The Classification of Garbage Based on Chinese Garbage Polity

Chenyang Wen, Chenyu Tang, Jiacheng Guo

Abstract

In this project, we want to develop a model that can classify garbage photos according to Shanghai garbage classification standards. For those new to Shanghai, such as us new NYU students, facing various of garbage, it is more troublesome to frequently check the standard. We choose three neural networks to build our model, CNN, WasteCNN and Res18. And Res18 has the best performance. And the accuracy of Res18 model reach to 90%.

Introduction

In Shanghai, garbage classification has been implemented for more than 20 years. Meanwhile, the classification standard has changed many times. Published on July 1, 2019, Shanghai Municipal Household Garbage Management Regulations divides the classification standard into four categories: recyclable waste, residual waste, household waste and hazardous waste. The strict implementation with high quality of garbage classification cannot be separated from the accurate classification by every citizen in Shanghai. In this project, we want to develop a framework to automatic garbage categorization.

Similar to how people identify garbage categories, we want our framework to be able to output garbage categories when it sees the garbage (i.e., the input image). The first step in this process requires the ability to recognize the content of the input image. In the field of image identification, machine learning, especially convolutional neural networks [2], has huge advantages. After get the preprocessed image, our method needs to classify it.

Dataset

In this project, dataset is images about garbage. We found two original datasets from garbage classification topic in the Kaggle platform [3,4]. The two datasets contain thousands of pictures of more than a dozen different types of garbage. We reassign four labels to the source dataset: recyclable trash, dry trash, wet trash and poison trash. For example, scrap metal and waste plastic relabeled as recyclable trash. In this way, the dataset that meets the requirements of

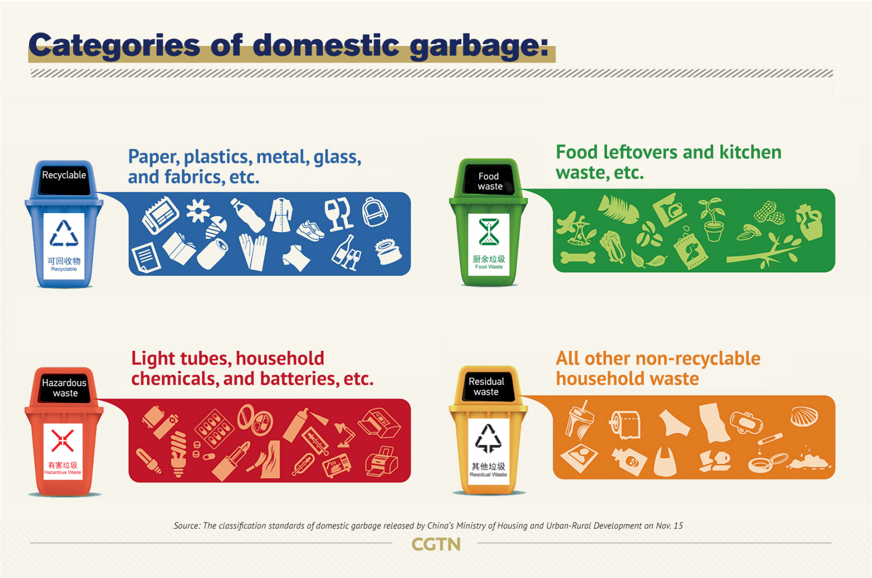

Categories of domestic garbage:
Recyclable
可回收物
Recyclable
Paper, plastics, metal, glass, and fabrics, etc.
Food waste
厨余垃圾
Food Waste
Food leftovers and kitchen waste, etc.
Hazardous waste
有害垃圾
Hazardous Waste
Light tubes, household chemicals, and batteries, etc.
Residual waste
其他垃圾
Residual Waste
All other non-recyclable household waste
Source: The classification standards of domestic garbage released by China's Ministry of Housing and Urban-Rural Development on Nov. 15
CGTN

Figure 1: Standard for waste classification from CGTN [1]

the project is obtained, which contains thousands of images and four labeled garbage types.

Solution

Data Pre-processing

For the ideal training efficiency and preventing over-fitting of the model, the following steps are used to preprocess the original data.

Image Resize

Original images have different resolution. We resize them into 128 * 128 pixel. It ensures a relatively fast training efficiency.

Image Normalize

In our model, input images are treated as matrix of the value for 3 color channels, R, G and B. Therefore, we need to normalize these values to a normal distribution N(0, 1).

Random Transform

In this project, various of image transformations are prepared. Each image will randomly select one of the transformations to prevent the model from over-fitting and enhance the anti-interference ability of the model. The following transformations are prepared.

Original	Flip horizontal	Flip vertical	Rotation randomly	Gaussian blur	Gray

Table 1: Value of precision, recall and F1-score

Data Split

After random transform, well-processed data is divided into training set, validation set and test set. The proportions of the three are respectively 72%, 18% and 10%. Then these three sets are read into the model in the format of Dataloader class, for the following model training and evaluation.

