IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 29, 2020, accepted July 14, 2020, date of publication July 20, 2020, date of current version August 11, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3010496

An Automatic Garbage Classification System Based on Deep Learning

ZHUANG KANG, JIE YANG, GUILAN LI, AND ZEYI ZHANG
School of Electrical Engineering and Automation, Jiangxi University of Science and Technology, Ganzhou 341000, China
Corresponding authors: Jie Yang (yangjie@jxust.edu.cn) and Zeyi Zhang (zhangzy@jxust.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61763016.

ABSTRACT Garbage classification has always been an important issue in environmental protection, resource recycling and social livelihood. In order to improve the efficiency of front-end garbage collection, an automatic garbage classification system is proposed based on deep learning. Firstly, the overall system of the garbage bin is designed, including the hardware structure and the mobile app. Secondly, the proposed garbage classification algorithm is based on ResNet-34 algorithm, and its network structure is further optimized by three aspects, including the multi feature fusion of input images, the feature reuse of the residual unit, and the design of a new activation function. Finally, the superiority of the proposed classification algorithm is verified with the constructed garbage data. The classification accuracy of the proposed algorithm is enhanced by 1.01%. The experimental results show that the classification accuracy is as high as 99%, the classification cycle of the system is as quick as 0.95 s.

INDEX TERMS Artificial intelligence, deep learning, image classification, neural networks.

I. INTRODUCTION

With the rapid development of economy and the improvement of people's living standards, the amount of garbage is increasing rapidly. According to the latest report of International Lianhe Zaobao, the global garbage volume will increase by 70% by 2050, and the task of garbage classification will be even more arduous. Scholars at home and abroad have done a lot of researches on garbage classification, but most of the proposed schemes are innovations of terminal recycling method [1]–[7]. In 2019, China started to require residential garbage classification, in which case the front-end collection is highly dependent on people's awareness. Therefore, the efficiency of the garbage classification still needs to be improved. It is of great academic value and practical significance to study an effective automatic garbage classification method.

For the commercialization of intelligent garbage bin, there are some attempts made by foreign companies, such as NASTAR in the United States, Zeton, and the Japanese somatosensory intelligent garbage bin. There are a lot of researches on the intelligence garbage bin in academia. Kano *et al.* [8] proposed a group-control robot system, which provides a decentralized control scheme for indoor garbage collection. But it depends on the cooperation between robots and takes a long time. Wesley *et al.* [9] proposed a low-cost intelligence garbage bin. It optimized the treatment of waste by taking the capacitance and spectrum of waste as the classification standard. However, this method is highly dependent on the surrounding environment, its classification still has limitations. The research of garbage classification system is mature, but the accuracy and speed of garbage classification still need to be improved. And there are few researches on garbage classification based on deep learning. At present, deep learning technology is widely used in image classification, and has some remarkable achievements. ResNet proposed by He *et al.* [10] achieved an accuracy of 93.03% and 95.51% on cifar-10 and Imagenet respectively. Zhong *et al.* [11] combined residual and Inception blocks, achieved the accuracy of 99.66%, 98.04%, and 95.32% on MNIST, SVHN, and cifar-10, respectively. In the public data set, using deep learning algorithm can get high classification accuracy, which still can be improved. It also provides a new idea for the research of garbage classification.

With the rapid development of artificial intelligence, image classification algorithms are gradually diversified. The mainstream algorithms are AlexNet, VGG, Inception and ResNet. Scholars implemented experiments using these

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

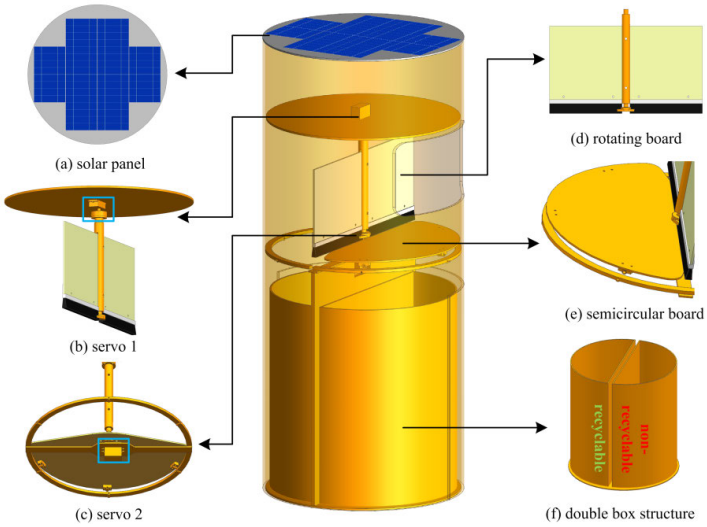


FIGURE 1. Structure of the trash bin.

methods and made attempts to modify in order to gain better results [12]–[17]. These are helpful for the improvement of network structure in this paper. Thanks to the nonlinear characteristics of activation function, neural network with improved activation function has shown good results [18]–[24]. In order to learn the distribution characteristics of the nonlinear data better, some improvements of the network are mainly focused on the network's depth. From AlexNet [25], VGG, GoogleNet [26] to ResNet-152, the problem of gradient dispersion becomes more and more prominent. Among most popular algorithms used to solve this problem, the effect of improving the activation function is outstanding. Reference [27] defined the activation function as a mapping which can be derived almost everywhere. The ReLU [28] (corrected linear unit) proposed by Nair has the characteristics of fast convergence speed, simple calculation and strong sparseness. Later, Leaky ReLU [29], ELU, and PReLU are emerging one after another, and all have achieved good results. RReLU is particular effective when trained with small data set. The SLU [30] is proposed, which effectively solved the problem of migration and vanishment of gradient. The elu-softsign is proposed in [31], which effectively alleviated the invalidity of negative distributed sample and improved the classification accuracy. These provide ideas for improving the activation function of image classification network.

This work is organized as follows. Sec. II briefly describes the hardware design of the system. Sec. III elaborates three modifications of the proposed classification algorithm, including the multi feature fusion of input images, the feature reuse of the residual unit, and the design of a new activation function. In section IV, the effect of three modifications are carefully analyzed. In section V, the automatic garbage classification system is integrated with the proposed classification algorithm and necessary hardware and tested. Finally, concluding remarks are addressed in Sec. VI.

II. HARDWARE DESIGN

A. STRUCTURE DESIGN

First, the garbage bin is modeled. The overall model is shown in Fig. 1. According to the structure design of the garbage bin, the product is made. The system height is 915 cm and the inner diameter is 390 cm. It is divided into three layers: upper, middle, and lower. The upper layer is equipped with solar panel, battery, light belt, raspberry pi and camera. The middle layer is equipped with mechanical transmission device, including two servos, rotating board, and semicircle board. Under the semicircle board, there are balls installed in order to reduce friction with the circular track. The lower layer is equipped with double boxes, including recyclable and non-recyclable boxes. As shown in Fig. 1, the specific device is as follows.

Solar power generation device: see Fig. 1 (a). The system uses solar energy for power supply. It's energy-saving and environmental friendly.

Servo: The servo module is ds3218. The torque is 20 kg·cm. The mass is only 60 g. And the working voltage is 5.0 $\sim$ 6.5 V.

Rotating board: it is connected with servo 1, as shown in Fig. 1 (b). It is fixed on the central column. When it rotates, it will bring the garbage identified as recyclable to the recycable box. The bottom is equipped with brush to avoid hardware friction, as shown in Fig. 1 (d). The height of the board is 150 cm and the width is 380 cm.

