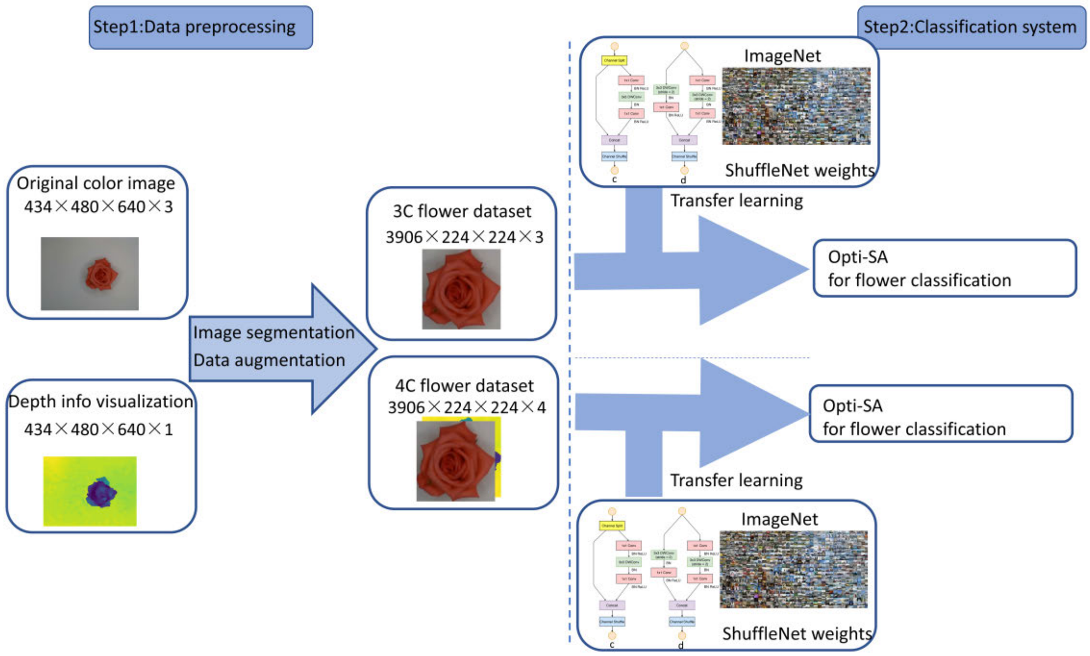

FIGURE 3. Flow diagram of flower data preprocessing and classification methods.

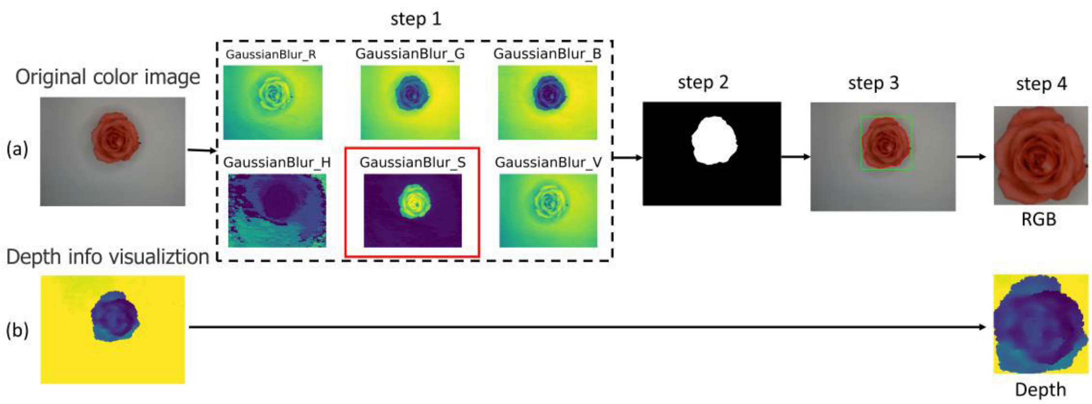

FIGURE 4. Results of image segmentation.

fewer interference factors, which could make the image segmentation more accurate.

Step 2: OTSU [31] is a very effective automatic global threshold segmentation algorithm for a single object in an image, which can accurately locate the discriminant region of flowers, separate the foreground and background of flowers, improve the feature extraction ability, and has the characteristics of fast calculation and accurate segmentation. In the experiment, the image was binarized by OTSU, dilated, and etched to obtain binary images.

Step 3: Detect the contour and edge of the flower to determine the smallest rectangular box that can wrap the detection line of the flower edge (the green box in Step 3 in Figure 4).

Step 4: The original color image is cropped according to the coordinates of the diagonal points of the rectangular box to obtain the RGB flower data for subsequent data augmentation.

In experiment (b), the depth information was visualized as a depth image, which was cropped with the smallest rectangular box, similar to the box in Step 3. Finally, one-dimensional depth information data was obtained.

2) DATA AUGMENTATION

In the experiment, we considered various possible situations for fresh-cut flowers during the actual classification process. For example, the diversity of flower images is caused by different postures and placement angles of the flowers. Image blur is caused by Gaussian noise, pepper and salt noise, and the mismatch of the camera frame rate and flower-moving speed on the production line. The inconsistent brightness of flower images is caused by the change in light intensity on-site. Owing to the limitation of flower data, in order to increase the robustness of model training in the later stage, improve generalization ability, reduce over-fitting phenomenon [32], and enhance the recognition and classification performance of fresh-cut flowers, data augmentation was performed on the flower data before training.

