Research on Garbage Classification Algorithm Based on Machine Learning

Yuxin Bai, Shaoru Li, Jingjing Fan(✉), and Jiana Zhao

Hebei University of Architecture, Zhangjiakou, China
565134037@qq.com

Abstract. With the development of urbanization and the improvement of consumption levels, garbage disposal has become one of the major challenges facing environmental protection. In order to classify garbage efficiently and accurately and promote the construction of "two-oriented society" and smart cities, this topic studies a garbage classification technology based on machine learning. Based on the PyTorch framework design, I compared the training and testing of the AlexNet model and the VGG model. Through testing the experimental results, it was found that using the VGG19 model and optimization method, the accuracy of garbage classification can reach 95%, which is higher than other models. Finally, the VGG19 model is used for training and transfer learning.

Keywords: Waste classification · machine learning · PyTorch · VGG model

1 Preface

In recent years, China's economy has developed rapidly, the quality of life of urban residents has improved significantly, and the amount of domestic waste generated has also increased. Traditional garbage classification mainly relies on manual operation. This method has low efficiency, large errors, and high resource consumption, which is not conducive to the popularization and promotion of garbage classification. Combining artificial intelligence and machine learning technology with traditional garbage classification methods, computers are used to simulate the human learning process to achieve automated garbage classification, which has the advantages of efficiency, accuracy, and automation.

At present, domestic research on garbage classification image recognition technology is mainly based on convolutional neural networks (CNN), and uses deep learning frameworks such as PyTorch and TensorFlow to design and optimize models. Foreign research mainly focuses on model innovation and optimization based on advanced technologies such as deep learning, transfer learning, and reinforcement learning. Today, my country's garbage classification image recognition technology still faces problems such as lack of data, inconsistent standards, immature algorithms, and single application scenarios. Based on this current situation, this topic seeks an efficient and accurate garbage classification algorithm by comparing and optimizing existing machine learning

W. Yu and L. Xuan (Eds.): MLICOM 2023, LNICST 518, pp. 27–37, 2024.
https://doi.org/10.1007/978-3-031-71716-1_3

algorithms, and verifies its classification accuracy and intelligence through experiments, providing reference for technology development of garbage classification field.

2 Experimental Data and Environment

2.1 Data Set Collection and Partitioning

This experiment downloaded 15515 pictures from the kaggle website, including pictures of used batteries, kitchen waste, plastic bottles, old cartons, discarded masks, etc. Then divided the pictures into the following four categories according to the traditional method: recyclable garbage, hazardous waste, kitchen waste and other garbage.

After collecting the data samples, 70% of the data in the samples are used for training and 30% of the data are used for testing. This experiment uses a training set to train the algorithm and a test set to evaluate the algorithm.

2.2 Experimental Environment Construction

This experiment uses Anaconda as the Python environment manager and Jupyter Notebook as the interactive programming and visualization tool.It also uses PyTorch as the deep learning framework, Python as the main programming language and PyCharm as the code editor (Table 1).

Table 1. Experimental environment

factor	property
OS	Windows 10
CPU	Intel(R) Core(TM) i7-4790
memory	16 GB
frequency	3.60 GHZ

3 Experimental Process and Result Analysis

3.1 Image Preprocessing

We need to convert the image into a tensor type that PyTorch can handle to facilitate data input and model training. When acquiring images, in order to obtain better results, the average filter is combined with other filters for image denoising.

This experiment uses standardization and normalization data preprocessing methods to convert the data set into a relatively uniform distribution.

The standardized mathematical formula is:

$$z = \frac{x_i - \mu}{\sigma} \tag{1}$$

Among them, x_i is the original data, μ is the mean of the data and σ is the standard deviation of the data. Through this formula, the original data can be standardized to eliminate the dimensional influence between different features and better perform data analysis and modeling.

The commonly used normalization method is maximum and minimum normalization. The formula for maximum and minimum normalization is:

$$\hat{x}_i = \frac{x_i - min(X)}{max(X) - min(X)} \tag{2}$$

Among them, x_i is the original data, min(X) is the minimum value of the data, max(X) is the maximum value of the data, and $\hat{x}_i$ is the normalized data.

