Garbage classification method based on deep learning

Xuemei Wang[1]

[1] Southwest University of science and technology
zty225@126.com

Abstract. Waste classification is an important link to realize waste reduction, harmlessness and recycling. Traditional waste classification is mostly carried out by manual sorting, which has the disadvantages of low sorting efficiency and high labor cost. With the continuous improvement of the level of intelligent equipment in China, it is possible to use computer vision intelligent equipment for waste sorting. However, traditional image classification algorithms use manual feature extraction and classification. When the type and quantity of garbage increase, the classification accuracy and efficiency of intelligent equipment will decline. The successful application of deep learning technology in the field of computer vision improves the accuracy and efficiency of image classification, which makes it a trend for deep learning technology to replace the traditional image classification algorithm for garbage classification. Therefore, it is of great significance to use deep learning technology for automatic classification of waste.

Keywords: Waste classification, Deep Learning, Image classification.

1 Introduce

With the continuous development of economy, the type and quantity of domestic waste are also increasing, which not only pollutes the environment, but also poses a great threat to people's physical and mental health. China has put forward it as early as the 1980s The concept of waste classification, but the effect is not significant. The main reason is that it is difficult for people to classify correctly because of the wide variety of waste and difficult to distinguish.

Deep learning has been applied to various fields, such as speech recognition, emotion recognition, driverless and so on, and achieved excellent results. This paper mainly studies the recyclable waste classification algorithm based on deep learning. Although various intelligent waste classification software came into being in recent years, there are still various deficiencies. Aiming at these problems, this paper improves the waste classification algorithm and improves the accuracy of recyclable waste classification[1].

The so-called deep learning is actually a more extensive machine learning based on artificial neural network. The deep learning architecture includes convolutional neural network, deep neural network and cyclic neural network. Deep learning is to imitate the human brain and work by learning data representation, which needs to be trained

EAI Urb-IoT 2021, December 20-21, Shenzhen, People's Republic of China

DOI 10.4108/eai.20-12-2021.2315017

through a large amount of data. Only through high-quality and sufficient data training can we learn more comprehensive image features, so as to enhance the robustness and generalization of the model. Therefore, the first step of deep learning is to find a large number of appropriate data.

Garbage classification can classify garbage through different standards. Different regions at home and abroad have different garbage classification standards. For example, Germany is generally divided into paper, glass, metal and plastic; Australia is generally divided into compostable waste, recyclable waste and non recyclable waste; Japan is generally divided into plastic bottles, recyclable plastics, other plastics, resource waste, large-scale waste, combustible waste, non combustible waste and harmful waste, etc. At first, it was mainly divided into recyclable and non recyclable in China. Since 2000, Beijing, Shanghai and other cities have become the first batch of pilot cities advocating waste classification. With the increase of the quantity and types of garbage, separate secondary classification can not meet the needs of garbage classification. Most of the current garbage classification is based on four classification. The four classification of garbage can deal with different garbage, so it helps to improve the utilization value of different garbage.At present, garbage classification is mainly carried out by manual sorting[2]. In the process of manual sorting, due to the wide variety of garbage and the limitation of manual sorting accuracy, it is easy to cause garbage sorting errors and reduce the efficiency, and finally reduce the recyclability and utilization efficiency of garbage.

Traditional garbage classification algorithms use manual extraction of image features. Once the feature extraction effect is not good, the recognition effect is not good. This paper uses deep learning algorithm to identify and classify garbage, which can not only improve the efficiency of garbage classification, but also improve the accuracy of garbage classification[3]. This paper mainly studies the four classification image classification model of garbage based on CNN, selects the network model suitable for garbage classification in this paper by using transfer learning and other methods, then uses model integration and other methods to increase the generalization ability of the model, and finally uses knowledge distillation technology to transfer the knowledge learned from the large model to the small model, while losing less accuracy Improve the prediction efficiency and accuracy of the model, and achieve the effect of quickly and accurately identifying waste types[4].

2 Model method

Compared with voice and image, text is more complex and abstract. Human beings can have an overall understanding of the text content after reading the text according to their own understanding ability. However, the semantics in natural language is difficult to be directly understood by computers[5]. Therefore, the text content must be expressed as forms that computers can understand and process, such as 0 and 1. Text representation model is to use numerical or symbolic vectors that can be expressed by computer to represent abstract and complex natural text. In order to better represent the text, it is necessary to extract the most representative features from the text data.

These features should have obvious statistical laws, which can reflect the text distribution in the feature space and minimize the computational complexity of text mapping to the feature space.

2.1 Convolutional neural network

Convolutional neural network is a kind of feedforward neural network with convolution calculation and depth structure. It is one of the classical representative algorithms of depth learning. It has the ability of representation learning[6]. The most important characteristics are "local perception" and "parameter sharing". It can classify the input information according to its hierarchical structure. It is also called "translation invariant artificial neural network".

