IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 27 May 2023, accepted 24 June 2023, date of publication 27 June 2023, date of current version 18 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3290026

RESEARCH ARTICLE

An Improved Res2Net-Based Model for Classifying the Appearance of Deer Antler Slices

DONGMING LI[1,2], RUI YAO[2], CHENGLIN YANG[2], CHUNXI ZHAO[2,3], AND LIJUAN ZHANG[1]
[1]College of Internet of Things Engineering, Wuxi University, Wuxi 214105, China
[2]College of Information Technology, Jilin Agricultural University, Changchun 130118, China
[3]Information Center, Jilin Agricultural University, Changchun 130118, China

Corresponding authors: Chunxi Zhao (zcx@jlau.edu.cn) and Lijuan Zhang (zhanglijuan@cwxu.edu.cn)

ABSTRACT Deer antler slices are highly valued in Chinese herbal medicine due to their medicinal properties. However, the current process for classifying these slices is time-consuming and subjective. To overcome this issue, we propose an intelligent classification and recognition model based on the Res2Net architecture. Our neural network utilizes an inverse bottleneck structure to enhance grouped convolution and reduce model parameters and computation time. Additionally, we integrate an improved grouped convolution into the Res2Net model and leverage the efficient channel attention (ECA) mechanism to improve feature extraction. Our model achieves an impressive 97.96% accuracy in classifying deer antler slices and outperforms other related models. This approach can accurately differentiate between different types of deer antler slices and is particularly suitable for small-scale datasets.

INDEX TERMS Attention mechanism, deer antler slices, feature extraction, image processing, deep learning, multi-scale backbone networks.

I. INTRODUCTION

Chinese medicine has a rich history in China, with deer antler being a highly valued herbal medicine [1] [2]. Deer antler, which refers to the unboned and antlerized young antlers of the red deer or sika deer, has a translucent texture and can appear either round and thin or with bone [3]. It contains various beneficial components for humans such as biogenic amines, peptides and amino acids. The content of these substances varies among different antler species. Nowadays, the quality identification of deer antler slices mainly relies on the experience judgment of professionals, and this method is highly subjective. At the same time, the differences in the appearance of deer antler slices are small, and identification by manual methods is costly and the accuracy cannot be guaranteed. There are also often easily mixed, inferior and counterfeit products in the market, which brings many challenges to the market supervision and causes a lot of trouble to the buyers. Therefore, an intelligent and efficient antler identification method is of great practical importance to reduce labor costs while avoiding errors caused by subjective judgment as much as possible.

In recent years, deep learning has been widely used in computer vision, and Convolutional Neural Networks (CNNs) have made significant contributions to image classification [5], [6], [7], [8]. For instance, Lee and Hong [9] proposed a method to classify 32 plant images based on leaves' main vein and frequency domain data, as well as the distance between leaf contour and center of mass using Fast Fourier Transform achieving an average recognition accuracy of 97.19%. Fangliang et al. [10] classified five Chinese herbal images using the AlexNet model with an average accuracy of 87.5%,.While Peisen et al. [11] identified and classified five types of chrysanthemums based on convolutional neural networks with an average accuracy of 95%. Lipeng et al. [12] combined ResNet101 network and migration learning techniques, added Dropout regularization and batch regularization to identify a large dataset of wild plants with an accuracy of 85.6%. Similarly, Hanying et al. [13] selected 11 kinds of Chinese herbal medicines, extracted their features using an improved deep CNN model, and achieved better performance than traditional classical networks, with an accuracy rate of 97.63%. Xin [14] based

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

on ResNet-50 classified five common Chinese medicine slice images in different environments and used Soft-NMS strategy for border filtering to improve the recognition accuracy of the network, achieving 94.98% detection accuracy for single species Chinese medicine slices images and 94.68% average detection accuracy for multiple Chinese medicine slices mixed together. Ting et al. [15] classified images of Tibetan medicinal plants based on ResNet34, improved the network structure and loss function, achieving an accuracy of 91.63% for recognizing images of Tibetan medicinal plants. Yi [16] utilized convolutional neural networks to recognize 100 powder microscopic images of Chinese herbs with an accuracy of 81.5%. Yiding et al. [17] performed classification recognition of 34 herbal powders based on ResNetv2-101, and by adding the channel attention module and spatial attention module after the residual block, the accuracy of the model reached 93.9%. Al-Badri et al. [18] proposed a hybrid CNN model with three different extractors on the network backbone, namely VGG-16, ResNet-50 and Inception-V3. Combining these networks into an integrated model improved extraction of deeper features, achieving 97.51% classification accuracy for Rumex. Lastly, Amano et al. [19]used CNN to classify three bacteria based on their phenotypic appearance with an accuracy of up to 97.19%.

Despite the importance of deer antler in Chinese medicine, there is currently no literature on the application of deep convolutional neural network techniques for classifying and identifying deer antler slices. Recognizing the challenge of the small size and difficulty of recognition of deer antler slices, this paper proposes an improvement to the Res2Net [20] network by introducing an improved group convolution and ECA attention mechanism to enhance the model's ability to process deer antler images and improve recognition accuracy significantly.

The remainder of this paper is organized as follows: the second section introduces the dataset and its preprocessing, the third section describes the details of the network improvement, the fourth section outlines the experimental environment and presents an analysis of the experimental results and the final section provides concluding remarks.

II. IMAGE DATA SET AND PREPROCESSING

In 2021, batches of deer antler slices were collected from a deer farm in Changchun and the College of Animal Science at Jilin Agricultural University. These samples were verified as genuine deer antler slices by experts specializing in antler research. To capture high-quality images, the deer antler slices were placed on a white background in the laboratory and photographed using a mobile phone with natural light. The experiment focused on five types of deer antler slices: sika deer wax slice, red deer wax slice, red deer powder slice, red deer powder slice and red deer bone slice. The dataset comprised a total of 2211 images of deer antler slices, including 480 red deer powder slice, 510 red deer blood slice, 410 red deer wax slice, 620 red deer blood slice and 191 sika deer wax slice. Figure 1a shows the images in the dataset.