Semicircle board: it is connected with servo 2, as shown in Fig. 1 (c). It is perpendicular to the rotating board. When rotating, the garbage which is identified as non-recyclable would fall into the non-recyclable box. The bottom is equipped with a bearing seat, connected with the circular track through several balls to reduce friction, as shown in Fig. 1 (e). The diameter of the semicircle board is 384 cm.

Double box: located at the lower part of the garbage bin, it is divided into recyclable and non-recyclable boxes,

as shown in Fig. 1 (f). The height is 150 cm and the diameter is 384 cm.

B. THE CHOOSE OF THE DEVELOPMENT BOARD

According to the design requirements of the system, the development board used shall include an image receiving unit, which can be connected with the camera to obtain the pictures. It shall include an image processing unit, which can process the collected images and finish category identification. It shall include a servo control unit, which is connected with servo 1 and servo 2. Through the control of the servos, the semicircle board and the rotating board can rotate to drop the garbage into the right box. Therefore, raspberry pi is selected as the development board of the system. The raspberry pi is a single-chip microcomputer equipped with Linux kernel, which is widely used and can realize many intelligent demands [32]–[34]. The raspberry pi 3B + has a 64 bit, 1.4 GHz four-core CPU, and is equipped with an integrated Broadcom Videocore-IV GPU. The serial interface of camera meets the requirements of the development board in the system.

III. ALGORITHM DESIGN

A. BASIC MODEL SELECTION

To achieve an accurate classification of garbage types for the collected pictures, excellent classification algorithm is needed. Therefore, we have carefully compared and selected the basic model. Current popular image classification algorithms are Inception, KNN, SVM, AlexNet, VGG, Migration Learning, and CNN, among which AlexNet, VGG, Inception V3, and ResNet are the typical models. To select an appropriate image classification model, we use the existing cifar-10 data set to compare LeNet, AlexNet, VGG, Inception V3, and ResNet models. 200 steps of iteration training are performed. After the training, the accuracy and loss of each method are compared as shown in Table. 1.

It can be seen from the table that after 200 steps of training, Inception and ResNet perform well in both train data set and val data set. However, the network structure of the Inception series is more complex than ResNet series, which can result in huge amount of parameter calculation. Hence, ResNet series is selected as the basic model. ResNet-34 and ResNet-152 show good performance in both accuracy and loss respectively. However, because the deeper the model, the slower the calculation speed. Therefore, ResNet-34 is chosen as the basic model.

TABLE 1. Comparison of accuracy and loss for different networks on cifar-10 data set.

Network	Train		Val	
	Accuracy	Loss	Accuracy	Loss
LeNet	0.757 8	0.740 6	0.739 5	0.842 1
AlexNet	0.864 4	0.267 3	0.755 0	1.566 7
VGG-11	0.972 0	0.122 5	0.842 7	1.412 1
VGG-13	0.967 5	0.183 4	0.873 7	0.968 9
VGG-16	0.977 7	0.176 4	0.883 1	0.959 9
VGG-19	0.981 7	0.175 2	0.889 2	0.814 9
Inception V1	0.897 8	0.326 5	0.827 6	0.563 4
Inception V2	0.969 9	0.182 3	0.911 4	0.482 3
Inception V3	0.975 4	0.254 0	0.915 7	0.476 3
ResNet-18	0.978 2	0.195 6	0.910 8	0.459 3
ResNet-34	0.987 4	0.170 9	0.922 6	0.451 0
ResNet-50	0.984 1	0.175 5	0.922 5	0.411 1
ResNet-101	0.985 2	0.171 6	0.920 4	0.420 5
ResNet-152	0.986 8	0.165 0	0.927 9	0.392 3

B. MODEL MODIFICATION

1) MODEL STRUCTURE MODIFICATION

ResNet-34 is proposed by He Kaiming, which originally has 34 layers. In addition to a 7*7 convolution layer and a FC layer, it arranges 3*3 convolution in the order of [3], [4], [6], [3], and performs a dimension reduction at each group junction. This network has a good effect in alleviating the gradient vanishing. Our improved algorithm is based on ResNet-34. The improvements include three aspects, which are multi feature fusion of input images, feature reuse of the residual unit and improved activation function. The overall improved model is shown in Fig. 2.

a: MULTI FEATURE FUSION

In this paper, the input of the network is optimized. The feature of the input is extracted in parallel, and the feature fusion is carried out at the end of the parallel structure. The specific method is to adopt three parallel routes. To ensure the integrity of the size of the input image, in all convolution operations of multi feature fusion, the step is set as 1. At the same time, in order to increase the number of feature layers and also apply to the first residual unit, the convolution kernel of multi feature fusion are 4*4. The first route carries out one convolution operation. The convolution kernel size is 1*1. The second route carries out two convolution operations. The convolution kernel sizes are 1*1 and 3*3. The third route carries out three convolution operations. The convolution kernel sizes are 1*1, 3*3, 3*3. Finally, the outputs of the three parallel routes are used for feature fusion to further extract image features, so as to improve the performance of the model.

b: RESIDUAL UNIT MODIFICATION

The residual structure has a good effect in preventing gradient exploding and vanishing. Feature reuse is beneficial to feature extraction. In this paper, the residual unit is improved. In the process of feature extraction, the first residual block outputs 224*224 feature information to the input of the third residual block through down sampling, and the input scale is changed to 112*112. In the same way, the output feature information of the first residual block is sent to the input and output of the fourth residual block through multiple down sampling, and the size of the feature are 56*56 and 28*28, respectively. Similar to the method used in the first residual block, the output scale of the second residual block is 112*112. The output is subsampled and sent to the input and output of the

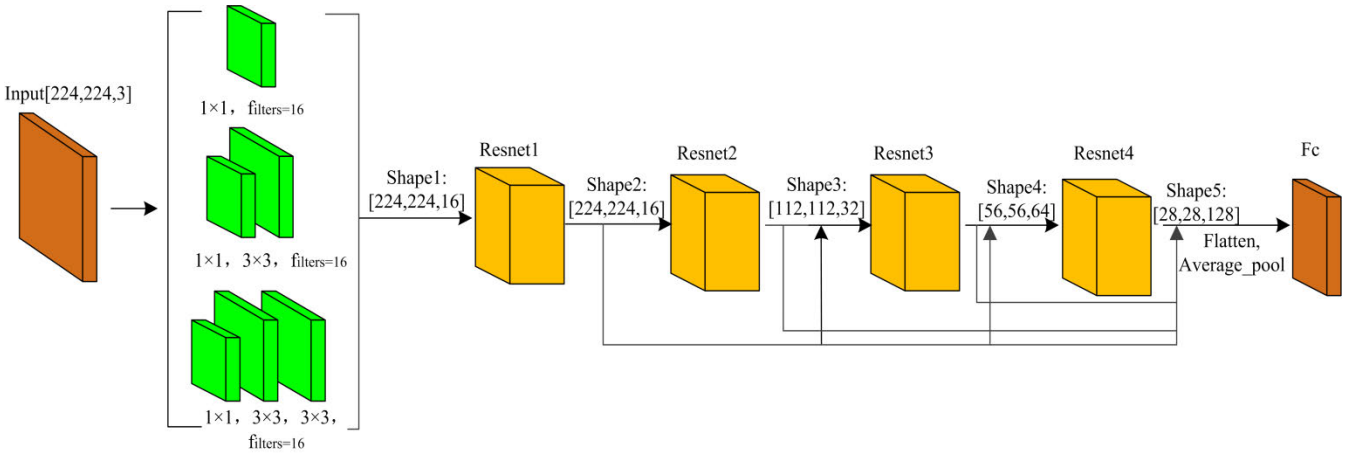


FIGURE 2. Network structure.

fourth residual block. In the same way, the output of the third residual block is subsampled and sent to the output of the fourth residual block.