3.1.1 Image Preprocessing Display

Image preprocessing can improve the training effect and accuracy of the model in image classification and recognition, because it can remove interference factors in the image and enhance the observability and detectability of effective factors. At the same time, it can also simplify data processing, which is very beneficial to the extraction and classification of image features.

3.2 Model Architecture and Data Processing

3.2.1 Convolutional Neural Network

CNN extracts features from images through convolution operations. The convolution operation can be seen as a process of scanning the image to be identified using a sliding window. The pixel value in a window is combined with a convolution kernel (also known as filters) are multiplied and added to obtain a feature quantity, and then the entire frame is removed to obtain a feature quantity. The convolution kernel can be regarded as a template for pattern recognition. The parameters of the convolution kernel must be obtained through training. Different convolution kernels can extract different features. Such as edges, textures, etc. Through repeated iterative training, the model can automatically learn the optimal convolution kernel.

The so-called different network structures are composed of convolutional layers, pooling layers, fully connected layers, etc. in CNN in a certain order to form different network structures, such as LeNet, AlexNet, VGG and ResNet, etc. Different network structures can be used for different image recognition and processing tasks. AlexNet and VGG are one of the most classic network structures and are widely used in the field of image recognition. In order to build a garbage classification model based on machine learning, this study adopted VGGNet and AlexNet as the basic models, and made appropriate modifications and adjustments to them to adapt to the data set and tasks of this study.

The AlexNet network has an eight-level architecture, and the first five levels are convolutional architectures. On this basis, the pooling layer, Dropout and RuLu activation functions are introduced, and LRN local response normalization is used to achieve effective extraction of deep features, reduce over-matching, and greatly improve recognition accuracy. After completing the five levels of feature extraction, the remaining three levels are used to train the obtained features, laying a good foundation for future recognition work (Fig. 1).

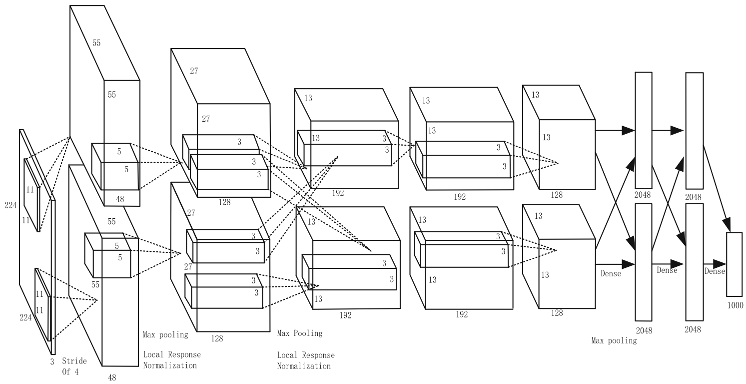

Fig. 1. AlexNet network model

An important feature of the VGG network structure is the use of 3x3 convolution kernels instead of larger kernels such as 11 × 11 and 5 × 5 in AlexNet. In addition, the VGG also uses two 3 × 3-level convolution kernels instead of the 5 × 5-level convolution

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
Input (224*224 RGB image)					
Conv3-64	Conv3-64 LRN	Conv3-64 Conv3-64	Conv3-64 Conv3-64	Conv3-64 Conv3-64	Conv3-64 Conv3-64
maxpool					
Conv3-128	Conv3-128	Conv3-128 Conv3-128	Conv3-128 Conv3-128	Conv3-128 Conv3-128	Conv3-128 Conv3-128
maxpool					
Conv3-256 Conv3-256	Conv3-256 Conv3-256	Conv3-256 Conv3-256	Conv3-256 Conv3-256 Conv1-256	Conv3-256 Conv3-256 Conv3-256	Conv3-256 Conv3-256 Conv3-256 Conv3-256
maxpool					
Conv3-512 Conv3-512	Conv3-512 Conv3-512	Conv3-512 Conv3-512	Conv3-512 Conv3-512 Conv1-512	Conv3-512 Conv3-512 Conv3-512	Conv3-512 Conv3-512 Conv3-512 Conv3-512
maxpool					
Conv3-512 Conv3-512	Conv3-512 Conv3-512	Conv3-512 Conv3-512	Conv3-512 Conv3-512 Conv1-512	Conv3-512 Conv3-512 Conv3-512	Conv3-512 Conv3-512 Conv3-512 Conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
Soft-max					