In Figure 5, the original cropped color and depth information images are processed. Image (1) is the result of the horizontal flipping operation on the original image, image (2) is the result of the vertical flipping operation on the original image, image (3) is the result of the horizontal and vertical flipping operations on the original image, image (4) is the result of a 90° rotation operation on the original image, image (5) is the result of adding random salt and pepper noise to the original image, image (6) is the result of randomly adding Gaussian noise to the original image, image (7) is the result of the blurring operation on the original image, and image (8) is the result of changing the brightness of the original image.

Because of the different sizes of the minimum rectangular boxes used for capture and cropping, the sizes of the images after cropping were different. To facilitate the subsequent model training and testing, the samples were adjusted to a uniform size, and a 224 × 224 3C flower dataset with 3906 samples was obtained after data augmentation. A total of 434 cropped RGB images and 434 one-dimensional images with depth information were integrated into a 224 × 224 4C flower dataset with 3906 samples after data augmentation.

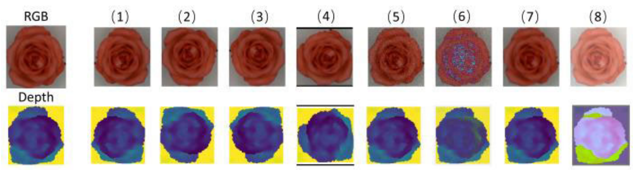

FIGURE 5. Data augmentation strategies for the rose dataset.

B. CLASSIFICATION METHOD

1) NETWORK ARCHITECTURE

The establishment of ShuffleNet V2 [16] model architecture is based on a large number of experimental verification, which started from the determination of the number of input channels and output channels, the use of grouping convolution, the fragmentation operation of multi-channels and branches, and the element-by-element operation, proves that the performance index of network architecture is not only floating-point operations (FLOPs), but also is necessary to consider the amount of memory access cost (MAC), calculation speed and other factors, so the ShuffleNet V2 has better performance in reasoning speed and accuracy than ShuffleNet V1, Xception [33], MobileNet V2 and other lightweight networks.

To achieve high classification accuracy in a fresh-cut flower production line, improve the speed of recognition and classification, and deploy the system on production line devices with limited computing power, ShuffleNet V2 [16] was used as the basic model in our experiment. Considering the limited sample size, the ShuffleNet V2 model was selected to carry out transfer learning on the weight trained on the ImageNet dataset, which could obtain general shallow information, such as contour and color, more quickly in the training of feature extraction on the flower dataset. It had a better fitting starting point and training basis, and could promote faster convergence in shallow feature extraction, reduce the computing power consumption and the cost of learning shallow information, and obtain an accurate convergence direction with high probability.

In the experiment, the flower dataset was normalized by 0.5 mean and 0.5 variance to accelerate the convergence of subsequent model iterations. In the network architecture of the model, convolution (ConV), batch normalization (BN [34]) and nonlinear ReLU activation functions were used as BasicConv2d (BC) modules. Three BC modules, MaxPool and Dropout [35], were combined into two blocks, named Block1 and Block3. Two BC modules, MaxPool and Dropout formed Block2. Pretrained_block was inserted between Block1 and Block2, and a total of 16 original blocks of stage2, stage3 and stage4 of ShuffleNet V2 were called in Pretrained_block, while the weight parameters trained in ImageNet were used in the model training. The effective connection between the shufflenet lightweight core module Pretrained_block and the three blocks was established, while ensuring sufficient light weight. The Conv layer in the self-built block was used to strengthen the effective extraction of flower data features. The core size of MaxPooling (2 × 2) was used for downsampling to reduce the number of training parameters and to improve the training speed of the model. BN and Dropout [35], [36] were used to reduce the over-fitting probability in the model training process, prevent gradient disappearance and the Dead ReLU problem effectively, and improve the model generalization ability. Each dropout layer used a probability of 0.25 to randomly discard neurons at the relevant connection layer.

The Fully Connected (FC) layer, which is used to integrate information [37] and output classification results, accounts for a large proportion of the total parameters. Therefore, when setting this layer, the number of neuron nodes of the FC layer should be reduced as much as possible to reduce the total number of architectural parameters to improve the training speed and reduce memory usage. Only 16 output neurons were used in the experiment.

In addition to ensuring fewer parameters and higher recognition speed, an appropriate attention mechanism (AM) was added between the Block3 and FC layers to increase attention to target objects, obtain more effective information, and improve the recognition and classification accuracy [38], [39]. Finally, using the Softmax function, the probability

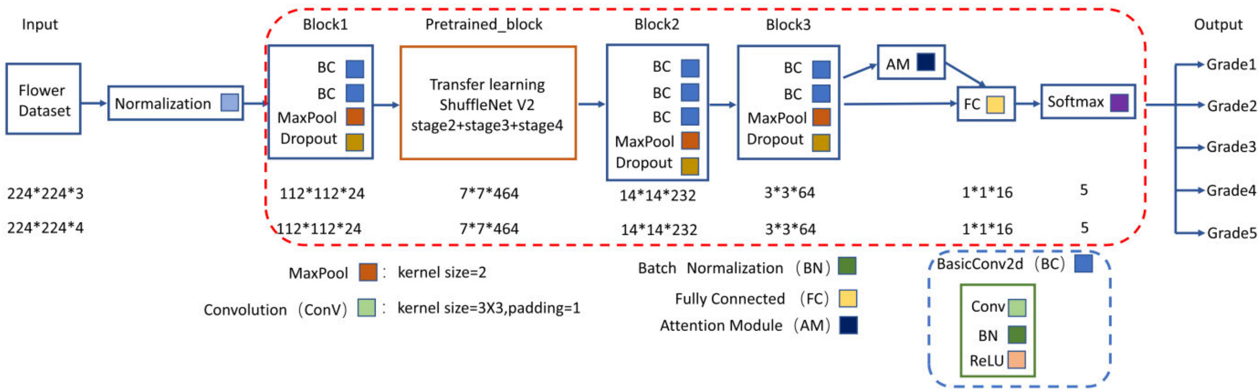

FIGURE 6. Structure of the Opti-SA for fresh-cut flower grading.

that flowers belong to each grade is output. The proposed model is named Opti-SA (Optimized ShuffleNet with Attention), and its structure is shown in Figure 6.

