Research on Garbage Image Classification Method Based on Convolutional Neural Network

Kangjian Tang
School of Communication
Engineering Chengdu
University of Information
Technology Chengdu, Sichuan

Zhan Wen
School of Communication
Engineering Chengdu
University of Information
Technology Chengdu, Sichuan

Yahui Chen
School of Communication
Engineering Chengdu
University of Information
Technology Chengdu, Sichuan

Wenzao Li
School of Communication
Engineering Chengdu
University of Information
Technology Chengdu, Sichuan

Abstract: The issue of garbage classification has aroused great concern for residents recently. The traditional garbage classification method has a low classification accuracy rate, high cost, and large personnel demand. With the increasing amount of garbage, it is difficult for traditional methods to effectively separate waste. Therefore, consider applying deep learning to the garbage classification problem. This paper uses two types of Convolutional Neural Networks (CNNs)--Inception V3 and Inception V4 to train Huawei's public garbage data set (Garbage Date) and establish a garbage classification model. After experiments, the classification results are compared and the performance of the model is tested. In this paper, by observing the changes in the accuracy rate and cross-entropy loss function of the two models on the training and test sets in the experiment, it is found that both models can obtain higher accuracy of garbage classification The network model using Inception V4 is more stable and accurate than the network model using Inception V3. The experimental results also show that this method can improve the accuracy of garbage classification in daily life and the efficiency of recyclable garbage collection.

Keywords: Deep learning; Convolutional neural network; Inception

1. INTRODUCTION

The problem of garbage classification and treatment has always been closely related to people's lives. An accurate and efficient method of garbage classification and treatment can promote the sustainable development of society and have a profound impact on saving resources and protecting the environment. Due to the lack of awareness of garbage classification and limited garbage classification equipment, traditional garbage classification methods have the characteristics of low classification efficiency, high cost, and large demand for personnel, and it is difficult to achieve accurate and efficient garbage classification, leading to garbage Classification efficiency and accuracy are very low, and environmental and resource issues are becoming increasingly serious. The method of combining artificial intelligence with garbage classification and processing problems can make up for the disadvantages of traditional garbage processing methods and has the advantages of high efficiency, high accuracy, and low cost.

This paper considers the use of deep learning to classify garbage images using convolutional neural networks. This paper mainly researches the garbage classification method based on convolutional neural networks. Two types of garbage classification models are established by using convolutional neural networks including Inception module--

Inception V3 and Inception V4, and the models are used in Huawei's open garbage dataset Garbage Date was trained and obtained higher accuracy and lower loss function values. The establishment of this model improves the efficiency and accuracy of waste classification and aids with subsequent waste classification processing.

2. RELATED WORKS

In recent years, CNN has been widely used in the field of image processing. Researchers have proposed many convolutional neural network models with superior performance. The proposal of the Inception series network is a milestone in the development history of CNN classifiers. Before the Inception network was proposed, in order to make the neural network's ability to extract event features better, in order to achieve better classification results, most popular universal CNN use stacking of convolutional layers in the network multiple times. Model [8-9]. For example, the AlexNet network proposed by Krizhevsky has extended the depth of its network structure and used network optimization methods such as ReLU and Dropout when constructing the network model [1]. Similarly, the VGG-Net network proposed by Simonyan et al. Also has a deeper network structure and uses a smaller convolution kernel than the AlexNet network [2]. However, the VGG-Net network has a good generalization performance, but also because of the deepening of the network structure, the parameters of its network are very large. Subsequently, a deep convolutional neural network called Inception was proposed. The main feature of the

network structure is that the utilization of computing resources inside the network is improved, and the width and depth of the network are increased while keeping the computing budget constant. The architecture decision is based on the Hebbian principle and multi-scale processing intuition to optimize network quality [3]. The Inception of deep convolution architecture was first introduced by Szegedy et al. And is called GoogLeNet or Inception V1 [3]. With the continuous improvement of the Inception architecture, Batch Normalization [4] was introduced by Ioffe et al., Which is called Inception V2. The Inception system was improved by adding the factorization idea, and the arrangement and size of the convolutional layers were changed to obtain Inception V3 [5]. To reduce the complexity of the modules in the network and make the network modules more consistent, a fast reduction was added to the Inception system. At the same time, a new series of network structures Inception V4 [6- 7].

3. EXPERIMENTAL METHODS AND VERIFICATION

3.1 Network model

The reason why Inception series networks have such excellent performance is that different types of Inception modules are used in the network. The Inception module has multiple size filters in parallel on the same level. Its main advantages include: it can well collect feature information in input images of different sizes, and it can well solve the easy overfitting caused by the network being too deep And gradient updates are difficult to transmit to the entire network, which can greatly reduce the computing resources consumed during the calculation.

