International Journal of Management Science Research, Vol.8, Issue 6, (Jun) 2025
ISSN 2536-605X
DOI: 10.53469/ijomsr.2025.08(06).05

A Garbage Image Classification Application Based on Transport Learning

Haoyang Lei[1], Li Zhou[2]

Jincheng College, School of Computer and Software, Sichuan University, Chengdu, Sichuan 611731

Abstract: ***With the continuous development of society, the continuous improvement of people's living standards and the great improvement of lifestyle, waste disposal issues have received more and more attention. If there is no taxonomy of science to deal with waste, on the contrary is to throw away, then the soil, water resources and other human survival environment will be a great degree of pollution. Waste sorting is a key part of effectively alleviating environmental pollution problems. Migration learning can effectively solve the problem of model training that results in adaptation when there is not enough data, This paper uses a migration learning method based on pre-trained networks to train the model to classify garbage images by using 80% of the garbage image dataset as a training set and the rest as a test set. By comparing the experimental data of different models under the same parameters, ResNet152 is selected as the pre-trained network model, and the test accuracy is up to 82%.***

Keywords: Transfer learning;ResNet152; Refuse classification.

1. INTRODUCTION

With the development of the social economy, people's income has increased, their consumption capacity has gradually increased, and the increase in purchasing power has been accompanied by more and more waste generated. However, if the generated domestic waste is discarded everywhere or disposed of in an improper manner at all times, it will cause great pollution to the soil, water sources, air and other environmental factors on which people depend for survival, so rational disposal of the generated waste is an effective means to reduce the pollution of the environment. At present, according to the Chengdu Municipal Living Waste Management Regulations, which were officially implemented in 2021, the waste is classified into four categories: recyclable waste, kitchen waste, hazardous waste and other waste. However, at present, the relevant domestic technology for image recognition and classification of these garbage categories is not yet mature, and the application of garbage classification algorithms is not widespread. There are many early studies on the basis of SEM model, Raspberry Pi and other technologies to achieve robot recognition of garbage images, but there are low precision, slow operation, difficult training and other problems [1]. With the increasing maturity of convolutional neural networks, deep learning has also been added to the application of image recognition, and although with the help of deep learning algorithms, regions-CNN, YOLO and other models have more or less solved some of the above early problems, but there are still difficult training problems. Therefore, this paper compares GoogLeNet and ResNet152, two easy-to-train neural network models, choosing one with higher accuracy as a pre-trained network model.

Pal et al. (2025) developed an AI-based credit risk assessment system with intelligent matching mechanisms for supply chain finance, enhancing decision-making accuracy in economic management [1]. Computer vision technologies have progressed substantially, with Guo et al. (2025) improving vehicle detection through an enhanced YOLOv8 network architecture [2], while Jin et al. (2024) achieved superior object detection and pose estimation by combining hybrid task cascade with high-resolution networks [3]. Zhang et al. (2025) contributed to data analytics by developing machine learning-based anomaly detection techniques for complex biomechanical datasets [4]. Smart infrastructure has benefited from Fang's (2025) adaptive QoS-aware cloud-edge architecture for real-time water service management [5], complemented by Tu's (2025) innovative approach to reliable vehicle platooning using redundant 5G link aggregation in smart road systems [7]. In environmental health, Ma et al. (2024) investigated the impacts of metal exposure on fetal development through maternal and cord blood analysis [8]. Financial applications continue to evolve, as shown by Jiang et al. (2025) who created Investment Advisory Robotics 2.0, leveraging deep neural networks for personalized financial guidance [9]. Industrial applications include Zhao et al.'s (2024) deep learning approach for optimizing steel production scheduling [10]. The financial sector further advanced with Yang et al.'s (2025) CNN-based method for stock market sentiment analysis and prediction [11].

2. THEORETICAL KNOWLEDGE

2.1 Convolutional Neural Networks

Convolutional neural networks are multilevel neural networks that are improved on traditional neural networks, and are also a class of forward neural networks consisting of convolutional computations and deep structures [2]. After continuous improvement and perfection, it has become one of the representative algorithms of deep learning. Each layer of a convolutional neural network consists of a single and independent neuron, and its implicit layer consists of convolutional, pooled, and fully connected layers. The fact that convolutional neural networks provide features with a certain degree of spatial immutability is the biggest difference between them and other neural networks. This difference is due to the convolutional and sampling operations of convolutional neural networks being based on artificial neural networks [3]. It was introduced. Due to the spatial immutability of the derived features, convolutional neural networks are very widely used in the field of image recognition and classification processing. And convolutional neural networks have fewer parameters to consider than other networks such as feeder neural networks [4]. The structure of convolutional neural networks is shown below, as shown in Figure 1.