2) ATTENTION MECHANISM

An attention mechanism is used to increase the weight of the target object, which has a significant influence on the entire network. By adding an appropriate attention mechanism to the network and increasing the weight of the effective information, the recognition and classification accuracy was improved.

Squeeze-and-excitation networks (SEnet) [38] focused attention on global channels, established the correlation link between channels, set the weights of each channel according to the importance of information, enhanced the attention to valid information, and improved the model classification performance. The efficient channel attention module (ECAnet) [30] is an improvement of SENet, which does not capture global channel information. ECAnet replaces the fully connected layer in SEnet with a one-dimensional convolution operation for an adaptive convolution kernel, and has a good ability to capture cross-channel information. The structure of the ECAnet block is shown in Figure 7.

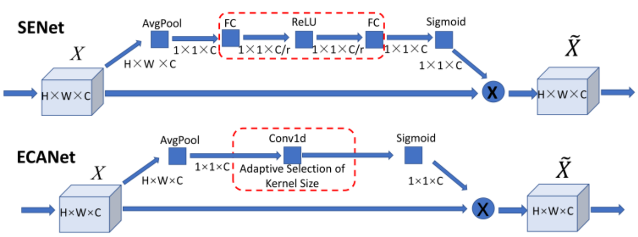

FIGURE 7. The structure of the ECAnet block and improvements compared to SENet.

The size of the adaptive convolution kernel in ECANet is

$$\mathrm{k} = \psi(C) = |\frac{\log_2(C)}{\gamma} + \frac{\mathrm{b}}{\gamma}|_{\mathrm{odd}} \tag{1}$$

where k is the size of the 1×1 convolution kernel, C is the number of channels, γ and b are set to two and one.

IV. RESULTS AND DISCUSSION

In the experiment, we obtained 3906 samples in 3C dataset and also 3906 samples in 4C dataset, after data augmentation of the original 434 original flower samples and the 3906 samples were randomly split into training and test sets at a ratio of 7:3, which means 2734 samples in training set and 1172 samples in test set.

The performance of the network depends on the design of the model structure and selection of hyperparameters. To obtain a satisfactory recognition performance, we initially used the Adam+True [40] optimizer on the model architecture and set the learning rate to 0.0001. Based on the existing model architecture, the batch size and attention were adjusted to observe the performance of the model on the 3C and 4C flower datasets.

All models were trained on the platform with an Intel i9-12900K processor, 2TB hard disk, 32GB RAM, and NVIDIA GeForce RTX3080Ti GPU. The experiment used the Jupyter notebook platform with PyTorch as the framework, and the programming language was Python 3.9.13.

1) PERFORMANCE OF ECANET WITH DIFFERENT BATCH SIZE

Batch size is an important parameter that affects the model performance in recognition and classification tasks [41], [42], [43]. In the experiment, 500 epochs were set, and the same attention mechanism, ECAnet, was used to test the model performance with different batch sizes. Test accuracy was used to measure the model performance for different batch sizes. As shown in Figure 8, the corresponding iteration optimization accuracy was higher when the batch size of the 3C and 4C flower datasets was set to 45. Therefore, to match the experimental dataset and obtain better performance, we set the batch size of the 3C and 4C flower datasets to 45 in subsequent experiments.

2) PERFORMANCE OF ECANET WITH DIFFERENT ATTENTION MECHANISM

The attention mechanism can assign different weights to different information in the image, improve the extraction of the region of interest of the target object, and enhance the

TABLE 1. Results of different methods for fresh-cut flower classification.

Methods	Flower dataset	Total parameters	Estimated total size (MB)	Test accuracy	Test time (s)
Opti-SA	3C	1 593 996	656.21	98.891	0.01622
	4C	1 594 212	658.22	99.915	0.01630
AlexNet-transfer	3C	57 024 325	273.57	86.092	0.01701
	4C	57 032 069	275.61	93.003	0.01857
DenseNet121-transfer	3C	6 958 981	1839.19	92.321	0.04664
	4C	6 962 117	1841.21	92.557	0.04665
SqueezeNet-transfer	3C	737 989	343.22	85.836	0.00424
	4C	742 693	345.25	91.894	0.00419
MobileNet V3-transfer	3C	4 208 437	727.33	94.539	0.00830
	4C	4 208 581	729.34	98.208	0.00856
MnasNet-transfer	3C	3 108 717	3.13	89.676	0.00694
	4C	3 109 005	3.17	96.928	0.00725

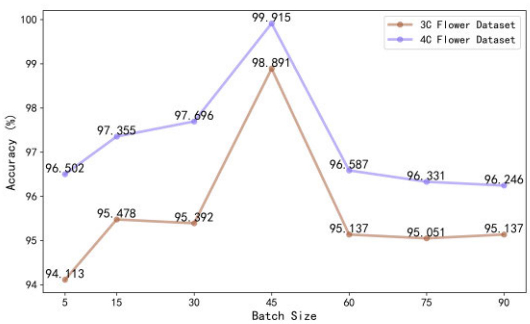

FIGURE 8. Accuracy of the proposed method with different batch sizes for fresh-cut flower grading.

classification accuracy [44]. In the experiment, we compared model performance when using SEnet, CBAM, ECAnet, and the model without the attention mechanism. As shown in Figure 9, when the attention mechanism was set to ECAnet, the test accuracy reached 98.891% on the 3C flower dataset, and 99.915% on the 4C flower dataset.

