IMAGE OF OUTDATED ITEMS CLASSIFICATION ALGORITHM BASED ON CONVOLUTIONAL NEURAL NETWORK.

Sharofiddin Allaberdiev[1*], Lochin Valiev[2*], Shakhriddin Allaberdiev[3], Husan Annakhulov[4]
[1*]Department of Electrical and Computer Engineering, Ajou University in Tashkent
[2*]Tashkent Branch of Nuclear Research National University MEPhI
[3]Tashkent State Transport University
[4]Termez Branch of Tashkent State Agrarian University
E-mail: [1*]Sh.Allaberdiev@mail.ru, [2*]valiyevl33@gmail.ru, [3]shahriddin1999@mail.ru

Abstract

One of the significant issues is resource recycling of outdated objects classification. It can effectively improve the efficiency of resource recycling and further reduce the harm caused by environmental pollution. By the gradual intellectualization of modern industries, traditional image classification algorithms no longer proper the requirements of garbage classification because there are lots of requirements for sorting equipment. This paper proposes to build outdated items' Classification Network "GCNet" based on a convolutional neural network. By constructing a realization mechanism, the model completes local and global feature extraction. It can complete productive feature information obtained. At the same time, through the feature combination mechanism, it's of different levels and sizes are fused to make more effective use of properties and avoid gradient disappearance. Experimental results prove that "GCNet" has achieved promising results on related outdated items' classification data sets. It has received an improvement in image identification of old-established items.

Keywords: Outdated items' image classification, convolutional neural network, image classification, realization mechanism, feature combination.

1. Introduction.

As the only way to develop a circular economy, garbage recycling is the key to eradicating pollution and improving the effectiveness of environmental governance. With the development of Uzbekistan's productivity level, domestic and industrial waste continues to increase. According to statistics, the waste covers an area of about 1,600 hectares in Uzbekistan and is "buried" in about 221 landfills. Currently, more than 80 million tons of squanders are disposed of in these areas too. Transport used to deliver the waste treatment plant after people put the garbage into the garbage bins. The current household waste treatment' collection is sorting by people, which is harmful to the health of the workers and the sorting efficiency is low. It can no longer meet a large amount of garbage disposal needs. In addition, the types of garbage are manually sorted and extremely limited the waste cannot be recycled, according to the result of thrown waste. With the development of deep learning technology, convolutional neural networks have greatly improved the accuracy and speed of image classification algorithms, allowing us to see the possibility of automatically sorting garbage with the help of vision technology. Shooting outdated items through a camera Picture, use a convolutional neural network to detect the category of waste, and then automatically complete the sorting task with the help of manipulators or push plates, which can reduce labor costs and improve sorting efficiency **[1]**. Therefore, it is significant to research garbage image classification algorithms' value.

2. Related work.

In the early days, scholars could only use classic image classification algorithms **[2–5]** to complete garbage image classification tasks by manually extracting image features and combining them with corresponding classifiers. Wu Jian et al. **[6]** used color and texture features to complete the waste and garbage identification. Due to the different data sets as the background, size, and quality are not the same. Based on the traditional' algorithm needs to extract complex features according to the corresponding data. The robustness of the algorithm is poor, and the processing method is complicated. It takes a long time to achieve real-time effects. According to the latest searches, the rapid development of the Convolution Neural Network