2) ACTIVATION FUNCTION MODIFICATION

ReLU can effectively solve the problem of gradient exploding and gradient vanishing with the constant gradient 1. However, for the case with large eigenvalue, it has no constraint ability, which may cause gradient exploding, thus affecting the learning ability of the model. For this problem, this paper refers to the use of swish, an effective activation function in deep learning. Swish [35] is chosen as the basic function. Then apply swish to function softplus(x), softsign(x) to build a new function, which can be described as

$$g(x) = x \cdot softsign(softplus(x)) \tag{1}$$

The detailed function is shown below:

$$g(x) = x \cdot \ln(e^x + 1)/(1 + \left|\ln(e^x + 1)\right|), x \in (-\infty, +\infty) \tag{2}$$

This new function combines the advantages of both softplus(x) and softsign(x). It can not only restrain the eigenvalues, but also reduce the calculation of parameters and speed up the convergence. The combined activation function $g(x) - h(x)$ is constructed. Reference [36] proposed a general form of the function:

$$g(x) - h(x) = \begin{cases} x \cdot \ln(e^x + 1)/(1 + \ln(e^x + 1)), & x \in (-\infty, 0); \\ k_{m+1}x + b_{m+1}, & x \in [0, x_{m+1}); \\ \vdots & \\ k_n x + b_n, & x \in [x_n, +\infty). \end{cases} \tag{3}$$

In the training process of convolution neural network, we observed that when the piecewise point of the activation function is set between 0 and 1, it has a great influence on the forward propagation of feature, the backward propagation of gradient and the curve change.

In this paper, the function is differentiable at point 0, and the slope of the function changes to 1 quickly, After many experiments, the piecewise of the function is set as 0.1. The function is as follows:

$$\begin{cases} x \cdot \ln(e^x + 1)/(1 + \ln(e^x + 1)), & x \in (-\infty, 0); \\ x \ln 2/(1 + \ln 2), & x \in [0, 0.1); \\ x + 0.1 \ln 2/(1 + \ln 2) - 0.1, & x \in [0.1, +\infty). \end{cases} \tag{4}$$

At the beginning of the experiment, the model was directly over fitted. It was found that the slope of the function changed too quick, and the slope of the curve could not be changed from $\ln 2/(1 + \ln 2)$ to 1 directly. So, a linear function was added in (0.1,1) to play the role of buffering. After many experiments, it was selected as (0.1,0.2).

And the modified function is

$$\begin{cases} x \cdot \ln(e^x + 1)/(1 + \ln(e^x + 1)), & x \in (-\infty, 0); \\ x \ln 2/(1 + \ln 2), & x \in [0, 0.1); \\ (2 - \ln 2/(1 + \ln 2))x & \\ \quad +0.2 \ln 2/(1 + \ln 2) - 0.2, & x \in [0.1, 0.2); \\ x, & x \in [0.2, +\infty). \end{cases} \tag{5}$$

The output mean value of the ReLU is compared with that of the new function, and the probability model of the parameter is set as $p(x; \alpha)$, the positive input is x^+, the negative input is x^-, α is the probability of the input x. After the nonlinear transformation, the output mean value of the new function is

$$E_{ours}(x) = \sum \alpha f_{ours}(x) = \sum \alpha x^+ + \sum \alpha x^- \tag{6}$$

In which:

$$\begin{aligned} \sum \alpha \mathrm{x}^+ = & \sum \alpha \mathrm{x}^+ + \sum \alpha \ln 2/(1 + \ln 2)x^+ \\ & + \sum \alpha \left[(2 - \ln 2/(1 + \ln 2))x^+ \right. \\ & \left. +0.2 \ln 2/(1 + \ln 2) - 0.2\right] \end{aligned} \tag{7}$$

$$\sum \alpha \mathrm{x}^- = \sum \alpha \mathrm{x}^- \cdot \ln(\exp[x^-]+1)/(1+\ln(\exp[x^-]+1)) \tag{8}$$

The output mean value of the ReLU is

$$E_{ReLU}(x) = \sum \alpha f_{ReLU}(x) = \sum wx^{+} + 0 \tag{9}$$

$E_{ReLU}(x)$ is always positive. The output value of the new function $E_{ours}(x)$ has both positive and negative values, which makes the mean value closer to zero. It can accelerate the update of parameters and the convergence of the model.

IV. DISCUSSION

A. EVALUATION INDEX

In order to verify the performance of the improved ResNet algorithm and activation function, evaluation indexes need to be set. There are generally recognized evaluation indexes in academia, such as accuracy, loss, IOU, and recall. This paper uses accuracy and loss as evaluation indexes.

Because there is softmax layer in the network used in the experiment, the evaluation index loss adopts the cross entropy loss function, whose expression is:

$$\mathrm{L} = -\sum_{\mathrm{K}=0}^{100} \mathrm{y_k} \log \mathrm{P}\,(\mathrm{y} = \mathrm{k}) \tag{10}$$

$$\mathrm{P}\,(\mathrm{y} = \mathrm{j}) = \exp[\alpha_\mathrm{j}] / \sum_{\mathrm{K}=1}^{100} \exp[\alpha_\mathrm{k}] \tag{11}$$

In the equation, $\mathrm{a_j}$ is the output value of the js-th neuron in the softmax classification layer; $\mathrm{P}\,(\mathrm{y} = \mathrm{j})$ is the classification probability of the j class; $\mathrm{y_k}$ represents the classification indicator value of the K class; if it belongs to K class, the corresponding probability output is 1, otherwise 0; L is the output value of the cross entropy loss of the 100 classes.

Accuracy is the proportion of the correct results to the total number of samples. Its function is:

$$\mathrm{Accuracy} = \mathrm{N_T}/\mathrm{N} \tag{12}$$

where, Accuracy is the classification accuracy and $\mathrm{N_T}$ is the number of correct results, N is the total number of samples.

When the accuracy and loss of the model tend to be stable in the later stage of training, they fluctuate in a small range. Therefore, the average value of the accuracy and loss of last 10 rounds is taken as the final accuracy and loss. The expression is as follows:

$$\mathrm{AVG_{Accuracy}} = \sum_{j=190}^{200} \mathrm{Accuracy_j}/10 \tag{13}$$

$$\mathrm{AVG_{loss}} = \sum_{\mathrm{j}=190}^{200} \mathrm{loss_j}/10 \tag{14}$$

In the formula, $\mathrm{Accuracy_j}$ is the classification accuracy after each round of training in the later stage, $\mathrm{AVG_{Accuracy}}$ is the average classification accuracy, $\mathrm{loss_j}$ is the loss after each round of training in the later stage, $\mathrm{AVG_{loss}}$ is the average loss value. The epoch of model training is set as 200. Since the data sets used in this paper are divided into train data set and val data set, they should be calculated separately, and then recorded as train acc, val acc, train loss, and val loss.

