Multilevel Characteristic Weighted Fusion Algorithm in Domestic Waste Information Classification

Min Li
School of Computer and Artificial Intelligence, Henan Finance University, Zhengzhou, 451464, China

***Abstracts*—The study of domestic waste image classification holds significant significance for fields like environmental protection and smart city development. To improve the classification efficiency of household waste information, a multi-feature weighted fusion method for household waste image classification is proposed. In this research, deep learning technology was applied to develop a multi-level feature-weighted fusion network model for domestic garbage image classification. The study first analyzed the VGG-16 architecture and created a garbage image dataset for domestic garbage according to the current Shenzhen garbage classification standard. Based on this, a multi-level feature-weighted fusion model for garbage image classification was constructed using VGG-16 as the backbone network. Furthermore, it was combined with the backbone feature extraction network as well as the content-aware and boundary-aware feature extraction networks. The performance of the classification model was tested, and it was found that the highest classification accuracy of the classification model can reach 0.98, and the shortest classification time is only 3s. The multi-level feature-weighted fusion garbage image classification model constructed in this research not only has better classification performance, but also can provide a new processing idea for the urban garbage classification problem.**

Keywords*—*Multi-feature; weighted fusion; image; deep learning; waste classification

I. Introduction

With urbanization accelerating, domestic waste disposal presents a growing challenge for city managers [1-2]. Waste classification (WC) can impact resource recovery, environmental protection, as well as the efficiency and quality of city operations. Successful WC entails fast, precise identification of various types of waste, a considerable challenge with the rising quantities of waste. Currently, the utilization of computer vision technology in WC is emerging as a popular research area [3-4]. Traditional garbage image classification methods rely on manually extracting features and using simple classifiers, which are limited in dealing with images that have complex backgrounds, diverse lighting, and different angles. Due to these limitations, the current garbage classification methods have the problems of low accuracy and low efficiency. Recently, convolutional neural networks (CNNs) in deep learning (DL) algorithms have displayed excellent performance for image classification tasks. However, current models encounter issues such as inadequate feature extraction and suboptimal adaptation to limited datasets [5]. In order to address the issue of low-quality garbage image classification, this study employs the VGG-16 model as the primary backbone network. A multi-feature weighted fusion image classification model of domestic waste is constructed by integrating the content and boundary perception model. Against this background, the research objective is to develop a new multi-level feature-weighted fusion network model (MFWFNM) to improve the accuracy and efficiency of domestic garbage image classification (DGIC), and to overcome the limitations of conventional models in handling complicated garbage images. The significance of this study lies in the use of this method to speed up garbage classification, facilitate users to perform garbage classification, create a good living environment, and provide certain ideas and methods for DGIC. The study comprises of six sections, first being the Introduction: Next section presents an analysis and summary of related research from other sources. The third section details the construction of the network model. The model's performance testing is covered in the fourth section, fifth section presents the discussion and the study's overall findings are presented in the last section.

II. Related Work

With the rapid pace of urbanization, the disposal of domestic waste poses a paramount challenge for urban management. Efficient waste management not only facilitates resource recovery and environmental protection, but also increases the productivity of waste disposal. In this regard, automatic classification of garbage images by computer vision technology is an important technical solution to improve the efficiency of waste management. Aiming at the shortcomings of traditional classification models, Ma et al. proposed a WC model that improves the structure of ResNet-50 network. The model was improved in two stages. First, by adding an attention module to the residual block and modifying the downsampling procedure, the input feature screening and information loss reduction were adjusted. Second, the model incorporated horizontal and vertical multi-scale feature fusion techniques to utilize the features more efficiently. The model's classification performance on tiny datasets was significantly enhanced by this restructure. The outcomes revealed that the model outperformed the original ResNet-50 model by 7.62% on the TrashNet dataset, demonstrating higher accuracy and robustness [6]. Wu et al. created a new publicly available benchmark dataset, in addition to the study a classification test was performed on this dataset using deep CNN. The results of the study showed that the created home spam image dataset was able to simulate different lighting, backgrounds, angles, and shapes, and the deep CNN was able to obtain a high classification accuracy (CA) in the data [7]. The rise of email as a means of official and personal communication has made it increasingly difficult to accurately identify and classify spam. Scholar Vivekanandam developed a novel hybrid machine learning approach to effectively detect spam. This research was significant in safeguarding data security against unauthorized access, while also offering fresh insights and techniques for spam detection [8]. In an effort to address the issues of overfitting, inadequate convergence, and decreased recall and accuracy in conventional image recognition algorithms, Li Y et al. presented a DL-based spam image identification system. The algorithm employed the ReLU activation function to handle the gradient dissipation problem in neural network training, the Adagrad adaptive method to modify the parameters of the deep neural network, and the Dropout strategy to prevent overfitting [9].

In recent years, the use of lightweight networks is pervasive in image segmentation and face recognition due to their advantageous characteristics, including minimal resource consumption and rapid reasoning speed. The utilization of lightweight networks for the purposes of classification and image recognition has emerged as a significant area of research. For use in the facial expression picture identification task, Zou et al. introduced the multi-feature fusion CNN, a lightweight network architecture. According to experimental data, the suggested model performed better on average in terms of recognition accuracy than other compared algorithms [10]. In the face of COVID-19 epidemic, the problem of high dropout rate in catechism classes was becoming more and more prominent. Because of this, Yujiao et al. presented a dropout prediction model that combined support vector machines with feature fusion from behavioral data. The final fused features were created by applying varying weights to various behavioral features in a model that was based on Pearson's principle. The suggested model had a higher accuracy of dropout prediction, according to experimental results [11]. Zhang et al. aimed to develop an intelligent detection model for strabismus based on corneal light reflection photographs. To enhance strabismus detection performance, a multi-feature fusion model is suggested that integrates both ratio and depth features. According to the experimental data, the model successfully increased the accuracy and reliability of strabismus diagnosis, achieving an accuracy of 97.17% in strabismus identification, which was much superior to the single-feature model [12]. A unique multi-feature fusion improved transformer for the creation of image descriptions was proposed by Zhang et al. According to test results on the MSCOCO dataset, the suggested model performed better in picture description than other cutting-edge techniques [13].

