International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 09 Issue 10 ǁ 2021 ǁ PP. 94-99

Research on Garbage Classification Model Based on the Fusion Model of CNN and SVM

Yanhe Na[1],Zhan Wen*[2],Wenzao Li[3],Haoning Pu[3]

[1, *2, 3] *School of Communication Engineering, Chengdu University of Information Technology, Chengdu, China*
Corresponding Author: Zhan Wen

Abstract

Garbage classification has become an urgent problem that needs to be solved today. The convolutional neural network (CNN) is a class of deep learning neural networks. It is widely used in area of image recognition like garbage classification. But the CNN model still has certain problems. For example, the complexity of the CNN model structure has deepened, the model training time has also become longer, and how to obtain stronger generalization capabilities has become the primary problem in image recognition today. In this paper,we propose the CNN+SVM image recognition model to garbage classification and train the model through transfer learning to reduce training time and improve the generalization ability of the model. Based on this model, we design a CNN+SVM model control experiment of garbage classification. The experimental results showed that the generalization ability of the Resnet50+SVM model was better, and then selected this model for final evaluation. The effect of the Resnet50+SVM model is evaluated by simulating real domestic garbage classification scenarios, and it is predicted that the garbage classification recognition rate of this model can reach 87%, which can be applied to domestic garbage classification scenarios.

Keywords: *Transfer Learning; Convolutional Neural Network (CNN); Support Vector Machine (SVM); Garbage classification*

Date of Submission: 20-09-2021 Date of acceptance: 05-10-2021

I. INTRODUCTION

With the vigorous economic development of various countries in the world, the quality of human life has also been significantly improved. People are no longer satisfied with maintaining a simple daily life, and their demand for material consumption has greatly increased, leading to a rapid increase in global waste production. Taking China as an example, the output of domestic waste has reached more than 400 million tons, and the pressure on the treatment of urban domestic waste has increased sharply.

Traditional garbage classification and identification methods mainly adopt manual identification. The specific method is to manually sort the garbage on the conveyor belt by workers. This method has many drawbacks, such as: time-consuming and labor-intensive, low work efficiency and high recognition error rate. In order to solve these problems, convolutional neural networks are introduced to recognize garbage images. This technology is an important research direction in deep learning. It has the advantages of strong anti-interference ability and strong feature learning ability, and is widely used in various fields. For example, Shailaja et al. proposed a linear discriminant regression classification algorithm based on deep learning. The experimental results of YALE and ORL databases show that this algorithm can effectively improve the accuracy of the face recognition system and has better performance[1]. Chang et al. proposed a traffic sign recognition algorithm based on the CNN model, which can realize real-time monitoring of vehicle speed with high accuracy [2].

The image recognition technology of the CNN model has developed rapidly and achieved remarkable results, but there are still some problems. How to obtain stronger generalization ability during CNN model training has become the primary problem in image recognition[3]; the high computational complexity of the network and long data training time are also inevitable problems of the current CNN model[4]. The support vector machine (SVM) has good robustness. For the addition or deletion of vector data, the stability of the model remains unchanged, and it performs well in today's mainstream image recognition and classification applications[5].

Therefore, we improved the Resnet50 network structure in convolutional neural network image recognition[6]. We proposed a Resnet+SVM fusion model to reduce over-fitting problems and improve the

The Network and Data Security Key Laboratory of Sichuan Province, UESTC (NO.NDS2021-7).

generalization ability of the model. After comparison and verification, our proposed model reduces the model training time and effectively improves the recognition accuracy.

II. MODEL DESIGN

2.1 Model development

The garbage classification model is selected based on the image classification and recognition algorithm of the combination of convolutional neural network and SVM, and further analysis and comparison are needed for the selection of CNN structure. VGG16 has a simple structure, but contains a large number of parameters; Resnet50 has a complex structure, but the number of parameters is lower than that of VGG16. In this regard, we use two network structures to build training models and analyze and evaluate the models through verification results, and select the convolutional neural network structure that performs better among the two as the garbage classification model. The two models are the VGG16+SVM model and the Resnet50+SVM model. In addition to the different pre-training model construction and parameters, other conditions including the source of the data set, data processing, training process and verification process remain the same. Figure 1 shows the overall design structure of the garbage classification model.