Fig. 2. VGG network model

kernel, and the 11 × 11-level convolution core is replaced by five 3x3-level convolution kernels. In this way, the VGG network is constructed using convolution cores of the same size, making the construction of the entire network simple. Another feature is the use of 2 × 2 instead of 3 × 3. When the depth is deeper and the height and width of the image become narrower, this method will reduce the size of the image by half and double the number of channels, thus effectively reducing the parameters in the learning process. The VGG neural network is repeatedly superimposed on a 3 × 2 small-scale convolution core and 2 pools, giving it stronger nonlinear computing capabilities and better fitting performance. This experiment uses a combination of softmax and cross-entropy functions in the model to effectively optimize model performance and accelerate the training process (Fig. 2).

3.2.2 Transfer Learning

Transfer learning refers to a machine learning technology that transfers existing data from one area to another area to achieve better processing results. This method utilizes the relevant or common features existing in multi-source heterogeneous information, and effectively transfers existing information to new problems through effective processing of existing information, thereby accelerating and improving learning efficiency. In deep learning, transfer learning is widely used in various tasks such as image classification, target detection, speech recognition, and natural language processing. There are correlations between many tasks, so existing knowledge and data can be leveraged through transfer learning to accelerate learning and improve performance. It improves learning by applying knowledge from one field to another. Transfer learning can use massive data in one domain to train a model and apply it to another domain, thus avoiding the huge time and resources required to retrain the model in a new domain.. The core idea of transfer learning is to transfer existing knowledge and data to new tasks.The core idea of transfer learning is to transfer existing knowledge and data to new tasks. Specifically,

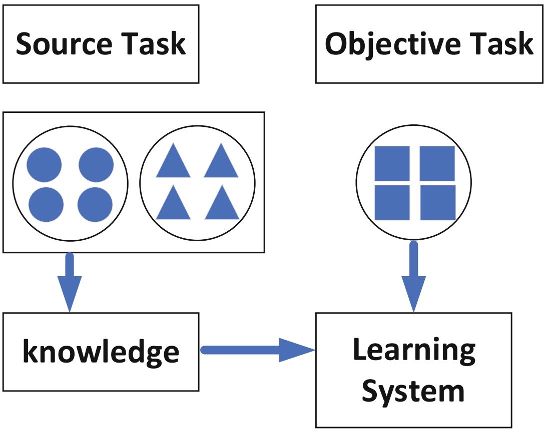

Fig. 3. Schematic diagram of transfer learning

transfer learning usually includes three steps: pre-training the model, feature extraction, and adjusting the model.

The advantage of transfer learning is that it can use existing knowledge and data to accelerate learning and improve performance, especially when the amount of data is insufficient. In addition, transfer learning can also help solve problems such as overfitting and vanishing gradients, while also reducing the cost and time of model training. This experiment will use transfer learning technology to improve model performance (Fig. 3).

3.2.3 The Adam Optimizer

Adam (Adaptive Moment Estimation) optimizer is an optimization algorithm widely used in the field of deep learning. It combines two adaptive learning rate methods: Momentum and RMSProp, thus performing superiorly in many tasks.

The explanation of algorithm:

The core idea of Adam is to maintain two exponential moving averages for each parameter, which are the mean of the gradient (momentum term) and the mean of the square of the gradient (adaptive learning rate term). During each iteration, Adam first calculates the gradient and then updates the two exponential moving averages. Then, these two values are corrected to eliminate the initialization deviation. Finally, the model parameters are updated based on the corrected average values.