TABLE 1. Partitioning of the deer antler slice dataset.

Types of deer antler slices	Training set	Validation set	Totals
Red deer powder slice	384	96	480
Red deer blood slice	408	102	510
Red deer wax slice	328	82	410
Red deer blood slice	496	124	620
Sika deer wax slice	153	38	191

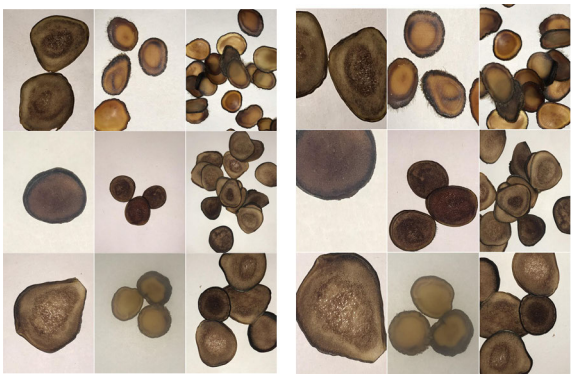

Figure1 (**a**) Figure1 (**b**)

FIGURE 1. Deer antler slices dataset (a) Original images of different kinds of deer antler slices (b)Processed images of different kinds of deer antler slices.

To eliminate experimental chance, simulate complex realistic scenarios, reduce the dependence of deep learning image classification models on certain feature attributes of images and improve network robustness and generalization ability [21], [22], we randomly divided the entire dataset using a Python scripting program. We used 80% of the dataset as the training set and the remaining 20% as the validation set. Due to the small number of samples, we did not create a separate test set. Instead, we used the validation set as a test set to verify the results. The dataset was divided into 1769 images for the training set and 442 images for the validation set, as shown in Table 1. To consider the uncertainties in realistic environments and the presence of numerous interference factors [23], the training set was randomly flipped and cropped. Figure 1b shows the processed images.

III. BUILDING THE MODEL

A. RES2NET MODEL

The Res2Net model, proposed by Gao [20] et al. in 2019, is an improved version of the ResNet model [24]. Compared to ResNet, Res2Net has enhanced generalization performance. The model includes hierarchical small residual blocks in the

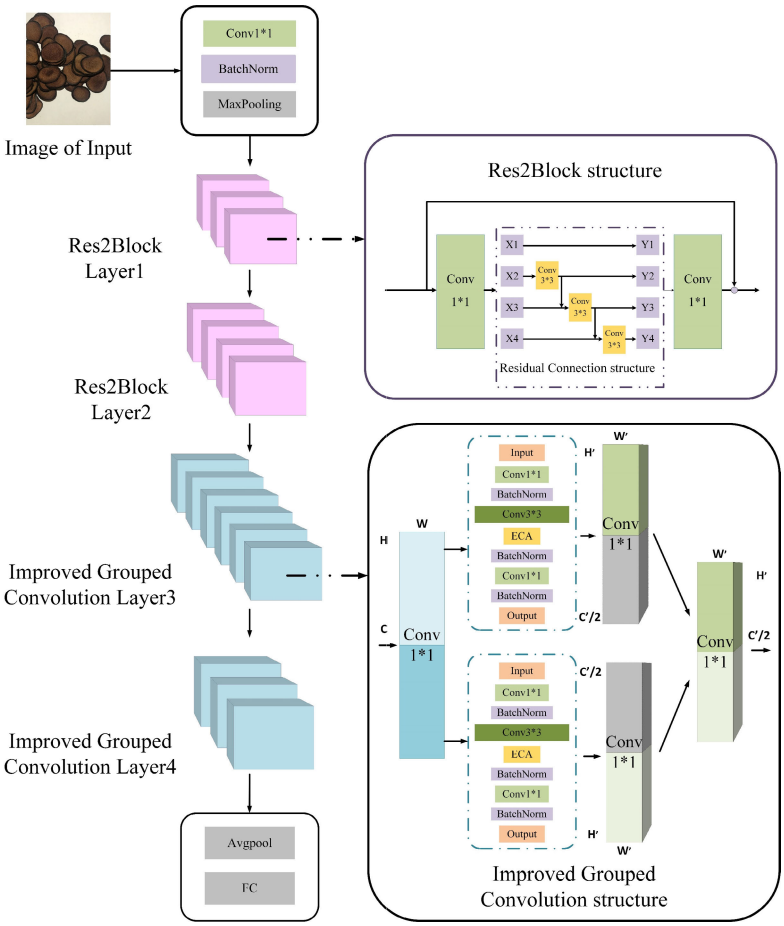

FIGURE 2. Improved res2net model Note: "W" represents the width of the input image, "H" represents the height of the input image, and "C" represents the number of channels.

residual cell structure, which increases the effective sensory field of each layer and expands the feature extraction performance of the whole network. Considering the small size and difficulty of recognizing deer antler slices, we selected the Res2Net model as the backbone network for the antler slice classification task due to it is potential for extracting useful features from such images.

B. CLASSIFICATION MODEL OF DEER ANTLER SLICES

The deer antler slice classification model comprises two Res2Block layers at the beginning, with each Res2Block utilizing a different number of 3 × 3 sized convolutional kernels connected hierarchically to form multiple channels by splitting the feature map. This approach helps to extract detailed feature information from the input image.

The Res2Net model increases the network load and time cost of model learning when the number of convolutional layers is increased [25], [26], [27]. To address this problem, an improved grouped convolution module was introduced in the later stage of the model, which reduces the amount of operations and the number of parameters of the model. The inverse bottleneck structure was borrowed from the MobileNet V2 model [28] to strengthen the feature extraction ability of the grouped convolutional module. This allows for the extraction of more feature information while reducing the model complexity.To ensure effective recognition of the small sized and closely appearing sample images of deer antler slices, an efficient channel ECA attention mechanism was introduced after the 3 × 3 convolution in the inverse bottleneck module [29]. This enables the model to effectively capture fast channel interaction information while avoiding dimensionality reduction [30], [31]. The improved overall network framework can be seen in Figure 2.

