

# *Garbage Recognition Genius: Implementation of an Intelligent Waste Classification System and Interactive Platform Based on Enhanced ResNet50*

Chenlu Guo, Youli Chen, Haoze Yu, Qiang Qu\*

*School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan, 114051, China*

*\*Corresponding author*

**Keywords:** Intelligent Garbage Classification; ResNet50; Channel Attention Mechanism; Image Classification; PyQt5

**Abstract:** In response to the actual demands of automated sorting of urban domestic waste, and in view of the shortcomings of the traditional ResNet50 model in extracting features in complex environments and the low accuracy of small target waste recognition in the garbage classification task, this paper proposes an enhanced ResNet50 garbage classification recognition model that incorporates a multi-scale feature fusion module and a channel attention mechanism. By optimizing the structure of the residual module and the training strategy, the classification accuracy and generalization ability of the model are improved. At the same time, a corresponding visual interaction platform based on the PyQt5 framework is developed to realize the full process functions of garbage image upload, model inference, and real-time display of classification results. Experimental results show that on the public garbage classification dataset, the Top-1 classification accuracy of the proposed enhanced ResNet50 model reaches 96.27%, which is 3.15 percentage points higher than that of the original ResNet50 model. The precision, recall rate, and F1 value have also been significantly improved. The ablation experiments fully verify the effectiveness of each enhancement module. The inference and result display response time for a single garbage image of the developed interaction platform is less than 200ms, demonstrating good practicality and user experience, and providing technical support and solutions for the intelligent garbage sorting in urban sanitation scenarios.

## 1. Introduction

With the continuous acceleration of China's urbanization process, the generation and transportation volume of urban domestic waste have been continuously increasing[1]. The classification and treatment of domestic waste has become a key link in promoting the construction of "zero-waste cities" and implementing the "dual carbon" strategic goals. However, the traditional manual sorting mode has problems such as low sorting efficiency[2], high labor costs, harsh working environment, and unstable implementation of classification standards, which have become difficult to meet the needs of large-scale and standardized urban waste treatment. In recent years, the rapid

development of computer vision and deep learning technologies has provided mature technical support for intelligent waste sorting[3]. Therefore, conducting research on high-precision, strong generalization, and easy-to-apply classification recognition models and supporting interactive systems is of great practical significance and engineering application value for promoting the intelligent upgrade of the sanitation industry and improving the efficiency of the entire chain of urban domestic waste treatment[4].

Currently, the classification recognition technology has completed the technical iteration from traditional manual feature engineering to end-to-end recognition based on deep learning[5]. The traditional machine vision methods based on artificial features such as color, texture, and shape have poor robustness in recognizing complex scenes and deformed waste, and are difficult to adapt to the complex working conditions of actual waste classification[6]. While convolutional neural networks, with their powerful automatic feature extraction ability, have become the mainstream technical framework for image classification tasks. The ResNet50 model, through the residual connection mechanism, effectively solves the problem of gradient disappearance and degradation in deep networks[7], and has excellent feature extraction ability while maintaining high computational efficiency. It has become the mainstream basic model in the field of waste classification recognition. Existing research based on ResNet in waste classification mostly focuses on the direct transfer application of basic models, and the optimization for problems such as small class differences, strong background interference, and variable target shapes in garbage images is insufficient, resulting in limited recognition accuracy and generalization ability of the model in actual scenarios[8]. Moreover, most studies only focus on algorithm-level optimization and lack the design of a visual interactive system for practical applications, making it difficult to achieve convenient and practical application of the technical achievements.

In response to the shortcomings of existing research and the actual engineering requirements of urban waste classification, this paper aims to improve the accuracy of waste classification recognition and the practicality of the system as the core goal, proposing an enhanced ResNet50 model with a multi-scale feature fusion module and channel attention mechanism. Through the optimization of the basic network architecture and training strategy, the model's ability to extract key features and adapt to complex scenarios is strengthened. At the same time, based on the PyQt5 visualization development framework, a human-computer interaction platform integrating image upload, model inference, and real-time display of classification results is designed and implemented, completing the entire process design from algorithm optimization to system implementation.