Neural Network & Model Selection

Neural network has already taken a place for many years. Based on the practical experience and experiment result, it performed pretty well, even better than the traditional machine learning algorithm. And the convolutional neural network is one of the main categories to do computer vision tasks both in recognition and classification. Since the project is based on the image classification, the most common way to classify the image for now is to use deep neural network. And this is also the most mature way in computer vision, so we here specify three different network for training. The basic CNN with 2 convolution layers and 3 fully connected layers that for garbage classification tasks is as the base line for the improvement of other models. And compare with the basic CNN, wasteCNN add one convolutional layer addition for improvement; and another neural network is Res18, inspired by the original ResNet [5], cascading with previous layers. This kind of skip connection let the network able to preserve the feature from previous layers, make the network much easier to preserve the gradient.

Models Construction & Training

And as for the model chosen process, if the model is much deeper, then the time consuming for training process will be much longer, so the res18 is the deepest network we choose. It is true that the Res50 will have much deeper network and similar amount of parameter with the network that applied with ResNet. So it is also the network we want to realize by ourselves, but these are sufficient models to classify the garbage by these models and choose the best one. As for network parameter, all the convolutional kernels including the first layer of

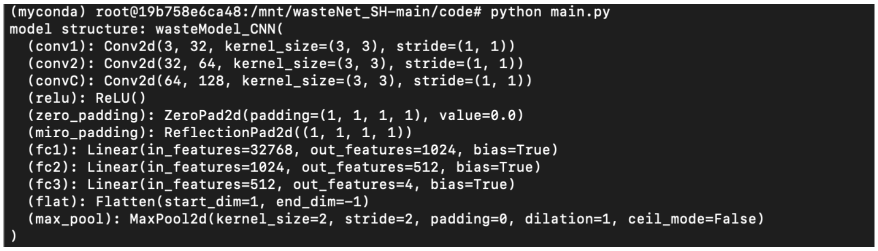
Res18 are using 3×3 large to reduce the parameter number, stride set default as 1 except some layers within Res18 for downsampling.

CNN

This is the most basic model applied in the project. The model's input, which is the image size is 128. And adjust the image size will also affect the model's first fully connected layer's neuron's number. Since the structure is similar as wasteCNN, so we won't specific shown the structure figure here.

WasteCNN

The model structure is shown in Figure 2 with the image size is 128, adjusting the image size will affect the model's first fully connected layer. Basic backbone is as same as CNN. The difference is that there is 3 convolutional layers and larger fully connected layer.



```
(myconda) root@19b758e6ca48:/mnt/wasteNet_SH-main/code# python main.py
model structure: wasteModel_CNN(
  (conv1): Conv2d(3, 32, kernel_size=(3, 3), stride=(1, 1))
  (conv2): Conv2d(32, 64, kernel_size=(3, 3), stride=(1, 1))
  (convC): Conv2d(64, 128, kernel_size=(3, 3), stride=(1, 1))
  (relu): ReLU()
  (zero_padding): ZeroPad2d(padding=(1, 1, 1, 1), value=0.0)
  (miro_padding): ReflectionPad2d((1, 1, 1, 1))
  (fc1): Linear(in_features=32768, out_features=1024, bias=True)
  (fc2): Linear(in_features=1024, out_features=512, bias=True)
  (fc3): Linear(in_features=512, out_features=4, bias=True)
  (flat): Flatten(start_dim=1, end_dim=-1)
  (max_pool): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
)
```

Figure 2: wasteCNN model structure

Res18

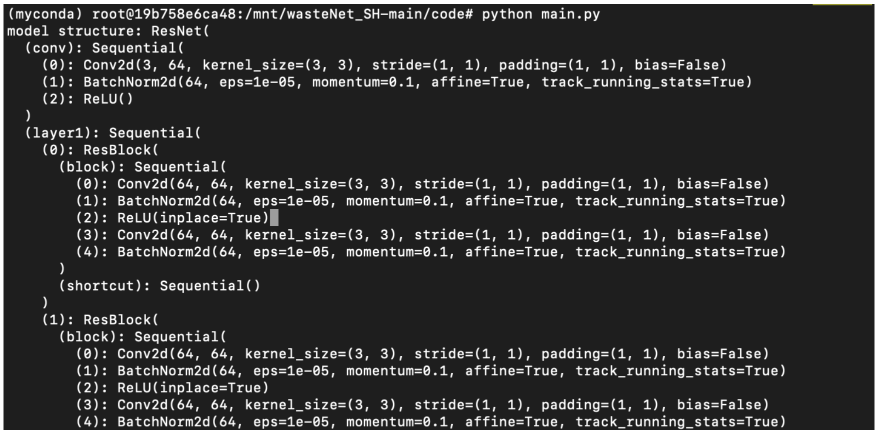
Res18 has 18 convolutional layers. First thing is to construct the ResBlock, which is the basic component for the ResNet. ResBlock is built by two convolutional layers and a shortcut, which represent the combination of current Conv layer output and previous output. Then connect with 4 blocks and addition 1 Conv layer in the front and 1 fully connected layer at the end. And the part of ResNet is also shown as Figure 3.