C. IMPROVED RES2NET MODEL

1) GROUPED CONVOLUTION

Figure 3 shows group convolution, which groups the different feature maps of the input layer and splits the filter into different groups. It uses different convolution kernels to do convolution in the corresponding groups. The feature map is divided into two groups for convolution. "W" represents the width of the input image, "H" represents the height of the input image and "C" represents the number of channels. The parameters for this are as follows (divided into

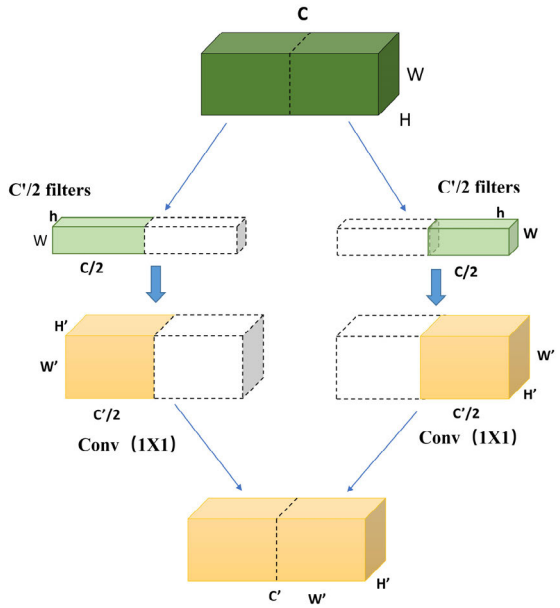

FIGURE 3. Grouped convolution Note "W" represents the width of the input image, "H" represents the height of the input image, and "C" represents the number of channels.

two groups as an example):

$$\text{Size of input image} = H \times W \times \frac{C}{2} \tag{1}$$

$$\text{Size of output image} = H' \times W' \times \frac{C'}{2} \tag{2}$$

$$\begin{aligned}\text{Number of participants} &= H \times W \times \frac{C}{2} \times \frac{C'}{2} \times 2 \\ &= H \times W \times C \times \frac{C'}{2}\end{aligned} \tag{3}$$

$$\begin{aligned}\text{Volume of operations} &= H \times W \times \frac{C}{2} \times \frac{C'}{2} \times 2 \times H' \times W' \\ &= H \times W \times C \times C' \times W' \times \frac{H'}{2}\end{aligned} \tag{4}$$

Compared to standard convolution, grouped convolution reduces the amount of parameters to half of the original amount. If the feature maps of the input layer are divided into four groups, the number of parameters will be reduced to one-quarter of the original amount. This reduction effectively improves computational speed and increases diagonal correlation between the filters, it also helps avoid overfitting problems during experiments.

2) IMPROVED GROUPED CONVOLUTION

This paper introduces an inverse bottleneck structure when using grouped convolution, as shown in Figure 4. The inverse bottleneck structure achieves dimensionality enhancement through a 1×1 convolutional layer, followed by feature extraction through a 3×3 convolutional layer. Finally, a 1×1 convolutional layer is used to restore the dimensionality of the input feature information. This structure

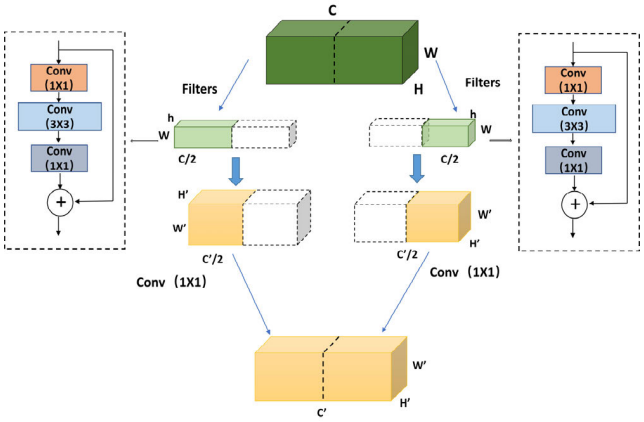

FIGURE 4. Improved grouped convolution Note:"W" represents the width of the input image, "H" represents the height of the input image, and "C" represents the number of channels.

has two small ends and a large middle, which effectively improves the expression capability of the inverse bottleneck structure compared to the traditional convolutional process of reducing and then increasing the dimensionality of the residual network.The inverse bottleneck structure enhances the expressiveness of the network.

Let the input be $h \times w \times k$, the output is $h \times w \times (tk)$,after 1×1 convolutional layer up-dimensioning. At this time, $h \times w \times (tk)$ is used as the input of 3×3 convolutional layer, and the output is $\frac{h}{s} \times \frac{w}{s} \times (tk)$. The $\frac{h}{s} \times \frac{w}{s} \times (tk)$ is used as the input of the next layer, and the output is $\frac{h}{s} \times \frac{w}{s} \times k'$ after the final 1×1 convolution layer, where the convolution kernel size is k, stride is s, and expansion factor is t.

Although the number of channels has increased significantly, the use of grouped convolution limits the increase in the number of parameters. This results in significant improvement in the model's ability to learn space, greatly expanding the model's ability to extract features.

3) ATTENTION MECHANISM

Attention mechanisms are modeled after the way humans process information, filtering efficiently and accurately from a multitude of information to valid information [32]. They are widely used in the field of deep learning and are extremely helpful for various aspects of image processing tasks [33].The network model introduced in this paper utilizes the Efficient Channel Attention (ECA) mechanism as a channel attention mechanism. The structure of the ECA attention mechanism can be seen in Figure 5.

The coverage of cross-channel information interactions used by this attention mechanism (the kernel size of the one-dimensional convolution k) should also be proportional to the channel dimension C. In other words, there may be a mapping between k and C φ

$$C = \varphi(k) \tag{5}$$

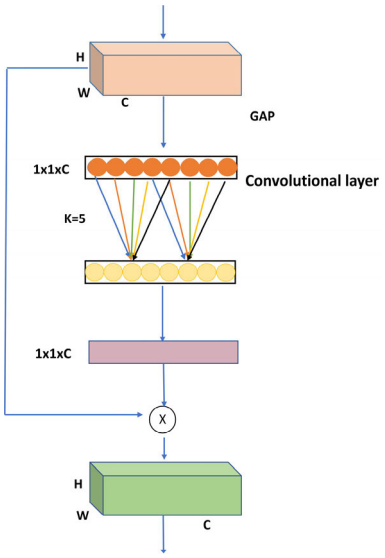

FIGURE 5. ECA Attention Mechanism model Note:"k"represents the kernel size, represents the coverage of local cross-channel interaction, "W" represents the width of the input image, "H" represents the height of the input image, and "C" represents the number of channels.

