Clean Our City: An Automatic Urban Garbage Classification Algorithm Using Computer Vision and Transfer Learning Technologies

Yuke Chen\textsuperscript{1, a\textsuperscript{†}, Wendi Han\textsuperscript{2, b\textsuperscript{†}, Jizhi Jin\textsuperscript{3, c\textsuperscript{†}, Hao Wang\textsuperscript{4, d\textsuperscript{†}, Qiaochu Xing\textsuperscript{5, e\textsuperscript{†}, Yiwen Zhang\textsuperscript{6, f\textsuperscript{†}}

\textsuperscript{1}Wuhan University of Technology, Wuhan 430000, Hubei province, China, \\
\textsuperscript{2}University of Notthingham, Triumph Road, Nottingham NG7 2TU, United Kingdom, \\
\textsuperscript{3}University of Liverpool, Liverpool L69 3BX, United Kingdom, \\
\textsuperscript{4}Dalian University of Technology, Dalian 116000, Liaoning province, China, \\
\textsuperscript{5}King’s college London, Strand, London WC2R 2LS, United Kingdom, \\
\textsuperscript{6}Tianjin University, Tianjin 300000, Tianjin City China

\textsuperscript{a}chenyuke@whut.edu.cn, \textsuperscript{b}wendihan493@outlook.com, \textsuperscript{c}1428649489@qq.com, \textsuperscript{d}1270388560@qq.com, \textsuperscript{e}1339810989@qq.com, \textsuperscript{f}Yiwenzhang1019@163.com

\textsuperscript{†}These authors contributed equally.

Abstract. To improve the quality of human life in the city, the first thing to solve is the problem of urban garbage. So far, the best way to solve this problem is garbage classification. At present, many algorithms have been put forward one after another. Previous research proposed some computer vision systems to solve the problem of urban garbage classification. In recent years, with the development of computer hardware and large-scale data sets, the algorithm based on depth learning has shown superior performance in the field of image classification. Thus, the features designed by traditional methods are gradually replaced, which far exceed these traditional image classification algorithms in classification accuracy. This study proposes an algorithm based on InceptionV3 networks and test the algorithm on a large-scale garbage classification data-set. The data set was divided into 80% training sets, 10% validation set, and 10% test set and use the transfer learning approach. The model achieved an accuracy of 93.125\%, which solved image garbage classification very well. What is more, the algorithm can play an important role in the medical area and help control the mechanical arm.

1. Introduction
With the accelerating economic growth, the amount of city garbage has also increased at a prodigious speed and even threatened the further development in the economy especially in the major and medium cities. As one of the largest developing countries, China, occupied with the largest population globally, there is a total of 300 kg household garbage generated individually per year [1]. And in 2019, the garbage in 202 major and medium cities weighs a total of 235.602 billion tons in China [2]. More seriously, the figure is increasing at a rate of 8\%-10\% per year [3]. Nowadays, in China, landfilling and incineration
are the most common approaches to processing garbage. However, both of them have serious shortcomings: landfilling would contaminate surface water, environment and even produce toxification, and incineration would produce air pollution such as some decomposed heavy metal and threaten soil resources [4, 5]. Therefore, China’s government has called the public to implement garbage classification. Shanghai, as the most developed city in China, has conducted several related actions and set a series of rules [6]. For example, there are a total 4 categories of garbage when classified: dry garbage, wet garbage, recyclable garbage, and hazardous garbage [7]. The benefits of garbage classification are wide-ranging: helping to protect water and land resources, reduction in costs in garbage processing, generating power for electricity, promoting recycling, and accelerating economic development [8]. For example, the amount of daily dry recycled garbage reaches about 5605 tons in 2019, increased by 5 times as 2018, and the amount of wet recycled garbage reaches about 9008 tons in 2019, increased by 1.3 times as 2018 after implementing garbage classification[1]. Nevertheless, this task is not only tired but also time-consumed greatly. Hence, people urgently need to find a new method that can replace human classifying work and improve the accuracy of classification.