## 2. Enhanced ResNet50 Architecture and System Overall Design

### 2.1 Overview of Basic Theories

Convolutional Neural Network (CNN) is the core foundation model for computer vision tasks. Its core mechanism involves the stacking of convolutional layers, activation layers, and pooling layers to achieve end-to-end automatic extraction of image features from low-level details to high-level semantic features. Compared to traditional manual feature engineering, it possesses stronger feature representation capabilities and generalization performance[9].

Two-dimensional convolution is the core operation of CNN. For an input feature map  $X \in \mathbb{R}^{H \times W \times C_{in}}$  (where  $H$  and  $W$  are the height and width of the feature map, and  $C_{in}$  is the number of input channels), assuming the convolution kernel  $K \in \mathbb{R}^{k \times k \times C_{in} \times C_{out}}$  (where  $k$  is the size of the convolution kernel and  $C_{out}$  is the number of output channels), the numerical calculation formula for the value at position  $(i, j, c)$  in the output feature map  $Y$  is:

$$Y(i, j, c) = \sum_{m=0}^{k-1} \sum_{n=0}^{k-1} \sum_{d=0}^{C_m-1} X(i+m, j+n, d) \cdot K(m, n, d, c) + b_c \quad (1)$$

In the formula,  $b_c$  represents the bias term of the  $c$  th output channel. After the convolution operation, a linear rectifier unit (ReLU) is usually introduced as the activation function to introduce nonlinear fitting capability to the network, and its expression is:

$$\text{ReLU}(x) = \max(0, x) \quad (2)$$

The pooling layer reduces the feature dimension by downsampling the feature map, while retaining the key features. This reduces the computational load of the model and commonly uses the maximum pooling operation to compress features by taking the maximum value of a local region, thereby further enhancing the feature invariance of the model.

As the depth of the network increases, traditional CNNs tend to encounter problems such as gradient disappearance and network degradation - that is, after the number of network layers increases, the training accuracy reaches saturation or even decreases, and it is impossible to improve the model performance by deepening the network. The ResNet series of models solve the degradation problem of deep networks by proposing the residual learning mechanism, and have become the mainstream backbone network for tasks such as image classification and object detection.

The core idea of residual learning is to reconstruct the mapping target of the network. For the stacked nonlinear network layers, let the expected ideal mapping they are supposed to fit be  $H(x)$ , the network no longer directly learns this mapping, but learns the residual mapping  $F(x) = H(x) - x$ , and the final output mapping is:

$$H(x) = F(x) + x \quad (3)$$

In the formula,  $x$  represents the input feature of the residual block, which is directly passed to the output end through the shortcut connection. When the ideal mapping becomes the identity mapping, it is sufficient to set the weights of the residual branch  $F(x)$  to 0 to achieve this, without the need to fit complex identity transformations. This effectively solves the problem of gradient disappearance and degradation in deep networks. The original ResNet50 network architecture is shown in Figure 1.

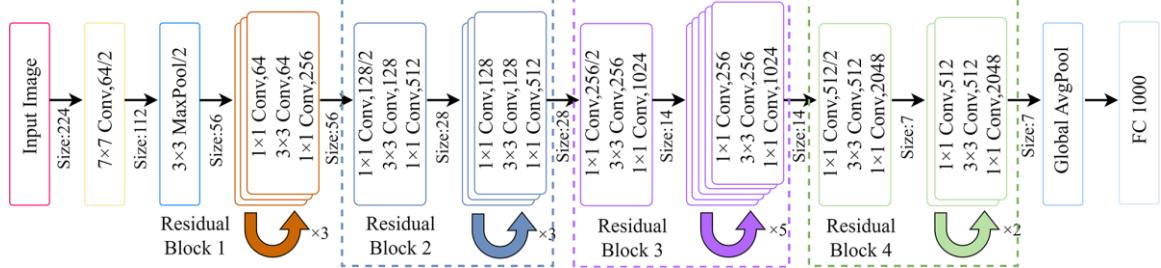


Figure 1: The original ResNet50 network architecture

## 2.2 Enhanced ResNet50 Architecture Design

In order to enable the model to adaptively strengthen the feature channels that contribute significantly to garbage classification and suppress the ineffective channels corresponding to background noise, this paper embeds the squeeze-excitation (Squeeze-and-Excitation, SE) channel attention module into the bottleneck residual block of ResNet50 to achieve adaptive re-calibration of feature channels and enhance the model's ability to focus on key features[10].