In general, the channel dimension C (i.e., the number of filters) is usually set to a multiplicative operation of 2. We introduce a possible solution by extending the linear function $(k) = \gamma * k - b$ to a nonlinear function, i.e.

$$C = \varphi(k) = 2^{(\gamma * k - b)} \tag{6}$$

Given the channel dimension C, the kernel size k can adaptively determine the convolution size K size

$$k = \varphi(C) = |\frac{\log_2(C)}{\gamma} + \frac{b}{\gamma}|_{odd} \tag{7}$$

The introduction of the ECA attention mechanism greatly improves the ability to extract valid information from a large number of features. This is especially important due to the similar appearance of deer antler slices and the complex background of the dataset [34], [35], which make image recognition difficult. The ECA attention mechanism avoids dimensionality reduction and enables fast capture of cross-channel interactions. The coverage of local cross-channel interactions is determined by adaptively selecting the kernel size of the one-dimensional convolution. This brings significant performance gains with the addition of only a small number of parameters.

IV. EXPERIMENTS AND ANALYSIS OF RESULTS

A. EXPERIMENTAL ENVIRONMENT PARAMETER SETTING

The configuration environment for this antler classification experiment is: processor: AMD Ryzen9 5900HX with RadeonGraphics 3.30GHz, graphics card: Nvidia GeForce RTX 3070 LaptopGPU, operating system: windows11 (home edition), Python3.10.1 based Pytorch1.10 deep learning framework built on Python3.10.1 programming language, software configuration installed as Anaconda3-2021.11-windows version.

TABLE 2. Training hyperparameter information.

Parameter	Value or name
Training epochs	120
Batch size	32
Learning rate	0.001
Optimizer	SGD
Input size/pixels	224×224

Information on the specific experimental parameters used in the training of the new network model proposed in this paper is shown in Table 2.

B. COMPARISON EXPERIMENTS OF DIFFERENT MODELS

The effectiveness and advancement of the deer antler tablets classification model were verified through the use of four convolutional neural network models, namely ResNet50, ConvNext_t, RepVggNet_b0 and IresNet50. all comparison models followed the same model framework, training strategy and parameter setting methods as described in this study. The final experimental results presented in this paper are the average of five trials to ensure statistical significance and reliability.

Classification experiments were performed using various metrics to evaluate the spatial complexity of each model. The metric "Params" indicates the total number of parameters for the model, while "Memory" represents the size of memory occupied by it. "FLOPs" stands for the computational power of the floating-point operation of the model and "MAdds" indicates the computational power of the dot product operation. Inference time is defined as the minimum time it takes for the model to analyze a single image, whereas "FPS" denotes the rate of image transmission per second while the model is active. By evaluating these metrics, we can determine the relative performance of different models and select the most efficient and effective one for our purposes.

After analyzing the comparative experimental results, we can conclude that the proposed deer antler slice classification model in this paper achieves the best accuracy of 97.96%. It also has the lowest Memory value of 43.13MB, the lowest FLOPs value of 2.26GMac, the lowest MAdds value of 3.53GMAdd and an inference time of only 7.03ms. The visualized images of the comparison experiments also show that the deer antler slice classification model proposed in this study tends to smooth out in about 30 iterations, achieving fast convergence and a more compact and lightweight model. In contrast, the other models have large curve jitter, slow convergence, lower recognition accuracy for antler slice image samples and weak generalization ability. In summary, the network model proposed in this study has better performance for the classification task of deer antler slices. The experimental results are presented in Table 3, Figure 6 and Figure 7.

TABLE 3. Comparison experiments of different models.

Model	Accuracy	Params	Memory	FLOPs	MAdds		FPS
ResNet50	91.93%	23.52M	109.68MB	4.12GMac	8.72GMAdd	8.39ms	119.10fps
ConvNext_T	87.78%	27.82M	100.71MB	4.47GMac	8.07 GMAdd	7.19ms	140.47fps
RepVggNet_B0	96.83%	13.06M	77.43MB	3.06GMac	6.82 GMAdd	11.16ms	89.61fps
IresNet50	94.96%	23.52M	104.13MB	4.16GMac	8.29 GMAdd	8.32ms	120.07fps
Our model	97.96%	16.30M	43.14MB	2.26GMac	3.53 GMAdd	7.03ms	142.25fps

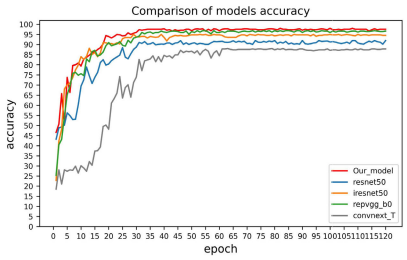

FIGURE 6. Visualization results of the accuracy of different model comparison experiments.

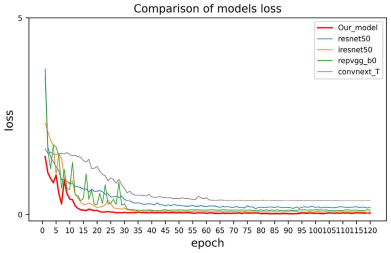

FIGURE 7. Visualization results of the loss of different model comparison experiments.

C. ABLATION EXPERIMENTS

To evaluate the impact of grouped convolution, improved grouped convolution and ECA attention mechanism on the model performance, we conducted ablation experiments using Res2Net as the base network. The results are presented in Table 4, it is evident that there was a significant improvement in model performance when the improved group convolution was used in combination with the ECA attention mechanism. The experimental results show that the accuracy of the deer antler slice classification model reaches 97.96% and the F1 score of the model reaches 97.8%.

