Waste Management 162 (2023) 123–130

Contents lists available at ScienceDirect

Waste Management

journal homepage: www.elsevier.com/locate/wasman

Garbage detection and classification using a new deep learning-based machine vision system as a tool for sustainable waste recycling

Shoufeng Jin [a], Zixuan Yang [a], Grzegorz Królczykg [b], Xinying Liu [a], Paolo Gardoni [c], Zhixiong Li [d,e,*]

[a] College of Mechanical and Electrical Engineering, Xi'an Polytechnic University, Xi'an 710600, China
[b] Faculty of Mechanical Engineering, Opole University of Technology, Opole 45-758, Poland
[c] Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA
[d] Donghai Laboratory, Zhoushan 316021, Zhejiang, China
[e] Yonsei Frontier Lab, Yonsei University, Seoul 03722, South Korea

ARTICLE INFO

Keywords:
Waste management
Garbage classification
Deep learning
Machine vision

ABSTRACT

Waste recycling is a critical issue for environment pollution management while garbage classification determines the recycling efficiency. In order to reduce labor costs and increase garbage classification capacity, a machine vision system is established based on the deep learning and transfer learning. In this new method, an improved MobileNetV2 deep learning model is proposed for garbage detection and classification, where the attention mechanism is introduced into the first and last convolution layers of the MobileNetV2 model to improve the recognition accuracy and the transfer learning uses a set of pre-trained weight parameters to extend the model generalization ability. In addition, the principal component analysis (PCA) is employed to reduce the dimension of the last fully connected layer to enable real-time operation of the developed model on an edge device. The experimental results demonstrate that the proposed method generates 90.7 % of the garbage classification accuracy on the "Huawei Cloud" datasets, the average inference time is 600 ms on the raspberry Pi 4B microprocessor, and the model volume compression is 30.1 % of the basic MobileNetV2 model. Furthermore, a garbage sorting porotype is designed and manufactured to evaluate the performance of the proposed MobileNetV2 model on the real-world garbage identification, which turns out that the average garbage classification accuracy is 89.26 %. Hence, the developed garbage sorting porotype can be used a effective tool for sustainable waste recycling.

1. Introduction

With the increasing improvement of people's living standards, the types of household garbage are becoming increasingly diversified (Qin et al., 2022). Effective waste classification is of great significance to the later treatment and reuse (Majchrowska et al., 2022; Chen et al., 2022). However, the implementation of garbage classification is not optimistic; one main reason is that the residents' awareness of garbage classification is not strong, and it is difficult to accurately classify too many garbage types. This is probably because the garbage classification standards in different regions are not consistent due to different culturs and opinions. Relying on manual garbage classification is not only of high-labor cost and low classification efficiency, but also adverse to human health (Tong et al., 2020). In order to address this issue, the machine vision and machine learning technologies have been introduced into intelligent garbage management (Li et al., 2021; Al Mamun et al. 2014).

Machine learning has made remarkable progress in the field of computer vision. Comprehensive researches on the garbage identification and classification have been carried out based on the popular machine learning algorithms, such as the artificial neural network (ANN) (Cai et al., 2022). For example, Yuan et al. (2021) proposed a lightweight residual network MAPMobileNet-18 to address the garbage classification problem, and the evaluation of the garbage classification accuracy, detection speed and edge equipment was performed. Fu et al. (2021) proposed a garbage classification method based on the MobileNetV3 model to achieve 92.62 % classification accuracy; however, the heavy network structure made the computing speed very slow. Chen

* Corresponding author.
E-mail address: zhixiong.li@yonsei.ac.kr (Z. Li).

https://doi.org/10.1016/j.wasman.2023.02.014
Received 12 September 2022; Received in revised form 27 January 2023; Accepted 13 February 2023
Available online 28 March 2023