For the input feature map  $U \in \mathbb{R}^{H \times W \times C}$ , perform global average pooling along the spatial dimension,

compressing each two-dimensional channel feature into a global representation scalar, and obtaining the channel-level global statistical feature  $z \in \mathbb{R}^C$ . The compression value of the  $c$ -th channel is calculated as follows:

$$z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W U_c(i, j) \quad (4)$$

A nonlinear gating mechanism is constructed through two fully connected layers to learn the dependency relationships between channels and generate the weight coefficients for each channel. Firstly, the channel number is reduced from  $C$  to  $C/r$  (where  $r$  is the compression rate, and in this paper,  $r=16$ ) through a dimensionality reduction fully connected layer. After ReLU activation, it is restored to  $C$  through a dimensionality increase fully connected layer. Finally, the channel weight vector  $s \in \mathbb{R}^C$  within the range of 0 to 1 is generated through the Sigmoid activation function. The formula is:

$$s = \sigma(\text{FC}_2(\text{ReLU}(\text{FC}_1(z)))) \quad (5)$$

In the formula,  $\sigma(\cdot)$  represents the Sigmoid activation function, and  $\text{FC}_1$  and  $\text{FC}_2$  are the dimensionality reduction and dimensionality increase fully connected layers, respectively.

Multiply the generated channel weights  $s$  element-wise with the corresponding channels of the original input feature map  $U$  to complete the weighted calibration of the feature channels, enhance the effective features and suppress the invalid noise, and output the calibrated feature map  $\tilde{U}$ :

$$\tilde{U}_c = s_c \cdot U_c \quad (6)$$

In the formula,  $\tilde{U}_c$  represents the feature map after calibration for the  $c$ th channel, and  $s_c$  is the corresponding weight coefficient for that channel. The three feature fusion structures are shown in Figure 2.

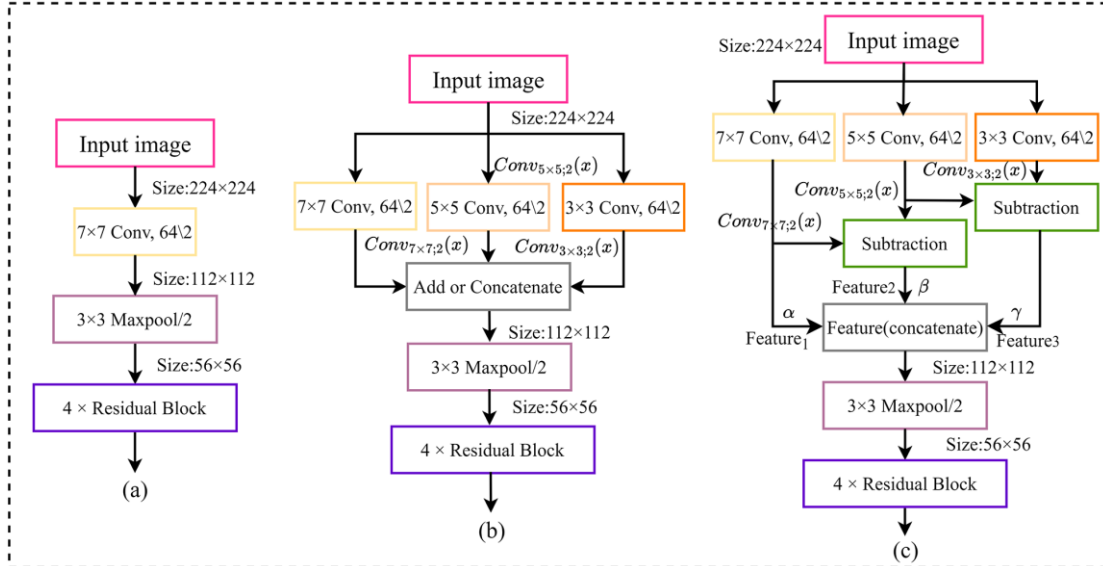


Figure 2: Three-feature fusion structure

Figure 2(a) shows the feature fusion module of the standard ResNet50, Figure 2(b) shows the multi-scale feature fusion module, and Figure 2(c) is our enhanced version of the feature fusion module of ResNet50. For the input feature map  $F_{in} \in \mathbb{R}^{H \times W \times C}$ , the features output by each branch are respectively denoted as  $F_1, F_2, F_3, F_4$ . Among them, the global average pooling branch first obtains the