3.3 Model Training and Performance Analysis

During the model training process, this experiment uses the convolutional neural network method to automatically extract features from domestic waste. On this basis, transfer learning is used to allocate existing model weights. After training the model, you can use the test set to evaluate it. We will focus on testing the accuracy of graph classification, the epoch required to achieve the optimal solution of the model, changes in learning rate, and changes in loss. If the accuracy of the model is high, it means that the model can effectively classify the garbage data set.

For the classification problem of garbage image recognition, this experiment used three models of AlexNet, VGG16, and VGG19 for training and recognition (Figs. 4, 5, 6, 7).

The experimental results of the three models are compared as follows (Tables 2, 3, 4):

After testing VGG19, the confusion matrix obtained using this model method is shown in Fig. 8. It can be seen that there are many and complex types of image data objects in the waste category, and there are many misclassified data. This will also be Issues we need to focus on in the next phase.

We see that as the training period increases, both the training accuracy and the test accuracy of the three models continue to improve, indicating that the performance of the models continues to improve. According to experimental results, the accuracy of the VGG19 model has always been higher than other models in garbage classification tasks. On the training set, although its initial accuracy is low, as the number of training times increases, the accuracy increases rapidly, and by the 8th cycle of training, the accuracy can reach 95%. On the test set, the accuracy of both the VGG16 model and the AlexNet

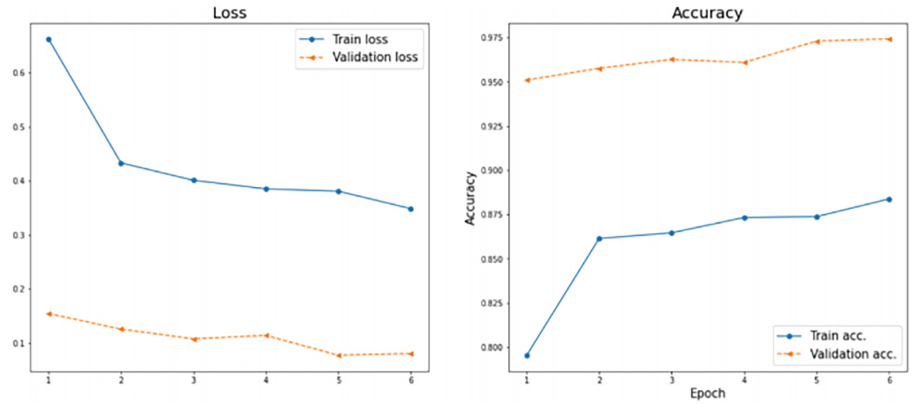

Fig. 4. AlexNet model training process diagram

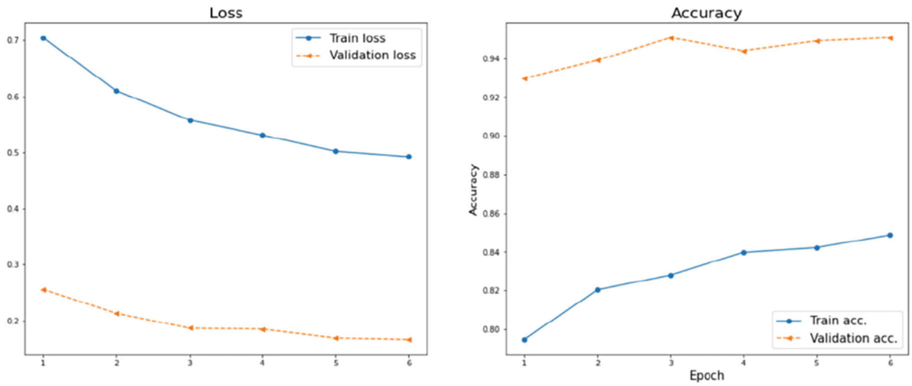

Fig. 5. VGG16 model training process diagram

model is lower than that of the VGG19 model. This shows that in this experiment, the VGG19 model performed better in terms of accuracy and training error.

The VGG19 model has shown good performance in terms of accuracy and robustness of garbage classification. In terms of image classification ability, after training and testing, it has reached a high level. At the same time, it also meets the actual requirements. Needs and has certain practical value. However, there is some room for improvement, such as increasing the dataset size, using higher-level deep learning models and optimizing hyperparameters to improve the accuracy and generalization ability of the model.

