WasNet: A Neural Network-Based Garbage Collection Management System

ZHIHU YANG AND DAN LI

Department of Computer Science and Technology, Northeast Forestry University, Harbin 150040, China

Corresponding author: Dan Li (ld@nefu.edu.cn)

ABSTRACT In response to the increasing pollution caused by unseparated garbage, classification systems for garbage separation have become very popular. First, we constructed a complex data augmentation combination for model training. Second, we designed a novel lightweight neural network garbage classification system called WasNet. This proposed network's 1.5 million parameters on the ImageNet dataset are one-half of mainstream neural networks, while at 3 million floating point operations per second (FLOPs) it is one third of mainstream neural networks that have obtained the best performance among known lightweight neural networks. The accuracy on the ImageNet data set is 64.5%, on the Garbage Classification dataset it is 82.5%, and on the TrashNet dataset it is 96.10%. Furthermore, we transplanted the model to the hardware platform and assembled an intelligent trash can; we developed a garbage recognition application to facilitate users to directly identify and receive platform information; we built a visualization and decision support platform to help managers monitor traffic in real time. We combined the intelligent trash can, application, visualization and decision-making platform into a system, which is the most complete and effective system among the known research works. The results of the test we conducted on our platform using our extended dataset showed that our scheme is very reliable. At the same time, we also open source our extended datasets for use by other researchers.

INDEX TERMS Pattern classification, artificial neural network, smart city, smart environments, urban computing.

I. INTRODUCTION

Garbage classification has become an increasingly popular topic in recent years. A report published by the World Bank in 2018 [1] shows that in 2016, 242 million tons of plastic waste was generated worldwide, accounting for 12% of the total solid waste. It is estimated that by 2050, the world will produce 3.40 billion tons of waste every year, a significant increase from today's 2.01 billion tons. In India, scavengers and water workers play a vital role in the recovery of municipal solid waste. However, extremely hazardous working conditions cause infections of the skin, respiratory system, and gastrointestinal tract, such that the incidence of disease is very higher. In China, waste sorting is highly valued by the government and the public. In a processing plant that handles a large amount of waste, separation and release are the first steps in the treatment process. Manual pipeline sorting of waste is used but has the disadvantages of a harsh working

environment, high labor intensity, and low sorting efficiency. Furthermore, faced with a large amount of garbage, manual sorting can only sort out a minimal amount of recyclable and harmful garbage. Most of the garbage goes to landfill, resulting in a waste of resources and environmental pollution.

In order to solve various problems in garbage classification and recycling, make the whole process more efficient and save resources, we have proposed a novel lightweight neural network WasNet and garbage collection management system. We did the following in our study.

- In terms of the dataset, we used Huawei Cloud's open source Garbage Classification dataset [2] with a total of 18079 pieces in 40 categories. At the same time, we used multiple data enhancements to construct a larger amount of data. We matched the garbage classifying dataset through search engines such as Google and Bing, which expanded it by 4175 pieces. See Section III-A for details.

- WasNet is designed by stacking multiple times in depth and width, and embeds the attention mechanism, taking into account the use and accuracy of the terminal. Compared with

---

The associate editor coordinating the review of this manuscript and approving it for publication was Jonghoon Kim.

existing lightweight neural networks, it performs better in all respects. See Section IV-B for details.

- We deployed the model to the terminal, completed the overall architecture of cloud-side collaboration, and obtained good test data. At the same time, real-time data integration architecture was employed to prevent data shifts caused by data expansion in actual production, which could lead to a deterioration of the model's performance. See Section V-B for details.

- A system equipped with the WasNet model was developed by us to solve the various replacement problems of garbage collection. The platform includes a smart trash can, mobile phone application, data visualization and decision-making platform. For other contents of the system architecture set, see the Section V-C for details.

II. EXISTING WORK

In the digital world, we can use more advanced modern technologies and the Internet to simplify our work and increase efficiency. However, in some countries, it takes a lot of effort and money to complete garbage collection. The waste collector must arrive at the designated place on time, which is undoubtedly the main disadvantage of the system [3], and the waste cannot be sorted and recycled at the source. In other countries, the system collects trash from fixed waste bins. The waste bins are often overfilled, causing problems such as bacterial growth, animal foraging for food, and insect reproduction. It also does not solve the problem of sorting and recycling garbage at the source, and even loses more time and energy to deal with garbage [4].

There has been considerable research in the field of garbage collection. Reference [5] proposed the use of a short-wave infrared hyperspectral imaging system for automatic waste sorting. In this system, waste is scanned by a short-wave infrared (SWIR) hyperspectral imaging system and is classified into six plastic waste types (polyethylene terephthalate, high density polyethylene, polyvinyl chloride, low density polyethylene, polypropylene, and polystyrene) as well as into paper, cardboard, metal, and glass. As far as we know, the price of infrared hyperspectral imagers is expensive, which results in high cost of the system, and the garbage classification standards adopted by the system cannot meet the requirements of real life. Before the smart trash mainly focused on the use of sensor methods. The dustbin [6] proposed is assisted by an ultrasonic sensor connected to the Arduino UNO to check the amount of garbage filled in the dustbin. Reference [7] uses IoT and wireless sensor solutions, and applies machine learning techniques such as decision forest regression to the use of sensor data. Once the garbage is filled, an alert will be sent to the municipal web server. The "IoT-based waste management system." proposed by [8] uses flame sensors to detect fire and moisture sensors to separate dry and wet waste. This system is helpful for storing dry and wet waste separately, so different types of waste should be treated differently (including composting, recycling, incineration). In [9], garbage is divided into

three categories: metal, dry garbage and wet garbage. The dry garbage is further divided into paper and plastic. The garbage classification system allocates a sensor for each kind of garbage to identify the type of waste through the sensor. In [10], a garbage can composed of sensors is introduced to measure the weight of garbage and the level of garbage in the garbage can. In [11], the author proposes intelligent garbage can with sensor Internet of things model, which is mainly used to measure the garbage level in the garbage can to help workers optimize the collection path. These methods of using sensors to build garbage recycling devices have the problems of few types of recyclable garbage, low accuracy, and complicated assembly.