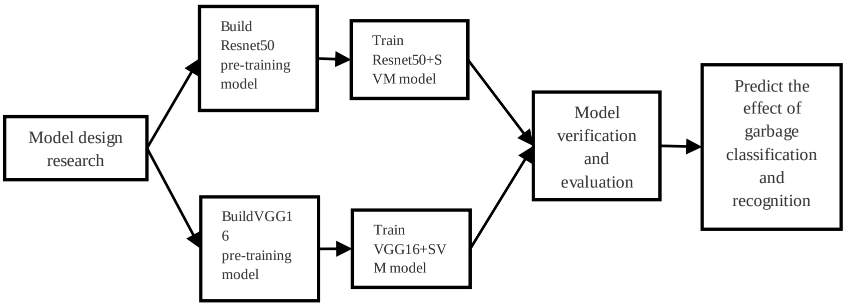

Figure 1: Fusion model design

2.2 Pre-training model construction for tranfer learning

In CNN model training, when the training effect is not ideal, the training samples are not rich enough, and it is troublesome to re-adjust the parameters to build the CNN model, you can consider using the transfer learning method. In the image classification problem, transfer learning can not only learn the low-level features such as edges and colors, but also learn the deep features of the data, which can effectively improve the classification effect of the network. Compared with the traditional CNN model, the effect of using the tranfer learning network to fit the data is better, and the recognition accuracy and convergence speed of the network will also be improved.

For the construction of the two pre-training models, the content of tranfer learning is used, and the weight of the trained network structure is used to speed up the work progress. By calling torchvision.models in Pytorch, using the model that has been trained with a large amount of data, this paper selects the pre-training model corresponding to the network structure, and uses some network institutions as the feature extraction layer, and re-according to the layers after the feature extraction layer To predict the number of categories, create your own layer structure; freeze the parameters of the previous feature extraction layer, and only train the layer parameters created after the update. For the Vgg16 pre-training model, after loading the model, freeze the 13-layer convolutional layer as the feature extraction layer, modify the fully connected layer, and change the classification category in the last layer from 1000 to 6; define forward During the propagation process, the output is named x to save memory, and the data obtained by the feature extraction layer is converted into a 1-dimensional vector through torch.flatten and sent to the fully connected layer. For the Resnet50 pre-training model, after loading the model, the previous layer is frozen as the feature extraction layer, and the input is sent to a fully connected layer with 256 output units, and then the Relu layer and the Dropout layer are connected. Then there is a 256*6 fully connected layer, the output is a 6-channel softmax layer; the forward propagation process is defined, the output is named x to save memory, and the data obtained by the feature extraction layer is converted into a 2048 column vector through out.view The data is then sent to the fully connected layer.

2.3 Data set selection and processing

We selected a data set with moderate processing capacity, which was selected and deleted based on the 40 categories of the Huawei Cloud garbage classification data set. The data set is divided into 6 categories, glass, cardboard, metal, paper, plastic and trash. It contains 2527 pictures of domestic garbage, such as Table 1 shows:

Table1: Garbage classification data set

Index	Name	Data size
1	glass	501
2	paper	594
3	cardboard	403
4	plastic	482
5	metal	410
6	trash	137

Create cardboard, glass, metal, paper, plastic and trash folders in the project root directory. We divide the garbage classification data set into a training set and a validation set. The ratio of the validation set is 0.33, and the random_date parameter is fixed to 42 to ensure that the training set and validation set are the same each time.

In the convolutional neural network model, the problem of overfitting is prone to occur. In order to reduce the probability of this problem, we have adopted data enhancement processing on the garbage classification data training set. For the verification set, in order to be closer to real life scenes, we did not perform The data enhancement operation only performed image scaling and standardization processing.

2.4 Model training

For the construction of the two pre-training models, the content of tranfer learning is used, and the weight of the trained network structure is used to speed up the work progress. By calling torchvision.models in Pytorch, using the model that has been trained with a large amount of data, this paper selects the pre-training model corresponding to the network structure, and uses some network institutions as the feature extraction layer, and re-according to the layers after the feature extraction layer.To predict the number of categories,create your own layer structure; freeze the parameters of the previous feature extraction layer, and only train the layer parameters created after the update.

Read the built pre-training model VGG16, set the initial training accuracy rate and the initial state of the loss rate to 0, and record the training start time. Set the number of training times to 20 epochs (one epoch is the process of training all data samples in the data set once), select the loss function of this model as MultiMarginLoss, the optimization function as Adam Optimizer, and the learning rate as 0.0003.

Read the training data set that has been preprocessed, and select the batch gradient descent method (BGD) according to the size of the training data set. Ensure that the data in the batch read in the same order between each epoch is different. The gradient is initially set to 0, the data is transmitted to the GPU to obtain the predicted value, and the loss function “torch.nn.MultiMarginLoss” is used to calculate the loss of the predicted value and the actual value. The gradient is backpropagated and calculated, the gradient is clipped and normalized, and all parameters are updated.