Loss Function

Since the project is classification problem, so the cross-entropy loss is chosen as the loss function. And here is the loss function.

$$J(w) = -\Sigma_{i=1}^{N}\{y_i logp(x_i; w) + (1 - y_i)log(1 - p(x_i; w))\} \tag{1}$$

Other sets of training are: 'batch size': 32, 'validation data's percentage': 0.2, 'learning rate': 1e-3, 'epoch': 50, and if the code is running on GPU is much quicker.

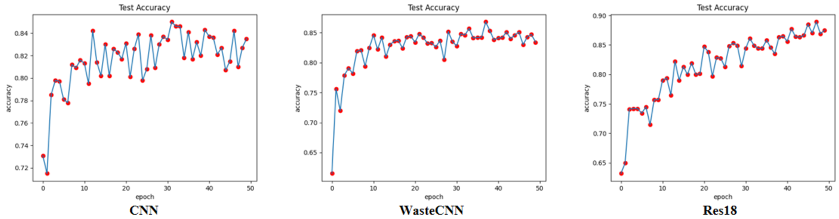


```
(myconda) root@19b758e6ca48:/mnt/wasteNet_SH-main/code# python main.py
model structure: ResNet(
  (conv): Sequential(
    (0): Conv2d(3, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
    (1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (2): ReLU()
  )
  (layer1): Sequential(
    (0): ResBlock(
      (block): Sequential(
        (0): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (2): ReLU(inplace=True)
        (3): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      )
      (shortcut): Sequential()
    )
    (1): ResBlock(
      (block): Sequential(
        (0): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (2): ReLU(inplace=True)
        (3): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
```

Figure 3: Part of ResNet with 18 convolutional layer

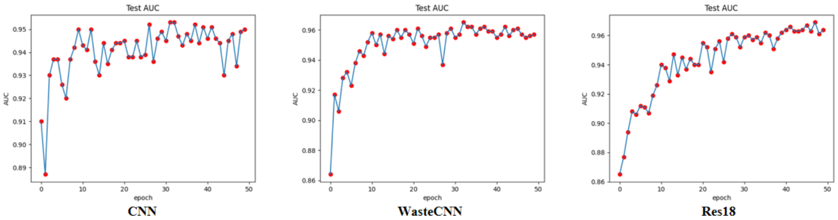
Results and Discussion

The project separately evaluates these three kinds of models by using many evaluation methods and obtained the following figures and results: the variation curve of Test Accuracy with the number of epochs shown as Figure 4. And the variation curve of AUC value with the number of epochs as Figure 5.



CNN WasteCNN Res18

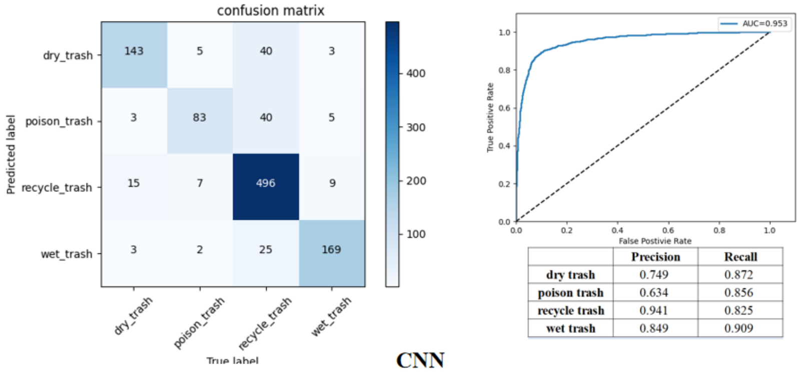
Figure 4: Accuracy-Epoch Curve



CNN WasteCNN Res18

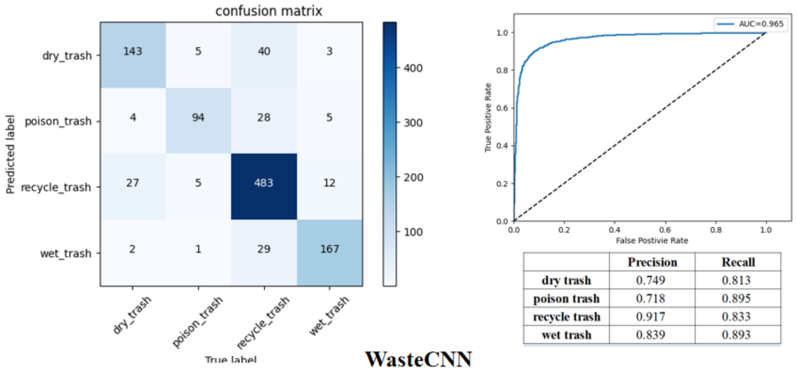
Figure 5: AUC value-Epoch Curve

According to figures 4 and 5, the project selects the iteration that performs better in the test dataset and get the confusion matrix and ROC curve (after averaging the four categories) under that iteration. Then calculate the Precision, Recall for each category.