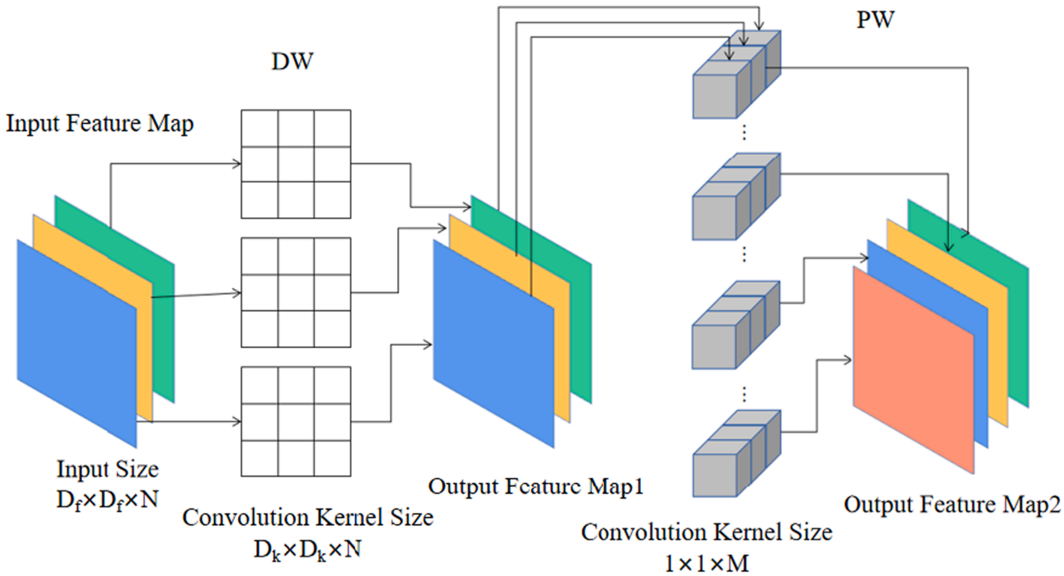


Fig. 1. Schematic diagram of the depth-separable convolution.

et al. (2021) selected the MobileNet-v2 backbone for distillation of a garbage detection model, which reduced the number of the model parameters and improved the detection accuracy, but the generalization ability of the present model was not considered. Feng et al. (2021) used a 23-layer convolutional neural network (CNN) to improve the garbage detection accuracy, but complex layers significantly increased the calculating amount and decreased the model real-time ability. Zhao et al. (2022) proposed an intelligent garbage classification system by combining the MobileNetV3 with the long-short term memory (LSTM), but a practical platform for actual implementation of the proposed classification method. Kang et al. (2020) employed the ResNet-34 backbone for multi-feature fusion of the input garbage images. A new activation function was adopted to reuse the features to improve the ResNet-34 model for small garbage detection; however, real-time garbage classification was a problem due to large calculation of the ResNet-34 network. Gupta et al. (2022) compared the performance of different pre-trained neural networks in garbage classification using auxiliary hardware components (PiCam, raspberry pi, infrared sensors, etc.), but real-time garbage sorting was not achieved. Shi et al. (2020) modified the Xceptionl network layers to solve the negative effects of the back propagation; although good garbage classification accuracy was obtained, the computational cost was very high. In order to resolve the real-time issue, the edge equipment is introduced for the garbage classification systems (Shen et al. 2021; Yang et al. 2021; Liu et al., 2022; Lv et al., 2021). However, most of existing researches focus on the selection of ANNs, but ignore the high computational cost problem, which makes it very difficult to embed the ANN-detection model into the hardware of the edge equipment. As a result, real-time garbage classification remains a challenging task; and to our best knowledge, real-time garbage classification devices/systems are seldom reported in open literature.

In order to address the aforementioned challenging task of real-time garbage classification, this work proposed a new deep learning-based machine vision system. In this new system, the MobileNetV2 is adopted as the backbone network to significantly reduce the model computational cost; the embedding channels and the spatial attention modules are used to enhance the feature extraction ability of the MobileNetV2 model; and the transfer learning is introduced to optimize the model parameters to improve its generalization capacity when dealing with unlabeled/new data. More importantly, in order to realize real-time implementation of the proposed model, the PCA is used to reduce the model parameters. The experimental tests and demonstrate satisfactory real-time garbage detection and classification performance of the proposed system.

The remainder of this work is arranged as follows. Section 2 introduces the proposed improved MobileNetV2 detection model. Section 3 established the real-time implementation system to evaluate the performance of the garbage detection and classification. The main conclusions are drawn in Section 4.

2. Garbage detection and classification method

As a typical lightweight network, the MobileNetV2 (Sandler et al. 2018) is the second generation of the Google MobileNetV1 (Zoph et al. 2018). Compared with the MobileNetV1, the number of the model parameters of the MobileNetV2 is reduced by 20 % and the computation speed is faster, while the feature extraction ability is somehow weaker. As a result, the MobileNetV2-based classification system can fulfil the real-time operating requirement but its feature extraction ability is not strong enough to identify different garbage types. To this end, this study introduces the convolutional attention module into the MobileNetV2 structure (Sandler et al. 2018) to enhance its feature extraction ability. Meanwhile, the transfer learning is used to improve the generalization capacity of the MobileNetV2 model.

2.1. Improved MobileNetV2 network

The proposed model is mainly composed of four parts, as shown in Appendix 1, including one MobileNetV2 backbone, one convolutional attention module (CBAM), one PCA dimensionality reduction module, and one fully connected classification layer. Firstly, the input garbage images are normalized to the requested size of the first convolutional layer Conv2d_1; then, in the CBAM layer the feature map is generated by the spatial attention process. Furthermore, the feature map goes through the second convolutional layer Conv2d_2 and the second CBAM layer to enhance the feature extraction ability. After processed by the global mean pooling and global maximum pooling at the Avgpool layer, the features are further processed using the element-wise summing operation in the first full connection (FC) layer. Lastly, the PCA is used to reduce the dimension of the features to improve the garbage identification efficiency in the second FC layer. The parameters of each module in Appendix 1 are listed in Appendix 2.