global feature through global average pooling, adjusts the channel number through  $1 \times 1$  convolution, and then restores the spatial size consistent with the input feature map through bilinear interpolation to ensure the dimension matching of the multi-branch features. Subsequently, the output features of the four branches are concatenated in the channel dimension, and then channel fusion and dimension compression are completed through  $1 \times 1$  convolution to restore the channel number to the input feature's channel number  $C$ , obtaining the final fused feature  $F_{out}$ , and the calculation formula is:

$$F_{cat} = \text{Concat}(F_1, F_2, F_3, F_4) \quad (7)$$

$$F_{out} = \text{Conv}_{1 \times 1}(F_{cat}) \quad (8)$$

In order to reduce the computational cost of the model and improve the real-time inference speed of the interaction platform, this paper makes a lightweight improvement to the bottleneck residual block of ResNet50. The  $3 \times 3$  standard convolution in the residual branch is replaced with a depthwise separable convolution (Depthwise Separable Convolution, DSC) to ensure the feature extraction capability while significantly reducing the number of parameters and computational cost of the model.

Depthwise separable convolution decomposes the standard convolution into two steps: the depthwise convolution uses a single convolution kernel for spatial dimension feature extraction for each input channel, and the pointwise convolution realizes information interaction in the channel dimension through a  $1 \times 1$  convolution. For a convolution operation with an input feature map size of  $H \times W \times C_{in}$ , an output feature map size of  $H \times W \times C_{out}$ , and a convolution kernel size of  $k \times k$ , the computational cost of the standard convolution is:

$$\text{Flops}_{std} = H \times W \times C_{in} \times C_{out} \times k \times k \quad (9)$$

And the computational cost of depthwise separable convolution is:

$$\text{Flops}_{dsc} = H \times W \times C_{in} \times k \times k + H \times W \times C_{in} \times C_{out} \quad (10)$$

The ratio of the computational amounts of the two is:

$$\frac{\text{Flops}_{dsc}}{\text{Flops}_{std}} = \frac{1}{C_{out}} + \frac{1}{k^2} \quad (11)$$

When the kernel size  $k = 3$ , the computational cost of depthwise separable convolution is only 1/8 to 1/9 of that of standard convolution, which can significantly reduce the computational complexity of the model and improve the inference speed.

To address the problems of class imbalance among samples and model overfitting in the garbage classification dataset, this paper simultaneously optimizes the training strategy: First, a composite data augmentation strategy is adopted, including random horizontal flipping, random rotation, brightness/contrast perturbation, Mixup and CutMix mixed augmentation methods, etc. These methods are introduced during the training stage to expand the diversity of training samples and improve the generalization ability of the model; Second, a cosine annealing learning rate scheduling strategy is used, with the initial learning rate set at  $1 \times 10^{-4}$ . During the training process, the learning rate decays according to a cosine law with the number of training rounds, improving the convergence effect of the model; Third, a cross-entropy loss function with label smoothing is adopted to alleviate the problem of model overfitting and overconfidence. Its expression is:

$$L_{CE} = - \sum_{c=1}^C \left[ (1 - \varepsilon) \cdot y_c + \frac{\varepsilon}{C} \right] \cdot \log(p_c) \quad (12)$$

In the formula,  $C$  represents the total number of garbage categories,  $y_c$  is the one-hot encoded form

of the true label,  $p_c$  is the category probability predicted by the model, and  $\epsilon$  is the label smoothing coefficient. In this paper,  $\epsilon$  is set to 0.1.

### 2.3 The overall architecture of the intelligent garbage classification system

This study presents an intelligent waste classification system built upon an enhanced ResNet50 architecture, implementing a closed-loop pipeline encompassing front-end interaction, back-end preprocessing, model inference, and result feedback. The system comprises four core functional modules: an image input module enabling user upload and preview of local waste images in standard formats; an image preprocessing module performing resizing to  $224 \times 224$  pixels, pixel value normalization, and channel dimension conversion to generate model-compliant tensor inputs; a model inference module as the central component loading trained enhanced ResNet50 weights to execute forward propagation and output Top-1 classification results with confidence scores; and a result visualization module displaying classification categories, confidence levels, and disposal recommendations on the front-end interface with synchronous original image presentation. The operational workflow proceeds through user image upload triggering preprocessing, subsequent back-end model inference, and final front-end result presentation to achieve integrated visual waste classification. The overall functional link of the system is shown in Figure 3.