This experiment is based on machine learning algorithms and combined with image processing technology to propose a garbage classification method based on convolutional neural networks. The experiment was carried out on the existing sample set to preprocess the data and classify the proposed model. By comparing with the existing sample set, the universality of the proposed model was judged. In the deep learning algorithm, it focuses on the basic theory and characteristics of convolutional neural networks, and

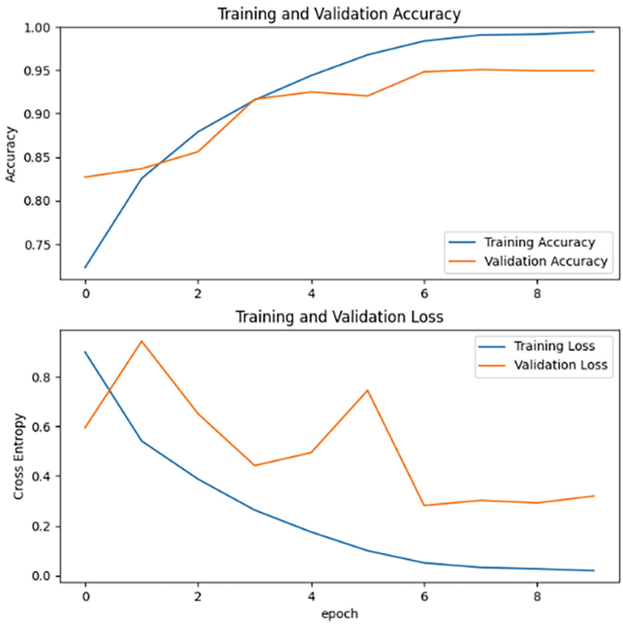

Fig. 6. VGG19 model training process diagram

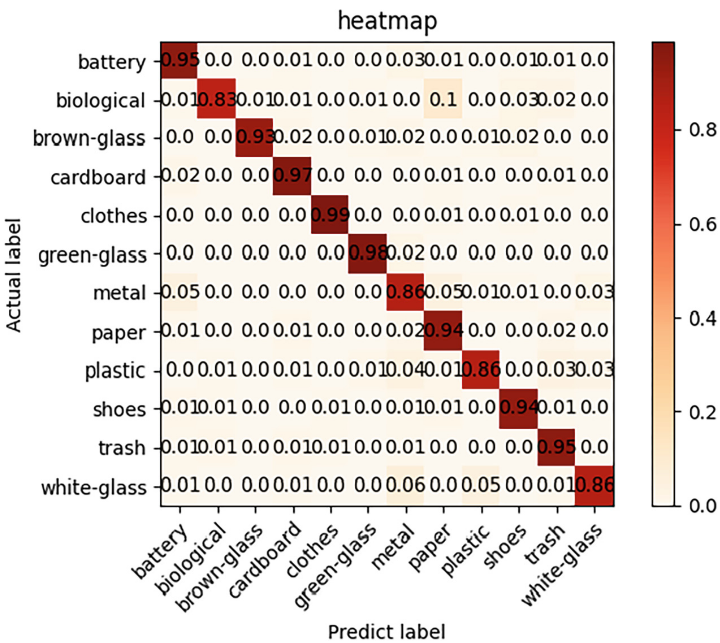

Fig. 7. Confusion matrix

Table 2. VGG19 classification accuracy

cycle	training accuracy	test accuracy
1	0.65	0.72
2	0.73	0.76
3	0.78	0.79
4	0.82	0.82
5	0.85	0.84
6	0.89	0.86
7	0.92	0.91
8	0.95	0.94

Table 3. VGG16 classification accuracy

cycle	Training accuracy	test accuracy
1	0.56	0.55
2	0.62	0.62
3	0.69	0.66
4	0.75	0.75
5	0.83	0.81
6	0.89	0.86
7	0.91	0.88
8	0.93	0.92

Table 4. AlexNet classification accuracy

cycle	Training accuracy	test accuracy
1	0.65	0.63
2	0.73	0.68
3	0.70	0.69
4	0.78	0.75
5	0.80	0.80
6	0.86	0.85
7	0.88	0.86
8	0.90	0.89

makes a detailed analysis of two typical convolutional neural network models - AlexNet and VGGNet. This experiment explores the concepts and methods of transfer learning. It is believed that transfer learning can play a better role in garbage classification problems. It also explores the impact of using different network optimization technologies such as ReLu activation function, LRN and Dropout on model performance. Results show that these techniques can improve model accuracy and performance.