In summary, numerous scholars have conducted research on junk image classification and multi-feature fusion algorithms, achieving significant results. The extant research on garbage image classification indicates that, while the results are effective, the CA and efficiency remain at a relatively low level. It is imperative to identify an effective method to enhance the classification of garbage images. Therefore, further exploration of this area appears promising. To enhance CNN's performance in solving the DGIC problem, this study utilizes a multi-feature weighted fusion DGIC model with VGG-16 as the backbone network. This model is combined with content and boundary-aware models to improve the WC model's CA and provide fresh solutions to the real-world WC problem.

III. Domestic Garbage Image Classification Method Based on Multi-Feature Weighted Fusion

To enhance the CA of domestic garbage images (DGI), this research firstly introduces the VGG-16CNN model in DL algorithm and the DGI dataset used in this research. Based on this, the boundary features and contents of the garbage images are fused, and a multilevel feature-weighted fusion algorithm model for DGIC is further proposed.

A. VGG-16 Modeling and Design of Domestic Garbage Images Dataset

The VGG-16 model of CNN is adopted in this study for domestic waste treatment. CNN belongs to the part of artificial intelligence research field, because it has a certain degree of autonomous learning, and can simulate the processing mechanism of the human brain visual nerve, so it is often used in the analysis of visual images [14]. The input layer (IL), convolutional layer (CL), pooling layer (PL), fully connected layer (FCL), and output layer make up the majority of a typical CNN. Fig. 1 depicts its overall structure.

The primary function of the CL in Fig. 1 is to extract features using the convolutional kernel. The retrieved data comes from the input of the IL, and the image features that each layer of the CL can extract are different. CNN backpropagation optimizes the weight coefficients and bias of the convolution kernel, and the size of the convolution kernel controls the complexity of the features that can be recovered from the convolution layer. The primary task of the PL is to select a large number of features to simplify the complexity of the model operation. The PL often comes after the CL. After feature extraction in the CL, a large number of image features will be sent to the PL. PL operation is generally divided into maximum pooling and average pooling, both are a large amount of feature information is divided into a number of small regions. The difference is that the former outputs the maximum value in each small region, while the latter outputs the average value of all feature information in each small region. In addition to the input and output layers, the FCL also plays an important role in CNN. The FCL, which is at the end of the CNN, is responsible for nonlinearly combining all the feature information before feeding it to the output layer. The CL and PL, on the other hand, extract and select image features, respectively.

The VGG-16 model is a popular DL network architecture named after the institution where its development team is located, i.e., the visual geometry group (VGG) at the University of Oxford. VGG-16 refers to the fact that this network structure contains 16 network layers with weights, including convolutional and FCLs. VGG-16 is a unique CNN that is often utilized in domains like image recognition and classification because of its outstanding performance and straightforward architecture. Fig. 2 depicts the VGG-16 network model's structure.

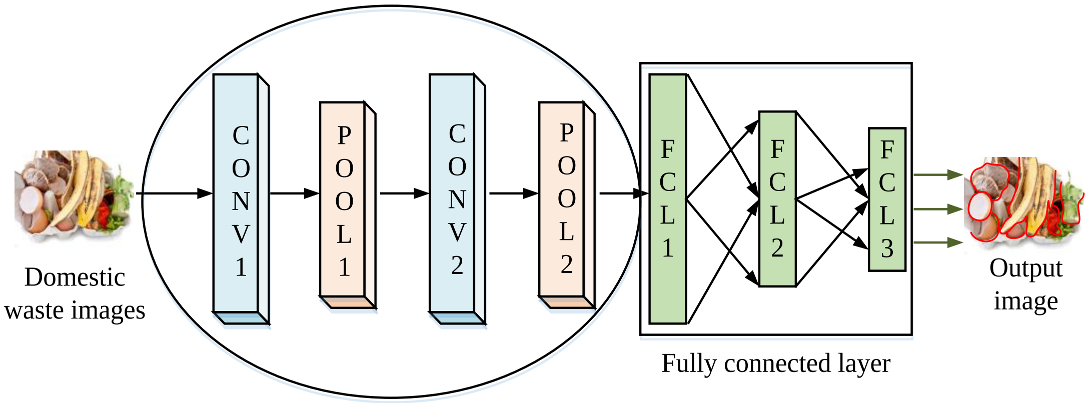


Fig. 1. Basic structure of CNN model.

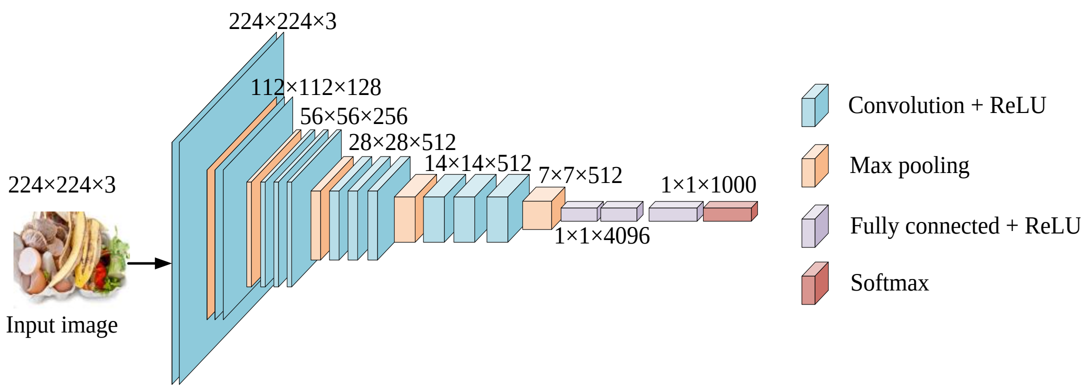


Fig. 2. VGG-16 network model structure diagram.