(CNN), deep learning is widely used in image recognition. As a data-driven algorithm, CNN has a powerful feature fitting ability, which can effectively and automatically extract image features, and has more Fast running speed. In 2012, AlexNet **[7]** won the ImageNet image classification competition, marking the rise of deep learning. In the following years, GoogleNet **[8]**, VGGNet **[9]**, ResNet **[10]**, and other algorithms improved the accuracy of image classification. They have successfully applied in many fields, such as face recognition and vehicle detection. Image of classification method also achieved great success using deep learning algorithms. Stanford University's staff and others have established the TrashNet Dataset for the public. The dataset contained six categories and a total of 2527 pictures. Ozkaya et al. **[11]** compared the classification capabilities of different CNN networks. They made a neural network and called its name a TrashNet, and fine-tuned its parameters. It achieved 97.86% success on the TrashNet dataset. In terms of non-public datasets, Mittal et al. **[12]** self-made 2561 garbage image datasets GINI, using the GarbNet model, and got 87.69% accuracy, Zheng Hailong et al. **[13]** used the SVM method to research the classification of construction waste. Xiang Wei et al. **[14]** used the classification network CaffeNet to adjust the size of the convolution kernel and the network depth to make it suitable for surface waste classification. It achieved a 95.75% recognition rate on its self-made 1500 image data set. In 2019, Huawei held an image classification competition and constructed a data set with a sample size of more than 10,000 images which further promoted the development of this field. The classification standards of domestic waste in various regions of our country are different. They can roughly divide into four categories: recyclable waste, hazardous waste, kitchen waste, and other waste. Research on garbage image recognition according to such classification standards is still in its infancy. Existing graphic classification algorithms, They used in the different fields of garbage treatment. They have the shortcomings of insufficient accuracy, poor generalization performance, and low processing efficiency. Aiming at the shortcomings of existing methods, we proposed to work on a garbage image classification algorithm based on a convolutional neural network (Garbage Classification Net, GCNet), which combines the attention mechanism module and feature fusion module in the network structure. It improves the accuracy and robustness of the model on garbage classification tasks.

3.1. Algorithm design.

Model structure. The GCNet model constructed in this paper includes two parts: a feature extractor and a classifier. The overall construction has shown in **Figure 1.** The feature extractor in the figure is composed of Resnet101 as the fundamental part. It included five bottlenecks then we added the attention mechanism module. Feature fusion performed on the features extracted by modules to extract the feature information F_1 of the image from the input:

$$F_1 = M_e(x) \qquad (1)$$

Among them, M_e represent the feature extractor. The classifier consists of two fully connected layers and a Softmax classifier, which classifies the extracted feature information F_1 to obtain the final score y_i of the image in each category:

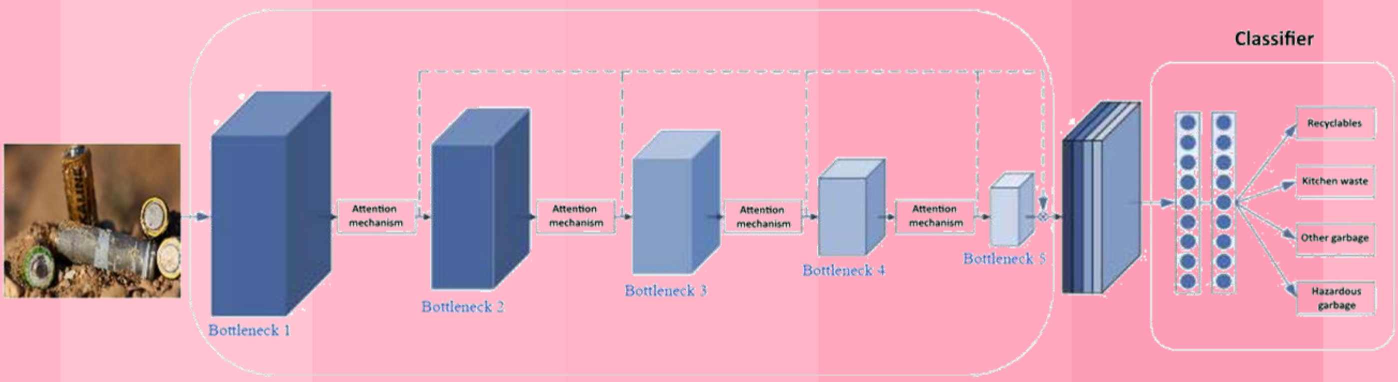
$$y_i = M_c(F_1), i \in \{1, 2, \cdots, n\} \qquad (2)$$

Among them, M_c represents the classifier.