B. EXPERIMENTAL ENVIRONMENT AND DATA SET

1) EXPERIMENTAL ENVIRONMENT

In this paper, ResNet-34 is selected for the experiment, which is verified on cifar-10 and MNIST data set. The computer configuration of the experiment platform is i5-6500 CPU, 8gram, 64 bits Windows 7, and GPU which includes two Tesla-P100.

2) DATA SET PREPARATION AND PRODUCTION

The constructed garbage data set of this paper comes from 4 168 pictures taken online and in real life. See Table. 2.

TABLE 2. Garbage data set.

Recyclable data set		Non-recyclable data set	
Book	217	Plastic	425
Bottle	417	Bulb	206
Towel	340	Packing bag	296
Carton	205	Battery	276
Metal	278	Banana peel	343
Broken glass	376	Leaf	209
Paper ball	286	Orange peel	294
Book	217	Plastic	425

C. ACTIVATION FUNCTION

1) PARAMETER SETTING

Configuration of experimental training parameters: the batch size is 32, and the initial learning rate is 0.001. In this paper, the number of training epoch is 200, and the learning rate decreases gradually according to the number of steps. It is 0.000 1 after 80 steps, 0.000 01 after 120 steps, 0.000 001 after 160 steps, and 0.000 000 1 after 180 steps. The learning rate setting is shown in Fig. 3 below.

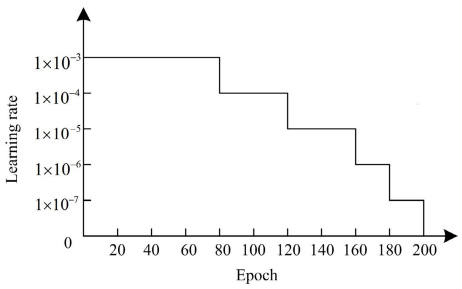


FIGURE 3. Learning rate - epoch curve.

For the design of the new activation function, the idea of piecewise function is adopted, and the selection of the piecewise point have great influence on the experiment. The tests are conducted on MNIST and cifar-10. Finally, 0.1 is determined as the piecewise point. The results are shown in Table. 3 and Table. 4 below.

2) EXPERIMENT IN MNIST DATA SET

MNIST contains 10 kinds of handwritten Arabic numerals, including 60 000 training data and 10 000 testing data. The image gray level is 8 and the resolution is 28*28. The test results are shown in Table. 5.

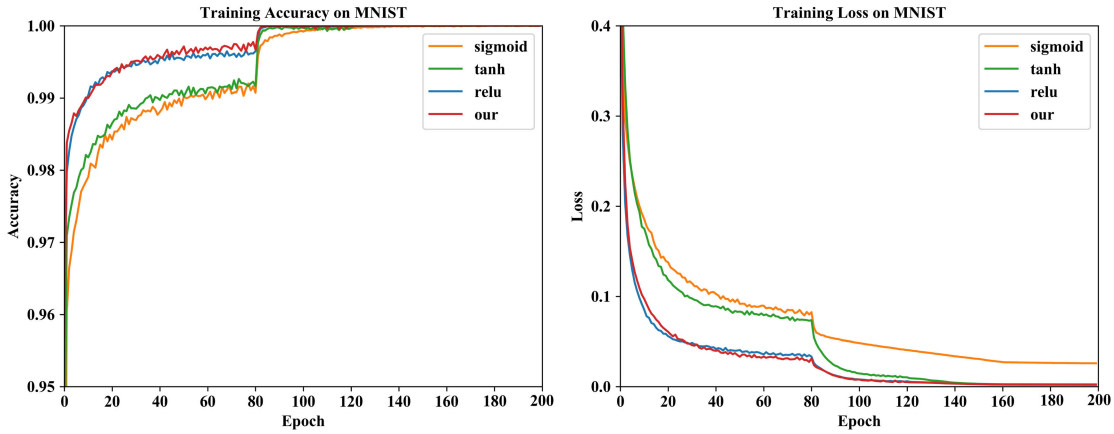


FIGURE 4. Training diagram on MNIST data set.

TABLE 3. The experiment of different piecewise points of activation function on MNIST data set.

Value		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Train	Acc	**0.999 8**	0.994 4	0.993 6	0.993 9	0.994 3	0.995 4	0.992 4	0.992 3	0.993 8
	Loss	**0.002 0**	0.002 3	0.002 6	0.002 8	0.002 5	0.002 4	0.003 0	0.003 1	0.002 9
Val	Acc	**0.995 1**	0.994 4	0.993 7	0.994 2	0.993 8	0.993 6	0.992 0	0.992 1	0.993 5
	Loss	**0.026 6**	0.030 3	0.038 6	0.040 2	0.039 6	0.038 0	0.042 0	0.039 5	0.039 6

TABLE 4. The experiment of different the piecewise points of activation function on cifar-10 data set.

Value		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Train	Acc	**0.991 3**	0.986 9	0.985 4	0.983 6	0.989 1	0.989 5	0.987 4	0.979 9	0.975 4
	Loss	**0.157 7**	0.165 2	0.161 8	0.160 3	0.174 0	0.168 6	0.166 1	0.171 8	0.172 6
Val	Acc	**0.925 1**	0.913 5	0.913 8	0.909 1	0.910 6	0.916 5	0.909 2	0.902 7	0.903 6
	Loss	**0.456 5**	0.580 9	0.525 6	0.538 4	0.464 4	0.467 5	0.461 9	0.468 5	0.460 7

TABLE 5. Test results of various activation functions on MNIST.

Activation function	Train		Val	
	Accuracy	Loss	Accuracy	Loss
Sigmoid	0.994 3	0.026 0	0.993 5	0.059 9
tanh	0.996 8	0.002 1	0.994 4	0.038 2
ReLU	0.998 9	0.002 5	0.993 1	0.032 3
ours	0.999 8	0.002 0	0.995 1	0.026 6

It can be seen from Table. 5 that after the activation function designed in this paper is trained on MNIST, the accuracy is 99.98%, which is 0.09% higher than the current maximum result, and the loss is lower than other activation functions; on the val data set, the accuracy is 99.51%, and the loss value is 0.026 6, showing good performance.

At the same time, the intermediate results of training are plotted as shown in Fig. 4. It can be seen from Fig. 4 that the accuracy and loss curves after improving the activation function tend to be gentle in the later training period, only fluctuate in a small range.

It shows that the function converges successfully on ResNet-34. The effect of the improved activation function on MNIST is slightly better than the current optimal ReLU. And it is quite stable after about 80 rounds. The convergence speed of tanh and sigmoid functions is much slower than that of the improved function. Accuracy of the improved function is similar to ReLU, whereas the loss is lower than tanh, sigmoid and ReLU. It can be seen that the improved activation function has better performance in the convergence speed and has better accuracy when used to train the neural network model.

3) EXPERIMENT ON CIFAR-10 DATA SET

Train ResNet-34 on the cifar-10 data set. The cifar-10 data set contains 10 types of objects, ranging from 0 to 9. It is composed of 50 000 training data and 10 000 testing data. The resolution of the picture is 32*32. See Table. 6 for the test results.