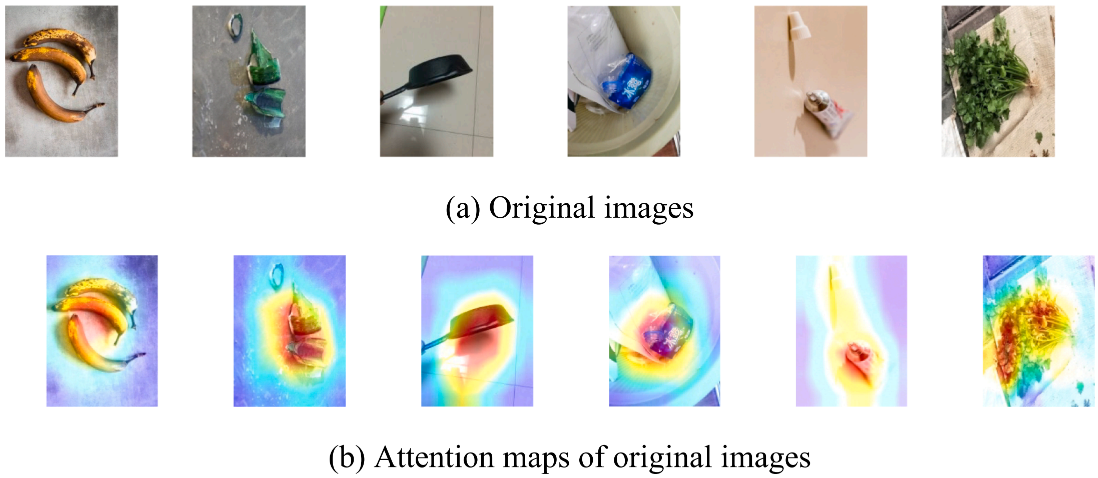


Fig. 2. Visualization of CBAM performance: (a) The origin images randomly selected from dataset; (b) The attention maps extracted from the origin images, in which the feature extraction layer will focus on the red areas.

Table 1
Compares the raw data with PCA for dimension reduction.

Test No.	CNN without PCA		CNN with PCA		Time improvement
	Time (ms)	Precision	Time (ms)	Precision	
1	1,023	84.2 %	655	84.4 %	35.97 %
2	932	84.0 %	632	83.8 %	32.18 %
3	997	82.3 %	589	82.5 %	40.92 %

2.2. *Depth-Separable convolution of MobileNetV2*

The core unit of the MobileNetV2 is the depth-separable convolution, including the deep convolution (DW) and point-by-point convolution (PW) operations. Fig. 1 depicts the schematic diagram of the depth-separable convolution.

The DW uses the filter to convolute each input channel, and then, the PW uses the convolution kernel operation to produce the output feature map.