1) EFFECT OF ATTENTIONAL MECHANISMS ON NETWORK MODEL PERFORMANCE

The attention mechanism is modular and can be added theoretically after any feature layer of the network to improve its ability to extract effective image features. In this experiment, we introduced different attention mechanisms into the proposed antler slice classification model for comparison purposes. We used the channel attention SE (Squeeze-Excitation) model to consider the interdependence between the channels of the network model. The CBAM (Convolutional Block Attention Module) model was utilized to consider the dual attention of channel and spatial dimensional tandem. Additionally, we employed the CA (Channel Attention) model that embeds location information into channel attention for comparison. To compare these attention mechanisms with the ECA (Efficient Channel Attention) attention mechanism module, which is a non-dimensioned local cross-channel interaction strategy used in this study, we used the original model (Res2Net) for image recognition. The results of these experiments are presented in Table 5, it is evident that the ECA attention mechanism provided the best performance compared to other attention mechanisms. It achieved the highest accuracy while maintaining low computational complexity. Therefore, we chose the ECA attention mechanism as the optimal choice to improve the antler slice classification model's performance.

Table 5 presents a comparison of results after introducing different attention mechanisms into the antler slice classification model. The results demonstrate that the average accuracy of the model improved by 1.35%, 0.71%, 1.65% and 3.14% after introducing the SE, CBAM, CA and ECA attention mechanisms, respectively, compared to the original model. The introduction of attention mechanisms significantly improved the model's performance, particularly the ECA attention mechanism, which demonstrated the most outstanding performance in recognizing deer antler slices. Therefore, we chose the ECA attention mechanism to optimize the antler slice classification model.

To visually analyze the effectiveness of the improved antler slice classification model, we utilized Grad-CAM [36], a visualization tool. We generated thermal feature visualization images for the last improved grouped convolutional layer after introducing the ECA attention mechanism module. These images highlight critical regions in the images used for prediction and provide insight into the model's decision-making process. The experimental results demonstrate that the addition of the ECA attention mechanism module

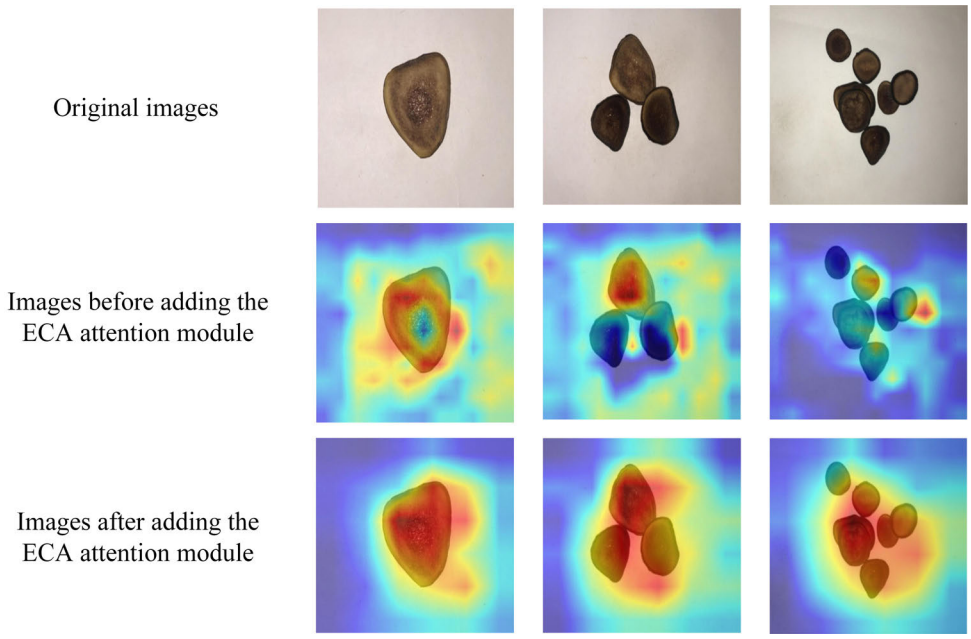

FIGURE 8. Visualization results of the new network thermal characteristic map before and after improvement.

TABLE 4. Combination experiment of different modules.

Number	Res2Net	Grouped convolution	Improved Grouped convolution	ECA Attention Mechanism	Accuracy	F1 Score
1	√				91.40%	90.7%
2	√	√			95.37%	93.2%
3	√		√		96.48%	95.7%
4	√			√	95.74%	95.9%
5	√	√		√	97.12%	96.5%
6	√		√	√	97.96%	97.8%

TABLE 5. Recognition results comparison of different attention mechanism models.

Model	Loss	Accuracy	Convergence/ epoch
No	0.54	94.82%	55
Res2Net+SE	0.45	96.17%	51
Res2Net+CBAM	0.50	95.53%	54
Res2Net+CA	0.47	96.47%	49
Res2Net+ECA	0.43	97.96%	35

enhances the antler slice classification model's feature extraction capability, making it more accurate in locating valuable regions in antler slice images. This improvement leads to increased accuracy in predicting the class of an image. The visualized experimental results are presented in Figure 8.

In the deer antler slice classification model, we integrated the ECA attention mechanism into the inverse bottleneck module. To determine the optimal placement for the ECA attention mechanism, we performed ablation experiments on it. The results showed that when placed in the 1×1 up-convolution and 1×1 down-convolution, the accuracy of the model was 96.27% and 95.84%, respectively, the loss values were 0.49 and 0.52, respectively. However, placing the ECA attention mechanism in the 3×3 convolutional layer improved the average accuracy to an optimal value of 97.96%. This finding indicates the superior feature extraction ability of the 3×3 convolutional layer in the inverse bottleneck module. As a result, we decided to place the ECA attention mechanism after the 3×3 convolutional layers in the inverse bottleneck module to maximize its performance improvement on the overall model. Table 6 outlines the experimental results.

TABLE 6. Attentional mechanism ablation experiments.