An FCL, a Softmax layer, an activation function, an IL, a CL, and a PL make up the entire VGG-16 network model shown in Fig. 2. The IL is a fixed-size, 224*224 RGB picture. The whole model has a total of 13 CLs, each CL has a convolutional kernel size of 3*3 with a step size of 1. Using a small convolutional kernel enables the VGG-16 network to have more weight layers, which improves the depth and complexity of the network. Each CL is followed by a ReLU activation function to increase the nonlinearity of the network. The network has a total of five PLs with a filter size of 2x2 and a step size of 2. Behind the convolutional and PLs, there are three FCLs that are connected to a Softmax layer. The Softmax layer outputs the probability values for image classification. Since the VGG-16 network model has a simple operational framework, powerful feature extraction capability and good generalization ability, this study uses the VGG-16 network as the backbone network to build a garbage image classification model.

Due to the wide variety of household garbage categories and different classification standards, the data used in this study is the DGI dataset that Huawei has previously made public in the WC Challenge Cup. This dataset contains more categories of garbage and a richer database than foreign garbage image datasets such as TrashNet and GINI, and contains a total of more than 10,000 DGIs, and the garbage in this dataset is categorized according to the WC standards of Shenzhen City. Additionally, Table I displays its classification standards.

TABLE I. CLASSIFICATION STANDARDS FOR THE DOMESTIC WASTE DATA SET

Garbage classification standards	Contains content
Hazardous waste	Batteries, expired medicines, ointments, etc.
Recyclable trash	Beverage bottles, glasses, cans, pillows, plush items, cardboard boxes, etc.
Kitchen waste	Fruit peels, tea leaves, bones, eggshells, vegetable leaves, leftovers, etc.
Other garbage	Disposable tableware, cigarette butts, toothpicks, plastic products, etc.

As shown in Table I, the waste image dataset used in this study is classified into four major categories: hazardous waste, recyclable waste, food waste, and other waste according to the WC standard of Shenzhen. Due to the large number and variety of waste images in the dataset, it is easy for the existence of similar appearance of different categories of household waste to lead to classification errors. To reduce the experimental errors, 80% of the images in the dataset are randomly and mutually exclusive divided into a training set before the experiment, leaving only 20% of the images as a test set, and the distribution of the dataset does not change during the entire experiment.

B. *Design of Multi-Level Feature Weighted Fusion Network Model for Domestic Garbage Image Classification*

Classifying DGIs quickly and accurately is a crucial and difficult task in the field of WC and processing nowadays.

Traditional image classification methods often face the problems of poor accuracy and lack of robustness when dealing with spam images with complex backgrounds, diverse lighting conditions and variable shapes. The utilization of multi-feature fusion methods to enhance image classification performance has become a new research trend with the development of DL and computer vision technologies. However, the effective integration of different levels of image features, such as color, shape, and texture, into beneficial information for classification remains a problem to be solved [15]. In this research, the VGG-16 network is used as the backbone network, and four different feature extraction networks are fused to design the DGIC-oriented MFWFNM. Fig. 3 displays the overall categorization model's structure.

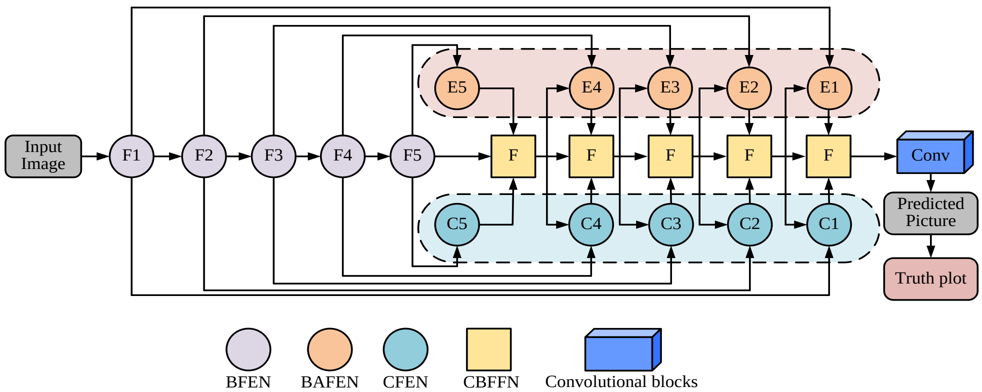


Fig. 3. Structure diagram of multi-level feature weighted fusion network model.

The multilevel feature-weighted fusion network model in Fig. 3 incorporates a total of four feature extraction networks, namely backbone feature extraction network (BFEN), boundary-aware feature extraction network (BAFEN), content-aware feature extraction network (CFEN), and content boundary feature fusion network (CBFFN). The garbage images from the domestic waste dataset are selected and input into the classification model, firstly, the multi-level features are extracted by using BFEN in VGG-16 network, then the effective content is extracted from the main features by using CFEN, and the edge information is extracted from the main features by using BAFEN. Then CBFFN is used at each stage to screen and fuse the features. Lastly, the model's ability to categorize the garbage photos is tested using a bespoke loss function. Training is stopped whenever the model is able to do so. Fig. 4 depicts the precise CFEN structure.