Nowadays, in most areas of China, residents still do not need to do garbage sorting by themselves. Instead, they throw garbage together in the trash can, which is recycled by special staff who manually sort them at the garbage disposal station. This is a waste of time and manpower for the government. However, although this approach has invested a lot of money and manpower, it has not achieved very good results. The living garbage has dispersive, complexity, the harm of the secret [9]. Currently, garbage classification is particularly important. In 2019, key cities in China, such as Shanghai, began experimenting with garbage sorting policies [10]. This is indeed good for the environment, but it also exposes many problems. For example, many people do not know how to classify their household garbage when the new policy publishes. Many government reports are also put forward: “Internet + garbage classification model mainly depends on the advanced Internet, cloud computing and big data frontier technology, implementation of the new model, not only can promote the efficiency of garbage classification, still can make garbage classification information more transparent” [11]. Since the neural network can recognize the digital image system, can it also be applied to household waste classification? If this kind of intelligent assistance can improve the accuracy rate to a very high value, it will help people achieve garbage classification.

In recent years, convolutional neural networks (CNNs) have shown remarkable successes in image classifications. No combination of feature maps (Ncfm) is applied for better performance using MNIST datasets, and its network results in a 0.24% error rate [12]. Under the premise of good performance in image classification using CNN, the garbage classification problem could be regarded as an image classification problem and use a CNN architecture to identify the garbage category. Previous researchers focus on recognize wastes made of special materials. Polyth-Net uses deep learning to distinguish polythene bags from other wastes, and another model uses R-CNN to classify the waste as biodegradable and non-biodegradable [13, 14]. These methods can help garbage dump segregate special material from others and dispose of it separately to recycle. However, one shortcoming is that instead of classifying the waste into different categories, they can only decide whether the garbage is made of special materials or not. WasteNet uses transfer learning to solve this problem [15]. They use VGG-16 as a base model to train, and the model achieves a high accuracy of 97%. The dataset in their model contains six categories with a pure white background: plastic, paper, glass, metal, cardboard, and others. This paper will focus on distinguishing 12 different types of garbage, and most of the images in the dataset have a messy background that is in line with daily life.

So, a highly efficient deep neural network model is to be designed to do waste classifications using photos from cellphones’ cameras. And this model is expected have high accuracy and high efficiency even if the background of the images is complex. This model can help the residents who are not familiar with how to classify the waste, improving waste management efficiency. This will help reduce the cities’ burden on waste management and protect the environment [16]. This paper presents a deep neural network model of high accuracy and high efficiency to do waste classification in complex image backgrounds. Our model achieves a state-of-the-art accuracy score of 91% on an open dataset from the
Kaggle website [18]. A recently published automatic image-based waste classification model WasteNet has shown an accuracy of 97% on the TrashNet dataset [15]. This is also a large improvement on the CNN approach that achieved an accuracy of 22% [17]. A recently published automatic polythene bags classification model Polyth-Net has gained an accuracy of 95.52% [13]. Moreover, this model has shown a good performance even if the image has a complex background, shown in part 2. Our training database is collected by ourselves, which consist of 12 common waste in the daily life. The background of the images is complex. It makes the model fit the daily life application, which the previous models cannot do.

2. Methods

2.1. Data Collection

The dataset is a open dataset from the Kaggle website [18]. It contains 12 classes of waste images, battery, biological, brown-glass, cardboard, clothes, green-glass, metal, paper, plastic, shoes, trash, and white glass. There are 945 battery images, 985 biological images, 607 brown glass images, 891 cardboard images, 5325 clothes images, 629 green glass images, 769 metal images, 1050 paper images, 865 plastic images, 1977 shoe images, 697 trash images, and 775 white glass images. The data is split into three parts, using 80% as training data, 10% as validation data, and 10% as test data. Some examples in our dataset are shown in figure 1.
2.2. **Data Augmentation**

To gain a better performance, Random flip, random rotation, and random zoom these data augmentation techniques are used. Some example images in the dataset and the process of data augments are shown as follows. Original images are shown in figure 2a. Random flip will flip some of the images in random. Images after random flip are shown in figure 2b Random rotation will rotate some of the images in random. Images after random rotation are shown in figure 2c Random zoom will zoom some of the images in random. Images after random zoom are shown in figure 2d.