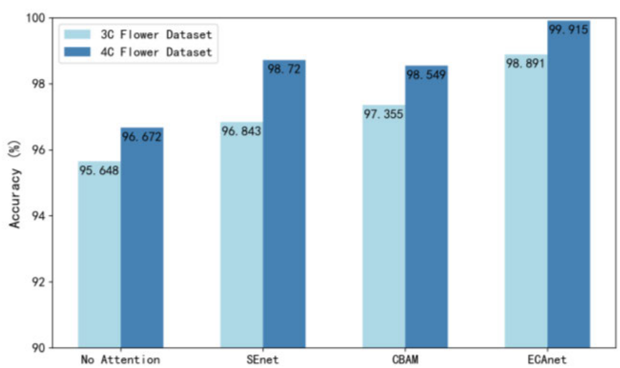

FIGURE 9. Accuracy of the proposed method with different attention mechanisms for fresh-cut flower grading.

After a series of optimization, we finally decided to use the Adam+True optimizer, set the learning rate to 0.0001 and the batch size to 45, and use ECAnet, multiple training optimization iterations were performed until the model converged smoothly and a model with a classification performance advantage was obtained. The accuracies of Opti-SA on the 3C and 4C flower datasets during training and testing at 500 epochs are shown in Figure 10. The best test accuracy of the 4C flower dataset was 99.915%, and that of the 3C flower dataset was 98.891%. The starting point and convergence speed of training and testing on the 4C flower dataset during the iterative process were higher than those of the 3C flower dataset. In the first 20 epochs, the model quickly entered a state of good performance with stable convergence and good robustness, which further proves the feasibility of the proposed model. After 20 epochs, it can be seen that the model on the 4C flower dataset performed significantly better than the 3C flower dataset on the test set by approximately 1%. The test accuracy on the 4C flower dataset fluctuated between 99% and 100% with small fluctuations. The robustness of Opti-SA on the 4C flower dataset was superior to that on the 3C flower dataset.

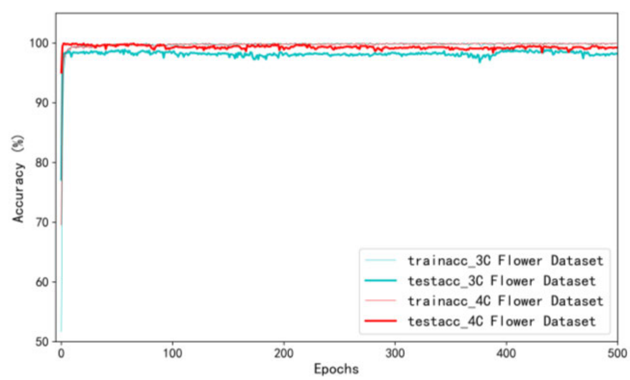

FIGURE 10. Accuracy of the proposed method of testing and training on 3C and 4C flower datasets.

3) PERFORMANCE OF DIFFERENT METHODS FOR FRESH-CUT FLOWER CLASSIFICATION

AlexNet [35], DenseNet121 [45], SqueezeNet, MobileNet_v3 and MnasNet [46] were used for transfer learning in the

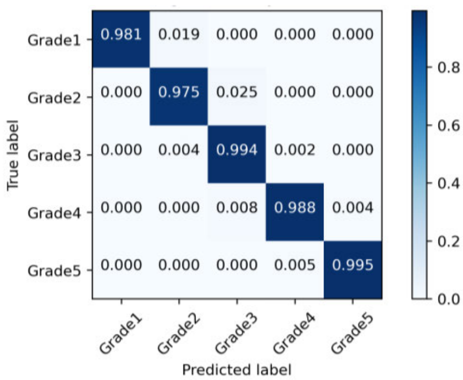

FIGURE 11. Confusion matrix on the 3C flower dataset.

experiment, and their comprehensive performance was compared with Opti-SA on the 3C and 4C Flower datasets. The results are listed in Table 1.

As shown in Table 1, our proposed model Opti-SA has obvious comprehensive advantages in terms of the total number of parameters and estimated total size, and can obtain 98.891% and 99.915% test accuracy on two datasets (3C and 4C flower datasets), respectively. Compared with AlexNet, the Opti-SA architecture does not have an advantage in the estimated total size, but Opti-SA only requires 1/35 of the total number of parameters and achieves a significant improvement in test accuracy of 12.799% and 6.912%, respectively. Compared to DenseNet121, the test accuracy on the 3C flower dataset and the 4C flower dataset was increased by approximately 7%, the running speed was increased by approximately 1.8 times, the number of parameters was reduced by more than three times, and the estimated total size was reduced by approximately 2 times. Compared with the classic lightweight network architectures SqueezeNet, MobileNet_v3, and MnasNet, the Opti-SA architecture has absolute competitive advantages in test accuracy, with a maximum improvement of 13.055% on 3C flower dataset and 8.066% on the 4C flower dataset, although the test speed is slightly slower.

The rose-series flower classification and processing system of BERCOMEX can detect and process 9000 roses per hour, which means that the processing time for each fresh-cut flower is approximately 0.4 seconds. The processing speed of the fastest tulip grading system in HAVATEC is about 0.2 s/flower (18,000 tulips per hour), Comparing the processing time of the classification lines of BERCOMEX and HAVATEC, our flower classification and grading process with processing speed of 0.020s per flower (including image preprocessing time) will not become a bottleneck in the whole flower classification system. By optimizing other processes and the speed of the classification line, we can effectively improve the efficiency of the flower grading system, reduce the overall time of flower grading and processing, reduce losses, and maintain the quality of flowers with an efficient processing speed.