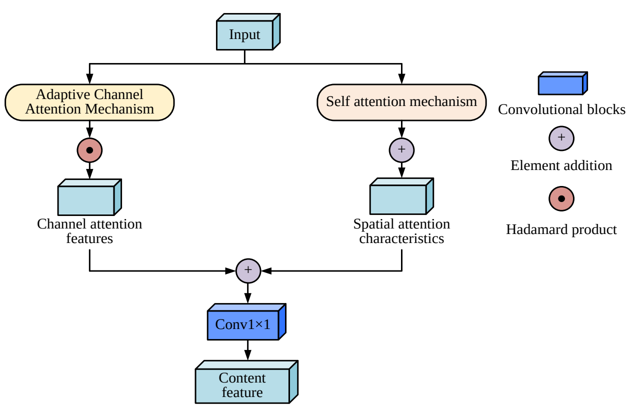


Fig. 4. CFEN structure diagram.

In Fig. 4, the CFEN is mainly composed of the channel attention feature extraction module and the spatial attention feature extraction module. For layer i, the fusion feature S_{i+1} from the previous layer of CBFFN is used as the input of CFEN, and the channel attention mechanism and spatial attention mechanism are utilized to complete the fusion of the features, so as to obtain the new features. Assuming that the feature of the output channel is F_C, the calculation equation of the channel attention feature extraction module is obtained as shown in Eq. (1).

$$F_C = S_{i+1} \oplus CA(S_{i+1}) \tag{1}$$

In Eq. (1), CA denotes the adaptive channel attention mechanism. $\oplus$ denotes the inter-pixel addition operation. Eq. (2) displays the computational equation for the feature extraction module of spatial attention.

$$F_S = S_{i+1} \oplus SA(S_{i+1}) \tag{2}$$

In Eq. (2), F_S denotes the output spatial features. SA denotes the adaptive spatial attention mechanism. The extracted channel dimension and spatial dimension features in Eq. (1) and Eq. (2) are fused to obtain the equation of the fused features as shown in Eq. (3).

$$C_i = conv(F_C \oplus F_S) \tag{3}$$

In Eq. (3), $conv(\times)$ denotes the convolution operation and the size of the convolution kernel is 1*1. C_i denotes the content fusion feature.

In addition to using the CFEN model to fuse content features, it is also necessary to extract boundary features from garbage images. This study employs the BAFEN network to supplement edge information in order to extract more meaningful edge characteristics, therefore increasing the classification model's accuracy. Fig. 5 displays the particular structure of BAFEN.

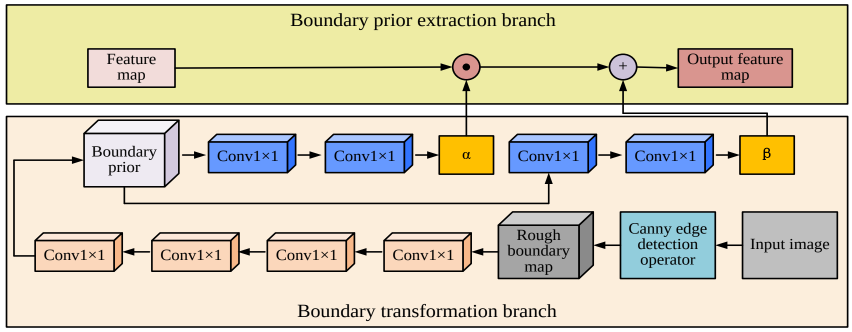


Fig. 5. BAFEN structure diagram.

In Fig. 5, the BAFEN structure mainly consists of two parts: the boundary prior extraction branch and the boundary transformation branch. Extracting features from boundary prior knowledge to serve as guidance for further processing is the primary task of the boundary prior extraction branch. The function of the boundary transformation branch is to transform the guidance information obtained from the boundary prior extraction branch into specific boundary features. First, Canny edge detection is performed on the input image, and then the detection result is used as a rough boundary map and a prior is applied to the boundary map. By extracting prior boundary knowledge and utilizing boundary transformation branches to generate the final boundary features. The mathematical calculation model of BAFEN is shown in Eq. (4).

$$BAFEN(F_i) = F_i \odot \alpha + \beta \tag{4}$$

In Eq. (4), F_i represents the output characteristics of the backbone network. α represents the boundary prior factor. β represents the boundary conversion factor. $\odot$ represents Hadamard product. The range of values for F_i is shown in Eq. (5).

$$F = F_i \quad i \in \{1,2,3,4,5\} \tag{5}$$

In Eq. (5), F represents the set of output features. $i \in \{1,2,3,4,5\}$ represents the five CLs in BFEN. The calculation equation for α is shown in Eq. (6).

$$\alpha = \mathrm{Re}LU\Big(Conv\Big(\mathrm{Re}LU\big(Conv\big(E_{prior}\big)\big)\Big)\Big) \tag{6}$$

In Eq. (6), $\mathrm{Re}LU$ represents the activation function. E_{prior} represents the initial boundary map features. The calculation equation for β is shown in Eq. (7).

$$\beta = \mathrm{Re}LU\Big(Conv\Big(\mathrm{Re}LU\big(E'_{prior}\big)\Big)\Big) \tag{7}$$

In Eq. (7), E'_{prior} represents the boundary map features after boundary transformation. By combining Eq. (6) and (7), the final boundary perception fusion feature can be obtained as shown in Eq. (8).

$$E_i = T_{conv}\big(Canny(F_i)\big) \tag{8}$$

In Eq. (8), T_{conv} represents continuous convolution operation. E_i represents boundary perception fusion features.