![Image after Data Augmentation](c) Image after Random Rotation (d) Image after Random Zoom

**Figure 2. Image after Data Augmentation**

### Table 1. The Architecture of The Model

| Type                     | Number of Nodes/batch Size | Input Size     |
|--------------------------|----------------------------|----------------|
| Pre-trained InceptionV3  | 299*299*3                  | 299*299*3      |
| Fully connected layer    | 512                        | 2048           |
| Fully connected layer    | 512                        | 512            |
| Batch Normalization      | 128                        | 512            |
| Fully connected layer    | 256                        | 512            |
| Fully connected layer    | 256                        | 256            |
| Batch Normalization      | 128                        | 256            |
| Fully connected layer    | 12                         | 256            |

### Table 2. The Architecture of Inception V3

| Type       | Patch Size/stride or Remarks | Input Size     |
|------------|------------------------------|----------------|
| Conv       | 3*3/2                        | 299*299*3      |
| Conv       | 3*3/1                        | 149*149*32     |
| Conv padded| 3*3/1                        | 147*147*32     |
| Pool       | 3*3/2                        | 147*147*64     |

2.3. **Model Structure**

The architecture of our model is shown in Table 1. First, data augmentation techniques are used to process the data. A pre-trained model InceptionV3 is used in the model [19]. The pre-trained InceptionV3 helps to extract the features from the images. Then several fully connected layers are connected to learn the features from the data. The fully connected layers in the model will fit the features, which is to “learn” from the data. And there are some Batch Normalization layers to reduce the possibility of overfitting.

**Table 2. The Architecture of Inception V3**

| Type       | Patch Size/stride or Remarks | Input Size     |
|------------|------------------------------|----------------|
| Conv       | 3*3/2                        | 299*299*3      |
| Conv       | 3*3/1                        | 149*149*32     |
| Conv padded| 3*3/1                        | 147*147*32     |
| Pool       | 3*3/2                        | 147*147*64     |
2.4. Transfer Learning

A pre-trained model InceptionV3 from the TensorFlow hub is used [20]. The architecture of InceptionV3 is shown in figure 3 and Table 2 [19]. The model is pre-trained on the dataset ImageNet. The pre-trained InceptionV3 in the model is loaded and frozen. The inception can help to extract features from the images. The transfer learning technology can help to reduce the possibility of overfitting and decrease the time consumption of the training process.

![Original Inception Model](image)

**Figure 3. Original Inception Model**

2.5. Fine-tuning

First, the inception layer is frozen and the model is trained with the training dataset. After the model is trained, fine-tuning is applied to adjust the model. Freeze all the layers except the last three fully connected layers and use a 1% learning rate to train the model with the training dataset. The whole training process is shown in figure 4.

| Conv | 3*3/1 | 73*73*64 |
|------|-------|---------|
| Conv | 3*3/2 | 71*71*80 |
| Conv | 3*3/1 | 35*35*192 |
| 3*Inception | 35*35*288 |
| 5*Inception | 17*17*768 |
| 2*Inception | 8*8*1280 |
| Pool | 8*8 | 8*8*2048 |
| Linear | logits | 1*1*2048 |
| Softmax | classifier | 1*1*1000 |

| Table 3. Meaning of Abbreviation |
|----------------------------------|
| Positive | Negative |
| True     | TP       | TN       |
| False    | FP       | FN       |