3.2 Attention mechanism.

The mechanism of attention stems from the study of human vision. Humans will select specific areas in the retina to focus on according to actual needs and be allocate limited processing resources to significant parts. Since the characteristic representations of the same category of garbage may be quite different, it is not conducive to the correct classification of pictures which requires attention to the salient areas in the image. Inspired by this idea, by constructing an attention mechanism module, the network model focuses on the feature areas conducive to classification to achieve better. The feature extraction function, its specific structure are shown in **Figure 2.**

Figure 1. GCNet network structure diagram



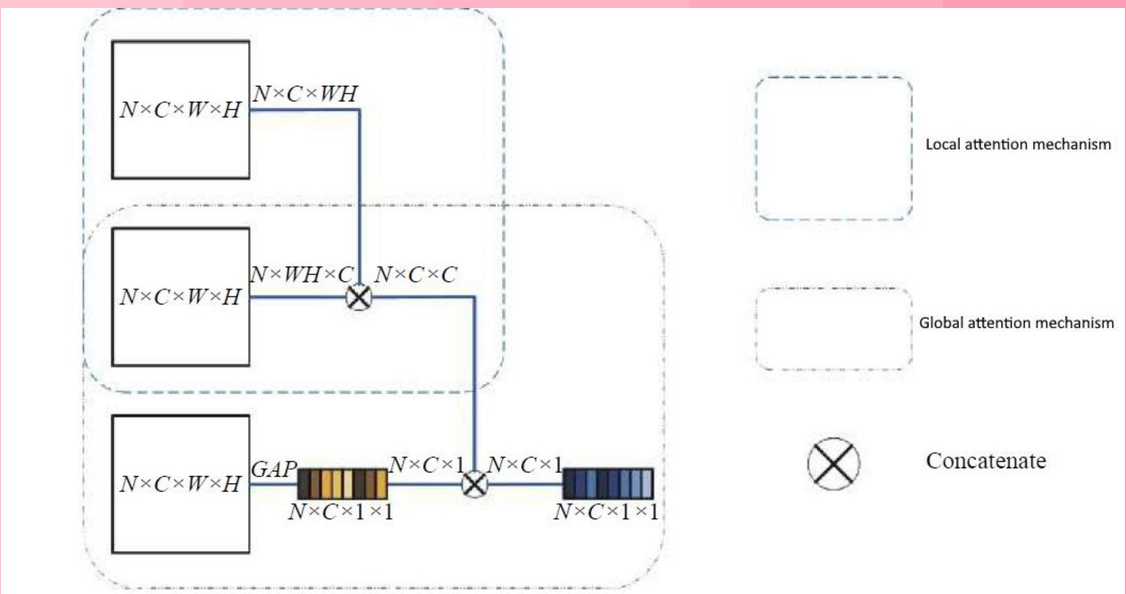


Figure 2. Schematic diagram of attention mechanism

The feature extracted is represented by each bottleneck F_i, and its size is $N \times \mathrm{C} \times \mathrm{W} \times \mathrm{H}$. By using the Gram matrix **[15]** to construct a local attention mechanism module, and multiply F_1 and its transpose F_i^T to obtain a size of $N \times \mathrm{C} \times C$'s local features F_{local}^i:

$$F_{\mathrm{local}}^i = F_i^{\mathrm{T}} \cdot F_i \qquad (3)$$

This operation can obtain the correlation between each element in the feature map. A large element and a small element with a value are necessary to highlight each part. That it is helpful to judge the characteristics of the category. That will be at the same time to suppress the features that affect the judgment. Subsequently, the global average pooling operation (Global Average Pooling, GAP) **[16]** performed on F_{local}^i, which can retain the spatial information and semantic information extracted by the feature F_i, and obtain a size of $N \times \mathrm{C} \times 1 \times 1$ Global features F_{global}^i :

$$F^i_{\text{global}} = GAP(F_i) \quad (4)$$

F^i_{global} represents the most prominent feature in the feature F_i. Finally, the local feature of F^i_{local} is multiplied by the global feature of F^i_{global} to obtain the overall featureF^i_{Fusion} contains both local and global information.