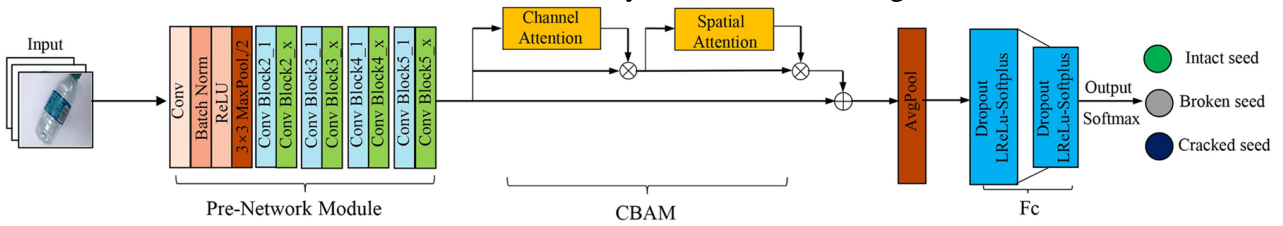


Figure 3: Overall functional link of the system

To lower the usage threshold of the model and facilitate the practical implementation of intelligent garbage classification technology, this paper, based on the PyQt5 visualization development framework of Python, designs and develops a lightweight human-computer interaction platform, providing users with an easy-to-operate and responsive garbage classification recognition interaction interface.

## 3. Experimental results and analysis

### 3.1 Experimental environment and dataset

All experiments were conducted under unified hardware and software environments comprising Intel Core i9-12900K central processing units, NVIDIA GeForce RTX 3090 graphics processing units with 24 gigabytes memory, 32 gigabytes DDR4 system memory, and 512 gigabytes solid-state storage, with Ubuntu 20.04 long-term support operating system, Python 3.8.10, PyTorch 1.12.1, CUDA 11.3, cuDNN 8.3.2, and PyQt5 5.15.7. The public TrashNet dataset containing 2527 balanced images across six waste categories including glass, paper, cardboard, plastic, metal, and trash was partitioned into 1769 training, 505 validation, and 253 testing samples via 7:2:1 random splitting. The six types of public datasets are shown in Figure 4. Training set preprocessing incorporated composite data augmentation comprising random horizontal flipping with probability 0.5, random rotation between  $-15$  and  $15$  degrees, brightness and contrast perturbation of  $\pm 10$  percent, and Mixup blending with coefficient 0.2, followed by resizing to  $224 \times 224$  pixels and ImageNet normalization using means of 0.485, 0.456, 0.406 and standard deviations of 0.229, 0.224, 0.225, while validation and testing sets received only resizing and normalization without augmentation to ensure evaluation authenticity.



Figure 4: Six types of public datasets

### 3.2 Evaluation indicators and experimental design

To comprehensively evaluate the classification performance of the model, this paper adopts accuracy (Accuracy), macro-average precision (Macro-Precision), macro-average recall (Macro-Recall), and macro-average F1 value (Macro-F1) as the core evaluation indicators. The calculation formulas for each indicator are as follows:

Accuracy refers to the proportion of correctly classified samples to the total number of samples, reflecting the overall classification accuracy of the model. The formula is:

$$\text{Accuracy} = \frac{\sum_{c=1}^C TP_c + TN_c}{\sum_{c=1}^C TP_c + TN_c + FP_c + FN_c} \quad (13)$$

Precision refers to the proportion of samples predicted as belonging to a certain category that actually belong to that category. Recall refers to the proportion of samples that actually belong to a certain category that are correctly predicted. The F1 score is the harmonic mean of precision and recall. In this paper, we adopt the macro-average method to calculate the overall metrics for multi-classification tasks, which involves calculating the metrics for each category first and then taking the arithmetic mean. The formulas are as follows:

$$\text{Macro-Precision} = \frac{1}{C} \sum_{c=1}^C \frac{TP_c}{TP_c + FP_c} \quad (14)$$

$$\text{Macro-Recall} = \frac{1}{C} \sum_{c=1}^C \frac{TP_c}{TP_c + FN_c} \quad (15)$$

$$\text{Macro-F1} = \frac{1}{C} \sum_{c=1}^C \frac{2 \cdot \text{Precision}_c \cdot \text{Recall}_c}{\text{Precision}_c + \text{Recall}_c} \quad (16)$$

In the above formula,  $c$  is the total number of garbage categories (in this paper,  $c = 6$ ), and  $TP_c$ ,  $TN_c$ ,  $FP_c$ ,  $FN_c$  are the real cases, true negative cases, false positive cases and false negative cases of category  $c$  respectively.