It can be seen from Table. 6 that after the activation function designed in this paper is used to train on cifar-10 data set, the accuracy of the train data set reaches 99.13%, which is 2.66% higher than the highest record of classical activation function, and the loss is 0.157 7, which is 0.14 lower than the lowest loss of classical activation function. In the val data set, the accuracy is 0.85% higher than that of classical activation function, and the loss is 0.020 2 lower.

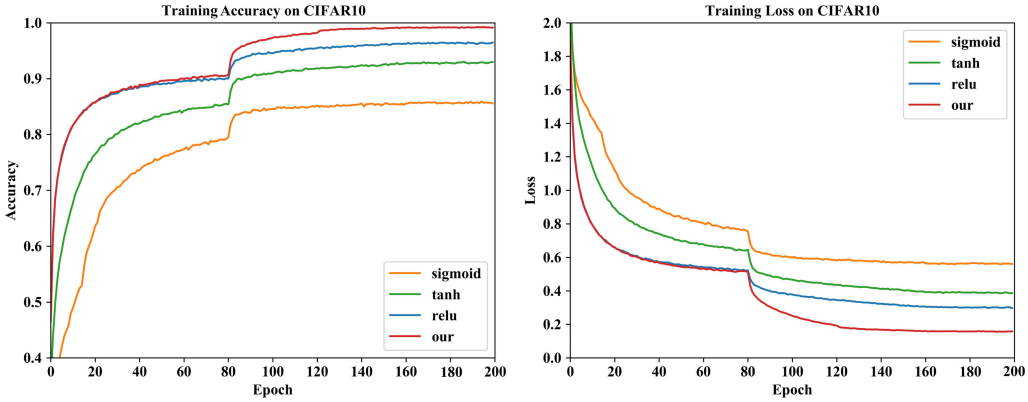


FIGURE 5. Training diagram on cifar-10 data set.

TABLE 6. Test results of various activation functions on cifar-10.

Activation function	Train		Val	
	Accuracy	Loss	Accuracy	Loss
Sigmoid	0.855 7	0.560 8	0.832 5	0.628 1
tanh	0.929 8	0.386 5	0.885 1	0.541 7
ReLU	0.964 7	0.297 7	0.906 6	0.496 7
ours	0.991 3	0.157 7	0.925 1	0.456 5

The improved activation function shows good performance in cifar-10, and effectively avoids over-fitting. At the same time, the whole intermediate results of training are plotted and shown in Fig. 5.

It can be seen from Fig. 5 that the accuracy and loss curves with the improved activation function tend to be gentle in the later training period, only fluctuate in a small range. It shows that the function converges successfully on ResNet-34. The convergence speed on cifar-10 data set is similar to that of ReLU function, and it is nearly stable after 80 rounds, and it has higher accuracy and lower loss than that of ReLU, sigmoid and tanh. It shows stronger robustness and generalization ability.

D. ALGORITHM EXPERIMENT

This paper mainly introduces three improvements based on ResNet-34. First, there are two structural improvements, A and B. Second, this paper optimizes the activation function, which is named method C. Method ALL integrates the above three methods. It is shown in Table. 7. The four changes are tested respectively.

1) EXPERIMENT ON MNIST DATA SET

In this paper, ResNet-34 and the improved four algorithms are respectively used to conduct 200-step training on MNIST data set. The stable accuracy and loss on the train data set and val data set after training are shown in Table. 8.

It can be seen from Table. 8 that method A has better performance on loss, method B performs better in the train data set. Method C performs better in the val data set. And

TABLE 7. Modification method.

Name	Method
ResNet-34-A	Multi feature fusion
ResNet-34-B	Feature reuse
ResNet-34-C	Modified activation function
ResNet-34-ALL	Integrate ResNet-34-A, ResNet-34-B, ResNet-34-C

TABLE 8. Experimental results on MNIST data set.

Method	Train		Val	
	Accuracy	Loss	Accuracy	Loss
ResNet-34	0.997 5	0.034 0	0.994 8	0.040 5
ResNet-34-A	0.998 7	0.019 3	0.995 2	0.030 5
ResNet-34-B	0.999 0	0.018 5	0.995 1	0.032 3
ResNet-34-C	0.997 7	0.031 0	0.995 7	0.038 2
ResNet-34-ALL	0.999 3	0.016 0	0.996 6	0.028 2

Method ALL has the best performance. Compared with the original algorithm, the loss of method ALL in the train set is 0.018 lower, the accuracy in train data set is 18% higher, the loss in val data set is 0.012 3 lower, and the accuracy is 18% higher. To observe the performance in the training process, the training accuracy and loss are shown in Fig. 6.

It can be seen from the Fig. 6 that when accuracy and loss are close to convergence point, the effect of method A and B is similar. Method C is slightly better than the original method, and method ALL has the best performance.

2) EXPERIMENT ON CIFAR-10 DATA SET

In this paper, ResNet-34 and the improved four algorithms are respectively used to conduct 200-step training and test on cifar-10. The stable accuracy and loss on the train data set and val data set are shown in Table. 9.

In Table. 9, the accuracy and loss of method A, B, and C are improved, whereas the combined method ALL is better than the above algorithm. Compared with the original algorithm, the accuracy of method ALL on the train data set is increased by 1.75%, the loss is 0.185 6 lower, the accuracy on the val

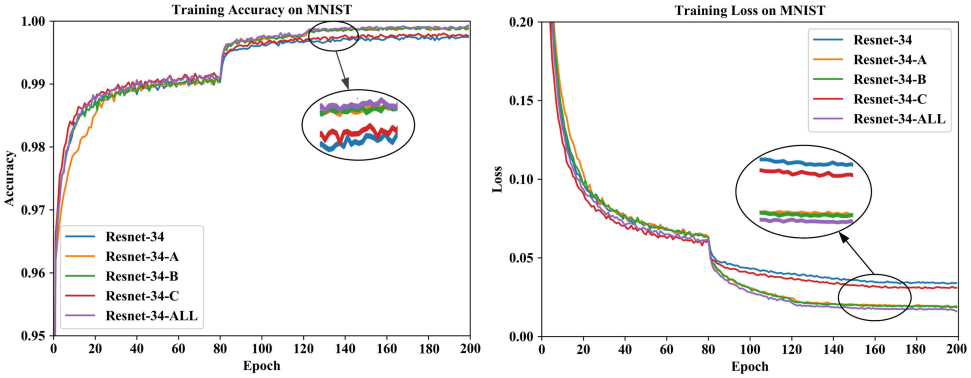


FIGURE 6. Training accuracy and loss on MNIST.

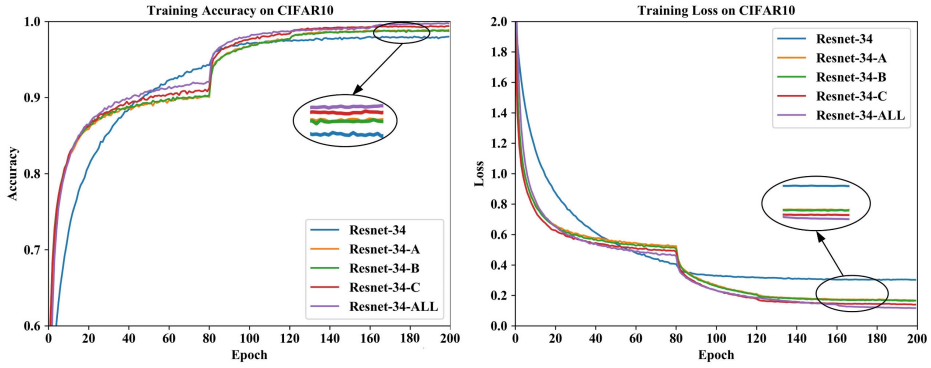


FIGURE 7. Training accuracy and loss on cifar-10.