Convolutional neural network is the main reason why deep learning can achieve breakthrough results in the field of computer vision in recent years. In the training process of the network, features can be automatically extracted from the image, and these features can be combined and classified. Compared with other models, convolutional neural network has better image recognition efficiency and prediction accuracy[7]. A standard convolutional neural network is mainly composed of core layers such as convolution layer, activation layer, pooling layer and full connection layer. The structure of convolutional neural network is shown in Figure 2.

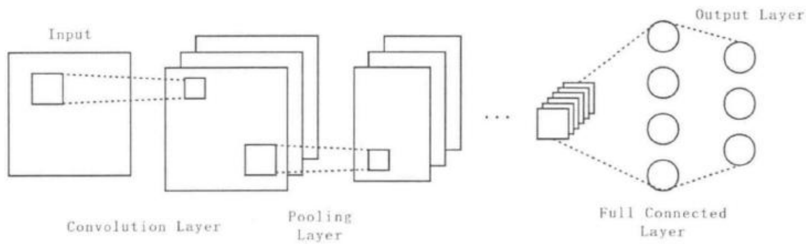

Fig. 1. The structure of convolutional neural network

Convolution layer is the core part of CNN, and most CNN includes multi-layer convolution layer. The convolution kernel in the convolution layer acts as a feature extractor, which is similar to the traditional manually designed feature extractor. The convolution kernel also extracts the features in the data by scanning the input image data for many times. Convolution operation has three important core ideas.

Sparse interactions. Sparse interaction idea uses convolution kernel to interact with the local region in the image. This region is called receptive field, which improves the parameters and efficiency compared with the full connection layer. For example, when processing a three channel picture, the pixels of the image may contain thousands of pixels, but when we only need to detect the edge information in the image, we don't need to connect the pixels of the whole picture, we only need to use the convolution kernel containing dozens or hundreds of pixels to detect. This calculation method not only improves the calculation efficiency, but also saves a large part of the parameter space.

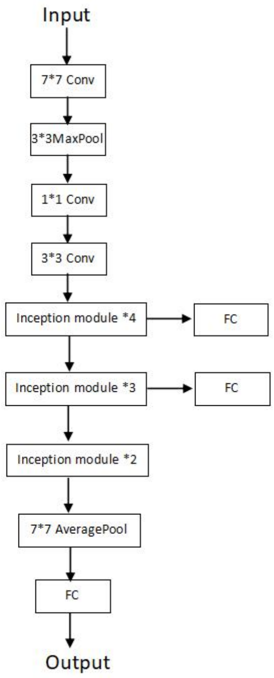

Fig. 2. The structure of convolutional neural network

Parameter sharing. Parameter sharing means that the parameters in the same convolution kernel are the same in the whole image, while the traditional full connection layer provides a learning parameter for each pixel. It is assumed that the characteristic graph parameter after the first layer convolution is 55 × fifty-five × 96, 290400 weight parameters after calculation. The convolution kernel is 11 × eleven × 3. 363 weights after calculation. 290400 if full connection layer calculation is used × 364 = 105705600 parameters, which is huge. However, if the convolution operation is used for operation, the parameter amount is 96 × eleven × eleven × 3 = 34848 parameters, so it reflects the advantages of convolution operation storage and efficiency.

Equivariant representations. The special form of parameter sharing makes the neural network have the property of translation and other variables. In the image, convolution produces a two-dimensional mapping to represent the position of some features such as edges. If we move the input object, its output position will also move by a corresponding amount. This property is useful for operators acting on the same small region.

2.2 Image classification model

Lenet convolutional neural network, which was first used to recognize handwritten digits, was written by Yann Lecun. Convolution layer block and full connection layer block are the main components of lenet network. There are mainly two operations in the convolution layer: convolution and pooling. Convolution identifies the edge and other features in the image. Pooling reduces the sensitivity of features to position through translation invariance. The convolution kernel size is 5 × 5. Use SIGMOD to activate the function. After the second convolution layer, the number of convolution cores is increased to 16. Because each convolution process will reduce the characteristic graph, the number of channels is appropriately increased to make the two convoluted parameter sizes similar. The maximum pooling is adopted in the pooling layer, and the size of the pooled convolution kernel is 2 × 2, step size is 2. The full connection layer uses flatten to transform each sample into a two-dimensional vector representation, and the vector length is the product of channel, height and width. The full connection layer has three layers, and the network outputs of each layer are 120, 84 and 10 respectively.

In 2012, alexnet convolution network was recognized by everyone for the first time in Imagenet image classification competition. It is the first real deep CNN. In the early stage of Imagenet competition, most image classification models were mainly based on traditional SVM and boost methods. He proved that an excellent image classification network can still be designed without using a manually designed feature extractor. Using an 8-tier network, alexnet won the Imagenet 2012 image recognition challenge. There are many significant improvements in the design of alexnet compared with lenet.

