Research on Recyclable Garbage Classification Algorithm Based on Attention Mechanism

Wei Chen

Tianjin University of Technology and Education,
Tianjin key laboratory of information sensing and intelligent control,Tianjin 300222, China
chenwei.2018[at]outlook.com

***Abstract:** With the rapid development of global productivity levels, the problem of garbage disposal is getting more and more serious. Garbage classification is an important step to realize garbage reduction, harmlessness and resource utilization. With the increase in types and quantities of garbage, traditional garbage classification image algorithms can no longer meet the accuracy requirements of garbage identification. This paper proposes a ResNet18 convolutional neural network model based on the attention mechanism for the classification of recyclable garbage. The attention module is added after convolution, so that the model can pay more attention to the important information in the feature map. The model can automatically extract the characteristics of garbage for classification, including: glass, metal, plastic and paper. Experimental results show that the algorithm has an accuracy rate of 92% in the classification of recyclable waste, which can effectively classify recyclable waste.*

Keywords: Recyclable garbage classification, Image classification algorithm, Convolutional Neural Network, Attention mechanism

1. Introduction

With the rapid increase in global productivity, the problem of waste disposal has also come along. Garbage classification is an important step to realize garbage reduction, harmlessness and resource utilization. The main benefits of garbage classification include: (a) Reduce floor space and improve land utilization. Some household garbage contains harmful substances, and that garbage that are not easily degraded cause serious damage to the land. If sorting can remove the recyclable garbage, the landfill quantity can be reduced by more than 60%. (b) Reduce waste pollution and protect the ecological environment. The current garbage disposal mostly uses landfill, which occupies a large amount of land area and causes unpleasant phenomena such as an increase in mosquitoes, a turbulent flow of sewage, and an unpleasant smell. In addition, the waste plastics in the soil will reduce crop yields. If eaten by animals or humans, it will endanger their health and even cause death. And it will pollute water resources, which is closely related to human health. Sorting and disposing of garbage can reduce such hazards. (c) Effective use of resources. 30%-40% of our domestic garbage can be recycled. Recycling this garbage can increase the recycling of resources. It can be seen that timely garbage classification, recycling and reuse is the best way to solve the garbage problem.

At present, the classification of garbage in various countries is mainly through manual screening, which has certain problems in efficiency and accuracy. In the early days, many scholars [1] manually designed feature extractors to classify and recognize garbage through traditional machine learning methods, but the recognition accuracy was limited. In recent years, the rapid development of deep learning technology has made convolutional neural networks more and more widely used in the field of image recognition. Stephenn L. Rabano [2] and others took the lead in using a lightweight convolutional neural network on the TrashNet Dataset, and achieved 87.2% image recognition accuracy. However, most scholars mainly distinguish garbage as recyclable garbage, non-recyclable garbage, kitchen waste, and hazardous garbage. This paper uses the convolutional neural network in deep learning to classify the recyclable garbage in the garbage and identify that the recyclables are glass Compared with traditional machine learning methods, which one of, metal, plastic, and paper, not only improves the efficiency of garbage identification in recyclables, but also improves the accuracy of garbage classification, which is of great significance to garbage classification algorithms.

2. Data Processing

2.1 Recyclable garbage data

This article summarizes the collected recyclable garbage picture data and the pictures collected on the Internet to form a recyclable garbage classification data set. The data set contains four categories, namely glass, metal, plastic, and paper, of which 500 are glass, 500 are metal, 500 are plastic, and 500 are paper. The data set is divided at a ratio of 8:1:1, which are training set, validation set and test set. The garbage picture is shown in **Error! Reference source not found.**.

Figure 1: Recyclable garbage image data

2.2 Data preprocessing

Before performing model training, preprocess the image data to increase the convergence speed and generalization ability of the model. The data preprocessing process is shown in

Error! Reference source not found..
A.Use the quadratic linear interpolation algorithm to reduce the size of the picture to a regular picture of 224×224×3. Since the collected pictures are of different sizes, this will cause great difficulties for the training of the model. Therefore, the use method of fixing the size of the pictures can increase the training speed of the model.

B. Use random rotation, random horizontal flip, and random vertical flip to increase the randomness of the training pictures, so that the pictures of the model are different each time, increase the generalization ability of the model, and the model can become more robust in prediction.

C. Normalize and standardize the pictures. The size of the picture data is scaled to between [0,1], and the mean value and standard deviation of the picture are scaled to around 0.456 and 0.224 respectively, which can increase the convergence speed of the model during training.