After obtaining fusion features and boundary perception fusion features, the study adopted a learning based approach to fuse the two features, thereby generating CBFFN. Firstly, pixel to pixel addition is used to preliminarily integrate information between C_i and E_i. Then, three CLs with 3 * 3 kernels and a step size of 1 are used for feature screening and extraction. Finally, S_{i+1} is overlaid on the output features of the last layer through residual learning to obtain a fused feature map. The calculation equation is shown in Eq. (9).

$$S_i = S_{i+1} \oplus T_{conv}(C_i \oplus E_i) \tag{9}$$

In Eq. (9), S_i represents the final fused feature map. To boost confidence, a loss function is added during the model training procedure. Eq. (10) displays the CFEN's content loss function.

$$L_C(P,T) = L_{SSIM}(P,T) + L_{IOU}(P,T) \tag{10}$$

In Eq. (10), L_C represents the content loss function of CFEN. L_{IOU} and L_{SSIM} represent the IOU loss function and SIIM loss function, respectively. P represents the classification result predicted by the model. T represents the true classification result of the model. Eq. (11) displays the BAFEN boundary loss function.

$$L_E(P,T) = L_{CBCE}(P,T_E) \tag{11}$$

In Eq. (11), L_E represents the boundary loss function of BAFEN. L_{CBCE} represents the CBCE loss function. T_E represents the boundary map of salient targets in real classification results. The fusion loss function of CBFFN is shown in Eq. (12).

$$L_F(P,T) = \lambda_1\big(L_{IOU}(P,T) + L_{SSIM}(P,T) + \lambda_2 L_{CBCE}(P,T)\big) \tag{12}$$

In Eq. (12), l_2 and l_1 both represent weight values. L_F represents the fusion loss function.

IV. Performance Testing and Application Analysis of Multi-level Feature Weighted Fusion Network Model for Domestic Garbage Image Classification

To verify the performance and application effectiveness of MFWFNM, this study compared three different image classification models: residual network (ResNet), faster region-based convolutional neural network (Faster R-CNN), and deformable convolutional networks (DCN). Using CA, classification error, and model iteration fitness values as performance evaluation indicators, it has been proven that the classification performance of MFWFNM is superior to other comparative models.

A. *Performance Testing of Multilevel Feature Weighted Fusion Network Model*

To demonstrate the benchmark performance of the model, this study chose a publicly available image dataset for simulation experiments. In order to avoid equipment errors during the experimental process, this study conducted comparative experiments in the same experimental environment. Table II displays the experimental setup and dataset details.

Table II provides the experimental environment and dataset information for this simulation experiment. Table II lists the three publicly accessible datasets that are chosen for this study's model performance testing: ImageNet, COCO, and TrashNet. Firstly, the iterative fitness changes of ResNet, Faster R-CNN, DCN, and MFWFNM are obtained as shown in Fig. 6.

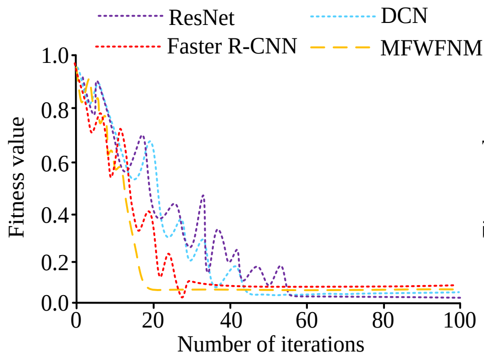
In Fig. 6, the iterative fitness values of four models, ResNet, DCN, Faster R-CNN, and MFWFNM, are shown in the ImageNet, COCO, and TrashNet datasets. As shown in Fig. 6(a), ResNet, DCN, Faster R-CNN, and MFWFNM can reach a stable state by iterating 56, 44, 28, and 17 times respectively in the dataset ImageNet. As shown in Fig. 6(b), ResNet, DCN, Faster R-CNN, and MFWFNM can reach a stable state after 61, 48, 41, and 18 iterations in the dataset COCO, respectively. In Fig. 6(c), ResNet, DCN, Faster R-CNN, and MFWFNM can reach a stable state in the dataset TrashNet after 49, 25, 27, and 13 iterations, respectively. In summary, MFWFNM can quickly iterate to a stable state in all three datasets, indicating that the processing efficiency of the model is higher.

TABLE II. EXPERIMENTAL ENVIRONMENT AND DATA SET INFORMATION TABLE

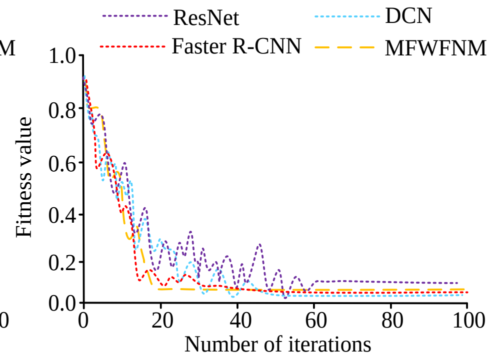
Project	Composition	Configuration
Lab environment	Processor	Intel Core i9
	Graphics processor	NVIDIA RTX 3080
	Memory	32GB RAM
	Storage	1TBSSD
	Operating system	Ubuntu 20.04
	Deep learning framework	TensorFlow
	Coding software	Python 3.8
Dataset information	ImageNet	1000+ image types
	COCO	More than 80 complex scene images
	TrashNet	6 image data sets for garbage classification

Fig. 7(a), 7(b), and 7(c) show the CA values of the four models in the ImageNet, COCO, and TrashNet datasets, respectively. Fig. 7 shows that while the CA of the MFWFNM model stays above 0.9, the CA values of ResNet, DCN, and Faster R-CNN all change to some amount as the number of samples grows. In Fig. 7(a), the highest CA of ResNet, DCN, Faster R-CNN, and MFWFNM models in the ImageNet dataset is 0.77, 0.84, 0.89, and 0.96, respectively. As shown in Fig. 7(b), the highest CA of ResNet, DCN, Faster R-CNN, and MFWFNM models in the COCO dataset is 0.78, 0.82, 0.91, and 0.95, respectively. As shown in Fig. 7(c), the highest CA of ResNet, DCN, Faster R-CNN, and MFWFNM models in the TrashNet dataset is 0.77, 0.79, 0.90, and 0.98, respectively.