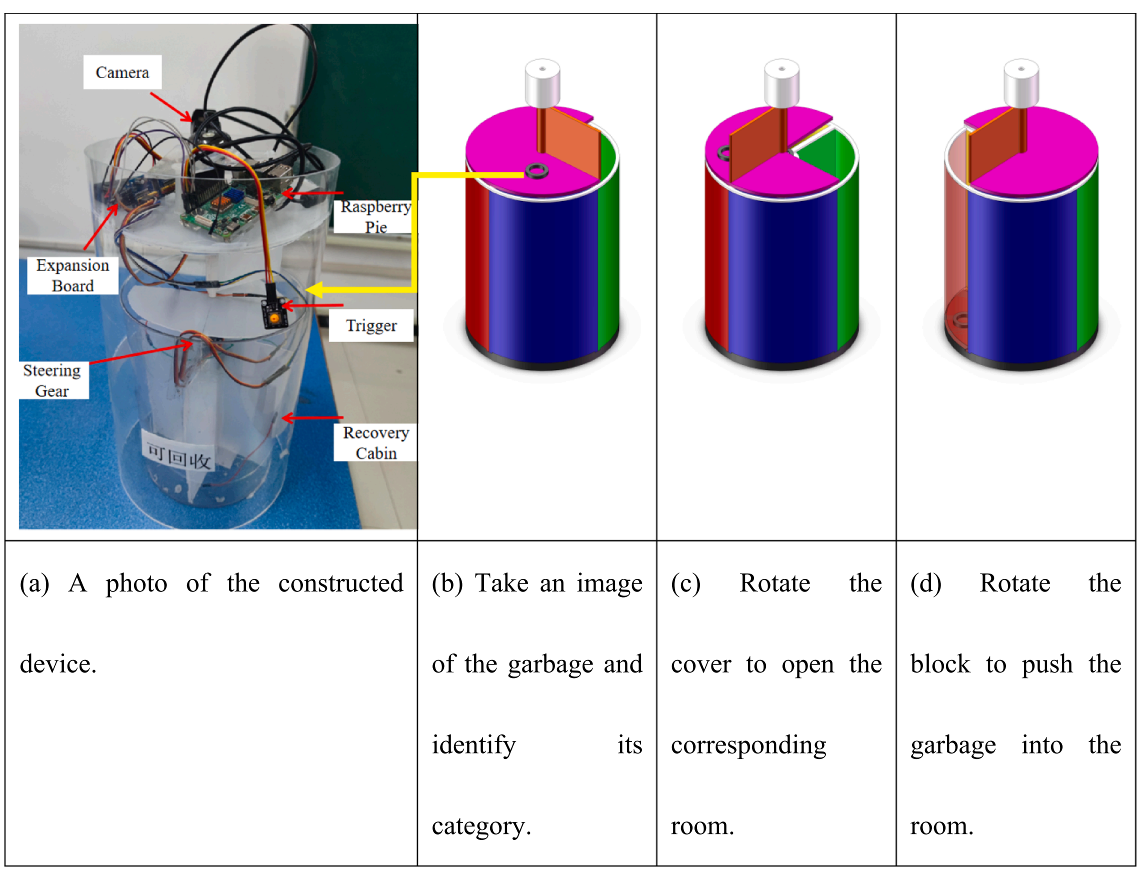


Fig. 3. The proposed garbage identification system.

Table 2
Experimental testing results.

Main network	Improvement factors			Overall precision	Parameters volume
	CBAM	PCA	Transfer learning		
Original MobileNetV2	–	–	–	71.4 %	14,110 KB
	√	–	–	83.8 %	22,414 KB
	–	√	–	81.7 %	9,857 KB
	–	–	√	79.4 %	14,110 KB
Improved MobileNetV2	√	√	√	90.7 %	10,127 KB

$$G_{i,j,m} = \sum_{w,h}^{W,H} K_{w,h,m} \times X_{i+w,j+h,m} \tag{1}$$

where G is the output feature map; K is the convolution kernel with the wide W and high H; X is the input feature map; m represents the m-th channel of the feature map, i and j represent the i-th and j-th coordinates of the output feature map on the m-th channel; w and h are the weight element coordinates of the convolutional kernel in the m-th channel.

Assuming that the input garbage image size is $D_f \times D_f \times M$ and the output garbage image size is $D_k \times D_k \times N$, according to the standard convolution, the calculation amount is $D_k \times D_k \times M \times N \times D_f \times D_f$, and DW and PW calculation amount are respectively $D_k \times D_k \times M \times D_f \times D_f$ and $M \times N \times D_f \times D_f$. Therefore, the calculation amount ratio between the depth-separable convolution and the traditional convolution can be expressed as

$$\frac{P_1}{P_2} = \frac{D_k \times D_k \times M \times D_f \times D_f + N \times M \times D_f \times D_f}{D_k \times D_k \times M \times D_f \times D_f \times N} = \frac{1}{N} + \frac{1}{D_k^2} \tag{2}$$

where P_1 is the calculation amount of the depth-separable convolution and P_2 is for the traditional convolution; D_f is length and width value of the input garbage image and D_k is for the output garbage image; M is the number of input garbage image channels and N is the output garbage image channels. The convolution operation usually uses 33 convolution kernels. Based on the model parameters in Appendix 2, the calculation amount ratio P_1/P_2 is 1/9, which indicates that the depth-separable convolution can effectively reduce the computing amount to improve the processing speed of the MobileNetV2 model. As a result, the computing burden of the hardware for real-time implementation can be significantly reduce.

2.3. Convolutional attention module (CBAM)

The attention mechanism (Park et al. 2020) can allocate computational resources to retain the garbage information and reduce the model interference. The CBAM searches the attention weights of the given intermediate features along the channel and space directions to adaptively adjust the features. Appendix 3 shows the channel attention mechanism (Park et al. 2020), where the input features undergo the maximum pooling and average pooling to generate two spatial features. Then, the FC layer weights the spatial features, and the sigmoid function is used to produce the channel attention weights.

The spatial attention mechanism is shown in Appendix 4, where the average pooling and maximum pooling operations are simultaneously performed along the channel direction, and then a convolution layer is used to calculate the feature vector. The sigmoid function is used to obtain the spatial attention weights.

In this study, the CBAM is embedded into the first convolutional layer and the last convolutional layer, respectively. The former can help the subsequent feature extraction layer to extract more important features, and the latter can help to find the key information for the FC layer to classify the garbage images. According to the visualization of several garbage images in Fig. 2, the attention mechanism can enable the MobileNetV2 to pay more attention to the target area and suppress unnecessary features. By adding the attention mechanism into the MobileNetV2, the training process for finding the network weights can converge in a short time. The fine-grained garbage image can improve the garbage classification accuracy.