TABLE 2. Performance of different models.

Methods	Year of publication	Flower dataset	Test accuracy
Opti-SA		3C	98.891
		4C	99.915
[9]	2019	3C	88.823
		4C	91.297
[27]	2021	3C	89.846
		4C	90.7
[47]	2021	3C	86.263
		4C	89.164
[48]	2022	3C	66.553
		4C	83.959
[49]	2022	3C	76.706
		4C	98.720
[50]	2018	3C	79.010
		4C	90.700
[51]	2020	3C	79.693
		4C	91.724

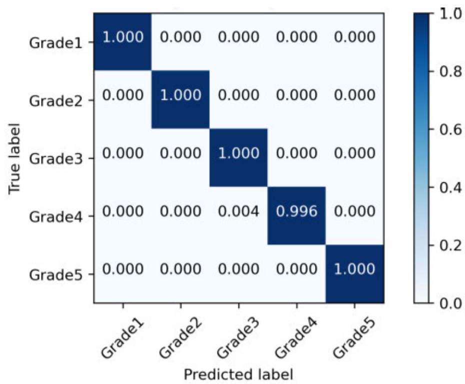

FIGURE 12. Confusion matrix on the 4C flower dataset.

In addition, we compared and tested the recognition and classification performance of the architectures in the existing literature in recent years on our dataset under 500 epochs iterations. As shown in Table 2, our network architecture Opti-SA obtains 9.045%-32.338% higher accuracy on the 3C flower dataset and 1.195%-15.956% higher accuracy on the 4C flower dataset than the other networks. At the same time, it can be observed from the experimental results that adding one dimension depth information is of great help in improving the accuracy of flower classification. As far as the self-built architecture is concerned, the test accuracy on the 4C flower dataset was approximately 1% higher than that on the 3C flower dataset.

4) DISCUSSION

The confusion matrix is a visual representation of the classifier's performance index [51]. We used it to evaluate the

TABLE 3. Classification performance measures (%).

		Accuracy	Precision	Recall	Specificity	F1-score
3C flower dataset	Grade 1	99.915	100	98.113	100	99.048
	Grade 2	99.317	98.462	97.462	99.692	97.959
	Grade 3	99.147	98.517	99.359	99.006	99.006
	Grade 4	99.573	99.194	98.795	99.783	98.994
	Grade 5	99.829	99.512	99.512	99.897	99.512
4C flower dataset	Grade 1	100	100	100	100	100
	Grade 2	100	100	100	100	100
	Grade 3	99.915	99.77	100	99.865	99.885
	Grade 4	99.915	100	99.617	100	99.808
	Grade 5	100	100	100	100	100

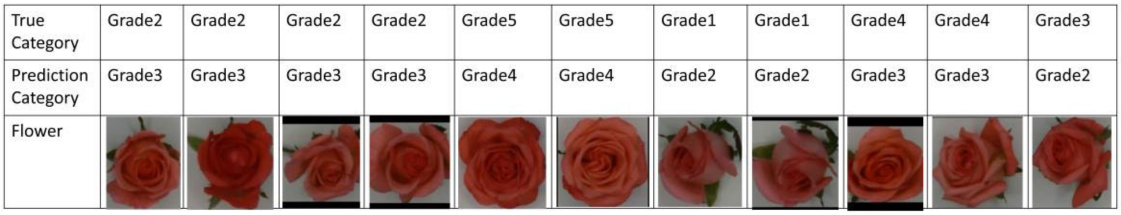

FIGURE 13. Sample images misclassified in the confusion matrix.

classification results of our model and obtain more classification information. Figure 12 shows the classification confusion matrix of the test results on the 3C flower dataset. The data on the diagonal in the figure represent the classification accuracy; the most incorrect predictions were made with a probability of 0.019 for Grade 1 being misclassified as Grade 2, 0.025 for Grade 2 being mispredicted as Grade 3, 0.004 for Grade 3 being mispredicted as Grade 2, 0.002 for Grade 3 being mispredicted as Grade 4, 0.008 for Grade 4 being mispredicted as Grade 3, 0.004 for Grade 4 being mispredicted as Grade 5, and 0.005 for Grade 5 being mispredicted as Grade 4, respectively. Figure 11 also shows that the misclassification of Opti-SA occurs only between adjacent grades and not between nonadjacent grades.

Figure 12 shows the performance of Opti-SA on the 4C flower dataset during the classification process. Only Grade 4 has a probability of 0.004 to be mispredicted as Grade 3, and the recognition performance is excellent in other grades.

We also used the parameters of TP(True Positive, an outcome where the model correctly predicts the positive flower grade), TN(True Negative, an outcome where the model correctly predicts the negative flower grade), FP(False Positive, an outcome where the model incorrectly predicts the positive flower grade), and FN(False Negative, an outcome where the model incorrectly predicts the negative flower grade). With TP, TN, FP and FN, we got the accuracy, precision, recall, specificity and F1-score to evaluate the performance of our model on 3C flower dataset and 4C flower dataset. The formulas are given between Eqs. (2)–(6), and Table 2 shows the performance metrics for each grade on 3C flower dataset and 4C flower dataset.