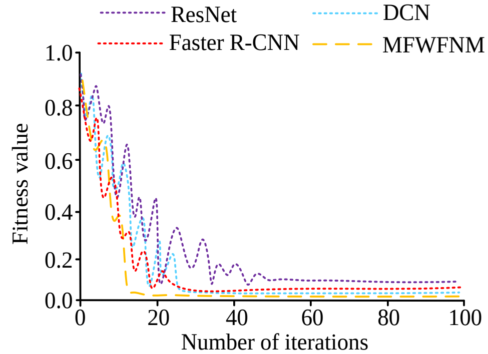
Fig. 8(a) and 8(b) show the mean square error (MSE) and mean absolute error (MAE) of the four models in the TrashNet dataset, respectively. In Fig. 8(a), the highest MSE values for the ResNet, DCN, Faster R-CNN, and MFWFNM models are 4.28, 3.11, 1.95, and 0.72, respectively. In Fig. 8(b), the highest MAE values for the four models ResNet, DCN, Faster R-CNN, and MFWFNM are 4.73, 3.81, 2.69, and 1.15, respectively. Overall, the MFWFNM model has better error performance when processing multi sample image data, and therefore can more accurately complete image classification tasks.



(a) Changes in fitness values of the four models under the ImageNet data set

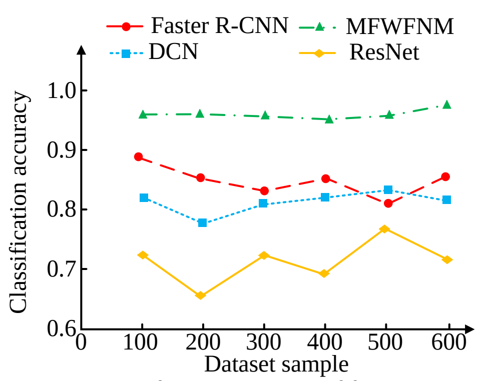


(b) Changes in fitness values of the four models under the COCO data set

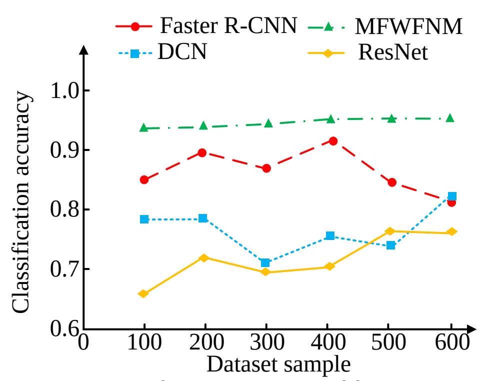


(c) Changes in fitness values of the four models under the TrashNet data set

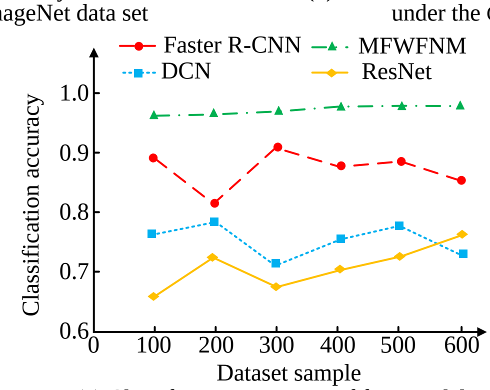
Fig. 6. Iterative fitness values of four models in three data sets.



(a) Classification accuracy of four models under the ImageNet data set

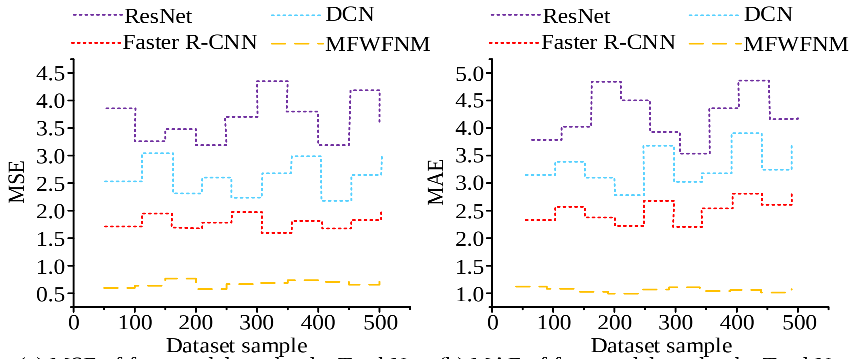


(b) Classification accuracy of four models under the COCO data set



(c) Classification accuracy of four models under the TrashNet data set

Fig. 7. Classification accuracy values of four models in three data sets.



(a) MSE of four models under the TrashNet data set (b) MAE of four models under the TrashNet data set

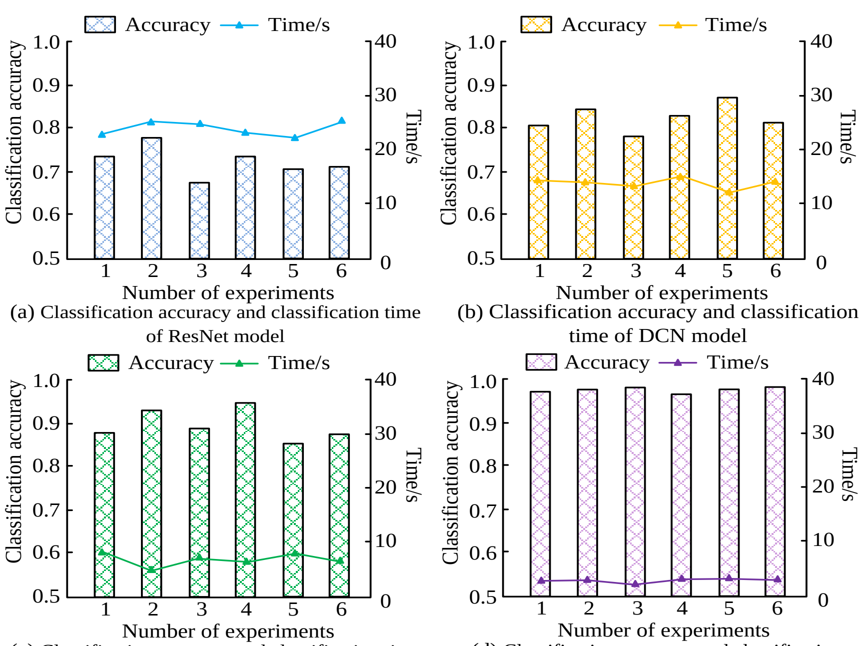
Fig. 8. Classification errors of the four models in the TrashNet data set.