$$F^i_{\text{fusion}} = F^i_{\text{local}} \cdot F^i_{\text{global}} \quad (5)$$

3.3 Feature fusion mechanism.

In the task of garbage sorting, the garbage belonging to the same category is often quite different. For example, the cardboard boxes and glasses included in the recyclables have large differences in appearance but belong to the same category. The discrimination increases a certain degree of difficulty. In addition, as the number of network layers deepens, a single image feature will lose information in some areas, which will lead to the deterioration of the classification performance of the model. In response to this problem, the commonly used method is through different convolutions The core pooling operation builds a multi-scale feature fusion module, but this operation will increase the computational complexity of the model and also make model training more difficult. Since the constructed feature extractor contains different pooling operations, it is only necessary to extract features under different bottlenecks to obtain features of different sizes. On this basis, this paper proposes an improved feature fusion mechanism. As shown by the dotted line, the overall feature F^i_{fusion} extracted by each bottleneck after the attention mechanism is fused:

$$F_1 = w_c * Concat(F^1_{\text{fusion}}, \cdots, F^5_{\text{fusion}}) + b_c \quad (6)$$

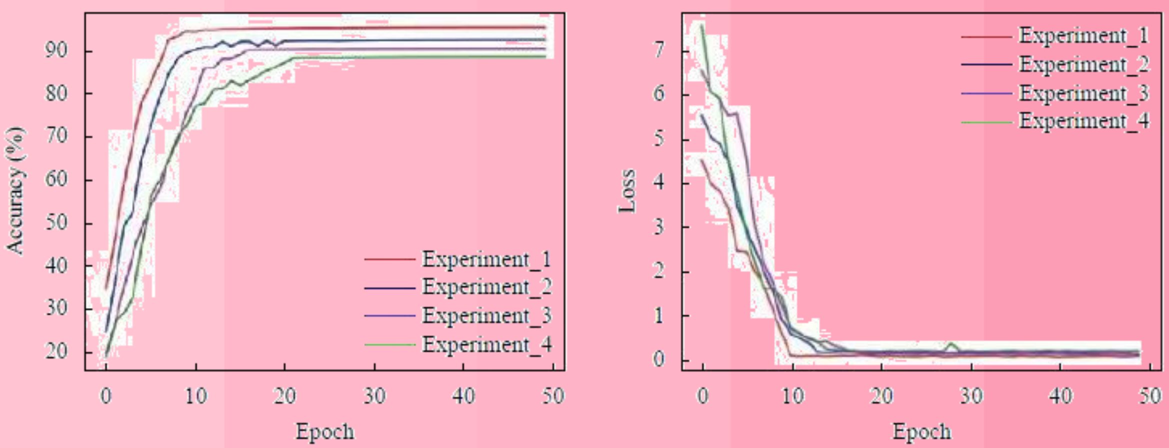
Among them, w_c is a 1x1 convolution kernel, b_c is a bias, and concatenation represents a fusion operation. This operation aims to use feature information of different scales to avoid information loss and further improve the robustness of the model.

4.1. Experiment and result analysis.

The experimental platform in the article is done under Ubuntu 16.04 system, using Python language and Pytorch deep learning framework. The hardware environment is CPU Intel 17-9700K, memory 32 GB, graphics card Nvidia GeForce RTX 2080Ti.

4.2. Experimental data.

This article uses the Huawei Waste Classification Challenge Cup dataset, which has all marked categories, including four categories of food waste, recyclables, other waste, and hazardous waste. Each category contains several sub-categories, such as a total of 40 small and a total of 14 683 images.



(a) Accuracy iteration curve (b) Loss function iteration curve

Figure 3. GCNet training process curve

We can see that from **Figure 3** that the accuracy rate increase and loss value decrease of experiment 1 is faster than the other three experiments. The first experiment training process converges faster. The speed of experiment 2 and experiment 3 are almost similar. Then the speed of experiment 4 is the slowest to verify the accuracy of the above training. The test was performed on the corresponding test set. The experimental results are shown in Table 1. Experiment 1, which includes the attention mechanism module and the feature fusion module, achieved an optimal accuracy rate of 96.73%. Each sub-category has achieved the highest accuracy rate, indicating that the model constructed in this article has good generalization ability. In the category of "other garbage" with large intra-class are the difference. Both the attention mechanism and the feature fusion mechanism can be significantly Improve the accuracy of the model. In the experiment, a 4:1 ratio is used to divide the data set, 80% is used as training data, and 20% is used as test data. In addition, to enhance the generalization ability and robustness of the model, data enhancement operations are also performed on the training samples included random rotation, random folding, random cropping, etc.