Model	Params	FLops	accuracy	Loss
First 1 × 1+ECA	16.3M	2.26GMac	96.27%	0.49
3× 3+ECA	16.3M	2.26GMac	97.96%	0.43
Last 1 × 1+ECA	16.3M	2.26GMac	95.84%	0.52

TABLE 7. Ablation experiments with improved convolution of grouped convolutions.

Number	Model	Params	FLOPs	Accuracy
1	$G_1R_2R_3R_4$	93.80M	3.68GMac	95.25%
2	$G_1G_2R_3R_4$	86.80M	2.27GMac	94.80%
3	$G_1G_2G_3R_4$	53.64M	1.33GMac	95.19%
4	$G_1G_2G_3G_4$	10.57M	2.63GMac	89.72%
5	$R_1G_2G_3G_4$	21.44M	3.29GMac	92.47%
6	$R_1R_2G_3G_4$	16.30M	2.26GMac	97.96%
7	$R_1R_2R_3G_4$	43.79M	3.62GMac	96.70%
8	$R_1R_2R_3R_4$	4.48M	94.43GMac	94.18%

Note: "G" represents the improved grouped convolution and "R" represents Res2Block layer. The network has four layers. Experiment 1 shows that the first layer is the improved block convolution, and the second, third and fourth layers are Res2Block layer.The following experiments are similar.

TABLE 8. Corresponding Labels And Types.

Label	Types of pilose deer antler slices
1	Red deer powder slice
2	Red deer blood slice
3	Red deer wax slice
4	Red deer blood slice
5	Sika deer wax slice

2) EFFECT OF GROUPED CONVOLUTION ON NETWORK MODEL PERFORMANCE

To attain the optimal performance of the deer antler slice classification model, we restructured the network layers. Through ablation experiments conducted on the reconstructed network layers, our findings revealed that number 6 reconstructed scheme attained an optimal value of 97.96% accuracy, surpassing other reconstructed schemes in terms of overall performance. As a result, we selected number 6 reconstructed scheme as the optimal network structure for the deer antler slice classification model.

D. COMPARISON OF RELEVANT INDICATORS

The confusion matrix is composed of columns that display the predicted category of antler tablets, with the total number of data instances predicted in each category. Each row represents the true attribution category of antler tablet data, with the total number of data instances in each category [37]. To facilitate visualization of the classification recognition effects of the new model and the pre-modified Res2Net model, we present the corresponding visual confusion matrix and the comparison graph of recognition accuracy rate of each category in Tables 8 and Figure 9. Compared to the original Res2Net model before improvement, the new model accurately predicted a greater number of images in the validation set for the red deer powder slice, red deer blood slice, red deer wax slice, red deer blood slice and sika deer wax slice image data.

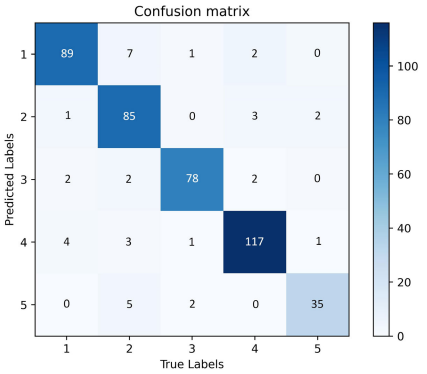

Figure 9（a）

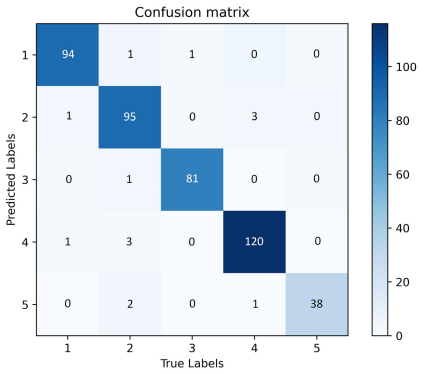

Figure 9（b）

FIGURE 9. Model confusion matrix visualization (a)Before improvement (b)After improvement.

The experimental results demonstrate that the improved model can significantly enhance the perceptual field of view of the deer antler slice image output and enrich feature information, enabling it to extract finer-grained color, texture, and other feature information of deer antler slices. This improvement effectively enhances the accuracy of deer antler slice species recognition.

In computer vision image classification studies, we usually use the precision, recall and F1 score as key evaluation metrics to study the classification recognition performance [38].

Precision is the probability that of all the samples predicted to be positive are actually positive samples for our predicted

TABLE 9. Comparison of model recognition performance evaluation metrics.

Types of pilose deer antler slices	Precision		Recall		F1 Score	
	Before	After	Before	After	Before	After
Sika deer wax slice	83%	92%	92%	100%	87%	96%
Red deer wax slice	92%	98%	95%	98%	93%	98%
Red deer powder slice	89%	97%	92%	97%	90%	97%
Red deer blood slice	92%	96%	94%	96%	93%	96%
Red deer bone slice	93%	96%	83%	93%	88%	94%

outcome.

$$\text{Precision} = \frac{\text{TP}}{\text{TF} + \text{FP}} \tag{8}$$

Recall refers to the probability that of the samples that are actually positive are predicted to be positive for the original sample.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{9}$$

F1 score is the average of the sum of the correct and recall rates. It reflects the stability and generalization ability of the model performance.

$$\text{F1} = \frac{2 \times (\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \tag{10}$$

After analyzing the results presented in Table 9, it can be observed that the improved sika deer wax slice has shown a significant improvement in precision by 9%, while recall and F1 score have been increased by 8% and 9%. Moreover, the red deer wax slice has also demonstrated an enhancement of 6% in checking accuracy, 3% in recall and 5% in F1 score. Similarly, the red deer powder slice has shown an increase in accuracy by 8%, recall by 5% and F1 score by 7%. Additionally, the red deer blood slice has demonstrated an improvement in accuracy by 4%, recall by 2% and F1 score by 3%. Finally, the red deer bone slice has exhibited an enhancement in accuracy by 2.6%, a 10% increase in recall and a 6% increase in F1 score.