In Table 3, we can conclude that Opti-SA on the 4C flower dataset has the satisfactory performance, according to the result of accuracy, precision, recall, specificity and F1-score [49], [52]. F1-score performs about 2% higher on 4C flower dataset than 3C flower dataset.

$$\mathrm{Accuracy}_{\mathrm{grade\_no}} = \frac{TP_{\mathrm{grade\_no}} + TN_{\mathrm{grade\_no}}}{TP_{\mathrm{grade\_no}} + TN_{\mathrm{grade\_no}} + FN_{\mathrm{grade\_no}} + FP_{\mathrm{grade\_no}}} \tag{2}$$

$$Precision_{\mathrm{grade\_no}} = \frac{TP_{\mathrm{grade\_no}}}{TP_{\mathrm{grade\_no}} + FP_{\mathrm{grade\_no}}} \tag{3}$$

$$\mathrm{Recall}_{\mathrm{grade\_no}} = \frac{TP_{\mathrm{grade\_no}}}{TP_{\mathrm{grade\_no}} + FN_{\mathrm{grade\_no}}} \tag{4}$$

$$\mathrm{Specificity}_{\mathrm{grade\_no}} = \frac{TN_{\mathrm{grade\_no}}}{TN_{\mathrm{grade\_no}} + FP_{\mathrm{grade\_no}}} \tag{5}$$

$$\mathrm{Specificity}_{\mathrm{grade\_no}} = \frac{TN_{\mathrm{grade\_no}}}{TN_{\mathrm{grade\_no}} + FP_{\mathrm{grade\_no}}} \tag{6}$$

We extracted 11 misclassified images from the test dataset in the 3C flower dataset, as shown in Figure 13. The main reasons for the wrong classification may be the following two points:1) It is stipulated in the flower classification standard that in the absence of obvious defects, when evaluating the degree of opening and maturity of the buds; if the flower with petals opened 3-5 pieces, it was classified as Grade 2, while the flower with petals opened more than 5 pieces in total was classified as Grade 3. There are no clear quantitative criteria or strict boundaries to distinguish the degree of petal opening and maturity among the standard taxonomic categories, and most of the data in the dataset are labeled manually, so there are overlapped labels. 2) The camera cannot cover the flowers in all directions because of the fixed shooting angle, and the features of the flowers facing away from the camera cannot be reflected in the image, resulting in a final judgment error.

V. CONCLUSION

This study proposed a high-quality classification model for fresh-cut flowers. To enhance robustness and generalization ability, a set of data preprocessing methods suitable for fresh-cut flowers was designed, including accurate edge detection, cropping, and specific data augmentation methods for flower images. The experimental results showed that under the condition of limited sample size and less hardware resources, a classification accuracy of 98.891% was obtained on the traditional color image dataset, and was up to 99.915% on the dataset including depth data. The classification speed could reach 0.020s per flower. Compared with other traditional and lightweight classical networks, our proposed model showed strong competitiveness and excellent classification performance in terms of the estimated total size, number of parameters, recognition accuracy, and image-detection speed.

In the future, we will conduct further research on the classification system of fresh-cut flowers, mainly for the classification system design of roses, tulips, gladiolas, carnations, chrysanthemums, and other different types of objects, to expand the application field and improve the practical application value of the classification system.

REFERENCES

[1] X. Xia, C. Xu, and B. Nan, "Inception-V3 for flower classification," in *Proc. 2nd Int. Conf. Image, Vis. Comput. (ICIVC)*, Jun. 2017, pp. 783–787, doi: 10.1109/ICIVC.2017.7984661.

[2] S. Dharwadkar, G. Bhat, N. V. S. Reddy, and P. K. Aithal, "Floriculture classification using simple neural network and deep learning," in *Proc. 2nd IEEE Int. Conf. Recent Trends Electron., Inf. Commun. Technol. (RTEICT)*, May 2017, pp. 619–622, doi: 10.1109/RTEICT.2017.8256671.

[3] A. Soleimanipour and G. R. Chegini, "A vision-based hybrid approach for identification of anthurium flower cultivars," *Comput. Electron. Agricult.*, vol. 174, Jul. 2020, Art. no. 105460, doi: 10.1016/j.compag.2020.105460.

[4] M. Malik, W. Aslam, E. A. Nasr, Z. Aslam, and S. Kadry, "A performance comparison of classification algorithms for rose plants," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–16, Aug. 2022, doi: 10.1155/2022/1842547.

[5] H. M. Zawbaa, M. Abbass, S. H. Basha, M. Hazman, and A. E. Hassenian, "An automatic flower classification approach using machine learning algorithms," in *Proc. Int. Conf. Adv. Comput., Commun. Informat. (ICACCI)*, Sep. 2014, pp. 895–901, doi: 10.1109/ICACCI.2014.6968612.

[6] S. Puttemans and T. Goedeme, "Visual detection and species classification of orchid flowers," in *Proc. 14th IAPR Int. Conf. Mach. Vis. Appl. (MVA)*, May 2015, pp. 505–509, doi: 10.1109/MVA.2015.7153241.

[7] H. Hiary, H. Saadeh, M. Saadeh, and M. Yaqub, "Flower classification using deep convolutional neural networks," *IET Comput. Vis.*, vol. 12, no. 6, pp. 855–862, Sep. 2018, doi: 10.1049/iet-cvi.2017.0155.

[8] Y. Wu, X. Qin, Y. Pan, and C. Yuan, "Convolution neural network based transfer learning for classification of flowers," in *Proc. IEEE 3rd Int. Conf. Signal Image Process. (ICSIP)*, Jul. 2018, pp. 562–566, doi: 10.1109/SIPROCESS.2018.8600536.

[9] A. Safar and M. Safar, "Intelligent flower detection system using machine learning," in *Proc. SAI Intell. Syst. Conf.*, 2019, pp. 463–472.

[10] F. Bozkurt, "A study on CNN based transfer learning for recognition of flower species," *Avrupa Bilim Ve Teknoloji Dergisi*, vol. 32, pp. 883–890, Jan. 2022, doi: 10.31590/ejosat.1039632.