Figure 1: Graph of the structure of convolutional neural networks

2.2 GoogLNet

GoogLeNet is the epoch-making deep Divine Network model based on the Inception module designed by giant internetworking company Google, which won the 2014 ImageNet competition. The innovation of GoogLeNet is that its framework is different from the linear model framework, but adopts the parallel branch structure [5]. It effectively solved the problem of insufficient features extracted from traditional network models.

Inception is to put a number of convolution or pooling operations together to form a network module, and then through the composition of the network module as a unit to form the entire network structure to design the Divine Network. In practical cases, different-scale pictures require different sizes of convolutional nuclei to make the network model perform better, and the Inception module provides multiple convolutional-size operations in parallel to meet the network's own requirements for choosing different sizes. Moreover, pooling operations are essential in the network, so pooling layers are also added to the network in parallel with different convolutional nuclei. The following is the Inception module structure, as shown in Figure 2.

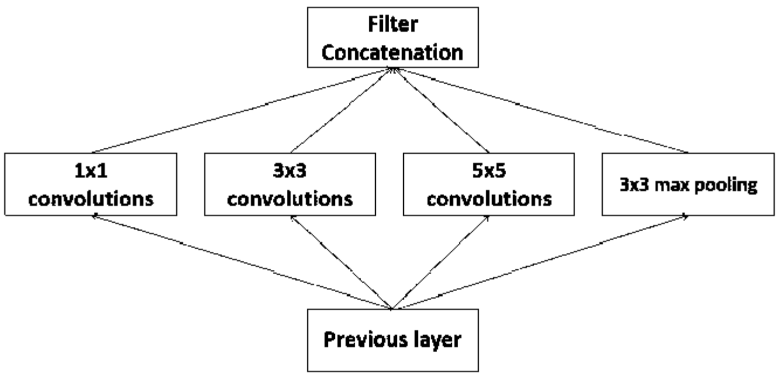

Figure 2: Inception module structure shows intent

2.3 ResNet

ResNet introduces residual learning to solve problems that are difficult to optimize deep networks, by Kaiming He, Xiangyu Zhang, Shao Qingren, Jian Sun [6]. I proposed it. It is a network model structure that gains deeper than ever before [7]. In general, we use two layers of convolutional and pooled layers to deepen the network structure or increase the width of the network by deepening the feature diagram in order to achieve greater robustness [8]. Features, but increasing width or depth requires an upper limit, otherwise it will cause gradient disappearance and gradient explosion when the network structure is too deep, or reduced computing efficiency when the networks are too wide. ResNet is proposed to solve the problem of network degradation caused by increasing the depth, making it possible to provide network performance simply by deepening the depth of the network. Figure 3 shows the intent of the ResNet152 network structure.