2.4. Reduction of feature dimension

In this study, the dimension of the feature map extracted by the MobileNetV2 is 1,280 × 1 × 1. The PCA is used to reduce the feature dimension. The features are first rearranged as a matrix $\mathbf{A}_{m\times n}$.

$$a_{ij} = \frac{(a_{ij} - \overline{a}_i)}{S_i} \tag{3}$$

where a_{ij} is the element of the first row of the matrix $\mathbf{A}_{m\times n}$,$\overline{a}_i$ is the mean of the first row and S_i is the standard deviation of the matrix.

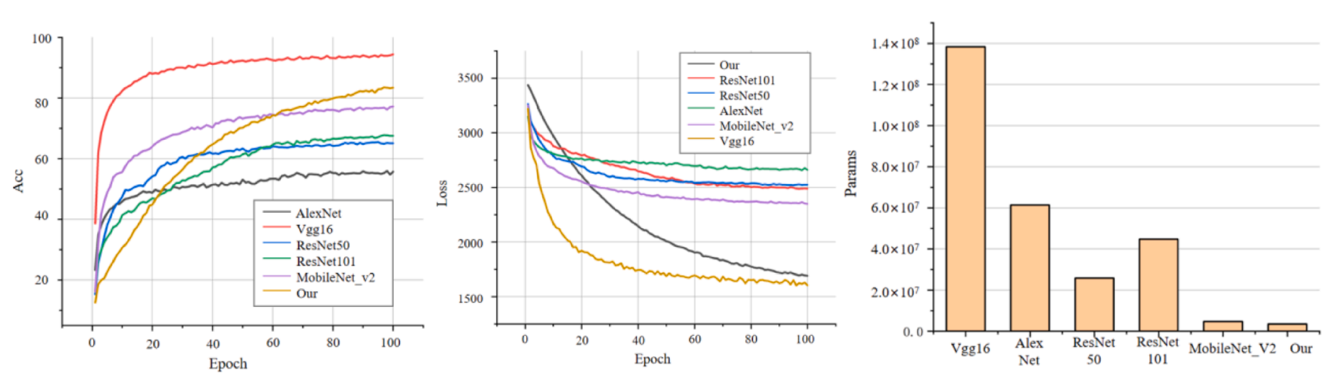
The covariance matrix is expressed as

$$C = \frac{1}{m}\sum_{k=0}^{j}(a_j - \overline{a})(a_j - \overline{a})^T \tag{4}$$

The eigenvalues of the covariance matrix are $\lambda_1, \lambda_2, \lambda_1, \cdots, \lambda_k$. Arranging the eigenvectors into rows of $\mathbf{A}_{m\times n}$ from top to bottom according to the corresponding eigenvalue size, the PCA contribution of each eigenvector is calculated by

$$e_i = \lambda_i / \sum_{i=1}^{k} \lambda_i \tag{5}$$

The relationship between the feature dimension and PCA contribution is shown in Appendix 5. By counting the cumulative contribution, the dimension number of 673 generates 95.02 % of the cumulative



(a) Accuracy (b) Loss value (c) different models

Fig. 4. Performance of different popular methods.

Fig. 5. Real-world garbage identification results.

Table 3
Garbage classification results.

Category	Original Huawei training data	GAN enhanced training data				
	Testing	Correct	Precision	Testing	Correct	Precision
Recyclables	130	120	92.31 %	130	122	93.84 %
Kitchen waste	33	27	81.82 %	33	29	87.88 %
Hazardous waste	53	46	86.79 %	53	47	88.67 %
Other waste	284	273	96.13 %	284	277	97.54 %
Total	500	466	89.26 %	500	475	91.98 %

contribution. Based on the 95 % selection standard, this work selects 673 features by the PCA.

In order to examine the appropriateness of the selected features, a brief test is conducted to evaluates the PCA performance in classifying the garbage images. The processing time and classification accuracy of the proposed deep-learning model with PCA and without PCA are tested using the Huawei image data. A total of 3,532 kitchen waste images are used to train and test the deep-learning model, and the results are shown in Table 1. As can be seen that using the PCA reduction the classification accuracy remains but the processing time is greatly saved. The average improvement of the processing time using the PCA feature reduction is 36.35 % against that without PCA reduction.

2.5. Transfer learning

Due to large number of garbage types, it is difficult/impossible to annotate every garbage dataset; consequently, the MobileNetV2 model has difficulty in identifying new garbage types. The transfer learning appropriately relaxes the assumption that the training data must be independent and identically distributed, thus it can address the problem of insufficient training data.

In this study, the pre-training + fine-tuning strategy is adopted for performing the transfer learning. Given a trained model, the transfer learning (Weiss et al., 2016) is able to utilize the existing learnt knowledge to deal with a new problem at a certain degree of avoiding another new training process. In the transfer learning processing, the MobileNetV2 model is used as the pre-trained model and a set of pre-trained weights are trained from the Huawei datasets for the MobileNetV2 model; and then, the pre-trained MobileNetV2 model is used for fine-tuning the weights at different network layers using an extended dataset. Since the MobileNetV2 belongs to a lightweight model and the number of model parameters is small, it is unnecessary to freeze the number of any network layers; in contrast, one can flexibly determine which layer(s) that need be loaded with the pretrained parameters.