The Inception V3 network introduces the idea of Factorization into small convolution, which decomposes a larger two-dimensional convolution into two smaller asymmetric one-dimensional convolutions, which saves many parameters. On the one hand, it speeds up the operation and reduces overfitting. At the same time, it also adds a layer of non-linear extended model expression ability. On the other hand, Inception V3 optimizes the structure of the Inception module and uses branches in the branches of the module [5]. This structure not only improves the training speed of the Inception V3 network but also improves the classification accuracy.

Looking back at the Inception series of networks, we limited the optimization and transformation of the network structure to the changing isolated network components and ignored the rest of the network. The unsimplified network part will make the network look more complicated. So, the Inception V4 network specifies the size of each grid, making a unified choice for the Inception module. Inception V4 network solves the unnecessary complexity of some modules, makes the modules more consistent, and uses more consistent modules to improve performance. Besides, the Inception V4 network also introduces a dedicated reduction block for changing the width and height of the grid, making the network model more concise [6].

Considering the high recognition efficiency and high accuracy of the Inception series network on image classification, this experiment uses Inception V3 and Inception V4 convolutional neural networks to construct a garbage classification model.

3.2 Experimental verification

The data set selected for the experiment is Huawei Garbage Date. The data includes 40 types of garbage, with a total of 11,087 pictures. In this experiment, Inception V3 and Inception V4 convolutional neural networks were used to train the Garbage Date dataset, and a garbage classification model was established. The performance of the classification model is determined by analyzing the accuracy of the garbage classification model and the value of the cross-entropy loss function. The closer the accuracy rate is to 1, the higher the accuracy of the model. The cross-entropy value describes the distance between two probability distributions. The smaller the cross-entropy value is, the closer the two are, and the closer the value is to 0, the more accurate the classification result. The training results were then tested using a single-class image dataset.

To compare the performance of Inception V3 and Inception V4, this experiment uses the same parameter settings for the basic parameters of the classification models trained by these two neural networks. Select the Garbage Date dataset as the model input, divide the training set and test set in a ratio of 3: 1, set the number of training times to 1,000, the learning rate to 0.001, and the batch-size of training to 64. The specific parameter settings are shown in Table 1 below. During training, the neural network first retrieves the category of the pictures in the data set to form a label. As a classified category, there are a total of 40 categories in this data set, so 40 labels will be formed. Then the pictures in the folder corresponding to each tag are retrieved in turn, and a txt text file corresponding to each picture is generated. When training, input a batch size image to the neural network each time, which can reduce the network training time. After 1000 iterations of training, the training results of the model are finally obtained.

Table 1 Basic parameter settings of the network model

Name of parameters	Value of parameters
Training percentage	75%
Testing percentage	25%
Training steps	1000
Learning rate	0.001
Batch size	64

In this section, two garbage classification models trained using Inception V3 and Inception V4 neural networks are compared. Every 32 iterations in the experiment, the accuracy of the training and testing sets of the model and the cross-entropy function value are output. Figures 3.1, 3.2, and 3.3 are the comparison results of training set accuracy, test set accuracy, and cross-entropy function values in two different
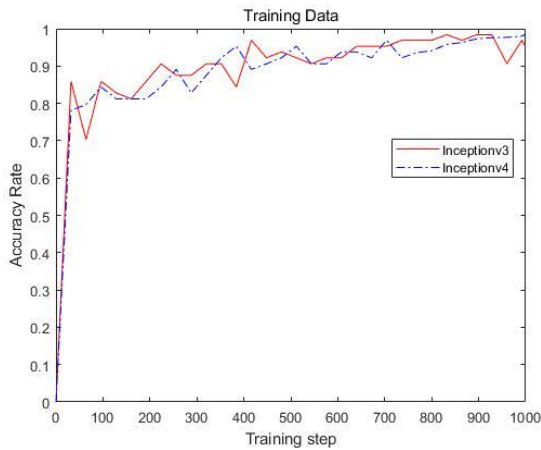
model.