2.6. Evaluation Methods

The model is expected to be a highly efficient deep neural network to do waste classifications using photos from cellphones’ cameras. And the model would have high accuracy and high efficiency even if the background of the images is complex. This model can help the residents who are not familiar with how to classify the waste, improving waste management efficiency. So, a higher precision of the prediction can help provide a more reliable classification reference. An F0.5 score is used to evaluate our model.
The meaning of TP, FP, TN, and FN is shown in Table 3, and they are used to evaluate our model. Precision is the fraction of retrieved documents that are relevant to the query, which is defined as:

\[
A_{\text{3}2} = A^{(+)} - A^{(-)}(I = \frac{3}{2})
\]  

(1)

Recall is the fraction of the relevant documents that are successfully retrieved, which is defined as:

\[
\phi_{\text{3}}(\vec{r}) = (2\pi)^{3/2} \exp\left(i\vec{k} \cdot \vec{r}\right)
\]  

(2)

Then, use the value of precision and recall calculating F0.5, which is defined as:

\[
\phi_{\text{3}}(\vec{r}) = (2\pi)^{3/2} \exp\left(i\vec{k} \cdot \vec{r}\right)
\]  

(3)

3. Result and Discussion

3.1. Performance Evaluation

| Dataset  | Accuracy | Precision | Recall | F0.5  |
|----------|----------|-----------|--------|-------|
| Training | 0.9585   | -         | -      | -     |
| Validation | 0.9077  | -         | -      | -     |
| Test     | 0.9313   | 0.9666    | 0.9970 | 0.9725|

Our model performed reasonably well on the validation dataset, and Table 4 shows the results for our classification method. In this experiment, the model achieved an accuracy of 93.1% on the test dataset and an f0.5-score of 0.97, which is the combination of the precision and recall values.

The blue line and the orange line in figure 5 show the accuracy and loss change over epochs at the training stage. Early stopping certainly can prevent our model from overfitting as the model stops at epoch 25 instead of epoch 200, which is the value which is pre-set to train the model. At around 16 epochs, the loss value and accuracy begin to stabilize and after 23 epochs show a downward trend in accuracy and an upward trend on loss.
3.2. Classification Histogram

Figure 6 and figure 7 show the classification result for 12 garbage categories by using our model on most categories. In figure 6, the histogram on the right side of the picture shows the classification results of this garbage category. The x-axis of the histogram represents different garbage types, and the y-axis shows the result distribution of these types. In addition, the blue line represents the correct classification, and the grey line represents the incorrect classification. The figure shows that the model has achieved a good prediction on most of the categories, such as boxes and clothes. However, for some very similar garbage categories, the model still hard to accurately distinguish all of them, such as plastic and white glass.
3.3. Classification Histogram

Using the dataset loaded from Kaggle containing 12 categories of garbage, a deep learning model is built to help people finish several kinds of garbage classification. Every category of garbage has images of more than 600, and there are 5325 images, especially of clothes. These images into three datasets: training set, validation set, and test set, respectively occupying 80%, 10%, and 10%.

By combining 5 fully connected layers, 2 batch-normalization layers and transfer learning together, the model have ultimately achieved relatively high accuracy in the training set, validation set, and test set: respectively 95.85%, 90.77%, and 93.125%. Lastly, F0.5 is used to evaluate our results.

To further evaluate the results with the traditional methods and find out some strengths and weaknesses, data from similar studies are collected for comparison. In Dong’s work [21], he has achieved very high accuracy, up to 96.52%, but he uses some complex networks, and it costs so much that it not realistic to be used in daily life when classifying garbage. P. Tiyajamorn’s work [22], they have also gained a relatively high accuracy, 94%. However, their dataset is a little small and only uses 5700 images. In R. Sultana’s work [23], although reaching an acceptable accuracy, their images are almost put in pure white backgrounds resisting the reality. In comparison, our method uses a simpler network, achieving a good result. The dataset is large enough composed of 15515 images, and the background is much closer to a realistic situation.