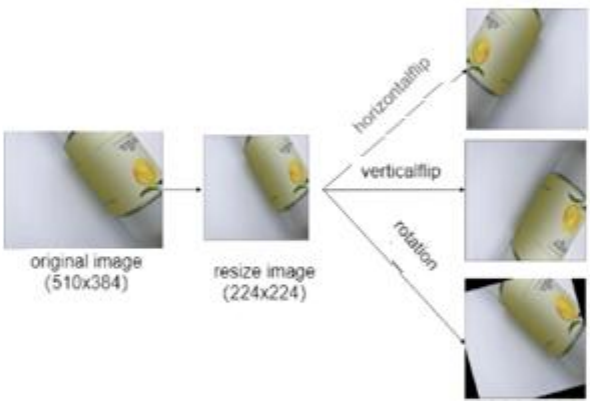

Figure 2: Data preprocessing process

3. Model building

3.1 Convolutional Neural Network

Convolutional Neural Network (CNN) is a neural network designed to process data with a similar grid structure. Usually refers to those neural networks that use convolution operations in at least one layer of the network instead of general matrix multiplication operations. This is one of the most successful applications of deep learning algorithms. CNN includes one-dimensional CNN, two-dimensional CNN and three-dimensional CNN. One-dimensional CNN is mainly used for sequence data processing, two-dimensional CNN is often used for image text recognition, and three-dimensional CNN is mainly used for medical image and video data recognition. Three important ideas in CNN to help improve the model are sparse interaction, parameter sharing and equivariant representation.

3.2 ResNet network

The ResNet[3] network was the champion of image recognition in the ImageNet competition in 2015, with an error rate of 3.57%, which exceeded the level of human image recognition. Prior to this, convolutional neural networks were all developing in a deeper direction, but as the depth of the network becomes deeper, the effect of the model becomes worse. ResNet uses a cross-layer connection method, which successfully alleviates the gradient dissipation problem in deep neural networks, and provides the possibility for thousands of layers of network training. In order to reduce the risk of model overfitting, this paper uses the ResNet18 network with a shallower model depth. The residual layer of the network is shown in **Error! Reference source not found.**.

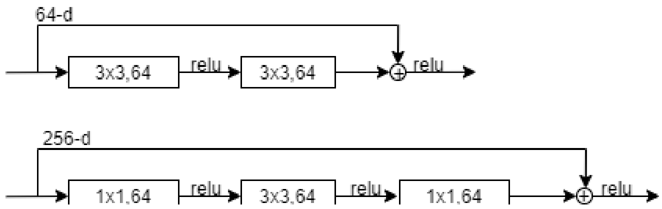

Figure 3: Residual layer of the network

3.3 Attention Mechanism

This paper uses the Convolutional Block Attention Module (CBAM) [4] attention mechanism. As shown in **Error! Reference source not found.**, this model combines the attention module of space and channel, which can get more attention than other attention mechanisms that only focus on channels or spaces. good result. The feature map after the convolutional layer first passes through a channel attention module, after obtaining the weighted result, passes through a spatial attention model and finally passes through the weighted result. Channel attention can compress the feature map in the spatial dimension, without considering the average pooling but also considering the maximum pooling. The result of average pooling has feedback for each pixel in the feature map, and maximum pooling uses only the feature map to respond to the largest gradient when performing gradient response propagation.

The formula for channel attention is:

$$\boldsymbol{M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F)))}$$

The formula for spatial attention is:

$$\boldsymbol{M_s(F) = \sigma\left(f^{7\times7}([AvgPool(F)\right);(MaxPool(F)]))}$$

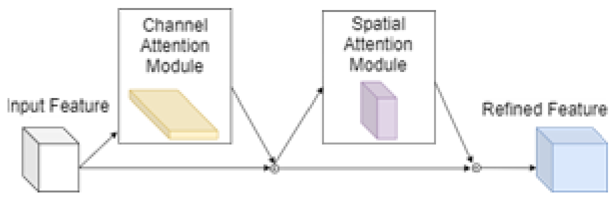

Figure 4: CBAM model diagram

3.4 Transfer learning

Transfer Learning is a machine learning method. Its main working idea is that the neural network learns knowledge or patterns from one task, and applies this knowledge to a different but related independent task to make the goal Tasks can achieve better learning results. In transfer learning, first train a basic network on the basic data set and task, and then readjust or transfer the learned features to the second network to train the target data set and task.

3.5 Resnet18 network based on attention mechanism

Considering the different forms of various garbage, the accuracy of the network may be reduced when using the network classification, so this paper proposes an improved ResNet18 convolutional neural network. Add the Attention module behind each Basic Block of ResNet, and use the CBAM attention module to give more weight to the main channels or graphs in the feature map extracted by the convolution module, so that the network pays more attention to those important features and ignores those Unimportant features.As shown in **Error! Reference source not found.**,Change the 1000 of the last output layers of the ResNet18 network to 4 layers to suit our classification task. Use pre-training weights when actually training the network, because pre-training weights are the result of training on ImageNet large data. Using pre-training weights can not only speed up the convergence of our model, but also make the model more effective.