The experimental data presented above demonstrates the high recognition accuracy of the improved network in deer antler slices image classification. This accuracy can be attributed to the network's multi-scale feature extraction capabilities, which effectively expand the sensory field of the images. Additionally, the group convolution with inverse bottleneck, along with the ECA attention mechanism, refine deer antler slices images, enhancing the network's ability to recognize details of textures in the later stages of feature extraction. These modifications ultimately strengthen the network's ability to recognize and classify deer antler slices images accurately.

V. CONCLUSION

Our study involved the collection of images of five different types of deer antler slices, including sika deer wax slice, red deer wax slice, red deer powder slice, red deer powder slice and red deer bone slice. We enhanced the original Res2Net model by incorporating grouped convolution through the inverse bottleneck structure in the later stage. By doing this, we greatly reduced the number of parameters and computation required by the model while maintaining its feature extraction capability.Furthermore, we added the ECA attention mechanism into the inverse bottleneck structure to ensure the full extraction of network feature information despite the small sample size of deer antler slices and their similar phenotypic features. Our results demonstrated a significant improvement over other advanced classification network models. The proposed network model converged quickly within 30 rounds during the training process, with an inference time of only 7.03ms. Ultimately, the accuracy reached an impressive 97.96%.

Due to current conditions, we were only able to collect five common and highly medicinal value deer antler slice images for our experiments. In future work, we plan to supplement our sample size with additional types of deer antler slices to enhance the model's accuracy. Additionally, we aim to introduce a lightweight model that can be combined with a portable mobile terminal to achieve fast, real-time classification of Chinese herbal medicine images.

REFERENCES

[1] X. Heli, *Extraction and Isolation of Active Components of Deer Antler and Its Antitumor Effect*. Daejeon, South Korea: Northwest Agriculture and Forestry University of Science and Technology, 2007.

[2] *Pharmacopoeia of the People's Republic of China: A*, China Pharmaceutical Science and Technology Press, National Pharmacopoeia Committee, Beijing, China, 2020, p. 156.