In order for a smooth transfer learning, it is critical to develop an extended image data for the fine-tuning. For this purpose, the training set of Huawei images are employed to construct the fine-tuning dataset. For a specific garbage object in the Huawei images, the original image of this object is processed through a certain angle of random jitter and random flip in horizontal and vertical directions to add different noise and disturbances to simulate different lighting conditions, saturation and brightness adjustment of the photo environment. For example, Appendix 6 shows the image extension for a waste can. After extended the Huawei training images, these new datasets are used to train the pre-trained model. By reading the information of the pre-training model, the weight of the pre-training is allocated according to the corresponding layer. Since the features extracted from the first convolution layer are more general, the network does not train the first convolution layer in the process of gradient updating, and other layers are updated according to the new datasets to enhance the classification ability of the MobileNetV2 model.

Appendix 7 compares the performance of the pre-trained MobileNetV2 model and the transfer learning model. It can be seen that the classification accuracy of the MobileNetV2 with the transfer learning is 83.52 %, which is higher than the pre-trained model.

3. Experimental validation and results

3.1. Garbage dataset

In this study, the garbage image dataset provided by Huawei Cloud Garbage Classification Competition is used to train and evaluate the proposed model. The Huawei Garbage Classification Challenge Cup dataset (HUAWEI-40) contains 14,683 images of four categories, including the recyclables (23 types), kitchen waste (8 types), hazardous waste (3 types), and other waste (6 types). All the images in the HUAWEI-40 are collected from people's daily lives through mobile phones. Appendix 8 lists a portion of these categories. It should emphasize that the 'other garbage' is one category stipulated by the Chinese government.

3.2. Experimental platform

Fig. 3 shows the developed garbage identification experimental system. The raspberry PI is used as the control center, where the proposed improved MobileNetV2 model has been embedded. The camera is connected to the raspberry PI to provide the MobileNetV2 model with real-time captured garbage images. The experimental server configuration is as follows: the CPU is Inter Core I7-1400 k; the GPU is Nvidia 2060ti; the operating system is deepin 20.4; the python version is 3.8.7; the deep learning framework is pytorch1.7.0; the edge device is based on the Raspberry PI 4B 4G memory version.

A four-gear steering device is used to control the device movements for waste classification management. As shown in Fig. 3(b)-(d), the storage tank is divided into four rooms, corresponding to the four garbage categories in the HUAWEI-40 data. The operation procedure is described as follows.

(1) The camera captures the image of the input garbage the cover of the storage rooms in real-time; the direction of the camera is vertical downward the cover of the storage tank.
(2) The proposed MobileNetV2 model detects and classifies the garbage from the obtained image to determine its category.
(3) The steering gear rotates the cover of the storage tank to open the room corresponding to the garbage category.
(4) Rotate the block to push the garbage into the corresponding room for garbage storage and future recycling.
(5) Repeat Step (1)-(4) for new input garbage.

The MobileNetV2 model is trained using the GPU training. During the training process, the flipping and random brightness transformation is applied to the original datasets to strengthen the generalization ability of the model. The model optimization selects Adam, the learning rate is 0.0002, the loss function uses cross-entropy loss function, the training cycle is 100, and eight (8) training images per batch.

The testing datasets adopts the "Huawei Cloud" garbage classification data. These characteristics of the "Huawei Cloud" significantly increase the garbage detection difficulty, which make it very suitable to evaluate the performance of the proposed MobileNetV2 model. According to the garbage classification standard, the "Huawei Cloud" data is divided into four categories: the recyclables, kitchen waste, hazardous waste and other waste. In the training process, the data is divided into the training set, verification set and test set using a ratio of 7:2:1; that is, according to Appendix 8, the training set including 6,967 images for Recyclables, 2,472 images for Kitchen waste, 445 images for Hazardous waste, and 392 images for Other waste.

3.3. Experimental tests

This section evaluates the performance of the proposed MobileNetV2 model for identifying the four garbage categories in the "Huawei Cloud" data. The experimental results are shown in Table 2. Compared with the original MobileNetV2 model, the proposed improved model produces much higher garbage identification accuracy. This is because the CBAM module can improve the identification accuracy of the original MobileNetV2 model by 12.4 %; the PCA module can reduce the parameter volume by 30.1 % and increases the identification accuracy by 10.3 %; the transfer learning helps the original MobileNetV2 improve 8.0 % identification accuracy. As a result, the integration of these three improvement factors in the proposed model, its parameter volume reduces by 28.2 % and the identification accuracy increases by 19.3 %.