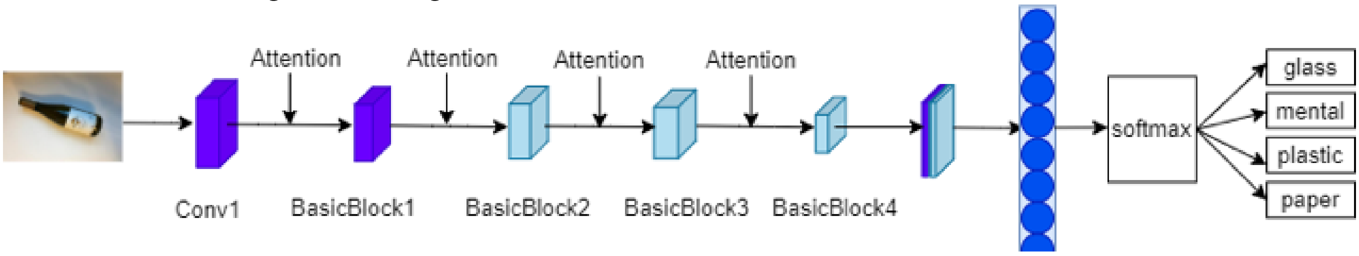

Figure 5: ResNet18 network structure based on Attention

4. Result

4.1 Evaluation indicators

This paper uses the recognition accuracy rate as the evaluation standard of the model, and the recognition accuracy rate formula is as follows

$$acc(f:D) = \frac{1}{m}\sum_{i=1}^{m} I(f(x_i) = y_i)$$

In the formula, m represents the total number of samples; I (·) represents the indicator function, and when the expressions in parentheses are true and false, the value is 1,0 respectively; $f(x_i)$ is the type predicted by the model; y_i is the sample Real category.

4.2 Result analysis

In the process of training the model, this article uses the validation set to monitor the fit status of the model. The validation set is used to evaluate the accuracy of each training generation, the best generation in the model training set is selected as the final model, and the accuracy of the test set is used as the evaluation Model indicators.

The experimental platform of this paper is under the Windows 10 system, using the Python language-based Pytorch framework to build the model structure. The hardware environment is CPU Intel E5-2640, the graphics card is NVIDIA Quadro M4000, the Adam [5] optimizer is used as the model optimizer, the cross entropy loss function is used as the loss function of the model, the initial learning rate is set to 0.001, and the weight attenuation factor is set 0.0001.

As shown in **Error! Reference source not found.**, it is the training and verification loss curve of the ResNet18 model. As shown in **Error! Reference source not found.**, it is the accuracy curve of model training and verification. It can be seen from the figure that the model has the highest accuracy in the verification set at the 44th generation, so the 44th generation model is used for testing, and the test accuracy rate is 89%.As shown in**Error! Reference source not found.**,the training verification loss curve of the model proposed in this paper. As shown in**Error! Reference source not found.**,the model's training and verification accuracy curve. From the figure, the verification accuracy rate of the model at the 49th generation is 92.5%. The 49th generation model is used as the final model for testing, and the test result is 92%. Comparing the training curve of ResNet18 with the model training curve proposed in this paper, the use of CBAM attention mechanism can make the training loss of the model lower and make the accuracy of the model higher.