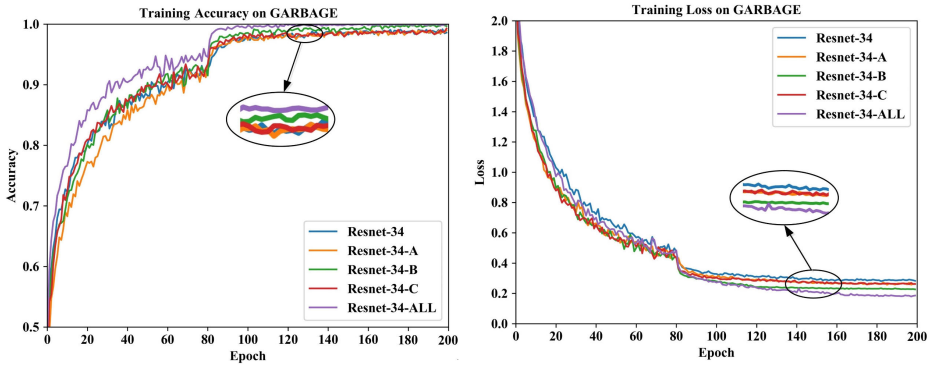


FIGURE 8. Training accuracy and loss on garbage data set.

data set is increased by 4.781%, and the loss is 0.259 lower. In addition, the process of the training accuracy and loss are shown in Fig. 7.

It can be seen from the Fig. 7 that when the accuracy and loss are close to convergence point, the effect of method A and method B is similar. Method C is slightly better than method A and method B, and among them, method ALL has the best performance.

3) EXPERIMENT ON GARBAGE DATA SET

In this paper, ResNet-34 and the improved four algorithms are respectively used to conduct 200-step training on garbage data set. The stable accuracy and loss on train and val data set are shown in Table. 10.

TABLE 9. Experimental results on cifar-10.

Method	Train		Val	
	Accuracy	Loss	Accuracy	Loss
ResNet-34	0.980 0	0.302 6	0.875 4	0.708 6
ResNet-34-A	0.987 8	0.166 7	0.918 6	0.476 4
ResNet-34-B	0.987 4	0.165 8	0.919 0	0.460 4
ResNet-34-C	0.993 9	0.139 7	0.919 7	0.477 0
ResNet-34-ALL	0.997 5	0.117 0	0.923 2	0.449 6

In Table. 10, the performance of method A and B are both better than method C. The combination of all three improvements is the best. Compared with the original algorithm, the accuracy of method ALL in train data set is increased by

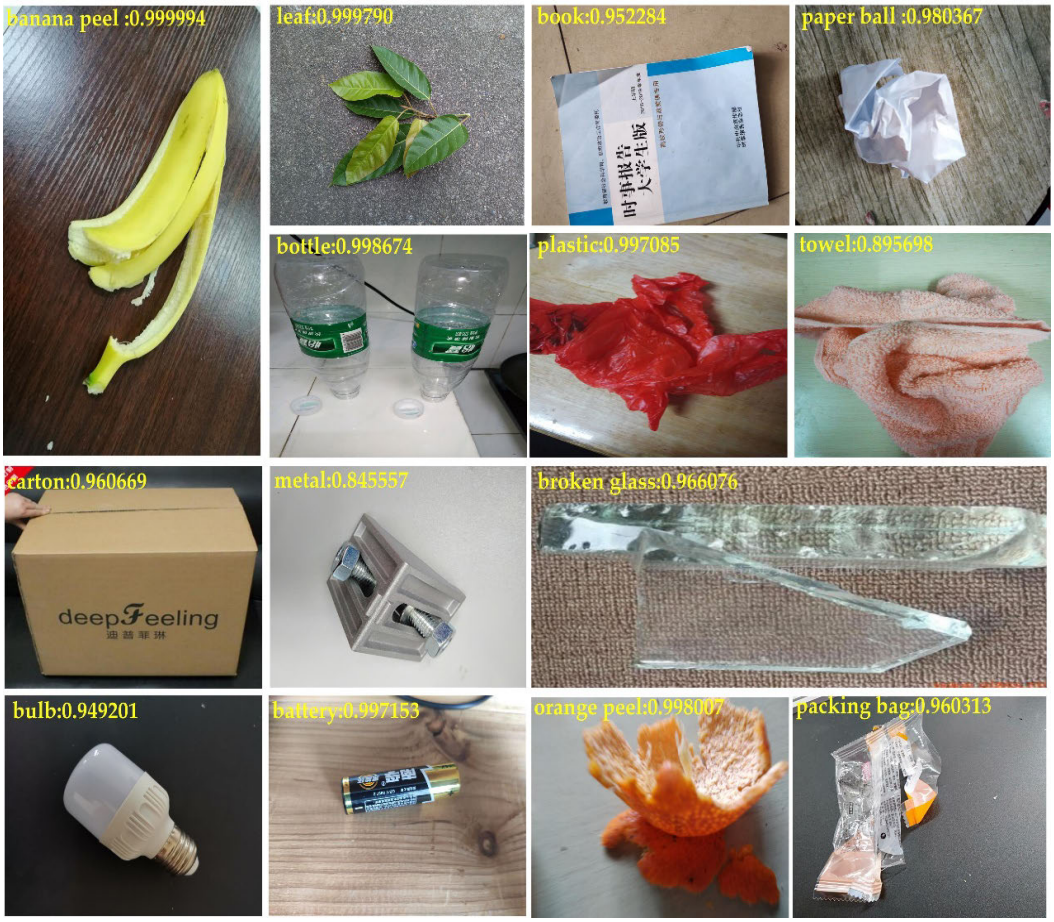


FIGURE 9. Test results.

TABLE 10. Experimental results on garbage data set.

Method	Train		Val	
	Accuracy	Loss	Accuracy	Loss
ResNet-34	0.985 9	0.263 1	0.854 0	0.852 6
ResNet-34-A	0.994 1	0.231 7	0.882 9	0.830 3
ResNet-34-B	0.999 5	0.226 7	0.889 6	0.900 6
ResNet-34-C	0.992 8	0.262 2	0.870 9	0.796 0
ResNet-34-ALL	0.999 6	0.222 9	0.899 6	0.680 6

1.37%, the loss is reduced by 0.040 2, the accuracy in val data set is increased by 4.56%, and the loss is reduced by 0.172. In addition, the process of the training accuracy and loss are shown in Fig 8.

It can be seen from Fig. 8 that when the accuracy and loss are close to convergence point, the effect of method A and C is similar. Method B is slightly better than method A and C, and method ALL is the best.

Looking at the above three experiments, method ALL has better generalization performance and robustness for garbage classification task, and can complete garbage automatic classification task well.

V. RESULTS AND ANALYSIS

A. HARDWARE TEST

14 kinds of garbage are identified, and the hardware system is tested. The servo can drive the rotating board and semicircle board well to realize the correct classification of the garbage. The app can also locate the position of the garbage bin and display the capacity of the recyclable and non-recyclable boxes. The solar panel is placed in the sun to charge the battery, and the battery supplies power to the system. The lamp is to provide the system with lighting at night. The results show that the solar panel can absorb the sunlight normally, and the hardware system can work stably.