[3] S. Yuping, L. Jianxun, and P. Ximing, "Species classification and appearance identification of Chinese deer antler," *Chin. Herbal Med.*, vol. 8, pp. 76–77, Jan. 2000.
[4] G. Xiaohan, C. Xianlong, L. Wenxi, L. Minghua, W. Feng, and M. Shuangcheng, "Multi-indicator quality grade evaluation of different sizes of deer antler tablets," *China Modern Traditional Chin. Med.*, vol. 23, no. 4, pp. 691–697,2021, doi: 10.13313/j.issn.1673-4890.20200318001.
[5] Z. Yu, T. Qian, W. Ren, and W. Yongjin, "Research on plant image classification method based on convolutional neural network," *Internet Things Technol.*, vol. 10, no. 3, pp. 72–75, 2020, doi: 10.16667/j.issn.2095-1302.2020.03.021.
[6] L. Yue, *Research on Plant Identification Method Based on Improved Capsule Neural Network*. Changchun, China: Jilin Agricultural Univ., 2020, doi: 10.27163/d.cnki.gjlnu.2020.000142.
[7] X. Zhouyi, F. Yazhi, H. Yanrong, and L. Hongjiu, "Rice pest image description generation based on ResNet18 feature encoder," *J. Agricult. Eng.*, vol. 38, no. 12, pp. 197–206,2022.
[8] T. Arias-Vergara, P. Klumpp, J. C. Vasquez-Correa, E. Nöth, J. R. Orozco-Arroyave, and M. Schuster, "Multi-channel spectrograms for speech processing applications using deep learning methods," *Pattern Anal. Appl.*, vol. 24, no. 2, pp. 423–431, May 2021, doi: 10.1007/s10044-020-00921-5.
[9] K. K. Lee, "An implementation of leaf recognition system using leaf vein and shape," *Int. J. Bio-Sci. Bio-Technol.*, vol. 2, pp. 57–66, May 2013.
[10] H. Fangliang, Y. Lei, S. Tongping, J. Li, X. Huanqing, and H. Xingyu, "Research and implementation of herbal plant image classification based on AlexNet deep learning model," *J. Qilu Univ. Technol.*, vol. 34, no. 2, pp. 44–49,2020, doi: 10.16442/j.cnki.qlgydxxb.2020.02.008.
[11] Y. Peisen, L. Wei, R. Shougang, and X. Huanliang, "Flower type and variety recognition of chrysanthemum based on convolutional neural network," *J. Agricult. Eng.*, vol. 34, no. 5, pp. 152–158,2018.
[12] L. Lipeng, S. Feipeng, and T. Wenbo, "Wild plant image recognition method based on residual network and migration learning," *Radio Eng.*, vol. 51, no. 9, pp. 857–863, 2021.
[13] W. Hanying, *Research on Identification Methods of Chinese Herbal Medicine Based on in-Depth Learning*. Guilin, China: Guilin Univ. Electronic Science and Technology, 2019.
[14] L. Xin, *Research on Image Recognition of Common Chinese Herbal Medicine Pieces Based on Convolution Neural Network*. Changsha, China: Hunan Agricultural Univ., 2021, doi: 10.27136/d.cnki.ghunu.2021.000496.
[15] Z. Ting, D. Jianqiang, Z. Yanchen, L. Jigen, and H. Dingxing, "Research on Tibetan medicine plant image classification algorithm based on improved ResNet," *Electron. Technol. Softw. Eng.*, vol. 3, pp. 164–169, Jan. 2022.
[16] Y. Yi, *Research on Micro Image Recognition of Chinese Herbal Medicine Powder Based on Fine Particle Size*. Beijing, China: North Univ. Technology, 2022, doi: 10.26926/d.cnki.gbfgu.2022.000585.
[17] W. Yiding, H. Chenyu, L. Yaoli, C. Shaoqing, and Y. Yuan, "Micro image recognition of small samples of Chinese herbal medicine powder based on deep learning," *Comput. Appl.*, vol. 40, no. 5, pp. 1301–1308, 2020.
[18] A. H. Al-Badri, N. A. Ismail, K. Al-Dulaimi, A. Rehman, I. Abunadi, and S. A. Bahaj, "Hybrid CNN model for classification of rumex obtusifolius in grassland," *IEEE Access*, vol. 10, pp. 90940–90957, 2022, doi: 10.1109/ACCESS.2022.3200603.
[19] M. Amano, D. Mai, G. Sun, T. N. Vu, L. T. Hoi, N. T. Hoa, and K. Ishibashi, "Deep learning approach for classifying bacteria types using morphology of bacterial colony," in *Proc. 44th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2022, pp. 2165–2168, doi: 10.1109/EMBC48229.2022.9870986.
[20] S.-H. Gao, M.-M. Cheng, K. Zhao, X.-Y. Zhang, M.-H. Yang, and P. Torr, "Res2Net: A new multi-scale backbone architecture," Dec. 2021, *arXiv:1904.01169*.
[21] L. Yating, *Research and Implementation of Plant Classification Recognition Method Based on Residual Network*. Hohhot, China: Inner Mongolia Univ., 2021, doi: 10.27224/d.cnki.gnmdu.2021.000367.
[22] X. Zhi, N. Wenchang, Z. Longyang, and M. Ruimin, "A fine-grained image classification method based on attention data augmentation," *J. Guilin Univ. Electron. Sci. Technol.*, vol. 41, no. 6, pp. 496–503, 2021, doi: 10.16725/j.cnki.cn45-1351/tn.2021.06.006.
[23] Z. Yi, W. Hua, and T. Shuqin, "A review and case study of computer vision-based classification techniques for Chinese medicine tablets," *Comput. Appl.*, to be published. [Online]. Available: http://kns.cnki.net/kcms/detail/51.1307.TP.20211203.0002.002.html
[24] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
[25] H. Xue, *Research on Fine-Grained Image Classification Method Based on Deep Learning*. Beijing, China: Beijing Forestry Univ., 2020, doi: 10.26949/d.cnki.gblyu.2020.000975.
[26] L. Dongming, T. Peng, Z. Lijuan, L. Yu, and L. Shuangli, "Identification of doxasticity of antifungal herbs using improved dense connectivity network," *J. Agricult. Eng.*, vol. 38, no. 3, pp. 276–285, 2022.
[27] P. Dexin, "Research on apple tree disease image classification technology based on convolutional neural network," Jilin University, 2022, doi: 10.27162/d.cnki.gjlin.2022.006244.
[28] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4510–4520.
[29] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-Net: Efficient channel attention for deep convolutional neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 11531–11539.
[30] Z. Ting, D. Jianqiang, Z. Yanchen, L. Jigen, and H. Dingxing, "Research on improved ResNet-based image classification algorithm for Tibetan medicinal plants," *Electron. Technol. Softw. Eng.*, vol. 3, pp. 164–169, Jan. 2022.
[31] L. Xueyang and C. Jinda, "Research on image classification algorithm framework based on deep learning," *Packag. Eng.*, vol. 42, no. 21, pp. 181–187, 2021, doi: 10.19554/j.cnki.1001-3563.2021.21.025.
[32] S. Lei, L. Rong, J. Yitao, and S. Huaibo, "Seedling plant height measurement method based on ResNeXt monocular depth estimation," *J. Agricult. Eng.*, vol. 38, no. 3, pp. 155–163, 2022.
[33] J. Hu, Z. Chen, M. Yang, R. Zhang, and Y. Cui, "A multiscale fusion convolutional neural network for plant leaf recognition," *IEEE Signal Process. Lett.*, vol. 25, no. 6, pp. 853–857, Jun. 2018.
[34] X. Yaohua, Z. Weichuan, R. Jie, and J. Junfeng, "Small-sample fine-grained image classification based on adaptive feature fusion," *Comput. Eng. Appl.*, to be published. [Online]. Available: http://kns.cnki.net/kcms/detail/11.2127.TP.20220824.1419.014.html
[35] L. Weiwei, Z. Haiping, Q. Xiaoyan, G. Jing, L. Jia, Z. Zhongzhan, and L. Chunyi, "Research progress on identification and quality control of deer antler herbs," *Chin. J. Traditional Chin. Med.*, vol. 42, no. 21, pp. 4110–4114, 2017, doi: 10.19540/j.cnki.cjcmm.20171016.002.
[36] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 618–626.
[37] J. Bowen, "An image classification method based on improved ResNet model," *Modern Inf. Technol.*, vol. 6, no. 12, pp. 83–85, 2022, doi: 10.19850/j.cnki.2096-4706.2022.012.021.
[38] Z. Yongjie, L. Caiming, Z. Yan, T. Yu, and N. Kanghui, "An anomaly analysis system based on open network dataset," *Modern Inf. Technol.*, vol. 4, no. 19, pp. 136–138,2020, doi: 10.19850/j.cnki.2096-4706.2020.19.035.

DONGMING LI received the Ph.D. degree from the Changchun University of Science and Technology, Changchun, China, in 2021.

He was a Visiting Scholar with CSIRO, Australia, from 2016 to 2017. He joined Wuxi University, Wuxi, where he is currently a Professor carrying out teaching and research on projects. His current research interests include computer vision, image analysis, and pattern recognition.

RUI YAO received the B.Mgt. degree from the Changchun University of Finance and Economics, Changchun, China, in 2021.

She is currently pursuing the master's degree in computer sciences with Jilin Agricultural University. Her research interest includes the computer vision image classification of Chinese herbs.

CHENGLIN YANG received the B.E. degree from Northeast Electric Power University, Jilin, China, in 2021. He is currently pursuing the master's degree in computer and science with Jilin Agricultural University.

His research interest includes the computer vision image classification of Chinese herbs.

CHUNXI ZHAO received the master's degree in education from Jilin Agricultural University, in 2008.

He is currently with the Center for Informatization, Jilin Agricultural University. His current research interests include computer networks and artificial intelligence applications.

LIJUAN ZHANG received the M.E. and Ph.D. degrees from the Changchun University of Science and Technology, Changchun, China, in 2004 and 2015, respectively.

She was a Visiting Scholar with CSIRO, Australia, from 2016 to 2017. She is currently a Professor with the College of Internet of Things Engineering, Wuxi University. Her research interests include image restoration, computer vision, and image analysis.