[11] M. Cibuk, U. Budak, Y. Guo, M. C. Ince, and A. Sengur, "Efficient deep features selections and classification for flower species recognition," *Measurement*, vol. 137, pp. 7–13. Apr. 2019, doi: 10.1016/j.measurement.2019.01.041.

[12] L. Liao and S. Zhang, "A flower classification method combining DenseNet architecture with SVM," in *Proc. 16th Int. Conf. Comput. Intell. Secur. (CIS)*, Nov. 2020, pp. 25–29, doi: 10.1109/CIS52066.2020.00014.

[13] B. R. Mete and T. Ensari, "Flower classification with deep CNN and machine learning algorithms," in *Proc. 3rd Int. Symp. Multidisciplinary Stud. Innov. Technol. (ISMSIT)*, Oct. 2019, pp. 1–5, doi: 10.1109/ISMSIT.2019.8932908.

[14] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 1–9, doi: 10.1109/CVPR.2015.7298594.

[15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[16] N. Ma, X. Zhang, H.-T. Zheng, and J. Sun, "ShuffleNet V2: Practical guidelines for efficient CNN architecture design," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, vol. 11218, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds. Cham, Switzerland: Springer, 2018, pp. 122–138, doi: 10.1007/978-3-030-01264-9_8.

[17] M. Aman, "Postharvest loss estimation of cut rose *(Rosa hybrida)* flower farms: Economic analysis in East Shoa Zone, Ethiopia," *Int. J. Sustain. Economy*, vol. 6, no. 1, pp. 82–95, 2014, doi: 10.1504/IJSE.2014.058518.

[18] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*.

[19] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4510–4520. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2018/html/Sandler_MobileNetV2_Inverted_Residuals_CVPR_2018_paper.html

[20] A. Howard, M. Sandler, B. Chen, W. Wang, L.-C. Chen, M. Tan, G. Chu, V. Vasudevan, Y. Zhu, R. Pang, H. Adam, and Q. Le, "Searching for MobileNetV3," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 1314–1324. [Online]. Available: https://openaccess.thecvf.com/content_ICCV_2019/html/Howard_Searching_for_MobileNetV3_ICCV_2019_paper.html

[21] X. Zhang, X. Zhou, M. Lin, and J. Sun, "ShuffleNet: An extremely efficient convolutional neural network for mobile devices," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 6848–6856.[Online]. Available: https://openaccess.thecvf.com/content_cvpr_2018/html/Zhang_ShuffleNet_An_Extremely_CVPR_2018_paper.html

[22] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size," 2016, *arXiv:1602.07360*.

[23] S. H. S. Basha, S. K. Vinakota, S. R. Dubey, V. Pulabaigari, and S. Mukherjee, "AutoFCL: Automatically tuning fully connected layers for handling small dataset," *Neural Comput. Appl.*, vol. 33, no. 13, pp. 8055–8065, Jul. 2021, doi: 10.1007/s00521-020-05549-4.

[24] Z. Wang, K. Wang, X. Wang, S. Pan, and X. Qiao, "Dynamic ensemble selection of convolutional neural networks and its application in flower classification," *Int. J. Agricult. Biol. Eng.*, vol. 15, no. 1, pp. 216–223, Feb. 2022, doi: 10.25165/j.ijabe.20221501.6313.

[25] D. H. Apriyanti, L. J. Spreeuwers, P. J. F. Lucas, and R. N. J. Veldhuis, "Automated color detection in orchids using color labels and deep learning," *PLoS ONE*, vol. 16, no. 10, Oct. 2021, Art. no. e0259036, doi: 10.1371/journal.pone.0259036.

[26] V.-T. Hoang and K.-H. Jo, "PydMobileNet: Pyramid depthwise separable convolution networks for image classification," in *Proc. IEEE 28th Int. Symp. Ind. Electron. (ISIE)*, Jun. 2019, pp. 1430–1434, doi: 10.1109/ISIE.2019.8781130.

[27] Y. Nagaraju, S. U. Hegde, and S. Stalin, "Fine-tuned MobileNet classifier for classification of strawberry and cherry fruit types," in *Proc. Int. Conf. Comput. Commun. Informat. (ICCCI)*, Jan. 2021, pp. 1–8, doi: 10.1109/ICCCI50826.2021.9402444.

[28] M. Qin, Y. Xi, and F. Jiang, "A new improved convolutional neural network flower image recognition model," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Dec. 2019, pp. 3110–3117, doi: 10.1109/SSCI44817.2019.9003016.