In recent years, neural network methods have been widely used in image classification, and have proved very effective at garbage classification. Reference [12] compared several convolutional neural network architectures: VGG, Inception, and ResNet. Among them, the Inception-ResNet model obtained the best classification results, reaching an accuracy of 88.6%. Reference [13] uses Softmax and SVM as classifiers, GoogleNet as a feature extractor, and employs a transfer learning strategy to achieve 97.86% accuracy on the TrashNet dataset. Reference [12], [13] method of the neural network, previous related studies used only the TrashNet dataset [14], as shown in Table 1, which does not involve multi-classes and multi-size. Furthermore, the neural network used has the problems of complex structure, many parameters and slow running speed, which makes it need a very expensive hardware platform to carry the model.

TABLE 1. There are six categories of TrashNet dataset with 4,175 pictures. The types in the dataset are too few, are not close to real life, and the amount of data is small. If the data enhancement work is not good, it can easily cause an overfit.

class	amount	class	amount	class	amount
glass	501	paper	594	cardboard	403
plastic	482	metal	410	trash	137

In addition to the study of garbage collection methods, some scholars have also done corresponding work in other areas. Reference [15] proposed a specific problem model, which can reduce investment costs, increase the number of residents who have installed the dustbin service and the accessibility of the system.

Summarize related work. Their system costs are high, the recognition types are few, and the accuracy is low. They cannot really solve the problem of the end-to-end garbage collection platform, but only conduct research and discussion at the method level. Compared with existing work, Huawei Cloud's open source garbage classification dataset is more representative of garbage classification tasks. WasNet has fewer parameters, shorter inference time and higher accuracy. Our system is integrated end-to-end and provides users with more information.

III. DATA

In this section, we mainly describe the work on data. The Section III-A is data collection and sorting; the Section III-B is data analysis; the Section III-C is data enhancement; the Section III-D is to verify the effectiveness of our work in the Section III-B and Section III-C through experimental results.

TABLE 2. There are as many as 30 categories of Garbage Classification dataset with 18079 pictures, which covers almost all domestic garbage. The amount of data is huge, which poses new challenges for improving the accuracy of the network.

class	amount	class	amount	class	amount
Lunch box	214	Peel	358	Cosmetic bottle	386
Plastic bag	370	Pulp	427	Plastic toys	308
Cigarette butts	277	Tea residue	384	Plastic bowl	364
Toothpick	84	Cauliflower root	711	Plastic hangers	312
Basins dishes	451	Eggshell	323	Courier paper	225
Chopsticks	288	Fishbone	408	Plug wire	652
Leftovers	388	Power Bank	355	Old clothes	376
Big bones	361	Bags	419	Cans	309
Pillow	318	Wine bottle	293	Plush toy	549
Shampoo bottle	320	Chopping board	416	Metal cans	323
Glass	488	Cardboard box	319	Pot	395
Leather shoes	377	Flavoringbottles	406	Cooking oil barrel	361
Drink bottle	223	Expired drugs	436	Ointment	387
Dry cell	322				

A. DATA COLLECTION

Our main data comes from the Huawei Cloud Garbage Classification Competition. As shown in Table 2, the dataset contains 18079 pictures in 40 categories. It is divided into recyclables, which refer to wastes suitable for recycling and resource utilization, including waste glass, metal, plastic, paper, fabrics, furniture, electrical and electronic products, and annual flowers; kitchen waste, which refers to perishable waste generated by households and individuals, including leftovers, leaves, peels, eggshells, tea residues, soup residues, bones, waste food and kitchen waste; hazardous waste, which refers to waste that directly or potentially harms human health or the natural environment and should be specially treated, including waste batteries and waste fluorescent tubes; and other garbage, which refers to domestic garbage such as diapers, dust, cigarette butts, disposable fast-food boxes, damaged flower pots and dishes, wallpaper, etc. Compared with the small data used in [5], [8], [12], [13], this dataset better fits the types of waste in real life scenarios, which increases the difficulty of model designing and training.

We also used the TrashNet dataset for the development of our newly developed lightweight network.

However, in order to ensure that our network system more accurately reflects actual practice, we searched Garbage Classification datasets through Google and Bing and subsequently expanded the dataset by 4175 pictures. This is shown in Section V-B and is intended to be an open source for other researchers to use.

B. DATA ANALYSIS

We counted the data sample distribution, size distribution, picture morphology, etc. Based on the analysis, we conducted targeted data preprocessing, which helped with designing and training the subsequent model. As can be seen from the data volume of each class in Table 2, the data volume of each category is uneven. Fewer data are "toothpick", "towel" and "drink boxe"; while more data are "leaf roo" and "plug wire".

Since the data distribution is unknown, the problem of dataset skewing needs to be addressed. During the initial model training, it was assumed that higher loss function values would be assigned to the categories with fewer data and the loss would be converted into a weighted average to tell the model to pay more attention to the lack of representing activeness. However, the problem of category imbalance in actual production is discussed in Section V-B.