4.3 Result analysis

The experiment selects the ADAM **[17]** optimization algorithm training model, the momentum coefficient is 0.9, a total of 50 iteration cycles are set, the initial learning rate is set to 0.01, the learning rate is attenuated by 0.1 times every 10 iteration cycles, and the exponential decay rate of the first-order moment estimation It is 0.99, and the exponential decay rate of the second-order moment estimation is 0.999. In addition, the cross-entropy loss function is used to train the optimization model. During the training process, the GCNet model is subjected to ablation experiments to verify the functions of the attention mechanism and the feature fusion mechanism respectively. Experiment 1 (Experiment_1) is a model containing an attention mechanism and feature fusion mechanism. Experiment 2 (Experiment_2) is a model containing only an attention mechanism, Experiment 3 (Experiment_3) is a model containing only a feature fusion mechanism, Experiment 4 (Experiment_4) is a model that does not hold attention mechanism and feature fusion mechanism. The iterative curve of the training process of each experiment is shown in **Figure 3.**

Table 1. Comparison of accuracy rate of GCNet model ablation experiment

Experiment number.	recyclable trash.	Hazardous garbage.	Kitchen waste.	Other garbage.	Average accuracy rate.
1	95.21	96.35	97.82	97.54	96.73
2	94.53	90.35	93.12	90.52	92.13
3	93.63	90.65	92.36	88.48	91.28
4	90.15	90.34	88.12	86.06	88.67

The accuracy comparison results of the model in this paper and other models are shown in Table 2. We can see that the GCNet constructed in this paper has the highest average accuracy rate and achieved the highest accuracy rate in each category, indicating the attention mechanism and characteristics explained. The fusion mechanism fully extracted features conducive to image classification, making the classification results more accurate. The average accuracy of TrashNet is slightly worse than that of GCNet, and CaffeNet has the worst outcome.

Table 2. Comparison of experimental accuracy of different models

Model.	recyclable trash.	Hazardous garbage.	Kitchen waste.	Other garbage.	Average accuracy rate.
GCNet	95.21	96.35	97.82	97.54	96.73
TrashNet	92.78	93.24	93.06	91.56	92.66
CaffeNet	90.53	89.19	93.81	87.43	90.24

Figure 4 shows the test results of 3 algorithms such as GCNet, TrashNet, and CaffeNet with the same four pictures predicted category and the probability of belonging to categories. GCNet achieved the best results in the samples of each classification.

Forecast category: Recyclables (score=0.9885) Hazardous garbage (score=0.0054) Food waste (score=0.0035) Other garbage (score=0.0026)

Forecast category: Hazardous garbage (score=0.9973) Recyclables (score=0.0010) Food waste (score=0.0009) Other garbage (score=0.0008)

Forecast category: Food waste (score=0.9897) Recyclables (score=0.0053) Hazardous garbage (score=0.0031) Other garbage (score=0.0019)

Forecast category: Other garbage (score=0.9912) Hazardous garbage (score=0.0044) Food waste (score=0.0032) Recyclables (score=0.0012)

Forecast category: Recyclables (score=0.9485) Hazardous garbage (score=0.0412) Food waste (score=0.0052) Other garbage (score=0.0051)

Forecast category: Hazardous garbage (score=0.9543) Recyclables (score=0.0311) Food waste (score=0.0092) Other garbage (score=0.0054)

Forecast category: Food waste (score=0.9379) Recyclables (score=0.0431) Hazardous garbage (score=0.0128) Other garbage (score=0.0062)

Forecast category: Other garbage (score=0.9357) Hazardous garbage (score=0.0356) Food waste (score=0.0151) Recyclables (score=0.0136)

Forecast category: Recyclables (score=0.9578) Hazardous garbage (score=0.0387) Food waste (score=0.0022) Other garbage (score=0.0013)

Forecast category: Hazardous garbage (score=0.9228) Recyclables (score=0.0696) Food waste (score=0.0043) Other garbage (score=0.0033)