III. EXPERIMENTS RESULTS

3.1 Model inspection

VGG16+SVM model verification: Read the trained VGG16+SVM model VGG16.pkl, set the initial verification accuracy and loss initial state to 0, and record the start verification time. Set the number of verifications to 20 epochs, and define the storage address of the verification log.

Read the preprocessed validation set data, select batch gradient descent (BGD), set batch_size to 36, read batches in order; set shuffle=True, shuffle the data set in each epoch, and do not change within the epoch. Bring the validation set data into the trained VGG16+SVM model VGG16.pkl for prediction, and get the predicted value. Call the loss function to calculate the verification loss and protect it, and compare the predicted value with the true value to get the accuracy rate. Saving the loss function value and accuracy rate to the verification log, this epoch ends. Proceed to the next epoch, and iterated 20 times to verify the end.

Resnet50+SVM model verification: Read the trained Resnet50+SVM model Resnet.pkl, and other state initialization is the same as the verification process.

3.2 Model inspection

3.2.1 Model test

After training and verifying the two models VGG16+SVM model and Resnet50+SVM model, as shown in Figure 2,3.We get the accuracy comparison chart and loss value comparison chart of the two models.

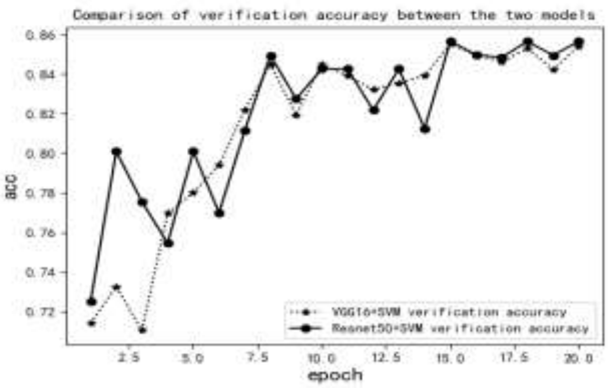

Figure 2: Comparison of verification accuracy between VGG16+SVM model and Resnet50+SVM model

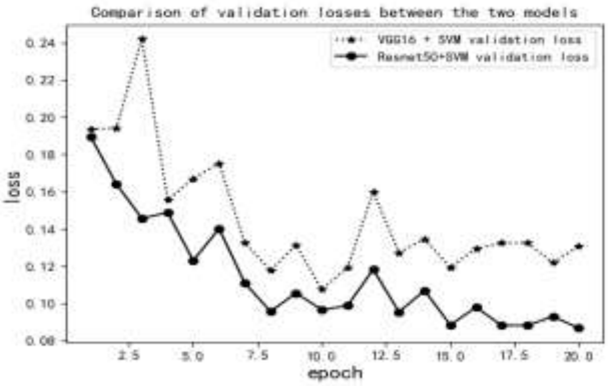

Figure 2:Comparison of validation loss between VGG16+SVM model and Resnet50+SVM model

Through Figure 2, we get that the accuracy rate of the Resnet50+SVM model on the verification set after the first epoch is 0.7250 higher than the accuracy rate of the VGG16+SVM model 0.7141; after the second epoch, the accuracy rate reaches 0.8009 obviously It is higher than the accuracy rate of 0.7326 of the VGG16+SVM model; after that, the relationship between the accuracy of the two models in the iterative process slightly changed, but in the end the Resnet50+SVM model reached a verification accuracy rate of 0.8565, which was slightly higher than the VGG16+SVM model. Through Figure 3, we get that the loss value of the verification function of the Resnet50+SVM model is generally lower than that of the VGG16+SVM model, and finally reaches 0.0867. Therefore, the recognition performance of the Resnet50+SVM model in the simulated domestic waste classification scene is stronger than that of the VGG16+SVM model.

3.2.2 Model analysis

We finally selected the Resnet50+SVM model with strong model generalization ability as the garbage classification model based on CNN and SVM. We brought the previously selected data set data into the Resnet50+SVM model after simple expansion and basic preprocessing. Perform generalization testing in.After testing, we get Table 2:

Table2: Resnet50+SVM generalization test

Predict Reality	cardboard	glass	metal	paper	plastic	trash
Cardboard	117	2	0	17	1	0
glass	0	158	4	0	1	0
metal	0	22	117	1	2	3
paper	1	3	3	182	3	4
plastic	0	19	0	3	137	2
trash	2	1	3	5	5	25

The classification evaluation index is calculated by the following formula:

$$\text{precision} = \frac{\text{TP}}{TP+FP} \quad (1)$$

$$\text{recall} = \frac{TP}{TP+FN} \quad (2)$$

$$F1score = \frac{2TP}{2TP+FP+FN} \quad (3)$$

Where,
FP = False Positive, predict the number of other classes as positive;
FN = False Negative, predict the positive class as the number of other classes;
TP=True Positive, predict the number of positive classes as positive classes.