Comparative experiments were conducted against classical computer vision architectures including VGG16 with uniform  $3 \times 3$  convolution kernels, lightweight MobileNetV2 employing depthwise separable convolutions and inverted residuals, and original ResNet50 with residual learning mechanisms, all trained under identical environments and strategies to ensure equitable

evaluation. Ablation studies were designed to quantify contributions of squeeze-and-excitation channel attention, multi-scale feature fusion, and depthwise separable convolution modules through five experimental configurations: baseline ResNet50, baseline with squeeze-and-excitation, baseline with multi-scale feature fusion, baseline with depthwise separable convolution, and the complete enhanced architecture integrating all three modules, with consistent training protocols and evaluation metrics across all groups.

### 3.3 Experimental results and analysis

The overall classification performance of the enhanced resnet50 model proposed in this paper is excellent on the test set, and each evaluation index has reached a high level. Specifically, the top-1 classification accuracy of the model was 96.27%, the macro average accuracy was 96.02%, the macro average recall rate was 96.18%, and the macro average F1 value was 95.94%, which fully verified the effectiveness and robustness of the model in garbage classification task.

The classification performance comparison results of different models on the test set are shown in Table 1. It can be seen from the table that the enhanced resnet50 model proposed in this paper is significantly better than other comparison models in all evaluation indexes: compared with the original resnet50, the accuracy of this model is improved by 3.15%, the macro average F1 value is improved by 3.26%, and the performance is significantly improved; Compared with vgg16, the accuracy of this model is improved by 6.94%, and the parameter is only 13.5% of vgg16, which realizes the double improvement of performance and efficiency; Compared with the lightweight model mobilenetv2, the accuracy of the model in this paper is still 4.97% higher, which ensures the lightweight while maintaining a higher classification accuracy.

Table 1: Comparison of classification performance of different models

Model	Accuracy(%)	Macro-Precision(%)	Macro-Recall(%)	Macro-F1(%)
VGG16	89.33	88.97	89.12	88.76
MobileNetV2	91.30	91.05	91.22	90.89
ResNet50	93.12	92.87	93.05	92.68
Proposed Model	96.27	96.02	96.18	95.94

Ablation results demonstrate that the squeeze-and-excitation module improves accuracy from 93.12 to 94.47 percent and macro-averaged F1-score from 92.68 to 94.02 percent by enhancing critical features and suppressing background noise, while the multi-scale feature fusion module further elevates accuracy to 94.86 percent and F1-score to 94.35 percent through capturing multi-scale characteristics of variable waste morphologies. Depthwise separable convolution substitution alone reduces accuracy marginally from 93.12 to 92.89 percent but decreases parameters from 25.56 million to 14.23 million and computational cost from 4.12 to 1.87 giga floating-point operations, achieving substantial lightweighting. The complete enhanced architecture integrating all three modules attains 96.27 percent accuracy and 95.94 percent macro-averaged F1-score, confirming synergistic effects that simultaneously achieve performance enhancement and model compression. The ablation results of each enhancement module are shown in Table 2.

Table 2: Ablation results of each enhancement module

Model	Accuracy(%)	Macro-Precision(%)	Macro-Recall(%)	Macro-F1(%)
Base	93.12	92.87	93.05	92.68
Base+SE	94.47	94.21	94.38	94.02
Base+MSFF	94.86	94.63	94.77	94.35
Base+DSC	92.89	92.65	92.81	92.44
Proposed Model	96.27	96.02	96.18	95.94

To more intuitively show the contribution of each module, this paper records the training process

of this model and the original resnet50, and draws the training loss curve and the verification accuracy curve, as shown in Figure 5. From the loss curve, we can see that the convergence speed of this model is faster, and the training loss decreases more smoothly; From the verification accuracy curve, we can see that the verification accuracy of this model is always higher than the original resnet50, and finally converges to a higher level, which fully verifies the effectiveness of the enhanced architecture.