For deep learning neural networks, the resolution and aspect ratio of the input pictures of the network will affect the performance of the network. In real life, the image data we collect or identify is often unknown in scale with an unknown aspect ratio. We found that the aspect ratio of the pictures in the Garbage Classification dataset and the extended dataset is different. From our statistics we found that most of the aspect ratios are concentrated between 0.8-1.2; in order to compare with existing neural networks when designing the model, the input size of our model was set at 1:1. Considering the number of parameters, the speed of the model, and the comparison of existing neural networks, we set the size of the input image to a commonly used 224 * 224.

We observed that the similarity between toothpick and chopsticks in the image data is too great, so that it may not be able to reflect more detailed features under the commonly used method of zooming pictures. We use algorithm 1 to make the original image based on the largest edge.

It was scaled proportionally, the missing part filled in with a certain value (the scaled side is 256/224 times the final input side length), and then it was cut. A comparison of the results is shown in Fig.1.
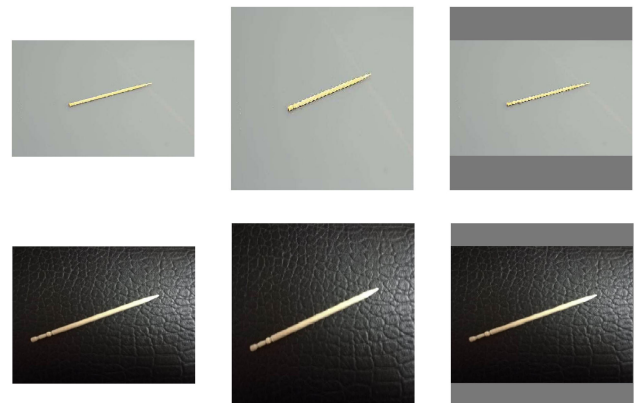
FIGURE 1. Two toothpicks are printed in the first column. It can be observed that the width is greater than the height. With conventional cropping methods, the second column of images is obtained, and the algorithm is used to obtain the third column of images. Compared with the second and third columns, our algorithm can reduce the feature loss of particular data such as chopsticks.

Algorithm 1 CenterImage

Input: *img* Image to modify, *s* To change the size of a picture, *fill* Value to fill

Output: Modified picture *background*

   function CenterImage(*img*, *s*, *fill*)

      M = Max(*h*, *w*)

      if *s == null* **then**

         $s \leftarrow M$

      end if

      $h \leftarrow$ *height of img*

      $w \leftarrow$ *width of img*

      scale $\leftarrow s \div M$

      image $\leftarrow$ *resize the img to* ($s * scale$, $s * scale$) *size*

      shape $\leftarrow$ *define a* (s, s) *size image*

      background $\leftarrow$ *fill the shape with value of fill*

      $c_x \leftarrow (s - w) \div 2$

      $c_y \leftarrow (s - h) \div 2$

      Local area with background from c_y *to* ($c_y +$ *h*) *and* c_x *to* ($c_x + w$) $\leftarrow$ *img*

      return *background*

   end function

C. DATA AUGMENTATION

For common data augmentation, we have selected a data augmentation combination of horizontal flip, vertical flip, and random panning [−0.1, 0.1].

There are many methods for image classification tasks that can reduce the risk of over-fitting and make the model resistant to occlusion by mixing the data to improve accuracy. Reference [16] described random erasing, randomly selecting rectangular areas in an image, and erasing pixels using random values. Reference [17] verified label smoothing and mixed training (Mixup). Mixup changed the way that previous enhancements only enhanced a single sample and modelled the domain relationship between different types of samples.

Reference [18] proposed Cutmix, a region discarding strategy to enhance the performance of convolutional neural network classifiers. Region discarding strategy refers to cutting and pasting patches between training images, where real labels are blended in proportion to the regions of the patch, by effectively using training pixels and preserving the regularization effect of region losses. Cutmix consistently outperforms the latest enhancement strategies on CIFAR and ImageNet classification tasks and ImageNet supervises localization tasks. We used these three hybrid data enhancement methods and integrated them into a complex set of data enhancers. Fig.2 shows the resulting visualization image.

D. DATA EXPERIMENT

This part mainly shows the effectiveness of our data modification and data enhancement strategy in the Section III-B and Section III-C. We use the ResNet50 [19] model, the optimizer is Adam, the learning rate is 1e-3, the pre-training parameters

of ResNet50 on ImageNet data are loaded, and the data as shown in Table 3 is obtained at 100 epoch. The abbreviations in the table represent: "h_f" is horizontal flip, "v_f" is vertical flip, "w_r" is width shift range, "h_r" is height shift range, "rr" is random rotation, "br" is brightness range, "ci" is center image, in the Section III-B has been expressed; "eraser", "Mixup", "Cutmix" has been described in the Section III-C. Val_acc refers to the accuracy of top-1. The model predicts a category, and when the predicted value is the same as the true value, the prediction is correct. Top-5 refers to the model predicting five categories, as long as any one of the five predicted values is the same as the true value, it means that the prediction is correct. Experiments show that the optimal combination of data enhancement is: horizontal vertical rotation, horizontal vertical translation, center image, "Mixup", "Cutmix", "Random Eraser". It is worth mentioning that we did not generate new data for the test data during the testing phase of this experiment and all subsequent experiments, nor did we adopt any data enhancement strategy.

IV. DEEP NEURAL NETWORKS

This section mainly describes our work in building a lightweight neural network. The Section IV-A is a brief review of the development of the most popular lightweight neural network; the Section IV-B describes the process of building WasNet; Section IV-C is to verify the advanced nature of WasNet through comparative experiments.