B. Application Effect of a Multilevel Feature Weighted Fusion Network Model in Actual Waste Classification

Using the daily WC requirements listed in Table I, this study collects over 10,000 common DGIs to further validate the good performance of the MFWFNM model in real-world WC situations. These images are then used to test the actual classification performance of different classification models. 100 images of hazardous waste, recyclable waste, kitchen waste, and other waste are selected from over 10000 datasets, and the CA and time of the four models under multiple classification tests are shown in Fig. 9.

In Fig. 9, the CA and time consumption of four models, ResNet, DCN, Faster R-CNN, and MFWFNM, in actual garbage image classification are presented. In Fig. 9(a), in the six classification tests, the ResNet classification model achieved the highest CA of 0.77 and the shortest time consumption is 22 seconds. In Fig. 9(b), in the six classification tests, the DCN classification model achieves the highest CA of 0.86 and the shortest time consumption is 13 seconds. In Fig. 9(c), in the six classification tests, the Faster R-CNN classification model achieved the highest CA of 0.95 and the shortest time consumption is six seconds. In Fig. 9(d), the CA of the MFWFNM classification model can reach up to 0.98, and the shortest time required is only three seconds. Compared to the other three WC models, the MFWFNM model can achieve better classification performance in multiple classification tests.

In Fig. 10, the evaluation scores of experts and residents for the four classification models are presented. Assuming a total of five points, the higher the score, the higher the satisfaction. Based on Fig. 10(a), (b), (c), (d), experts and users have the highest satisfaction with the classification model MFWFNM, with scores concentrated in the first quadrant. On the contrary, the satisfaction of experts and users with the other three models is relatively scattered, but overall, the Faster R-CNN model has higher satisfaction scores than the other two models.



(a) Classification accuracy and classification time of ResNet model
(b) Classification accuracy and classification time of DCN model
(c) Classification accuracy and classification time of Faster R-CNN model
(d) Classification accuracy and classification time of MFWFNM model

Fig. 9. Actual classification accuracy and classification time of the four models in the garbage image classification problem.

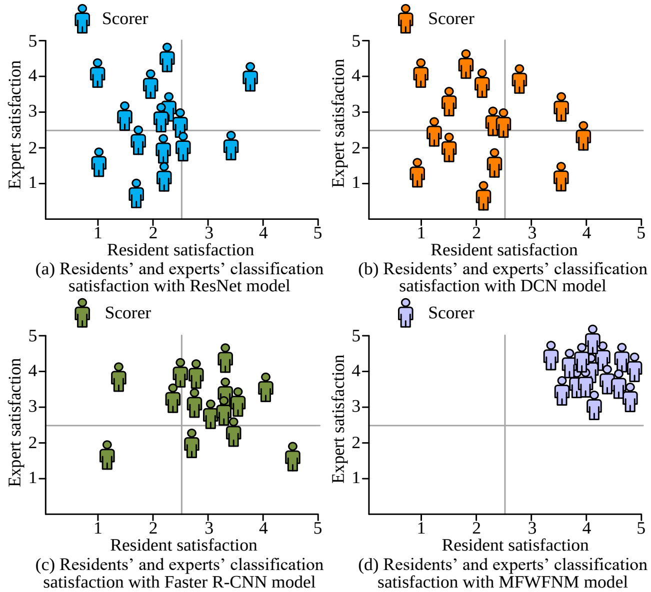


Fig. 10. Experts and residents' ratings of four classification models.

V. Discussion

In this study, the performance of MFWFNM model was analyzed, and the application effect of MFWFNM model in garbage classification was analyzed. The experimental results showed that MFWFNM model had significant advantages in CA, MSE and MAE. In the comparison experiment of MSE values, the highest MSE values of ResNet, DCN, Faster R-CNN and MFWFNM were 4.28, 3.11, 1.95 and 0.72, respectively. Moreover, the MFWFNM model proposed in this study had lower MSE values than the comparison model. The results showed that the introduction of VGG-16 network and the combination of four backbone networks improved the computational efficiency and accuracy of the model. Wang X et al. reached a similar conclusion in the relevant research of multi-feature fusion classification model [16]. Secondly, in the comparative analysis of CA and classification time, the highest CA and shortest classification time of MFWFNM model were 0.98 and 3s, respectively, which were lower than the comparison model. This result further indicated that, the fusion of BFEN, CFEN, BAFEN and CBFFN networks and the introduction of VGG-16 network improve the computational efficiency and image CA of MFWFNM model, thus improving the accuracy of garbage classification model. This conclusion coincided with the relevant conclusion proposed by Zheng Y et al in 2023 [17]. In terms of satisfaction evaluation, experts and users expressed the highest levels of satisfaction with the MFWFNM model proposed in the study, and its satisfaction ratings were significantly higher than those of the comparison model. This result was the same as the conclusion of the image classification model proposed by Xu X team [18]. The results showed that the proposed MFWFNM model had good practical value and was helpful to improve the efficiency and quality of garbage classification. In summary, compared with relevant studies, this study proposed the VGG-16 model as the main trunk network, combined with content and boundary perception model to build a multi-feature weighted fusion DGIC model. It has higher CA and precision in processing garbage images with complex backgrounds, diverse lighting conditions and changing shapes. This model not only provides convenience for accurately determining the types of domestic garbage, but also provides a certain idea and method for the theory of garbage image classification.