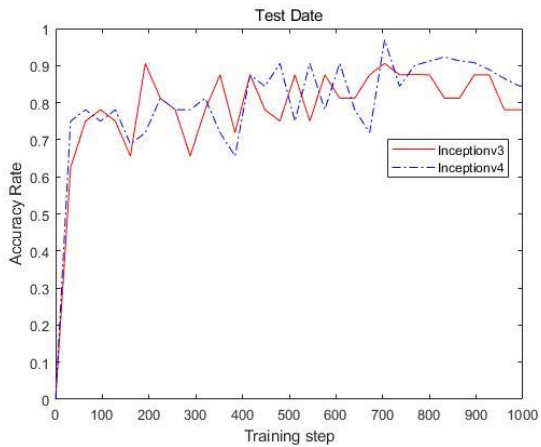Figure 3.1   Training set accuracy
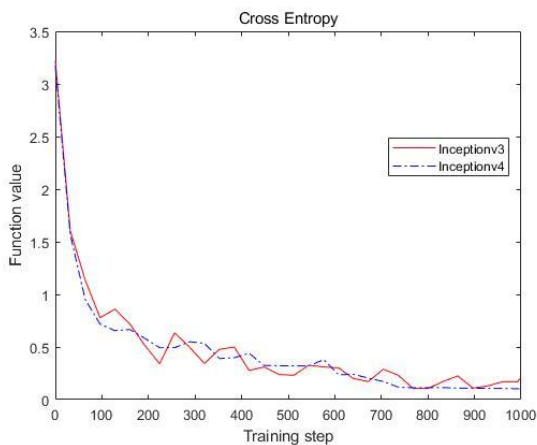

Figure 3.2   Test set accuracy

Figure 3.3   Loss function value

The red horizontal line in the figure above represents the model using the Inception V3 neural network and the blue dotted horizontal line represents the model using the Inception V4 neural network. From the three figures, we can know that with the continuous increase in the number of iterations, the accuracy of the two models on the training set and the test set

is increasing. When the number of iterations reaches about 50, the accuracy of the two networks is accurate. The rates increase rapidly and then fluctuate with the number of iterations. When the number of iterations reached about 750 times, the increase in accuracy on the training set and test set of Inception V4 leveled off, while the accuracy on the training set of Inception V3 still fluctuated slightly. The accuracy rate on the test set The fluctuations are still very obvious. After training, the accuracy of Inception v3 on the training set reached 95.3%, and the accuracy on the test set reached 80.2%；while the accuracy of Inception V4 on the training set reached 98.5%, and the accuracy on the test set The accuracy rate reached 84.5%. The cross-entropy function values of the other two networks also decrease with the number of iterations. The cross-entropy function value of Inception V3 fluctuates more than the cross-entropy function value of Inception V4. After training, the cross-entropy function value of Inception V3 is 0.206, and the cross-entropy function value of Inception V4 is 0.102. It can be seen that the garbage classification models using Inception V3 and Inception V4 networks can both obtain higher accuracy and lower loss function values, but the performance of the latter network is more stable than the former.

When the canned picture set is input to the two models to detect the accuracy of the two models, the top-5 results output by the two network models is shown in Table 2. It can be seen from Table 2 that from the top-5 results of classification, both networks have a high accuracy rate for the classification of canned picture sets. The accuracy rate of Inception V4 is 97.37%, and the accuracy rate of Inception V3 is 89.2%. The accuracy of Inception V4 is higher than that of Inception V3.

Table 2 Top-5 classification results of the two models

Top-5	Inception V3	Top-5	Inception V4
Cans	89.20%	Cans	97.37%
Drink bottle	8.66%	Drink bottle	0.97%
Dry cell	0.83%	Dry cell	0.63%
Metal can	0.42%	Metal can	0.36%
Wire plug	0.23%	Condiment bottle	0.17%

4. SUMMARY

Combining artificial intelligence with garbage classification is achievable. Compared with the traditional garbage classification method, the garbage classification method using deep learning convolutional neural network has higher accuracy, can achieve the purpose of garbage classification, and can save a lot of time and labor costs. This paper discusses two garbage classification models of convolutional neural networks. The results show that both Inception V3 and Inception V4 can achieve the purpose of garbage classification. Both network models can be used in deeper research in the field of garbage classification. After structural analysis and experimental comparison of the two networks, it is found that the stability and accuracy of Inception V4 are slightly higher than those of Inception V3.

6. REFERENCES

[1] Krizhevsky A, Sutskever II, Hinton G. Imagenet classification with deep convolutional neural networks//Proceeding of the Advances in Neural Information Processing Systems. LakeTahoe, USA, 2012: 1097 – 1105.

[2] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprintarXiv:1409.1556, 2014.

[3] Szegedy C, Liu W, Jia Y Q, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, Massachusetts, USA: IEEE, 2015. 1−9.

[4] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of The 32nd International Conference on Machine Learning, pages 448–456, 2015.

[5] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 2818-2826.

[6] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, Alex Alemi. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. arXiv:1602.07261,2016.

[7] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. arXiv preprint arXiv:1512.03385,2015.

[8] T. Tieleman and G. Hinton. Divide the gradient by a running average of its recent magnitude. COURSERA: Neural Networks for Machine Learning, 4, 2012. Accessed: 2015-11-05.

[9] I. Sutskever, J. Martens, G. Dahl, and G. Hinton. On the importance of initialization and momentum in deep learning. In Proceedings of the 30th International Conference on Machine Learning (ICML-13), volme 28, pages 1139–1147.JMLR Workshop and Conference Proceedings, May 2013.