Forecast category: Food waste (score=0.9216) Recyclables (score=0.0521) Hazardous garbage (score=0.0213) Other garbage (score=0.0050)

Forecast category: Other garbage (score=0.6275) Hazardous garbage (score=0.3561) Food waste (score=0.0095) Recyclables (score=0.0069)

(c) CaffeNet test results

Figure 4. Comparison of test results of various models

Especially In the 4th (other garbage) test, due to the difference within the image category of this category, the probability of CaffeNet identifying it as other garbage is only 62.75%. The test of incorrect as harmful garbage is higher at 35.61%. In some interference situations, this may cause judgment errors and have an impact on the classification results. The algorithm in this test successfully improved this problem, ensuring the accuracy of the more to distinguish categories.

5. Conclusion.

This paper constructs a convolutional neural network-based algorithm GCNet by aiming at the problem of image classification, which can effectively extract image features and reduce the impact of category differences by constructing an attention mechanism and feature fusion mechanism. In the experiment, an average accuracy has achieved a rate of 96.73% on the relevant data set, which has improved the accuracy rate by about 4% compared with the existing classification algorithm, which meets the actual application requirements and has a good application prospect.

References.

1. C. Bircanoğlu, M. Atay, F. Beşer, Ö. Genç and M. A. Kızrak, "RecycleNet: Intelligent Waste Sorting Using Deep Neural Networks," 2018 Innovations in Intelligent Systems and Applications (INISTA), 2018, pp. 1-7, DOI: 10.1109/INISTA.2018.8466276.
2. Abeywickrama T, Cheema MA, Taniar D. K-nearest neighbors on road networks: A journey in experimentation and in-memory implementation. Proceedings of the VLDB Endowment, 2016, 9(6): 492 –503. [doi: 10.14778 /2904121. 2904125]
3. Lowe DG. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 2004, 60(2): 91–110. [doi: 10.1023/B:VISI.0000029664.99615.94]
4. Harri C, Stephens M. A combined corner and edge detector. Proceedings of the 4th Alvey Vision Conference. Manchester, UK. 1988. 207–217.
5. Vapnik V. Statistical Learning Theory. New York: Wiley, 1998. 401–492.
6. Wu Jian, Chen Hao, Fang Wu. Research on Waste and Garbage Analysis and Identification Based on Computer Vision. Information Technology and Informatization, 2016, (10): 81–83. [doi: 10.3969/j.issn.1672-9528.2016. 10.020]
7. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems. Lake Tahoe, NV, USA. 2012. 1106–1114.
8. Szegedy C, Liu W, Jia YQ, et al. Going deeper with convolutions. arXiv: 1409.4842, 2014.
9. Simonyan K, Zisserman A. deep convolutional networks for large-scale image recognition. arXiv: 1409. 1556, 2015.
10. He KM, Zhang XY, Ren SQ, et al. Deep residual learning for image recognition. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. 2016. 770–778.
11. Ozkaya U, Seyfi L. Fine-tuning models comparisons on garbage classification for recyclability. arXiv: 1908.04393, 2019.
12. Mittal G, Yagnik KB, Garg M, et al. SpotGarbage: Smartphone app to detect garbage using deep learning. Proceedings of 2016 ACM International Joint Conference. Heidelberg, Germany. 2016. 940–945.
13. Zheng Longhai, Yuan Zuqiang, Yin Chenbo, et al. Research on the automatic classification system of construction waste based on machine vision. Mechanical Engineering and Automation, 2019, (6): 16-18. [doi: 10.3969/j.issn.1672-6413.2019.06.006]
14. Xiang Wei, Shi Jinfang, Liu Guihua, et al. Application of improved CaffeNet model in water surface garbage recognition. Sensors and Microsystems, 2019, 38(8): 150–152,156.

15. Zhang XK, Wang Y, Gou MR, et al. Efficient temporal sequence comparison and classification using gram matrix embeddings on a Riemannian manifold. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. 2016. 4498–4507.
16. M, Q, Yan SC. Network in network. arXiv: 1312.4400, 2014.
17. Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv: 1412.6980, 2017.