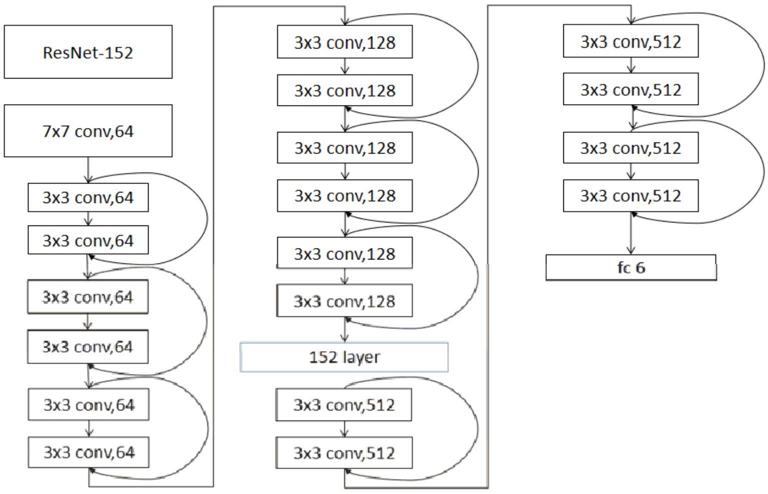

Figure 3: ResNet152 Network Structure Indicates Intent

3. APPLICATION IMPLEMENTATION

3.1 Migration model training process

First, the dataset contains all the image information as 80 percent as samples of the training set and 20 percent as examples of the test set to divide the entire dataset. The image information from the training and test sets is written to train.csv and test.csv files, respectively. Read the training set image information and test set image information separately according to the file name, and preprocess the corresponding image, including adjusting the image size to the same size. Enhance the image, etc., then load the ResNet152 pre-trained model, then freeze the network, and finally replace the last full-connectivity layer, put the model on the GPU, and then put the training data and test data on the TPU for training. The process of training the migration model is shown in Figure 4 below.

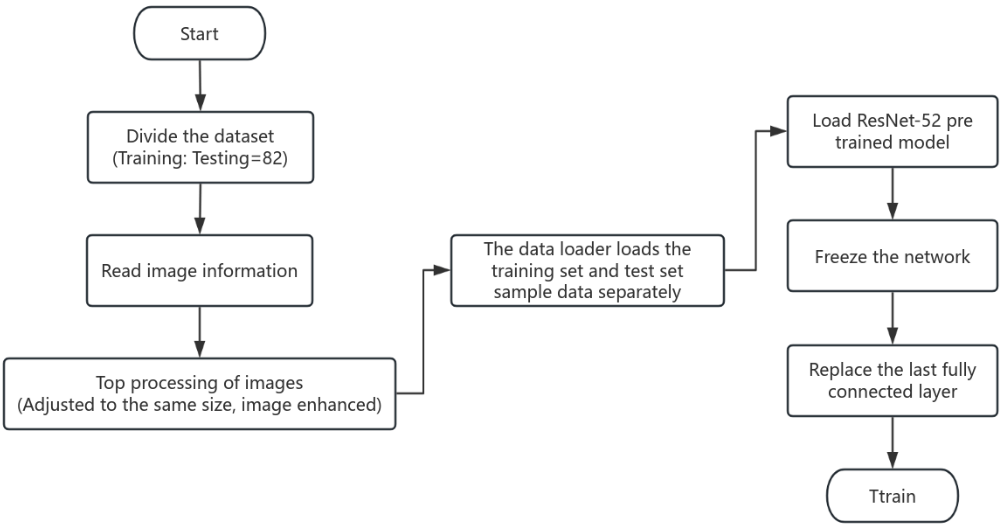

Figure 4: A migration model training flow chart

3.2 Experimental data and preprocessing

Huawei Garbage Classification Challenge dataset was used as the sample dataset. There were 14,802 sample images in the dataset, including other garbage, kitchen waste, recyclable waste and hazardous waste. There are six small categories, such as defaced plastics, eight small categories of kitchen waste including leftovers, fruit peel, etc., 23 small categories for recyclable waste includes chargers, cans, cardboard boxes, etc. and three small categories on hazardous waste which include dry batteries. Use 80 percent of the image dataset as a training set and the remaining 20 percent as a test set. In order to meet the experimental requirements, the analysis is mainly based on the waste generated in the daily life of the residents. Because the types of everyday waste are complex and diverse, and the color, shape, etc. of these waste vary greatly, we do not break down the categories of this waste, but divide it into four common categories based on the selected data set. The four types of waste are recyclable waste, hazardous waste, kitchen waste and other waste. The types of junk images in the dataset are shown in Figure 5.

Pre-processing the image of the selected training set and test set according to the experimental requirements. First, the size of the training set and the test set image is adjusted to 256x256, and the training set image is trimmed in the central area of the image. The image is enhanced by random horizontal flipping, modifying contrast-degree and degree of saturation, normalizing and normalizing.

Figure 5: Kinds of junk images in a dataset

3.3 Network comparison

Compare the results of GoogLeNet and ResNet152 experiments by comparing the training accuracy of the two network models with the same epoch, batch size, and learning rate parameters, and selecting the network model with the higher accuracy as the experiment. Table 1 below provides a comparison of the experimental data of the two network models.

Table 1: Comparison of GoogLeNet and ResNet152 experimental data

	Epoch	Batch_Size	Learning_Rate	Test_Acc
GoogLeNet	3	32	0.005	0.6499
ResNet152	3	32	0.005	0.7208

From the experimental data, it can be concluded that the test accuracy of ResNet152 is higher than that of GoogLeNet under the same three parameters, so ResNet152 is chosen as the pre-trained network model.