A. LIGHTWEIGHT IMAGE NETWORK

In the ILSVRC2012 competition image classification task, the convolutional neural network model proposed by [20], named AlexNet, won the first place with an error rate as low as 16.4%. AlexNet is an eight-layer convolutional neural network with the first five layers being convolutional layers and the last three layers being fully connected layers. The last layer of its fully connected layer uses the Softmax classification method and the ReLU function as a non-linear activation function. The Dropout method is applied to reduce the occurrence of overfitting. In 2015, [19] proposed a cross-layer connection method, making the network layer up to 152 layers, alleviating the problem of gradient dissipation in deep neural networks. In 2016, Xecption [21] uses deep separable convolution to maintain channel separation, instead of connecting deep convolutional structures to achieve spatial convolution, decoupling spatial information and depth information, can effectively use parameters to obtain better performance. So deep separable convolution can also be used in mobile devices. In 2018, ShuffleNet v2 [22] proposed that when the ratio of input channels to output channels is 1:1, the MAC is the smallest and the speed of the network is the fastest. The number of packets should be carefully selected based on the actual subject and the platform of the application; the packet volume should not be increased simply to increase the number of channels and thereby increase the accuracy.
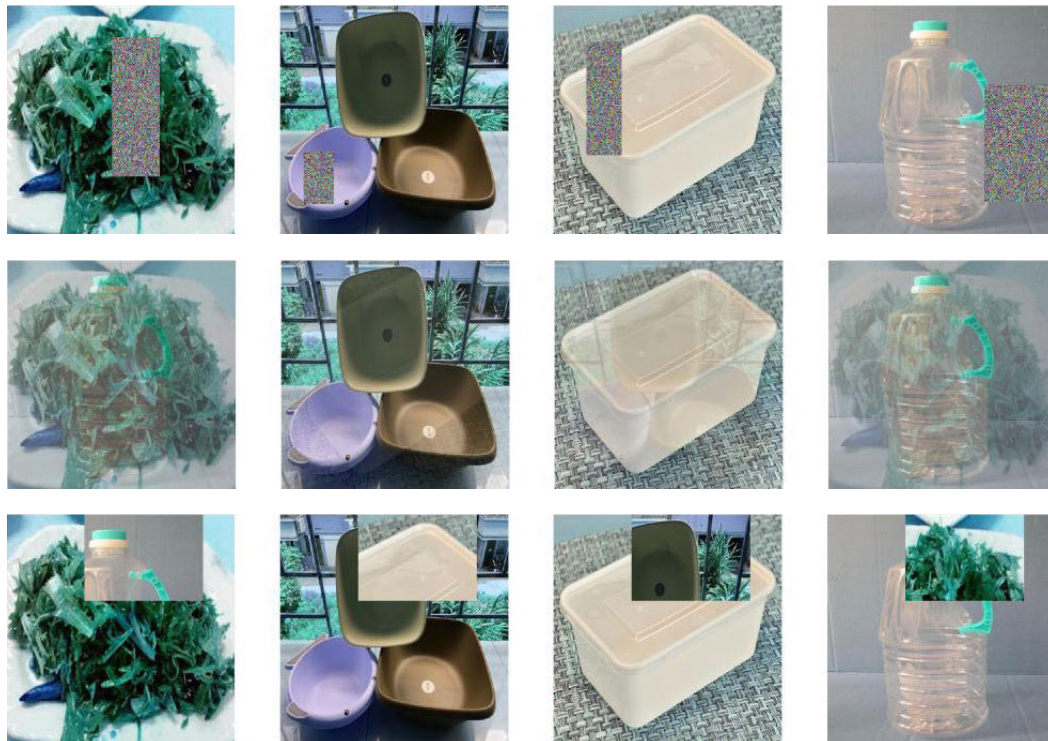
FIGURE 2. The first line is random erasure, the second line is Mixup, and the third line is Cutmix.

TABLE 3. Data modification and enhanced test data.

Enhanced method	Val Acc	Val Top 5	Val Loss
baseline	0.6629	0.8870	1.4570
h_f	0.7154	0.9196	1.1760
h_f,v_f	0.7337	0.9210	1.0990
h_f,v_f,w_r(0.05),h_r(0.05)	0.7487	0.9254	0.9846
h_f,v_f,w_r(0.10),h_r(0.10)	0.7766	0.9408	0.8543
h_f,v_f,w_r(0.15),h_r(0.15)	0.7743	0.9387	0.8856
h_f,v_f,w_r(0.10),h_r(0.10),rr	0.7157	0.9132	1.0840
h_f,v_f,w_r(0.10),h_r(0.10),br	0.7623	0.9356	0.9215
h_f,v_f,w_r(0.10),h_r(0.10),zr	0.7722	0.9397	0.8849
h_f,v_f,w_r(0.10),h_r(0.10),ci	0.7901	0.9501	0.8423
h_f,v_f,w_r(0.10),h_r(0.10),ci,cutmix(1.0)	0.8366	0.9677	0.7052
h_f,v_f,w_r(0.10),h_r(0.10),ci,mixup(0.5)	0.8291	0.9595	0.7139
h_f,v_f,w_r(0.10),h_r(0.10),ci,eraser(0.3)	0.8259	0.9595	0.6893
h_f,v_f,w_r(0.10),h_r(0.10),ci,mixup(0.5),cutmix(1.0),eraser(0.3)	0.8498	0.9646	0.6599