Recent tendency shows that countries and societies have paid more and more attention to environmental protection, and some related laws have also been promulgated. However, in automatic garbage classification, it is admittedly that there is still much space for us to explore. Luckily, with the development of deep learning, more and more delicate algorithms will be designed to solve this problem in the future.
3.4. Future Work

In the future, there is still much work to be done on the study of “domestic waste classification” based on neural network, which can be considered from the following four aspects:

- Since this paper only studies the classification of household garbage, in the future, the object of garbage classification can be extended to other fields, such as industrial waste, etc. If people want to identify other kinds of garbage, they need to do data augmentation.

- From the point of view of this paper, the result performance of the convolutional neural network is optimal. However, at present, the level of this research on convolutional neural networks is still very low. As more improvements are added, the model will yield more accurate results.

- In the case that household waste is still unable to be accurately classified, it is also a good direction to use neural network recognition and mechanical arm picking in garbage treatment plants. People need to combine the robot arm pickup program and neural network to achieve this.
• Suppose the study on garbage classification is too monotonous. In that case, the excellent performance of the convolutional neural network in the field of computer vision can be used to apply image classification to different fields. For example, on the medical side. According to the picture of different cases, you can tell whether the patient is sick and to what extent. Based on this paper, such an effect can be achieved by changing the data set and adjusting the model appropriately.

4. Conclusion
China is now implementing a garbage sorting policy in many major cities. However, many urban residents still cannot accurately classify a large amount of household garbage. This research has achieved the goal of using a neural network to classify domestic waste.

It used an open dataset from the Kaggle website, and a total of 15,515 images were used for the training. Trained on a computer equipped with Intel i7-9750H as CPU and Nvidia GTX 1660 Ti as GPU, our model uses random flip, random rotation, and random zoom to gain a better performance. The transfer learning technique is also used to reduce the possibility of overfitting and the time consumption of the training process. After using the fine-tuning to adjust the model, an F0.5 score is used to evaluate this model.

In the result part, the accuracy rate is as high as 0.91. Compared to other algorithm implementations in this field, many of images in the datasets are not pure white image backgrounds. But it also worked very well. This algorithm provides a possibility for residents to classify garbage and provides a way of thinking for other computer vision. It can be used, for example, in conjunction with robotic arms or in the medical field.

In future studies, the accuracy of this model can be improved by adding feedback methods. It is necessary to apply this model to other fields for further research.

References
[1] Y. Tong, J. Liu, and S. Liu, “China is implementing ‘Garbage Classification’ action,” Environmental Pollution, vol. 259. Elsevier Ltd, p. 113707, Apr. 01, 2020, doi: 10.1016/j.envpol.2019.113707.
[2] “2020 Annual Report on the Prevention and Control of Environmental Pollution by Solid Wastes in Large and Medium Cities of China.”
[3] Y. Kuang and B. Lin, “Public participation and city sustainability: Evidence from Urban Garbage Classification in China,” Sustainable Cities and Society, vol. 67, no. September 2020, p. 102741, 2021, doi: 10.1016/j.scs.2021.102741.
[4] M. Qasim et al., “Impact of landfill garbage on insect ecology and human health,” Acta Tropica, vol. 211. Elsevier B.V., p. 105630, Nov. 01, 2020, doi: 10.1016/j.actatropica.2020.105630.
[5] Y. Liang, D. Xu, P. Feng, B. Hao, Y. Guo, and S. Wang, “Municipal sewage sludge Incineration and its air pollution control,” Journal of Cleaner Production, vol. 295. Elsevier Ltd, p. 126456, May 01, 2021, doi: 10.1016/j.jclepro.2021.126456.
[6] S. Chen et al., “Carbon emissions under different domestic waste treatment modes induced by garbage classification: Case study in pilot communities in Shanghai, China,” Science of the Total Environment, vol. 717, p. 137193, May 2020, doi: 10.1016/j.scitotenv.2020.137193.
[7] T. Limei and W. Ruiyan, “Domestic waste classification policy in Shanghai from the perspective of policy text:historical evolution and trend analysis” pp. 7–9, 2020, doi: 10.16647/j.cnki.cn15-1369/X.2020.08.003.
[8] Legislation of Shanghai requires full implementation of the whole process of monitoring garbage classification to establish an incentive mechanism. https://www.chinacourt.org/article/detail/2019/08/id/4235479.shtml. (Aug 1, 2019), Accessed 1st Aug 2018 (in Chinese).
[9] Wang D D. Study on the legal policy of China’s municipal solid waste classification [J]. Journal of Taiyuan City Vocational and Technical College, 2021(03); 164-167
[10] Tong, Y., Liu, J., & Liu, S. (2020). *China is implementing “Garbage Classification” action*. In *Environmental Pollution* (Vol. 259). Elsevier Ltd. https://doi.org/10.1016/j.envpol.2019.113707