3.4 Framework

The model is trained and tested in Python's deep learning framework. Hardware environment: Intel (R) Core (TM) i7 10870H @ 2.90GHZ 2.90GHZ, NVIDIA GEFORCE RTX 2060. Software environment: Anaconda 3+cuda11.1+pytorch1.6+Windows10. In this experiment, Epoch (the number of times trained), Batch_Size (the number of network data submitted each time), Learning_Rate (learning rate) three parameters of migration model training.

Batch_Size: The number of network data submitted each time, adjusting the Batch_Size can effectively solve the problem of excessive data volume. When the BatchSize is too large, it causes the model to run slower. Therefore, the optimal Batch_Size can effectively balance capacity and efficiency, making the network model's performance and speed optimal [9].

Table 2: Table of experimental data

Epoch	Batch_Size	Learning_Rate	Train_Loss	Test_Loss	Training accuracy	Test accuracy
5	32	0.005	1.4737	1.7679	0.7128	0.7035
5	16	0.005	2.3462	2.7636	0.668	0.7004
5	8	0.005	3.879	3.6871	0.6051	0.696
5	32	0.001	0.7993	0.7283	0.7523	0.7912
5	32	0.01	2.8286	4.0709	0.6804	0.6677
5	32	0.0001	1.1781	1.0701	0.7202	0.7295
10	32	0.001	0.7012	1.041	0.7832	0.7345
20	32	0.001	0.6399	0.6135	0.8006	0.8267

3.5 Training results and analysis

In a trial under the ResNet 152 pre-trained network model, randomly divided each category of garbage images in a picture set according to 80% of the training set and 20% of the test set, and tested with different values for Epoch, Batch _ Size, and Learning _ Rate. The experimental data obtained are shown in table 2 below.

From the data in the table, you can see that ResNet152's test accuracy is 82%. Because increasing Epoch inevitably increases time, and the time spent training the network is not enough to be judged, the data combination (5,32,0.001) performs best when compared to the test accuracy of all Epoch = 5 data. Based on the comparison of experimental data, it can be shown that when Epoch, Batch_Size does not change, the Learning_Rate increases, the test accuracy decreases, and the inverse increases. However, if the value of the Learning_rate is too small when the data combination (5,32,0.0001), the test accuratecy will still decrease; When the Batch _ Size and Learning _ Rate are constant, the Epoch value increases, which will improve the test accuracy, and vice versa. When the Epoch, Learning _ Rate is constant, the smaller the Batch _ size value, as can be seen by the comparison of the three data combinations (5, 32, 0.005), (5, 16, 0.005) and (5, 8, 0.005). Although the test accuracy does not fluctuate much, the loss function value of the training and test sets will increase. Figure 6 is the loss curve of the data combination (5, 32, 0.001). Figure 7 is the training accuracy and test accuracy curve of the data combination (5, 32, 0.001).

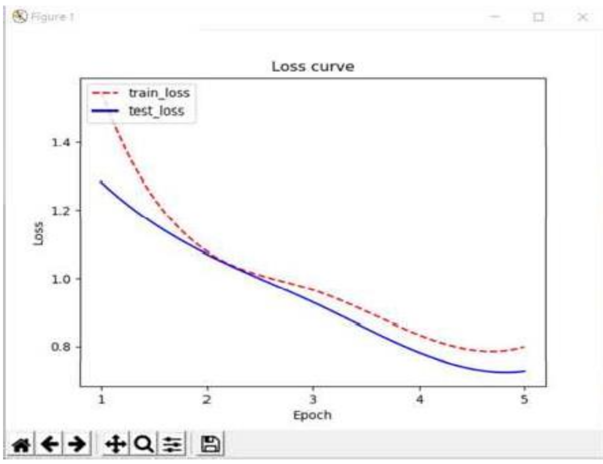

Figure 6: The loss curve of the data combination (5,32,0.001)