To highlight the effectiveness of the proposed method, different improvement factors are analyzed using the original MobileNetV2 model in the training process. The analysis results are illustrated in Appendix 9. As can be seen that the garbage identification accuracy of the original MobileNetV2 model is less than 75 %; if only use the CBAM module, the identification accuracy reaches more than 80 %; If only use the PCA module, at 120th training epoch the identification accuracy starts to exceed the that of the original MobileNetV2, which suggests that the PCA can select the most useful 673 features while discards the useless/redundant features to improve the training accuracy; If the CBAM and PCA modules are used at the same time, the training accuracy can further increase to more than 90 %.

Furthermore, the performance of the proposed method is compared with existing popular methods in the literature, including the classical AlexNet (Krizhevsky et al., 2017), Vgg16 (Simonyan, and Zisserman, 2014), ResNet50 (Xia et al., 2022), and ResNet101 networks (He et al. 2016). The training performance of these methods are portrayed in Fig. 4, which suggests that the Vgg16 generates the best training accuracy and smallest training loss among all these methods; while the present method produces the second-best performance. However, according to the comparison of the parameter volumes in Fig. 4(c), the number of model parameters of the proposed method is only 3.42 % of that of the Vgg16. Because the VGG16 requires a huge number of parameters, the average identification time of one image is about 5,000 ms, which cannot meet the requirement on real-time processing. The ResNet50 and ResNet101networks reduces the number of model parameters but the average identification time of each image is more than 1,500 ms. The average identification time of the original MobileNetV2 is 770 ms while the proposed MobileNetV2 model only needs 600 ms. As a result, the proposed method produces the best overall performance and can be used in the real-time implementation of the garbage detection and identification on the edge devices.

3.4. Edge-Device-based model deployment

In this section the proposed MobileNetV2 model is evaluated on the developed edge device in Fig. 3 for real-world garbage detection and identification. Firstly, the proposed MobileNetV2 model is trained utilized the HUAWEI-40 dataset; the training parameter setting is the same as in Section 3.2. After completing the model training, the well-trained MobileNetV2 model is embedded into the Raspberry PI hardware; then, the device is used to process real-world garbage, where the camera takes the images of arbitrary input garbage in real-time and the embedded MobileNetv2 model detects and classifies these real-world images for waste sorting, storage and future recycling.

During the edge device testing, the camera in the device randomly shot images of 130 recyclables, 33 kitchen, 53 hazardous, and 284 other wastes in actual implementation to verify the effectiveness of the proposed MobileNetv2 model. These images are new real-world images but do not belong to the HUAWEI-40 dataset. The garbage identification results are shown in Fig. 5 and Appendix 10, where the numbers under each garbage item are the confidence levels of the identification accuracy and 1.0 is the best value of the confidence level.

Fig. 5 depicts the identification results on the high-quality garbage images. As can been seen that, the confidence levels of the bag and chopstick are close to 1.0, while the confidence levels of some garbage items with complex background, such as the broken bowl and the shoes, are relatively low. In order to investigate the effect of background, Appendix 10 manifests the identification results on the low-quality garbage images, where the images are subject to strong background effect. It can observe from Appendix 10 that due to complex background, the confidence level in general decreases significantly, and the confidence level can be as low as 0.6701 for the leave identification.

The garbage classification results on the 500 real-world images are shown in Table 3. It should note that the 500 images mix high-quality and low-quality garbage images. One can find that the average classification accuracy is 89.26 % for all the 500 images; the highest accuracy 96.13 % is for the other waste category while the lowest is 81.82 % for the kitchen waste category among the four waste categories. By checking these real-world images, we found that the other waste images are generally with better resolution, clearer brightness and more accurate focus than that of the other images. As a result, in order to improve the performance of the developed device on the garbage detection and classification, it is critical to ensure the image quality. In order to address this issue, the training images of the Huawei Cloud data are processed by the Generative Adversarial Nets (GAN) (Goodfellow, et al. 2014). The GAN is very suitable for small-size images detection by improving the image quality. The garbage detection results using the proposed MobileNetv2 model trained by the GAN processed the Huawei data are shown in Table 3, where one can note that the identification precision of the kitchen waste is improved from 81.82 % to 87.88 % and the average classification accuracy is improved from 89.26 % to 91.98 %. These observations suggest that the image quality enhanced by the GAN can improve the garbage detection performance.

4. Conclusions

This work improves the original MobileNetv2 model to make the present garbage identification model applicable to edge devices. Experimental results show that the proposed method achieves high classification precision and low consumption cost. Using the MobileNetV2 network as the backbone network, the embedded convolutional attention module can improve the model accuracy and the PCA dimension reduction can greatly reduce the model parameters. The optimization of the initial weight parameters using the transfer learning is able to make the well-trained MobileNetV2 model adaptive to new garbage types. Compared with the traditional MobileNetV2 network, the garbage classification accuracy is improved by 19.3 %, the identification time is reduced by 170 ms, and the model parameter volume is compressed by 30.1 %. Hence, the present model has practical importance in garbage detection and classification.