In 2019 NAS technology [23] released two new-generation network structures, MobileNetV3-Large and MobileNetV3-Small, for different resource consumption scenarios. Constructed on the basis of its previous MobileNet and combined with the ideas of SENet [24], it added swish nonlinear activation. In order to reduce the cost of calculating traditional Sigmoid in swish, a hard Sigmoid was proposed.
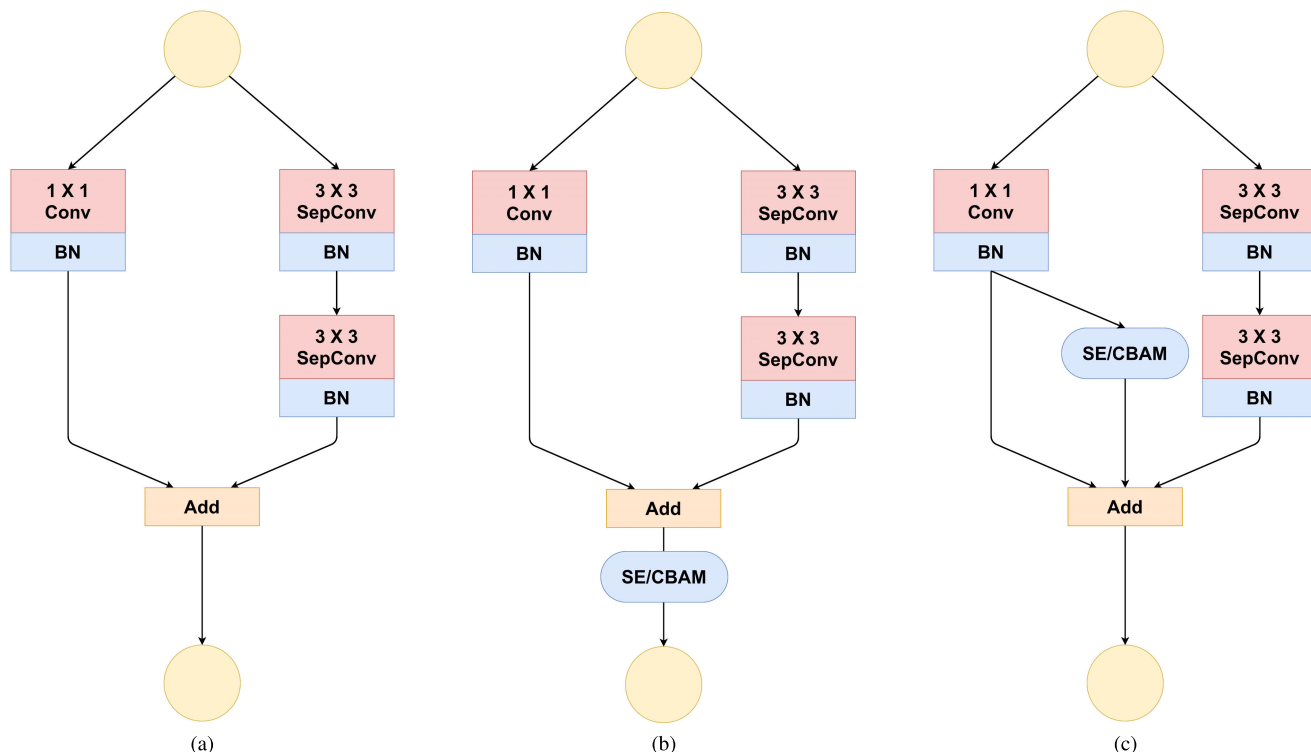
FIGURE 3. This figure was obtained by stacking convolutional layers and widening the convolution block. Fig.3-B and Fig.3-C show the diagrams of the two embedded representation mechanism modules we tested.

In addition, there is an Internal Covariate Shift phenomenon during the training of the neural network. The distribution of the input data of each layer changes with the change of the network parameters of the previous layer, which makes the training of deep neural networks more complicated and requires more cost time and resources to adjust the parameters. In order to solve this phenomenon, a Batch Normalization (BN) layer is designed in [25]. A BN layer is added before each input layer to do an improved normalization preprocessing for the input data of each layer. The normalization process will affect the features learned by the network layer, so the BN layer introduces two learnable parameters to improve the normalization operation so that the network can restore the distribution of features that the original network has to learn. The use of layers can effectively suppress the Internal Covariate Shift phenomenon.

B. WasNet

Our network referred to the design of existing high-precision neural networks. For the dataset, the experiment was continuously adjusted for the number of convolution layers and for the depth and width of the network in order to obtain the best network architecture as shown in Fig.3-A.

With the development of neural networks and the emergence of multi-tasking and multi-category classification tasks, people increasingly want neural networks to pay attention to sensitive areas. Accordingly, in order to improve the performance of the network, [24] invented SE and [26]

invented CBAM attention mechanism. We also include SE and CBAM into our neural networks. However, we were not sure whether adding conventional attention mechanisms would have a positive effect on results. Therefore, we experimented with two attention mechanism embedding methods as shown in Fig.3-B. and Fig.3-C. Table 4 shows the design and parameters of the entire network in detail.

TABLE 4. The design and parameters of the entire network.

Layer	Output size	KSize	Stride	padding	Output C
Image	224 X 224	-	-	-	-
ZeroPadding2D	230 X 230	-	-	3 X 3	32
Conv2D+BN+HS	228 X 228	3 X 3	1 X 1	-	32
ZeroPadding2D	230 X 230	-	-	1 X 1	32
Maxpooling2D	114 X 114	-	2 X 2	3 X 3	32
WasNet_Block_1	57 X 57	-	2 X 2	-	64
WasNet_Block_2	29 X 29	-	2 X 2	-	128
WasNet_Block_3	15 X 15	-	2 X 2	-	256
WasNet_Block_4	8 X 8	-	2 X 2	-	512
SepConv2D+BN+HS	8 X 8	3 X 3	1 X 1	-	512
GlobalPooling	1 X 1	-	-	-	-
FC	-	-	-	-	40/1000

C. EXPERIMENTS

When training the Garbage Classification dataset model, the parameters used are were as follows: the ratio of training data set and test data set is 0.8: 0.2, the optimizer was Adam, the learning rate was 1e-2, the val_loss was monitored, and

the learning rate was reduced by 0.5 times without dropping for ten rounds. Through experiments, we determined the nerves shown in Fig.3-C. As the last wasnet block, the network structure is best matched with the attention mechanism of CBAM. The experimental results are shown in the Table 5.