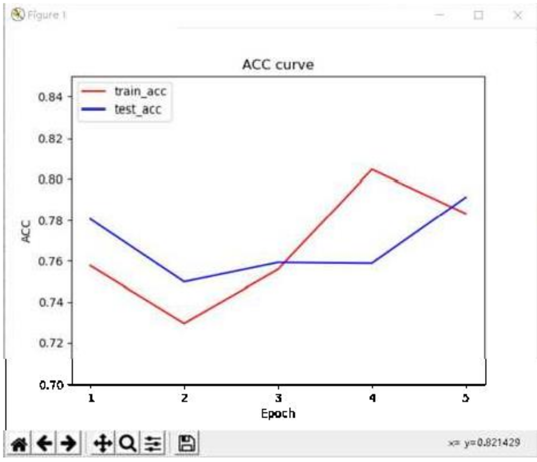

Figure 7: Training accuracy and test accuracy curve of the data combination (5,32,0.001)

4. CONCLUSION

4.1 Advantages and disadvantages

Advantages: According to the experimental data, the training accuracy of ResNet152 is relatively high, It effectively solved the problem of network degradation, which is that when the depth of the network increases, the performance decreases, and made forward and reverse propagation algorithms very smooth, making it significantly easier to optimize deeper models.

Disadvantages: According to the experimental results, the learning structure of ResNet152 is sensitive to fluctuations in network weights. Slight changes in network weights can significantly reduce training efficiency.

4.2 Application prospects

Waste sorting is an environmentally friendly industry with good prospects, and with the introduction of policies on garbage sorting by provincial governments, artificial intelligence, big data and other technologies continue to innovate and develop. Demand will bring a market, and the market for garbage sorting will be greatly expanded as a result. Artificial intelligence + garbage sorting will be a new development trend in the garbage classification industry, which can effectively solve many problems existing in the existing garbage separation industry, such as improving the efficiency of garbage cleaning and reducing the labor cost of garbage disposal. Using artificial intelligence for garbage sorting cannot be ambitious, and its effectiveness is not immediate, and this technology needs continuous innovation and continuous breakthroughs in order for the garbage separation industry to flourish.

REFERENCES

[1] Pal, P. et al. 2025. AI-Based Credit Risk Assessment and Intelligent Matching Mechanism in Supply Chain Finance. Journal of Theory and Practice in Economics and Management. 2, 3 (May 2025), 1–9. DOI:https://doi.org/10.5281/zenodo.15368771

[2] Guo, Haocheng, Yaqiong Zhang, Lieyang Chen, and Arfat Ahmad Khan. "Research on Vehicle Detection Based on Improved YOLOv8 Network." Applied and Computational Engineering 116 (2025): 161-167.

[3] Jin, Yuhui, Yaqiong Zhang, Zheyuan Xu, Wenqing Zhang, and Jingyu Xu. "Advanced object detection and pose estimation with hybrid task cascade and high-resolution networks." In 2024 International Conference on Image Processing, Computer Vision and Machine Learning (ICICML), pp. 1293-1297. IEEE, 2024.

[4] Zhang, Shengyuan, et al. "Research on machine learning-based anomaly detection techniques in biomechanical big data environments." Molecular & Cellular Biomechanics 22.3 (2025): 669-669.

[5] Fang, Z. (2025). Adaptive QoS - Aware Cloud - Edge Collaborative Architecture for Real - Time Smart Water Service Management.

[6] Ding, Y., Wang, X., Yuan, H., Qu, M., & Jian, X. (2025). Decoupling feature-driven and multimodal fusion attention for clothing-changing person re-identification. Artificial Intelligence Review, 58(8), 1-26.

[7] Tu, Tongwei. "Reliable Vehicle Platooning via Redundant 5G Link Aggregation in Smart Roads." (2025).

[8] Ma, Haowei, et al. "Maternal and cord blood levels of metals and fetal liver function." Environmental Pollution 363 (2024): 125305.
[9] Jiang, G., Yang, J., Zhao, S., Chen, H., Zhong, Y., & Gong, C. (2025). Investment Advisory Robotics 2.0: Leveraging Deep Neural Networks for Personalized Financial Guidance. Preprints. https://doi.org/10.20944/preprints202504.1735.v1
[10] Zhao, H., Chen, Y., Dang, B., & Jian, X. (2024). Research on Steel Production Scheduling Optimization Based on Deep Learning.
[11] Yang, W., Lin, Y., Xue, H., & Wang, J. (2025). Research on Stock Market Sentiment Analysis and Prediction Method Based on Convolutional Neural Network.