	Precision	Recall
dry trash	0.749	0.872
poison trash	0.634	0.856
recycle trash	0.941	0.825
wet trash	0.849	0.909

Figure 6: Confusion metrics and ROC curves of CNN

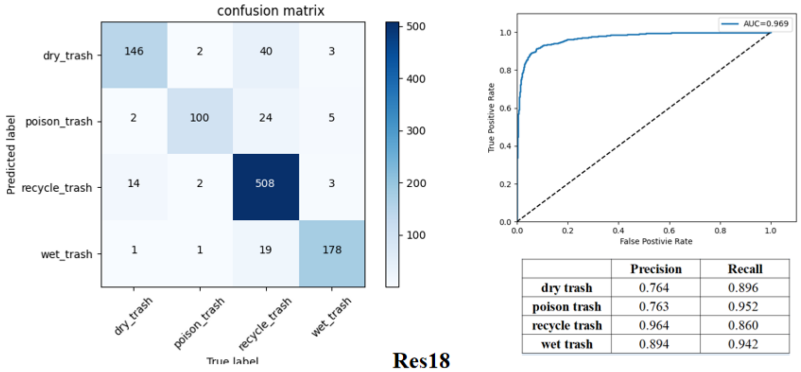


	Precision	Recall
dry trash	0.749	0.813
poison trash	0.718	0.895
recycle trash	0.917	0.833
wet trash	0.839	0.893

Figure 7: Confusion metrics and ROC curves of WasteCNN

Due to the uneven distribution of data in the test set, the project decides to calculate the macro average and micro average of precision, recall and F1-score of the three models respectively to compare the performance. The work result is shown at Table 2.

These results shows that Res18 model has the highest accuracy when predicting the test dataset, accuracy is up to 90% within 50 iterations, while the accuracy of CNN and wasteCNN models is about 83% and 85%. In addition, the average Precision, Recall and F1-score values of the Res18 model are also the closest to "1". The Precision, Recall and F1-score calculated by macro and micro average methods are basically the same, indicating the accuracy of classification is not affected by the number of samples in different categories. The AUC value of



Predicted label \ True label	dry_trash	poison_trash	recycle_trash	wet_trash
dry_trash	146	2	40	3
poison_trash	2	100	24	5
recycle_trash	14	2	508	3
wet_trash	1	1	19	178

	Precision	Recall
dry trash	0.764	0.896
poison trash	0.763	0.952
recycle trash	0.964	0.860
wet trash	0.894	0.942

Figure 8: Confusion metrics and ROC curves of Res18

Macro average			
	Precision	Recall	F1-score
CNN	0.793	0.865	0.828
WasteCNN	0.805	0.858	0.831
Res18	0.847	0.912	0.878

Micro average			
	Precision	Recall	F1-score
CNN	0.850	0.850	0.850
WasteCNN	0.846	0.846	0.846
Res18	0.889	0.889	0.889

Table 2: Value of precision, recall and F1-score

Res18 model is up to 0.97 within 50 iterations, which is also higher than in the wasteCNN and CNN models. Hence, it can be seen that Res18 model is the most accurate in the classification of garbage pictures among these three models. In the case of calculating the values of precision and recall by category of garbage, there is little difference in the value of recall for different categories in three models, but the value of precision for recyclable trash, which is up to 0.96 in Res18 model, is 0.2 more than that for dry trash and poison trash categories. This problem needs further study.

References

[1] "Explainer: China wastes no time in waste sorting - CGTN." [Online]. Available: https://news.cgtn.com/news/2019-12-12/Explainer-China-wastes-no-time-in-waste-sorting-MmLjlsvvP2/index.html

[2] Y. LeCun, Y. Bengio *et al.*, "Convolutional networks for images, speech, and time series," *The handbook of brain theory and neural networks*, vol. 3361,

no. 10, p. 1995, 1995.

[3] "Garbage Classification — Kaggle." [Online]. Available: https://www.kaggle.com/asdasdasasdas/garbage-classification/tasks?taskId=1283

[4] "Garbage Classification (12 classes) — Kaggle." [Online]. Available: https://www.kaggle.com/mostafaabla/garbage-classification/tasks?taskId=3361

[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.