TABLE 5. WasNet block selection experiment.

WasNet_Block	Top-1 Val Acc	Top-5 Val Acc
WasNet_Block_Fig.3(b)-SE	74.89%	91.22%
WasNet_Block_Fig.3(c)-SE	77.72%	92.77%
WasNet_Block_Fig.3(b)-CBAM	76.52%	92.56%
WasNet_Block_Fig.3(c)-CBAM	78.70%	94.00%

In order to prove the advanced nature of our neural network architecture, we did the following comparative experiments. Considering the resource problem of the hardware platform, the required model is a lightweight model with few parameters and fast inference. WasNet is a typical lightweight model, so our comparison network is selected from mainstream lightweight models. We compared WasNet with two mainstream lightweight networks, ShuffleNet-V2 and MobileNet-V3-small, on the dataset. Reference [27] research confirms that the transfer ability of each layer feature in the deep learning neural network. After the network trained on the ImageNet dataset and fine-tuning the target dataset, the generalization effect still exists. In order to enhance the accuracy, we use the transfer learning strategy on the ImageNet dataset. The parameters used for transfer learning were as follows: the optimizer was Adam, the learning rate was 0.1, and the val_acc was monitored. The learning rate was reduced by 0.7 times in the case of no increase in 3 rounds. A comparison of the final experimental results is shown in Table 6. Params are the data from the original researcher, while FLOPs are the data obtained in the experiment. WasNet* is a model for transfer learning training on the Garbage Classification dataset after training on the ImageNet dataset.

TABLE 6. Experiment and comparison on ImageNet dataset and Garbage Classification dataset.

	ImageNet			Garbage Classification	
	Val acc	Params	FLOPs	Top-1 val acc	Top-5 val acc
MobileNet	67.4%	2.9M	8.2M	66.7%	90.5%
ShuffleNet	69.4%	2.3M	9.7M	64.7%	90.1%
WasNet	65.4%	1.6M	3.0M	78.7%	94.0%
WasNet*				82.5%	97.3%

In order to ensure that our research is a leader in field, we compared it to other published articles relating to public garbage classification.

D. COMPARISONS

Through data comparison in the Tabel 6, we can find that WasNet has fewer parameters, smaller FLOPs and higher accuracy than ShuffleNet-V2 and MobileNet-V3-small.

According to Table 7, WasNet performs best on the public dataset TrashNet.

TABLE 7. Experiment and comparison on TrashNet dataset.

Author	Methodology	Accuracy
Yang and Thung [12]	AlexNet	22.40%
Awe et al. [28]	Augmented data on train R-CNN	68.30%
Kennedy [29]	OscarNet(based on VGG-19 pretrained)	88.42%
Ruiz et al. [30]	Inception-ResNet model	88.66%
Our	WasNet(pretrained)	96.10%

Through the comparison of experimental data, we can clearly obtain that WasNet is more efficient than the current mainstream lightweight neural network, and the effect is more obvious in the task of garbage classification.

V. SYSTEM
A. SYSTEM INTRODUCTION

We used trained deep learning models and other algorithms to build our end-to-end system. In terms of the intelligent trash can, we use Rapsbery Pi to carry our lightweight model, identify the throwing objects thrown into the trash can turntable, and transmit the identified signals to the hardware control turntable to put the throwing objects into the corresponding trash box, and at the same time transmit the relevant information to the data base. On the Android side, we pulled the Android sample application of TensorFlow Lite for modification and added our parts of the program. It can continuously classify the frames read from the device's rear camera, perform inference using TensorFlow Lite Java API, classify the frames in real time, and display the most likely classification. It allows users to choose between floating-point models or quantized models, choose the number of threads, and decide whether to run on a CPU, GPU, or NNAPI. It greatly reduces the degree of dependence on hardware and guarantees its user customization. We can enhance the data volume of the model through user feedback, and then train it offline or in real time.

B. ADJUSTMENT MECHANISM

The entire system consists of the platform, data, and model adjustment mechanisms. Considering that in different application environments, the implementation of model retraining will be affected by the imbalance of new data set categories, the following is suggested.