It should emphasize that the developed waste recycling device is only able to process a single object at each operation, where the camera shoots the object and the obtained image is analyzed by the embedded MobileNetv2 model to identify the garbage category to trigger the steering gear. As a result, a single garbage can be successfully stored in the recycling device. However, in reality a single image may contain multiple types (such as recyclables and kitchen waste), which significantly increases the processing difficulty for the correct garbage identification and recycling. It crucial to identify multiple objects in a single image, as well as in multiple images. To solve this challenging task, the Yolov5 deep learning model will be introduced into the present waste recycling device because that the Yolov5 model has been proven be effective to box multiple objects in a single image. In addition, we plan to further improve the accuracy of the deep-learning model by expanding the image dataset and improving the model structure using model pruning and distillation. A much more advanced/sophisticated recycling device will be developed in the next research plan, which will play a positive role in environmental protection.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This research is supported by the Natural Science Basic Research Program of Shaanxi (No. 2023-JC-YB-288) and the Science Foundation of Donghai Laboratory (No. DH-2022KF0302).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.wasman.2023.02.014.

References

Al Mamun, M., et al., 2014. Integrated sensing systems and algorithms for solid waste bin state management automation. IEEE Sens. J. 15 (1), 561–567.

Cai, X., et al., 2022. Towards lightweight neural networks for garbage object detection. Sensors 22, 7455.

Chen, Z., et al., Aug 2021. Garbage image classification algorithm based on improved MobileNet v2. J. ZheJiang Univ. (Eng. Sci.) 55 (8), 1490–1499.

Chen, Z., et al., March 2022. Garbage classification system based on improved shufflenet v2. Resour., Conserv. Recycling 178, 106090.

Feng, J., et al., Dec 2021. Garbage disposal of complex background based on deep learning with limited hardware resources. IEEE Sens. J. 21 (18), 21050–21058.

Fu, B., et al., Sep 2021. A novel intelligent garbage classification system based on deep learning and an embedded linux system. IEEE Access 9, 131134–131146.

Goodfellow, I., et al., 2014. Generative Adversarial Nets, NIPS 2014.

Gupta, T., et al., Feb 2022. A deep learning approach based hardware solution to categorise garbage in environment. Complex Intell. Syst. 8 (2), 1129–1152.

He, K., et al., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778.

Kang, Z., et al., June 2020. An automatic garbage classification system based on deep learning. IEEE Access 8, 140019–140029.

Krizhevsky, A., Sutskever, I., Hinton, G., 2017. Imagenet classification with deep convolutional neural networks. Commun. ACM 60 (6), 84–90.

Li, J., et al., Jan 2021. Automatic detection and classification system of domestic waste via multimodel cascaded convolutional neural network. IEEE Trans. Ind. Inf. 18, 163–173.

Liu, F., et al., May 2022. Depth-Wise Separable Convolution Attention Module for Garbage Image Classification. Sustainability 14, 3099.

Lv, Z., Li, H., Liu, Y., 2022. Garbage detection and classification method based on YoloV5 algorithm. Fourteenth International Conference on Machine Vision (ICMV 2021). Vol. 12084. SPIE, 2022.

Majchrowska, S., et al., Jan 2022. Deep learning-based waste detection in natural and urban environments. Waste Manag. 138, 274–284.

Park, J., et al., 2020. A simple and light-weight attention module for convolutional neural networks. Int. J. Comput. Vis. 128 (4), 783–798.

Qin, J., et al., Mar 2022. A robust framework combined saliency detection and image recognition for garbage classification. Waste Manag. 140, 193–203.

Sandler, M., et al., 2018. Mobilenetv2: Inverted residuals and linear bottlenecks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Shen, X., et al., 2021. An intelligent garbage sorting system based on edge computing and visual understanding of social internet of vehicles. Mob. Inf. Syst. in press.

Shi, C., Xia, R., Wang, L., Aug 2020. A novel multi-branch channel expansion network for garbage image classification. IEEE Access 8, 154436–154452.

Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556.

Tong, Y., Liu, J., Liu, S., May 2020. China is implementing "Garbage Classification" action. Environ. Pollut. 259, 113707.

Weiss, K., et al., Jan 2016. A survey of transfer learning. J. Big Data 3 (1), 1–40.

Xia, W., et al., June 2022. Application of machine learning algorithms in municipal solid waste management: A mini review. Waste Manag. Res. 40 (6), 609–624.

Yang, G., et al., 2021. Garbage classification system with YOLOV5 based on image recognition. In: 2021 IEEE 6th International Conference on Signal and Image Processing (ICSIP). IEEE, 2021.

Yuan, J., et al., Feb . 2021, Garbage image classification by lightweight residual network . ENVIRONMENTAL ENGINEERING , 39(2): , 110 – 115 . doi: 10.13205/j. hjgc.202102017.

Zhao, Y., et al., Jan 2022. Intelligent garbage classification system based on improve MobileNetV3-Large. Connect. Sci. 34 (1), 1299–1321.

Zoph, B., et al., 2018. Learning transferable architectures for scalable image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.