[29] M. Zhang, H. Su, and J. Wen, "Classification of flower image based on attention mechanism and multi-loss attention network," *Comput. Commun.*, vol. 179, pp. 307–317, Nov. 2021, doi: 10.1016/j.comcom.2021.09.001.
[30] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-Net: Efficient channel attention for deep convolutional neural networks," 2019, *arXiv:1910.03151*.
[31] A. Najjar and E. Zagrouba, "Flower image segmentation based on color analysis and a supervised evaluation," in *Proc. Int. Conf. Commun. Inf. Technol. (ICCIT)*, 2012, pp. 397–401, doi: 10.1109/ICCITechnol.2012.6285834.
[32] M. Tian, H. Chen, and Q. Wang, "Flower identification based on deep learning," *J. Phys., Conf.*, vol. 1237, no. 2, Jun. 2019, Art. no. 022060, doi: 10.1088/1742-6596/1237/2/022060.
[33] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1251–1258. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2017/html/Chollet_Xception_Deep_Learning_CVPR_2017_paper.html
[34] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. 32nd Int. Conf. Mach. Learn.*, vol. 37, 2015, pp. 448–456.
[35] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.
[36] N. Qu, Z. Li, J. Zuo, and J. Chen, "Fault detection on insulated overhead conductors based on DWT-LSTM and partial discharge," *IEEE Access*, vol. 8, pp. 87060–87070, 2020, doi: 10.1109/ACCESS.2020.2992790.
[37] C. Chen, Q. Yan, M. Li, and J. Tong, "Classification of blurred flowers using convolutional neural networks," in *Proc. 3rd Int. Conf. Deep Learn. Technol.*, Jul. 2019, pp. 71–74, doi: 10.1145/3342999.3343006.
[38] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2018/html/Hu_Squeeze-and-Excitation_Networks_CVPR_2018_paper.html
[39] S. Woo, J. Park, J. Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 3–19.
[40] S. J. Reddi, S. Kale, and S. Kumar, "On the convergence of Adam and beyond," 2019, *arXiv:1904.09237*.
[41] Z. Wang, K. Wang, X. Wang, and S. Pan, "A convolutional neural network ensemble for flower image classification," in *Proc. 9th Int. Conf. Comput. Pattern Recognit.*, Oct. 2020, pp. 225–230, doi: 10.1145/3436369.3437427.
[42] D. Masters and C. Luschi, "Revisiting small batch training for deep neural networks," 2018, *arXiv:1804.07612*.
[43] N. S. Keskar, D. Mudigere, J. Nocedal, M. Smelyanskiy, and P. T. P. Tang, "On large-batch training for deep learning: Generalization gap and sharp minima," 2016, *arXiv:1609.04836*.
[44] Y. Luo, M. Jiang, and Q. Zhao, "Visual attention in multi-label image classification," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2019, pp. 820–827.
[45] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4700–4708. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2017/html/Huang_Densely_Connected_Convolutional_CVPR_2017_paper.html
[46] M. Tan, B. Chen, R. Pang, V. Vasudevan, M. Sandler, A. Howard, and Q. V. Le, "MnasNet: Platform-aware neural architecture search for mobile," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 2820–2828.
[47] R. Lv, Z. Li, J. Zuo, and J. Liu, "Flower classification and recognition based on significance test and transfer learning," in *Proc. IEEE Int. Conf. Consum. Electron. Comput. Eng. (ICCECE)*, Jan. 2021, pp. 649–652, doi: 10.1109/ICCECE51280.2021.9342468.
[48] J. Zhou, Q. He, G. Cheng, and Z. Lin, "Union-Net: Lightweight deep neural network model suitable for small data sets," *J. Supercomput.*, pp. 1–16, Nov. 2022, doi: 10.1007/s11227-022-04963-w.
[49] A. Peryanto, A. Yudhana, and R. Umar, "Convolutional neural network and support vector machine in classification of flower images," *Khazanah Informatika, Jurnal Ilmu Komputer dan Informatika*, vol. 8, no. 1, pp. 1–7, Mar. 2022, doi: 10.23917/khif.v8i1.15531.
[50] M. V. D. Prasad, B. J. Lakshmamma, A. H. Chandana, K. Komali, M. V. N. Manoja, P. R. Kumar, C. R. Prasad, S. Inthiyaz, and P. S. Kiran, "An efficient classification of flower images with convolutional neural networks," *Int. J. Eng. Technol. (UAE)*, vol. 7, pp. 384–391, Jan. 2018, doi: 10.14419/ijet.v7i1.1.9857.
[51] S. Islam, M. F. A. Foysal, and N. Jahan, "A computer vision approach to classify local flower using convolutional neural network," in *Proc. 4th Int. Conf. Intell. Comput. Control Syst. (ICICCS)*, May 2020, pp. 1200–1204, doi: 10.1109/ICICCS48265.2020.9121143.
[52] M. Koklu, M. F. Unlersen, I. A. Ozkan, M. F. Aslan, and K. Sabanci, "A CNN-SVM study based on selected deep features for grapevine leaves classification," *Measurement*, vol. 188, Jan. 2022, Art. no. 110425, doi: 10.1016/j.measurement.2021.110425.

YEQI FEI was born in Suzhou, Jiangsu, China, in 1985. She received the master's degree in mechanical engineering from Nanjing Forestry University, China, in 2010, where she is currently pursuing the Ph.D. degree in mechanical engineering. Her research interests include artificial intelligence in data processing, image processing, and mechanical design.

ZHENYE LI received the bachelor's degree in engineering from Nanjing Forestry University, in 2019, where he is currently pursuing the master's degree. He is a Researcher in engineering with research interests in intelligent algorithms, image data processing, and hyperspectral analysis. His work focuses on developing innovative solutions to complex engineering challenges and applying intelligent algorithms to industry.

TINGTING ZHU received the Ph.D. degree in pattern recognition and artificial intelligence from the School of Automation, Southeast University, in 2019. She was a Visiting Student at the Department of Atmospheric and Oceanic Sciences, McGill University, Canada, from 2017 to 2018. She is currently an Associate Professor and working with the College of Mechanical and Electronic Engineering, Nanjing Forestry University, China. Her current research interests include machine learning, data processing and modeling, renewable energy generation forecasts, and climate feedback. She was jointly awarded by the Fonds de Recherche du Quebec—Nature et Technologies (FRQNT) and the China Scholarship Council.

CHAO NI was born in Nanjing, Jiangsu, China, in 1979. He received the B.S. degree in automation from the Nanjing University of Science and Technology, Nanjing, in 2001, and the Ph.D. degree in control theory and control engineering from Southeast University, Nanjing, in 2008. He was a Visiting Scholar at the University of Maryland, College Park, USA, from October 2017 to November 2018. He is currently a Professor with the Automation Department, Nanjing Forestry University, China. His research interests include artificial intelligence in industrial applications, data processing, and spectroscopic analysis.