The user first uploads pictures in the terminal device. Based on the type of model inference running on the terminal device, the user checks whether the inference prediction result is correct. If it is incorrect, the picture is stored as correct. If it is correct, a judgement is made as to whether the probability of its prediction pair exceeds a limited threshold; if it is exceeded, the changed picture is discarded, otherwise store the picture and label. After the storage reaches the limit, the two parts of the data are stored in Balancer for the final data loss weighted balance calculation. This part
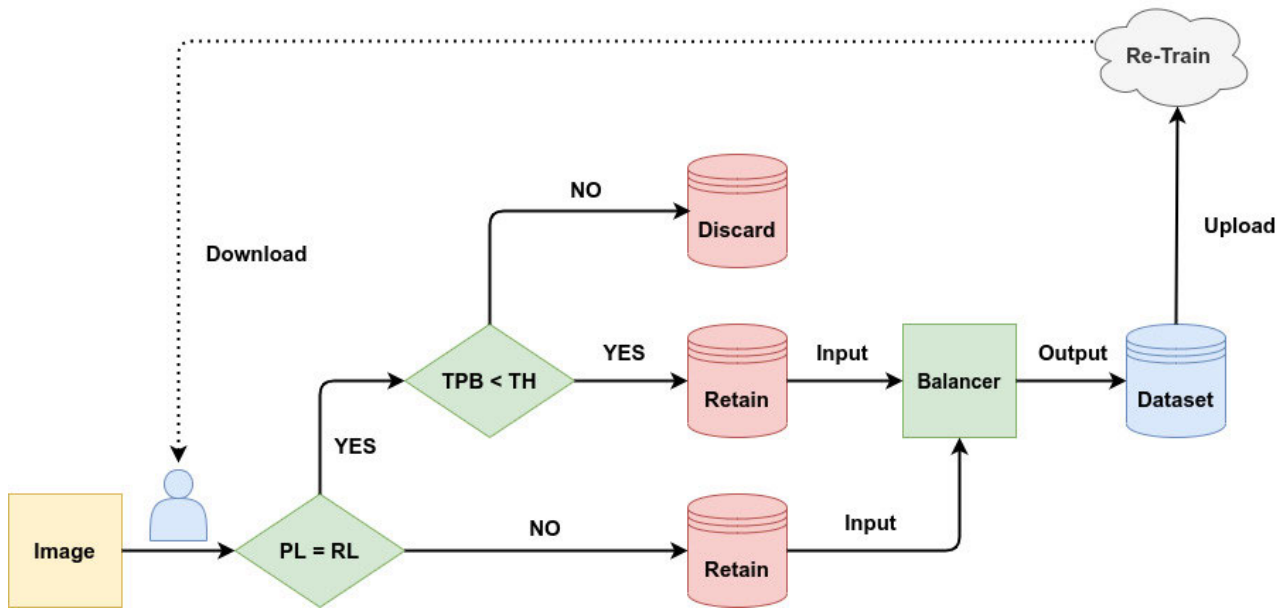
FIGURE 4. PL is the predicted value of the model, RL is the real value, TPB is the probability value of its accurate prediction when the prediction is correct, TH is the threshold value of the lowest probability value when the prediction is correct, Balancer is the data integrator, Re-Train is uploading the cloud or in the local model events in training.

FIGURE 5. Schematic diagram of the overall architecture of the system.

of the weighted calculation is composed of two parts: the prediction is incorrectly given a weight A, and the prediction is incorrectly given a weight 1-A. The method mentioned in Section IV-B is then used to perform a weight loss calculation to form the final data. Encode and upload to the cloud when required by the user or within a limited time period. After retraining in the cloud, bring it back to the terminal and resume work. The whole process is shown as a flow diagram in Fig.4.

C. PROJECT ARCHITECTURE

One of the main meanings of garbage collection is the reuse and treatment after garbage collection. In order to make the recycling and processing of garbage recycling more efficient, we need to grasp the data on the types of garbage,

the distribution of various types of garbage in various regions, and the quantity and rate of change of various types of overall garbage. The smart trash bin we designed can intelligently recycle trash, and the mobile phone application can help users correctly distinguish the types of trash and understand the latest news and policies on the classification of trash. These two parts ensure the treatment of trash at the source. At the same time, we need regional, city managers and decision makers to provide further measures to solve the problem after garbage collection.

Our platform collects messages sent from various local smart trash cans and data generated when people use mobile applications, integrates data, performs further data cleaning, and finally displays the required data on our platform. The entire project structure is as follows Fig.5 shows.

Our platform provides managers and decision makers with a clear data visualization, which is convenient for managers and decision makers to further process garbage after garbage collection, and can quantify the cost, economic benefits, and environmental costs of each treatment process. At the same time, it also provides data protection for managers and policy makers to formulate new civic codes, policies and laws.

VI. CONCLUSION

Although this research proposes a new lightweight neural network that has made a positive contribution to the relevant knowledge system of garbage classification and recycling, we can't just rely on this system. There are many other challenges. The first research direction is to solve the problem of sorting and recycling garbage in dumps in some underdeveloped areas. A special robotic arm can identify and locate garbage, and then sort and recycle it. The second research direction is to develop a robot that can automatically identify and recycle garbage that is discarded on the ground of streets, parks, schools, etc. It can reduce the pressure on sanitation workers and keep the city clean. The third research direction is the problem of multi-label garbage classification. As is common in real life, most of the garbage is mixed together. How to solve the problem of intelligent garbage sorting and recycling in this case is the key. There is also a need to pay more attention to environmental protection and pollution issues, and constant innovation is needed to develop more practical systems and tools to solve problems related to garbage classification and recycling.

ACKNOWLEDGMENT

REFERENCES

[1] S. Kaza, L. C. Yao, P. Bhada-Tata, and F. V. Woerden, "What a waste 2.0," *World Bank Publications*, vol. 30317. Washington, DC, USA: The World Bank, 2018. [Online]. Available: https://ideas.repec.org/b/wbk/wbpubs/30317.html

[2] H. Cloud. (2019). *Huawei Cloud's Open Source Garbage Classification Dataset*. [Online]. Available: https://ideas.repec.org/b/wbk/wbpubs/30317.html

[3] S. Mirchandani, S. Wadhwa, P. Wadhwa, and R. Joseph, "[IEEE 2017 international conference on big data, IoT and data science (bid) - Pune, India (2017.12.20-2017.12.22)] 2017 international conference on big data, IoT and data science (bid)–IoT enabled dustbins," Tech. Rep., 2017, pp. 73–76.

[4] S. Dugdhe, P. Shelar, S. Jire, and A. Apte, "Efficient waste collection system," in *Proc. Int. Conf. Internet Things Appl. (IOTA)*, Jan. 2016, pp. 143–147.

[5] A. C. Karaca, A. Erturk, M. K. Gullu, M. Elmas, and S. Erturk, "Automatic waste sorting using shortwave infrared hyperspectral imaging system," in *Proc. 5th Workshop Hyperspectral Image Signal Process., Evol. Remote Sens. (WHISPERS)*, Jun. 2013, pp. 1–4.