In 2014, vggnet won the second place in the Imagenet competition from the visual geometry group laboratory. Visual geometry group first proposed using the repetition of convolution module to increase the recognition accuracy. Similar to alexnet network, vggnet is also a convolution layer block, which is connected to the full connection layer for classification. Vggnet has 16 layers and 19 layers, including 550m parameters. VGG convolution layer uses padding of 1, convolution kernel size of 3, pool layer of 2, and 1 × The convolution layer of 1 uses multiple smaller convolutions to replace a larger convolution, which reduces the amount of parameters and increases the ability of nonlinear mapping.

In 2014, googlenet won the championship of Imagenet image recognition challenge. By designing a sparse network structure, but can produce dense data, it can not only enhance the performance of the neural network, but also ensure the computational efficiency of the network. The core module of googlenet is inception structure, which adopts parallel mode and contains four components: 1×1 convolution, 3×3 convolution, 5×5 convolution and 3×3. At last, the operation results of the four components are combined on the channel. The network in the convolution layer can extract various details in the input, and 5×5. The convolution kernel size of 5 can also cover most of the input of the receiving layer

3 Experimental results and analysis

3.1 Dataset

Before building a garbage classification dataset, you need to understand the explicit provisions of garbage classification. After the introduction of the waste classification policy, the waste classification policies of various cities have different provisions. Although the waste classification standards in each region are different from the names of various types of waste, the domestic waste is generally divided into four categories. Taking the names of various types of waste in Shanghai as an example, they are recyclable, hazardous waste, wet waste and dry waste. Since Shanghai officially began to implement waste classification on July 1, 2019, and the rules for waste types are relatively perfect, the waste classification data set of this subject is specified based on the waste classification rules of Shanghai. The existing public garbage classification data sets include Huawei garbage classification data set, trashnet garbage data set, etc. because there are few types of garbage classification data sets, this paper uses different methods to construct garbage classification data sets.

The garbage classification data set in this paper contains a total of 14087 garbage pictures, including 4 major categories and 60 sub categories respectively. The data set includes various common garbage categories in daily life, such as fruit peel, leftovers, vegetable leaves and roots in wet garbage; Recyclables include plastic bottles, glass bottles, books, etc; Plastic bags, disposable fast food boxes, cigarette butts, etc. in dry waste; It can be seen from the harmful waste traditional Chinese medicine tablets, LED bulbs, ointments, etc. that the data of most types of waste in the data set is about 200, of which the data of disposable cups and PE plastic bags are small, 113 and 114 respectively. The ointment has a large amount of data, with 532 figures.

3.2 Image enhancement technology

In the actual image garbage classification and recognition, the high quality of the image is often difficult to guarantee, because the real image will be affected by various factors, such as different illumination intensity, complex background, image occlusion and image color change. Therefore, if you want to obtain a model with strong generalization ability and strong anti-interference ability, you can use data enhancement technology to increase the changes of image morphology, color and brightness when training the model. This random change has been added to the model during training, which can increase the robustness of the model.

3.3 Model training and results

In this paper, several classification networks are used for comparative experiments to analyze the improvement of network accuracy. The results are shown in Table 1.

Tab. 1. Classification results

	Epoch	ACC
CNN	30	90.12%
LeNet-5	30	90.39%
VGG16	30	92.56%

From the table data, from CNN to VGG16, as the network depth and width of the model become deeper, the accuracy of the model has been improved to 92.56%, which shows that the improvement of the complexity of the model can improve the accuracy of waste classification.

4 Conclusion

Aiming at the problem of garbage classification, this paper uses the deep learning method to classify garbage, and realizes the automatic classification and recognition of garbage. This algorithm can be applied to various devices. The model constructed in this paper can only identify a single waste at a time, so it is more suitable for large shopping malls, office buildings, parks and other places. For some complex places, more complex solutions need to be adopted.

References

1. Li H . Characteristics of Domestic Waste and Suggestion for Waste Classification in Beijing Rural Area. Environmental Sanitation Engineering, 2008.
2. Andrews A . Radioactive Waste Streams: Waste Classification for Disposal. Library of Congress. Congressional Research Service. 2006.
3. Dimitrakopoulos R , Godoy M . Grade control based on economic ore/waste classification functions and stochastic simulations: examples, comparisons and applications. Transactions of the Institution of Mining and Metallurgy, Section A, 2014.
4. White G , Cabrera C , Palade A , et al. WasteNet: Waste Classification at the Edge for Smart Bins. 2020.
5. Islam M S , Hannan M A , Basri H , et al. Solid waste bin detection and classification using Dynamic Time Warping and MLP classifier. Waste Management, 2014, 34(2):281-290.
6. Simonyan K , Vedaldi A , Zisserman A . Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps. Computer ence, 2013.
7. Chapelle O , Haffner P . SVMs for Histogram-Based Image Classification. IEEE Trans on Neural Networks, 1999, 10(5):P.1055-1064.