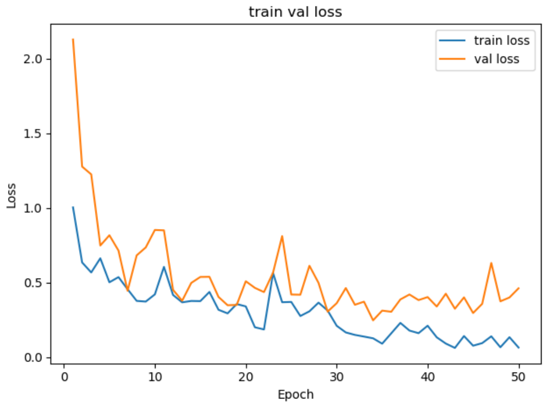

Figure 6: ResNet18 training verification loss

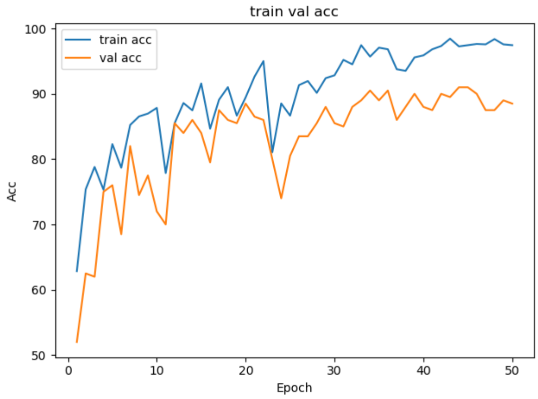

Figure 7: ResNet18 training verification accuracy

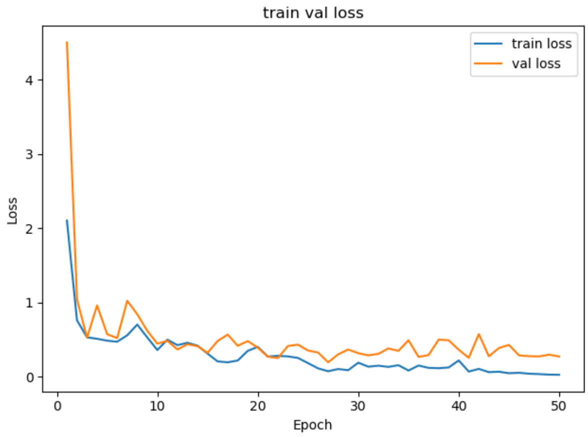

Figure 8: Attention-ResNet training verification loss

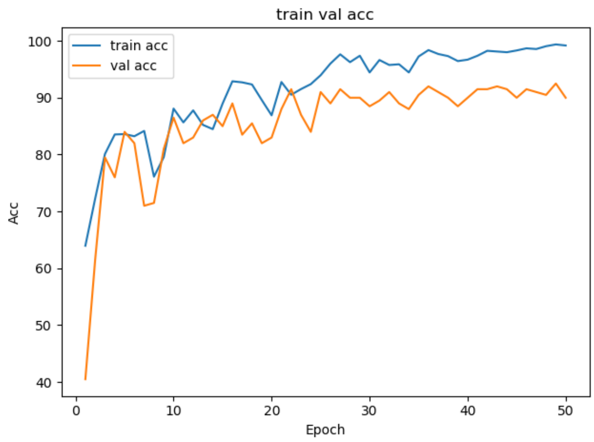

Figure 9: Attention-ResNet training verification accuracy

5. Conclusions

In this paper, four categories of recyclable garbage are classified, and the recyclable garbage is divided into glass, metal, plastic and paper, and a ResNet18 convolutional neural network based on the Attention mechanism is proposed. Compared with the traditional ResNet network, the accuracy of the model in classification can be improved by adding the CBAM attention module. The accuracy of the model is 88.4% in the ResNet18 network without an increased attention mechanism. In this paper, the accuracy of the attention-based ResNet18 neural network is 92%. The experimental results show that the model proposed in this paper can improve the classification accuracy of the traditional ResNet network.

References

[1] Yang M, Thung G. Classification of trash for recyclability status[J]. CS229 Project Report, 2016, 2016.
[2] Rabano S L, Cabatuan M K, Sybingco E, et al. Common garbage classification using mobilenet[C]//2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM). IEEE, 2018: 1-4.
[3] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.
[4] Woo S, Park J, Lee J Y, et al. Cbam: Convolutional block attention module[C]//Proceedings of the European conference on computer vision (ECCV). 2018: 3-19.
[5] Kingma D, Ba J. Adam: A Method for Stochastic Optimization[J]. Computer ence, 2014.
[6] Meng S, Zhang N, Ren Y. X-Dense Net: Deep Learning for Garbage Classification Based on Visual Images[J]. Journal of Physics Conference Series, 2020, 1575:012139.
[7] Adedeji O, Wang Z. Intelligent waste classification system using deep learning convolutional neural network[J]. Procedia Manufacturing, 2019, 35: 607-612.
[8] Zhang P, Zhao Q, Gao J, et al. Urban street cleanliness assessment using mobile edge computing and deep learning[J]. IEEE Access, 2019, 7: 63550-63563.
[9] Kang Z, Yang J, Li G, et al. An Automatic Garbage Classification System Based on Deep Learning[J]. IEEE Access, 2020, 8: 140019-140029.
[10] Ahmad K, Khan K, Al-Fuqaha A. Intelligent Fusion of Deep Features for Improved Waste Classification[J]. IEEE Access, 2020, 8(99):96495-96504.

Author Profile

Wei Chen, Tianjin key laboratory of information sensing and intelligent control, Tianjin University of Technology and Education, Tianjin 300222, China.