[6] N. S. Kumar, B. Vuayalakshmi, R. J. Prarthana, and A. Shankar, "IOT based smart garbage alert system using arduino UNO," in *Proc. IEEE Region Conf. (TENCON)*, Nov. 2016, pp. 1028–1034.

[7] J. Joshi, J. Reddy, P. Reddy, A. Agarwal, R. Agarwal, A. Bagga, and A. Bhargava, "Cloud computing based smart garbage monitoring system," in *Proc. 3rd Int. Conf. Electron. Design (ICED)*, Aug. 2016, pp. 70–75.

[8] S. A. Ghadage and N. A. Doshi, "IoT based garbage management (monitor and acknowledgment) system: A review," in *Proc. Int. Conf. Intell. Sustain. Syst. (ICISS)*, Dec. 2017, pp. 642–644.

[9] A. Sharanya, U. Harika, N. Sriya, and S. Kochuvila, "Automatic waste segregator," in *Proc. Int. Conf. Adv. Comput., Commun. Informat. (ICACCI)*, Sep. 2017, pp. 1313–1319.

[10] A. S. Wijaya, Z. Zainuddin, and M. Niswar, "Design a smart waste bin for smart waste management," in *Proc. 5th Int. Conf. Instrum., Control, Autom. (ICA)*, Aug. 2017, pp. 62–66.

[11] D. Nirmala, G. Bala, S. Kaviya, and S. Mary, "Smart waste management using Internet-of-Things (IoT)," *Int. J. Pure Appl. Math.*, vol. 119, pp. 203–205, 01 2018.

[12] M. Yang and G. Thung, "Classification of trash for recyclability status," CS229 Project Tech. Rep., 2016.

[13] U. Ozkaya and L. Seyfi, "Fine-tuning models comparisons on garbage classification for recyclability," 2019, *arXiv:1908.04393*. [Online]. Available: http://arxiv.org/abs/1908.04393

[14] M. Yang. (2019). *Trashnet Dataset*. [Online]. Available: https://github.com/garythung/trashnet

[15] J. Toutouh, D. Rossit, and S. Nesmachnow, "Soft computing methods for multiobjective location of garbage accumulation points in smart cities," *Ann. Math. Artif. Intell.*, vol. 88, nos. 1–3, pp. 105–131, Mar. 2020.

[16] Z. Zhong, L. Zheng, G. Kang, S. Li, and Y. Yang, "Random erasing data augmentation," 2017, *arXiv:1708.04896*. [Online]. Available: http://arxiv.org/abs/1708.04896

[17] T. He, Z. Zhang, H. Zhang, Z. Zhang, J. Xie, and M. Li, "Bag of tricks for image classification with convolutional neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 558–567.

[18] S. Yun, D. Han, S. Chun, S. J. Oh, Y. Yoo, and J. Choe, "CutMix: Regularization strategy to train strong classifiers with localizable features," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 6023–6032.

[19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[20] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

[21] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1251–1258.

[22] N. Ma, X. Zhang, H.-T. Zheng, and J. Sun, "Shufflenet v2: Practical guidelines for efficient cnn architecture design," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 116–131.

[23] A. Howard, M. Sandler, B. Chen, W. Wang, L.-C. Chen, M. Tan, G. Chu, V. Vasudevan, Y. Zhu, R. Pang, H. Adam, and Q. Le, "Searching for MobileNetV3," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 1314–1324.

[24] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141.

[25] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, *arXiv:1502.03167*. [Online]. Available: http://arxiv.org/abs/1502.03167

[26] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Sep. 2018, pp. 3–19.

[27] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, Jan. 2014, pp. 3320–3328.

[28] O. Awe, R. Mengistu, and V. Sreedhar, "Smart trash net: Waste localization and classification," 2017, *arXiv Preprint*.

[29] T. Kennedy, "OscarNet: Using transfer learning to classify disposable waste," Deep Learn., Stanford Univ., Stanford, CA, USA, Tech. Rep. CS230, 2018.

[30] V. Ruiz, A. Sánchez, J. F. Vélez, and B. Raducanu, "Automatic image-based waste classification," in *Proc. Int. Work-Conf. Interplay Between Natural Artif. Comput.* Springer, 2019, pp. 422–431.

ZHIHU YANG was born in Gansu, China, in 1998. He is currently pursuing the degree in computer science and technology with Northeast Forestry University.

His work directions are in computer vision and smart city. He received the Hand in Hand to Enhance Scholarship of Northeast Forecast University, the Excellent Award of Northeast Competition Area of Image Group of Future Cup AI Challenge, the Third Place in Heilongjiang Province of China University Student Computer Design Competition, the Third Prize in The Final of The Second National AI Innovation Competition, and the Fifth National Cloud Recognition Machine Image Competition by China Meterological Administration.

DAN LI is currently pursuing the Doctor of Engineering degree with the School of Information and Computer Engineering. He is also an Associate Professor and the Vice President with the School of Information and Computer Engineering. He is also a Visiting Scholar with the University of Georgia. He is also the Director of the Computer Branch, China Forestry Society. He is also the Director of the National Forestry and Grassland Internet of Things and Artificial Intelligence Application Technology Innovation Alliance. He has presided over seven projects, including the National Spark Plan, the National Forestry Administration Public Welfare Industry Special Project, the Heilongjiang Science and Technology Research Project, and the Heilongjiang Natural Science Foundation. He has published two teaching materials and more than 20 SCI and EI papers. His research interests include forest system modeling, forestry big data, and the forestry Internet of Things. He served as a Reviewer of articles in multiple journals.

● ● ●