B. OBJECT CLASSIFICATION TEST

To test the classification function of the model, the process is as follows: select 14 items as test objects, which includes book, bottle, towel, paper box, metal, broken glass, paper ball, plastic, light bulb, packing box, battery, banana skin, leaves and orange skin. Collect a large number of corresponding pictures. Each item has more than 200 pictures. They consist of a total of 416 8 pictures, which constitute the train data set of this paper. After retraining the model on the garbage data

TABLE 11. The system test result.

Type	Results	Semicircle board	Rotating board	Box
Book	Recyclable	Inaction	Action	Recyclable
Plastic	Non-recyclable	Action	Inaction	Non-recyclable
Bottle	Recyclable	Inaction	Action	Recyclable
Bulb	Non-recyclable	Action	Inaction	Non-recyclable
Towel	Recyclable	Inaction	Action	Recyclable
Packing bag	Non-recyclable	Action	Inaction	Non-recyclable
Carton	Recyclable	Inaction	Action	Recyclable
Battery	Non-recyclable	Action	Inaction	Non-recyclable
Metal	Recyclable	Inaction	Action	Recyclable
Banana peel	Non-recyclable	Action	Inaction	Non-recyclable
Broken glass	Recyclable	Inaction	Action	Recyclable
Leaf	Non-recyclable	Action	Inaction	Non-recyclable
Paper ball	Recyclable	Inaction	Action	Recyclable
Orange peel	Non-recyclable	Action	Inaction	Non-recyclable

set. The garbage pictures collected by the camera are tested to see whether the system can classify correctly. In the test, the objects in the picture are classified correctly. The specific test results are shown in Fig. 9. The yellow words in the upper left have the labels of the objects and corresponding accuracy. It can be seen that the classification effect is good. It can meet the requirements of the intelligent garbage bin well, and can complete the garbage classification accurately.

C. SYSTEM COMPREHENSIVE TEST

The comprehensive test results are shown in Table. 11. It can be seen from the table that the system can accurately classify the garbage, the mechanical structure of the system can operate normally and correctly. The system has good performance and can complete the garbage classification task well.

VI. CONCLUSION

Aiming at the problem of garbage classification, this paper proposes an improved algorithm based on ResNet-34 and three tailor-made modifications, including the multi feature fusion, the feature reuse of residual unit, and optimization of activation function. Firstly, experiments on common data sets have been implemented to determine the basic model and ResNet-34 shows best performance and is chosen. Secondly, ResNet-34 are modified to further improve the classification accuracy in three aspects. The modifications are tested on the garbage data set with 14 types of garbage. The accuracy of original ResNet-34 is 0.985 9. The accuracy of ResNet-34-A using multi feature fusion is increased to 0.994 1. The accuracy of ResNet-34-B using residual unit is improved to 0.999 5. And the accuracy of ResNet-34-C using modified activation function is increased to 0.992 8. And ResNet-34-ALL which combines all three modifications has the highest accuracy of 0.999 6. Finally, an automatic garbage classification system is integrated with the proposed algorithm and necessary hardware, and the system is effective with the classification accuracy as high as 0.999 6 and stable with the classification cycle as quick as 0.95 seconds on average.

In this paper, there are still some problems worthy of further exploration. Firstly, the classification of small target is not well and can be improved in future. Secondly, the classification criteria can be further extended, such as kitchen waste.

CODE AVAILABILITY

The source-code is available from https://github.com/XXYKZ/An-Automatic-Garbage-Classification-System-Based-on-Deep-Learning.git

REFERENCES

[1] J.-J. Wang, N.-N. Zhao, and J.-H. Li, "Current situation of marine microplastics pollution and prevention proposals in China," *China Environ. Sci.*, vol. 39, no. 7, pp. 3056–3063, Jul. 2019, doi: 10.19674/j.cnki.issn1000-6923.2019.0360.

[2] W.-B. Li, G. Ma, E.-Q. Yang, Y.-M. Cai, Z. Chen, R.-F. Gao, J.-H. Yan, X.-F. Cao, and E.-J. Pan, "Study on characteristics of electric dust removal fly ash and bag fly ash in circulating fluidized bed waste incineration system," *Proc. CSEE*, vol. 39, no. 5, pp. 1397–1405, Mar. 2019, doi: 10.13334/j.0258-8013.pcsee.181110.

[3] D. Porshnov, V. Ozols, and M. Klavins, "Thermogravimetric analysis as express tool for quality assessment of refuse derived fuels used for pyro-gasification," *Environ. Technol.*, vol. 41, no. 1, pp. 35–39, Mar. 2020.

[4] P. Kellow, R. J. P. C. Joel, D. Ousmane, D. A. Kumar, D.-A. C. V. Hugo, and K. A. Sergei, "A smart waste management solution geared towards citizens," *Sensors*, vol. 20, no. 8, pp. 1–15, Apr. 2020, doi: 10.3390/s20082380.

[5] Y.-G. Cheng, N. Chen, and H. Zhang, "Coal fly ash as an inducer to study its application in the production of methane gas from domestic waste," *Fresenius Environ. Bull.*, vol. 29, no. 2, pp. 1082–1089, 2020.

[6] Y. Ren, J.-J. Yang, H.-J. Cao, Q.-Y. Zhang, and Q. Liu, "All components resourcing system of rural garbage based on post-gather automatic sorting and disposal technology and its application," *Trans. Chin. Soc. Agricult. Eng.*, vol. 35, no. 4, pp. 248–254, Feb. 2019, doi: 10.11975/j.issn.1002-6819.2019.04.031.

[7] D.-E. Zhao, R. Wu, B.-G. Zhao, and Y.-Y. Chen, "Research on garbage classification and recognition on hyperspectral imaging technology," *Spectrosc. Spectral Anal.*, vol. 39, no. 3, pp. 921–926, 2019.

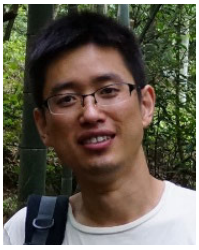
[8] T. Kano, E. Naito, and T. Aoshima, "Decentralized control for swarm robots that can effectively execute spatially distributed tasks," *Artif. Life*, vol. 26, pp. 243–260, Apr. 2020, doi: 10.1162/artl_a_00317.

[9] W. Pereira, S. Parulekar, S. Phaltankar, and V. Kamble, "Smart bin (waste segregation and optimisation)," in *Proc. Amity Int. Conf. Artif. Intell. (AICAI)*, Dubai, United Arab Emirates, Feb. 2019, pp. 274–279.