[11] Wang Ying. *Discussion on the future development trend of municipal waste classification in Jiangsu Province* [J]. *Time-brand Marketing*, 2021(03); 53-54

[12] Yin, Y., Wu, J., & Zheng, H. (2016). *Ncfm: Accurate handwritten digits recognition using Convolutional Neural Networks*. Proceedings of the International Joint Conference on Neural Networks, 2016-October, 525–531. https://doi.org/10.1109/ICNN.2016.7727244

[13] Singh, D. (2020). *Polyth-net: Classification of polythene bags for garbage segregation using deep learning*. ArXiv. https://doi.org/10.1109/sefet48154.2021.9375766

[14] Brintha, V. P., Rekha, R., Nandhini, J., Sreekarathick, N., Ishwaryaa, B., & Rahul, R. (2020). *Automatic Classification of Solid Waste Using Deep Learning*. In Proceedings of International Conference on Artificial Intelligence, Smart Grid and Smart City Applications (pp. 881–889). Springer International Publishing. https://doi.org/10.1007/978-3-030-24051-6_83

[15] White, G., Cabrera, C., Palade, A., Li, F., & Clarke, S. (2020). *WasteNet: Waste Classification at the Edge for Smart Bins*. ArXiv

[16] HOORNWEG D, BHADA-TATA P. *What a waste: a global review of solid waste management* [J]. 2012.

[17] YANG M., & Thung, G. (n.d.). *Classification of Trash for Recyclability Status*.

[18] MOHAMED M. *Garbage Classification (12 classes)* [M]. 2021.

[19] SZEGEDY C. VANHOUCKE V. IOFFE S. et al. *Rethinking the inception architecture for computer vision: proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition*. F. 2016 [C].

[20] GOOGLE. *TF2 Feature vectors of images with Inception V3 trained on ImageNet (ILSVRC-2012-CLS)*. [J]. 2021.

[21] Dong Ziyuan. *Design and implementation of garbage classification system based on deep learning* [D]. University of Chinese Academy of Sciences (Shenyang Institute of Computing Technology, Chinese Academy of Sciences), 2020.

[22] P. Tiyajamorn, P. Lorprasertkul, R. Assabumrungrat, W. Poomarin, and R. Chancharoen, “*Automatic Trash Classification using Convolutional Neural Network Machine Learning*,” Proceedings of the IEEE 2019 9th International Conference on Cybernetics and Intelligent Systems and Robotics, Automation and Mechatronics, CIS and RAM 2019, pp. 71–76, 2019, doi: 10.1109/CIS-ARAM.2019.9095775.

[23] R. Sultana, R. D. Adams, Y. Yan, P. M. Yanik, and M. L. Tanaka, “*Trash and Recycled Material Identification using Convolutional Neural Networks (CNN)*,” Conference Proceedings - IEEE SOUTHEASTCON, vol. 2020-March, 2020, doi: 10.1109/SouthEastCon44009.2020.9249739.