Finally, we verified the effectiveness of this method in garbage classification through an example. In practical applications, model parameters can be adjusted according to specific garbage classification scenarios and data set sizes to obtain better classification results; cameras or sensors can be used to capture images of garbage and input them into the model for classification. The classification results can be used to guide waste processing and further reduce labor costs, thus making a positive contribution to the protection and sustainable development of the urban environment.

3.4 Presentation

The model was tested using the test set and the expected results were achieved. We use the trained model for display, click to upload a picture, select a picture to upload, and click Identify after the upload is successful to classify and identify garbage.

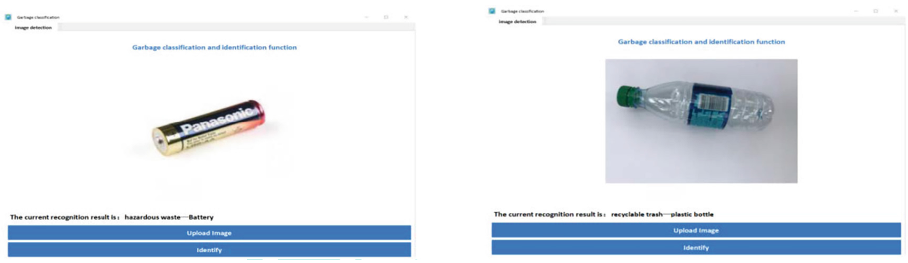

Fig. 8. Image Identification

In practical applications, cameras or sensors can be used to capture images of garbage and input them into the model for classification. The classification results can be used to guide waste disposal.

This experiment is based on the existing convolutional neural network, preprocessing the data on the existing sample set, and classifying the proposed model on the existing sample set, and judging by comparing with the existing sample set. Universality of the proposed model. In practical applications, model parameters can be adjusted according to specific garbage classification scenarios and data set sizes to obtain better classification results. Deep learning algorithms can also be used to build an efficient and accurate garbage classification system, improve the automation of garbage classification, further reduce labor costs, and thus make a positive contribution to the protection and sustainable development of the urban environment.

4 Conclusions

The experiment compared the training and testing of the ALexNet model, VGG19 and VGG16 models. It was found that using the VGG19 model and related optimization methods, the accuracy of garbage classification on the training set can reach 95%, which is better than the other two models. High accuracy.The experiment used transfer learning technology, which significantly shortened the training time and improved model performance in the process. The experiment used different network optimization technologies such as ReLu activation function, LRN and Dropout in the model. The results show that these techniques can improve the accuracy of the model.

References

1. Cheng, S.: Garbage image classification model and analysis based on machine learning. Guilin University of Electronic Science and Technology (2022)
2. Chunmei, Z., Yueqin, Z., Ding, S.: Research on handwritten digit recognition and application based on CNN under PyTorch. Comput. Digit. Eng. **49**(06), 1107–1112 (2021)
3. Lan, C.: Research on garbage detection and classification methods based on deep learning. South China University of Technology (2021)
4. Dingbang, F.: Research on Algorithm for Character Recognition of Handwritten Mathematical Formulas Based on Convolutional Neural Network. Huaqiao University, Fujian (2020)
5. Wang, Y.: Research on machine learning methods for garbage classification data. Heilongjiang University (2020)
6. Wang, A.: Research and application of image recognition based on deep learning. University of Electronic Science and Technology of China (2018)
7. Wang, R.: Research on garbage identification and classification based on machine learning. Lanzhou University of Technology (2021)
8. Lu, Z.: Research on garbage classification and detection methods based on deep learning. Nanchang University (2022)