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. CVPR*, Dec. 2015, pp. 770–778. [Online]. Available: http://arxiv.org/abs/1512.03385

[11] G. Zhong, W. Jiao, W. Gao, and K. Huang, "Automatic design of deep networks with neural blocks," *Cognit. Comput.*, vol. 12, no. 1, pp. 1–12, Jan. 2020, doi: 10.1007/s12559-019-09677-5.
[12] A. Azarmi Gilan, M. Emad, and B. Alizadeh, "FPGA-based implementation of a real-time object recognition system using convolutional neural network," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 67, no. 4, pp. 755–759, Apr. 2020, doi: 10.1109/TCSII.2019.2922372.
[13] Y.-X. Li, Y.-A. Li, X. Chen, and J. Wei, "Feature extraction and classification identification of ship radiated noise based on VMD and SVM," *J. Nat. Univ. Defense Technol.*, vol. 41, no. 1, pp. 89–94, 2019, doi: 10.11887/j.cn.201901013.
[14] L. Wen, X. Li, and L. Gao, "A transfer convolutional neural network for fault diagnosis based on ResNet-50," *Neural Comput. Appl.*, vol. 32, no. 10, pp. 6111–6124, May 2020, doi: 10.1007/s00521-019-04097-w.
[15] A. S. B. Reddy and D. S. Juliet, "Transfer learning with ResNet-50 for malaria cell-image classification," in *Proc. Int. Conf. Commun. Signal Process. (ICCSP)*, Melmaruvathur, India, Apr. 2019, pp. 0945–0949, doi: 10.1109/ICCSP.2019.8697909.
[16] H. Chen, X. Yuan, Z. Pei, M. Li, and J. Li, "Triple-classification of respiratory sounds using optimized S-transform and deep residual networks," *IEEE Access*, vol. 7, pp. 32845–32852, 2019, doi: 10.1109/ACCESS.2019.2903859.
[17] A. Mahajan and S. Chaudhary, "Categorical image classification based on representational deep network (RESNET)," in *Proc. 3rd Int. Conf. Electron., Commun. Aerosp. Technol. (ICECA)*, Coimbatore, India, Jun. 2019, pp. 327–330.
[18] K.-H. Liu, P.-S. Zhong, D.-F. Xu, Q. Xia, and M. Liu, "Tangent-based rectified linear unit," *Comput. Integr. Manuf. Syst.*, vol. 26, no. 1, pp. 145–151, Jan. 2020, doi: 10.13196/j.cims.2020.01.015.
[19] W.-F. Gong, H. Chen, Z.-H. Zhang, M.-L. Zhang, C. Guan, and X. Wang, "Intelligent fault diagnosis for rolling bearing based on improved convolutional neural network," *J. Vib. Eng.*, vol. 33, no. 2, pp. 400–413, Apr. 2020, doi: 10.16385/j.cnki.issn.1004-4523.2020.02.021.
[20] K.-L. Sun, J.-M. Yu, and G. Sun, "A convolutional neural network model based on improved Softplus activation function," *J. Fuy-ang Normal Univ. (Natural Sci.)*, vol. 37, no. 1, pp. 75–79, Mar. 2020, doi: 10.14096/j.cnki.cn34-1069/n/1004-4329(2020)01-075-05.
[21] H. Ide and T. Kurita, "Improvement of learning for CNN with ReLU activation by sparse regularization," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Anchorage, AK, USA, May 2017, pp. 2684–2691.
[22] G. Wang, G. B. Giannakis, and J. Chen, "Learning ReLU networks on linearly separable data: Algorithm, optimality, and generalization," *IEEE Trans. Signal Process.*, vol. 67, no. 9, pp. 2357–2370, May 2019, doi: 10.1109/TSP.2019.2904921.
[23] P. Hinz and S. van de Geer, "A framework for the construction of upper bounds on the number of affine linear regions of ReLU feed-forward neural networks," *IEEE Trans. Inf. Theory*, vol. 65, no. 11, pp. 7304–7324, Nov. 2019, doi: 10.1109/TIT.2019.2927252.
[24] S. Banerjee and S. S. Chaudhuri, "Total contribution score and fuzzy entropy based two-stage selection of FC, ReLU and inverseReLU features of multiple convolution neural networks for erythrocytes detection," *IET Comput. Vis.*, vol. 13, no. 7, pp. 640–650, Oct. 2019.
[25] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*. Lake Tahoe, Nevada: Curran Associates, 2012, pp. 1097–1105.
[26] 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. (CVPR)*. Boston, MA, USA: IEEE, Jun. 2015, pp. 1–9, doi: 10.1109/CVPR.2015.7298594.
[27] C. Gulcehre, M. Moczulski, and M. Denil, "Noisy activation functions," presented at the Int. Conf. Mach. Learn. (JMLR), 2016, pp. 3059–3068.
[28] V. Nair and G.-E. Hinton, "Rectified linear units improve restricted Boltzmann machines," presented at the 27th Int. Conf. Mach. Learn. (ICML). Haifa, Israel: Omnipress, Oct. 2010, pp. 807–814.
[29] A.-L. Maas, A.-Y. Hannun, and A.-Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," presented at the 30th Int. Conf. Mach. Learn. (JMLR), Atlanta, GA, USA, 2013, pp. 1–6.
[30] H.-Z. Zhao, F.-X. Liu, and L.-Y. Li, "A novel softplus linear unit for deep convolutional neural networks," *J. Harbin Inst. Technol.*, vol. 50, no. 4, pp. 117–123, 2018.
[31] H.-X. Wang, J.-Q. Zhou, and C.-H. Gu, "Design of activation function in CNN for image classification," *J. Zhejiang Univ. (Eng. Sci.)*, vol. 53, no. 7, pp. 1363–1373, May 2019.
[32] W. Guo, W. Li, B. Yang, Z. Zhu, D. Liu, and X. Zhu, "A novel non-invasive and cost-effective handheld detector on soluble solids content of fruits," *J. Food Eng.*, vol. 257, pp. 1–9, Sep. 2019, doi: 10.1016/j.jfoodeng.2019.03.022.
[33] R. D. Beason, R. Riesch, and J. Koricheva, "AURITA: An affordable, autonomous recording device for acoustic monitoring of audible and ultrasonic frequencies," *Bioacoustics*, vol. 28, no. 4, pp. 381–396, Jul. 2019, doi: 10.1080/09524622.2018.1463293.
[34] X. Lai, T. Yang, Z. Wang, and P. Chen, "IoT implementation of Kalman filter to improve accuracy of air quality monitoring and prediction," *Appl. Sci.*, vol. 9, no. 9, pp. 1–23 , May 2019, doi: 10.3390/app9091831.
[35] P. Ramachandran, B. Zoph, and Q. V. Le, "Searching for activation functions," 2017, *arXiv:1710.05941*. [Online]. Available: http://arxiv.org/abs/1710.05941
[36] Z. Chen and P.-H. Ho, "Global-connected network with generalized ReLU activation," *Pattern Recognit.*, vol. 96, Dec. 2019, Art. no. 106961.

ZHUANG KANG received the bachelor's degree from the Jiangxi University of Science and Technology, in 2018, where he is currently pursuing the Ph.D. degree. His main research interests include computer vision, artificial intelligence, and object detection.



JIE YANG was born in Anhui, China, in 1979. He received the Ph.D. degree in rail traffic safety from the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, China, in 2017. He joined the Jiangxi University of Science and Technology. He is also the Director of the Institute of Permanent Maglev and Railway Technology (IPMRT). His current research interests include computer vision, artificial intelligence, permanent maglev, and energy-efficient train control.

GUILAN LI received the bachelor's degree from the Jiangxi University of Science and Technology, in 2019, where she is currently pursuing the Ph.D. degree. Her main research interests include computer vision, image classify, and object detection.

ZEYI ZHANG received the bachelor's and Ph.D. degrees from The University of Hong Kong, in 2014 and 2018, respectively. He is currently a Lecturer with the Institute of Permanent Maglev and Railway Technology (IPMRT), Jiangxi University of Science and Technology. His research interests include permanent maglev and automatic control.