The specific meaning of accuracy is the proportion of data samples that are actually classified correctly (TP) in the data samples that are predicted to be correctly classified (TP+FP); the specific meaning of recall rate is in the actual data samples of this category (TP+FN), the proportion of data samples (TP) with the correct classification of recall; the higher the accuracy, the better, and the higher the recall rate, the better. F1 score considers both the accuracy and the recall rate, so that both can reach the highest and balance at the same time.. After calculation, we get Table 3.

Table 3:Resnet50+SVM model classification evaluation

Types of garbage	Precision	Recall	F1 score
Cardboard	0.97	0.85	0.91
glass	0.77	0.97	0.86
metal	0.92	0.81	0.86
paper	0.88	0.93	0.90
plastic	0.92	0.85	0.88
trash	0.74	0.61	0.67
Cardboard	0.97	0.85	0.91

We introduce "Micro Average", "Macro Average" and "Weighted Average" for the multi-classification problem to further analyze and process the data in Table 3, so as to specifically evaluate the Resnet50+SVM garbage classification model we have selected in the real life garbage classification scene Effect.

$$\text{Micro Average} = \frac{\alpha}{\Gamma} \quad (4)$$

$$\text{Macro Average} = \frac{1}{\sum \mathrm{i}} \sum \varphi_{(\text{category})} \quad (5)$$

$$\text{Weighted Average} = \sum \frac{\mathrm{i}_{(\text{category})}}{\Gamma} \varphi_{(category)} \quad (6)$$

Among them, α is the number of predicted correct data samples, and Γ is the total number of predicted data samples. φ is one of the three evaluation indexes (accuracy, recall and F1 score) of a certain type of garbage, and i is the quantity of a certain type of garbage.

The final calculation results are shown in Table 4.

Table4: Resnet50+SVM model evaluation index

	Precision	Recall	F1 score
Micro Average	0.87	0.87	0.87
Macro Average	0.87	0.84	0.85
Weighted Average	0.88	0.87	0.87

IV. CONCLUSION

The classification method of the CNN+SVM model we proposed could greatly reduce the problems of local optimization and over-fitting in deep learning. It has better generalization ability than traditional CNN based on softmax classification model.

However, we only selected some common household garbage pictures for the six types of garbage, and did not involve some other types of household garbage. Moreover, the garbage image data does not involve the mixing of multiple garbage, which is more common in real life. Therefore, we need more and more complete image data to perfect and improve the model.

REFERENCES

[1]. Yani Zhu,Chaoyang Zhu,Xiaoxin Li. Improved principal component analysis and linear regression classification for face recognition[J]. Signal Processing,2018,145.
[2]. K. Mirunalini,Vasantha Kalyani David. Traffic sign Detection using CNN[J]. (IJEAT),2021,10(3).
[3]. Ahmed Jawad A. AlBdairi,Zhu Xiao,Mohammed Alghaili,Chenxi Huang. Identifying Ethnics of People through Face Recognition: A Deep CNN Approach[J]. Scientific Programming,2020,2020.
[4]. Guha Riya,Das Nibaran,Kundu Mahantapas,Nasipuri Mita,Santosh K. C.. DevNet: An Efficient CNN Architecture for Handwritten Devanagari Character Recognition[J]. International Journal of Pattern Recognition and Artificial Intelligence,2020,34(12).
[5]. Sari Rochman Eka Mala,Oktavia Suzanti Ika,Imamah ,Ali Syakur Muhammad,Anamisa Devie Rosa,Khozaimi Ach.,Rachmad Aeri. Classification of Thesis Topics Based on Informatics Science Using SVM[J]. IOP Conference Series: Materials Science and Engineering,2021,1125(1).
[6]. J. Konečný, J. Liu, P. Richtárik and M. Takáč, "Mini-Batch Semi-Stochastic Gradient Descent in the Proximal Setting," in IEEE Journal of Selected Topics in Signal Processing, vol. 10, no. 2, pp. 242-255, March 2016.
[7]. J. Sun, Z. Wu, Z. Yin and Z. Yang, "SVM-CNN-Based Fusion Algorithm for Vehicle Navigation Considering Atypical Observations," in IEEE Signal Processing Letters, vol. 26, no. 2,pp. 212-216, Feb. 2019