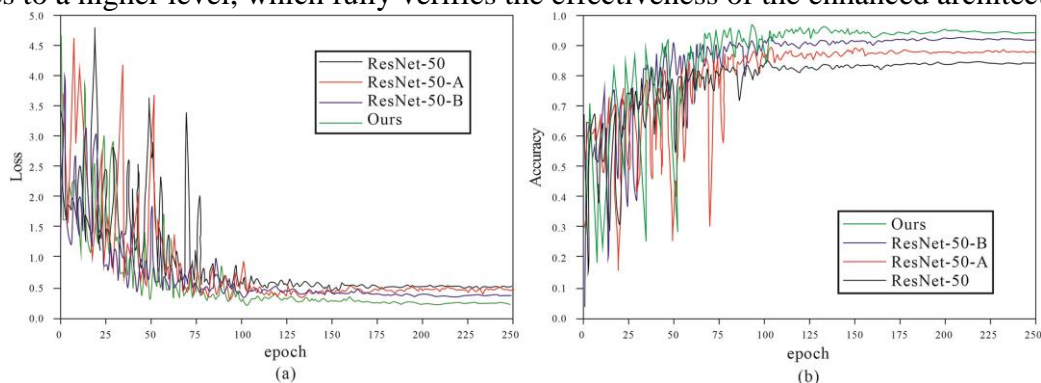


Figure 5: Training loss curve and verification accuracy curve

To further analyze the classification performance of the model in each category, this paper draws the confusion matrix of the model in this paper on the test set, as shown in Figure 6. From the confusion matrix, it can be seen that the classification accuracy of the model is very high in most categories, and there is only a small amount of confusion between "paper" and "cardboard", "glass" and "plastic", which is caused by the similarity of the visual characteristics of these types of garbage, but the overall number of confusion samples is very small, and the classification ability of the model is still at a high level.

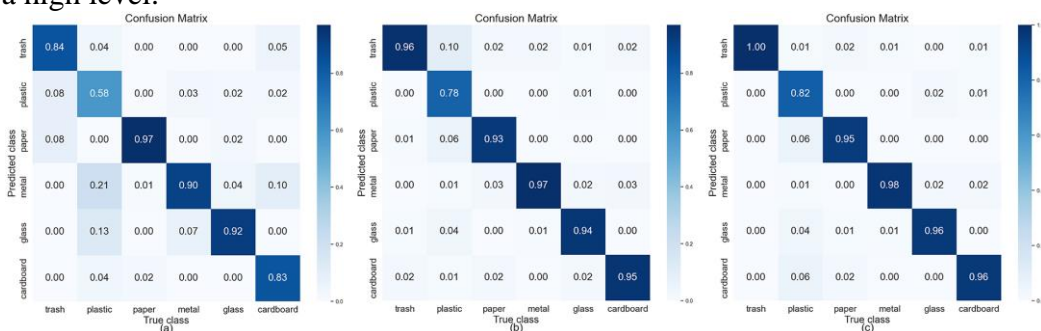


Figure 6. confusion matrix

In addition, this paper compares the parameters and calculations of different models, and the results are shown in Table 3. It can be seen from the table that the parameters of the model in this paper are 18.72m and the amount of calculation is 2.89gflops, which are reduced by 26.7% and 29.9% respectively compared with the original resnet50, realizing significant lightweight and more suitable for real-time reasoning deployment on the interactive platform.