VI. Conclusion

Human daily life inevitably generates waste, and effective waste sorting and recycling are currently major research topics in the field. The present study utilized the VGG-16 network as the backbone and combined it with BFEN, CFEN, BAFEN, and CBFFN to create MFWFNM for DGIC. The findings revealed that MFWFNM outperforms ResNet, DCN, and Faster R-CNN in achieving a stable state, requiring only 17, 18, and 13 iterations in ImageNet, COCO, and TrashNet datasets, respectively, compared to the other models. Notably, MFWFNM's CA in all three datasets reached up to 0.96, 0.95, and 0.98, respectively, surpassing the accuracy of the compared model. Additionally, the MFWFNM model exhibited smaller classification errors with a highest MSE value of 0.72 and a highest MAE value of 1.15, as compared to ResNet, DCN, and Faster R-CNN models with the highest MSE values of 4.28, 3.11, and 1.95, and highest MAE values of 4.73, 3.81, and 2.69, respectively. Four classification models were utilized to categorize the DGI dataset outlined in the article. The MFWFNM classification model exhibited the highest CA of 0.98, with the shortest classification time of merely three seconds, and outperformed the other three models across all metrics. In summary, the study's WC model exhibits excellent classification performance. However, it is yet to be determined whether the model can generalize to other image types. Future research should investigate the classification performance of the model in diverse settings. This study has two shortcomings. First, the number of garbage types used in the classification test is limited. In contrast, the actual environment features an increasing number of complex garbage categories. Second, although a garbage classification model has been proposed, it lacks a systematic design and implementation. On the one hand, the exploration direction of future research is to collect more image data of household garbage to increase the data set of household garbage images, and to include existing garbage types as much as possible, so as to provide data basis for training classification models. Conversely, a convenient and optimal operating system is designed in accordance with the model and user requirements, thereby facilitating convenient and expedient services for pertinent users. Moreover, a multi-feature fusion algorithm and image classification method are employed to segment and recognize medical images, thereby enabling the effective extraction of the diseased area of patients and the recognition of images. It is therefore the objective of future research to enhance the efficiency with which doctors utilize medical images for disease analysis, thereby improving the efficiency of subsequent treatment.

References

[1] Sharma T, Nair R, Gomathi S. Breast cancer image classification using transfer learning and convolutional neural network. International Journal of Modern Research, 2022, 2(1): 8-16.

[2] Hosseinalipour A, Ghanbarzadeh R. A novel approach for spam detection using horse herd optimization algorithm. Neural Computing and Applications, 2022, 34(15): 13091-13105.

[3] Guo Y, Mustafaoglu Z, Koundal D. Spam detection using bidirectional transformers and machine learning classifier algorithms. Journal of Computational and Cognitive Engineering, 2023, 2(1): 5-9.

[4] Ahraminezhad A, Mojarad M, Arfaeinia H. An intelligent ensemble classification method for spam diagnosis in social networks. International Journal of Intelligent Systems and Applications, 2022, 14(1): 24-31.

[5] Yi N, Luo W. Research on Water Garbage Detection Algorithm Based on GFL Network. Frontiers in Computing and Intelligent Systems, 2023, 3(1): 154-157.

[6] Ma X, Li Z, Zhang L. An improved ResNet-50 for garbage image classification. Tehnički vjesnik, 2022, 29(5): 1552-1559.

[7] Wu Z, Li H, Wang X, Wu Z, Zou L, Xu L, Tan M. New benchmark for household garbage image recognition. Tsinghua Science and Technology, 2022, 27(5): 793-803.

[8] Vivekanandam B. Spam Email Classification by Hybrid Feature Selection with Advanced Machine learning Algorithm-Future Perspective. Journal of Soft Computing Paradigm, 2022, 4(2): 58-68.

[9] Li Y, Liu W. Deep learning-based garbage image recognition algorithm. Applied Nanoscience, 2023, 13(2): 1415-1424.

[10] Zou W, Zhang D, Lee D J. A new multi-feature fusion based convolutional neural network for facial expression recognition. Applied Intelligence, 2022, 52(3): 2918-2929.

[11] Yujiao Z, Ang L W, Shaomin S, Palaniappan S. Dropout Prediction Model for College Students in MOOCs Based on Weighted Multi-feature and SVM. Journal of Informatics and Web Engineering, 2023, 2(2): 29-42.

[12] Zhang G, Xu W, Gong H, Sun L, Li C, Chen H, Xiang D. Multi-feature fusion-based strabismus detection for children. IET Image Processing, 2023, 17(5): 1590-1602.

[13] Zhang J, Fang Z, Wang Z. Multi-feature fusion enhanced transformer with multi-layer fused decoding for image captioning. Applied Intelligence, 2023, 53(11): 13398-13414.

[14] Huang L, Li M, Xu T, Dong S. A waste classification method based on a capsule network. Environmental Science and Pollution Research, 2023, 30(36): 86454-86462.

[15] Amin S N, Shivakumara P, Jun T X, Chong K Y, Zan D L L, Rahavendra R. An Augmented Reality-Based Approach for Designing Interactive Food Menu of Restaurant Using Android, Artificial Intelligence and Applications. 2023, 1(1): 26-34.

[16] Wang X, Duan L, Ning C. Global context-based multilevel feature fusion networks for multilabel remote sensing image scene classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021, 14(2): 11179-11196.

[17] Zheng Y, Li C, Zhou X, Chen, H., Xu, H., Li, Y. Application of transfer learning and ensemble learning in image-level classification for breast histopathology. Intelligent Medicine, 2023, 3(2): 115-128.

[18] Xu X, Huang L, Wu R, Zhang, W., Ding, G., Liu, L. Multi-feature fusion method for identifying carotid artery vulnerable plaque. IRBM, 2022, 43(4): 272-278.