Table 3: Comparison of parameters and calculation of different models

Model	Parameters (M)	Computational Complexity (GFLOPs)
VGG16	138.36	15.50
MobileNetV2	3.50	0.32
ResNet50	25.56	4.12
Proposed Model	18.72	2.89

The intelligent waste classification interactive platform developed in this paper based on pyqt5 has

a complete functional link of image upload, preview, recognition and result display. In order to verify the functional integrity of the platform, this paper selects 100 garbage images with different categories and resolutions for functional test. The test results show that the platform can stably support the upload and preview of mainstream image formats such as JPG and PNG. The "start recognition" button can normally trigger the background model reasoning. The classification results (including garbage categories, confidence, classification and disposal suggestions) can be displayed on the front-end interface in real time and accurately. All functions operate normally without abnormal err

## 4. Conclusion

According to the actual needs of automatic sorting of municipal solid waste, this paper proposes an enhanced resnet50 garbage classification and recognition model with embedded channel attention mechanism, multi-scale feature fusion module and lightweight improvement. At the same time, a supporting visual intelligent garbage classification interactive platform is developed based on pyqt5 framework, and the whole process design from algorithm optimization to system implementation is completed; The experimental results show that the top-1 classification accuracy of the enhanced resnet50 model proposed in this paper is 96.27%, which is 3.15% higher than the original resnet50 model, and the macro average F1 value is 95.94%. At the same time, the parameters and calculation amount of the model are reduced by 26.7% and 29.9% respectively compared with the original resnet50 model, which realizes the synchronous improvement of classification accuracy and reasoning efficiency. Ablation experiments fully verify the effectiveness of each enhancement module. The average response time of a single garbage image of the interactive platform developed in this paper is only 187ms, which has excellent real-time performance and practicability, and can provide reliable technical support for the intelligent garbage classification and landing of urban sanitation scenes. There is still some room for optimization in this study. In the future, we will further optimize the feature extraction ability of the model for the identification of complex outdoor scenes and multi-form mixed waste, expand the category coverage of waste identification, promote the extreme lightweight of the model and the deployment and adaptation of edge devices, optimize the multi terminal adaptation and batch identification function of the interactive platform, and promote the large-scale application of intelligent waste classification system in more actual scenes.

## Acknowledgements

This work was supported by the Undergraduate Innovation Training Program of University of Science and Technology Liaoning in 2026.

## References

- [1] Zhou, Y., Wang, Z., Zheng, S., Zhou, L., Dai, L., Luo, H., ... & Sui, M. (2024). Optimization of automated garbage recognition model based on ResNet-50 and weakly supervised CNN for sustainable urban development. *Alexandria Engineering Journal*, 108, 415-427.
- [2] Ghosh, D., & Goswami, A. (2025). Enhanced deep learning framework for efficient garbage classification in smart waste management systems. *Information Sciences*, 719, 122462.
- [3] Arul, N., & Vidya, R. (2026). An Enhanced Imbalanced Garbage Classification Model using Attention-based ResNet50 with Generative Adversarial Networks (GAN). In *A Circular Economy for Manufacturing Waste Management* (pp. 101-117). River Publishers.
- [4] Li, L., Wang, R., Zou, M., Guo, F., & Ren, Y. (2025). Enhanced ResNet-50 for garbage classification: Feature fusion and depth-separable convolutions. *PloS one*, 20(1), e0317999.
- [5] Fabijan, A., Zawadzka-Fabijan, A., Fabijan, R., Zakrzewski, K., Nowosławska, E., & Polis, B. (2024). Assessing the accuracy of artificial intelligence models in scoliosis classification and suggested therapeutic approaches. *Journal of Clinical Medicine*, 13(14), 4013.

- [6] Gupta, S. (2022). *Behavioural finance and investment decisions: Impact analysis and mitigation strategies in the context of COVID pandemic and DIY investing*. *21st century innovations in management, science and technology, education and social sciences*, 129.
- [7] Anand, R., Lakshmi, S. V., Pandey, D., & Pandey, B. K. (2024). *An enhanced ResNet-50 deep learning model for arrhythmia detection using electrocardiogram biomedical indicators*. *Evolving Systems*, 15(1), 83-97.
- [8] Wan, J., Li, B., Wang, K., Teng, X., Wang, T., & Mao, B. (2024). *An improved ResNet50 for environment image classification*. *Procedia Computer Science*, 242, 1000-1007.
- [9] Rajpal, S., Lakhyani, N., Singh, A. K., Kohli, R., & Kumar, N. (2021). *Using handpicked features in conjunction with ResNet-50 for improved detection of COVID-19 from chest X-ray images*. *Chaos, Solitons & Fractals*, 145, 110749.
- [10] Du, X., Si, L., Li, P., & Yun, Z. (2023). *A method for detecting the quality of cotton seeds based on an improved ResNet50 model*. *Plos one*, 18(2), e0273057.