Design and research of intelligent trash can based on convolutional neural network and transfer learning

Peng Ling a, Liu Tianyi b
School of Computer and Information City College of Dongguan University of Technology, Dongguan, China

a penny@ccdgut.edu.cn, b liuty@ccdgut.edu.cn

Abstract. With the improvement of living standards, the types and quantities of garbage produced by residents are increasing, and the garbage disposal methods need to be further improved. Standardize garbage disposal, garbage classification and garbage recycling will contribute to construction a green city. Garbage classification technologies also needs innovation. The intelligent garbage classification which depends on the Internet and intelligent technologies will overcome the disadvantages of traditional garbage classification. To this end, this paper proposes a smart trash can which based on convolutional neural network and transfer learning technologies. The smart trash can in this paper builds a smart trash monitoring network through NB-IOT, and automatically extract garbage features based on convolutional neural network, then trains the network through transfer learning technologies. The experimental results shows that the smart trash can designed in this paper has fast response speed and high accuracy.

Keywords: Intelligent garbage can, Convolutional neural network, Inceptionv3, The migration study.

1. Introduction
With the continuous progress of science and technology, the consumption level of Chinese residents is improved, and everyone can greatly improving the quality of life. However, some serious garbage disposal problems are hides in the development of economic. Garbage is underutilize resource, reduce waste production and strengthen garbage pollution control is an important part to build a moderately prosperous society in all respects and accelerating urban modernization, as well as an important guarantee for realizing sustainable economic and social development [1]. At present, although there are a variety of garbage bins in the street, the effect of garbage classification is not good under the condition of no supervision. Therefore, it is particularly important to design an intelligent garbage bin that can identify and classify garbage independently. So this paper proposes an smart trash can based on convolutional neural network and transfer learning. The smart trash can has such functions as garbage identification and classification, control the switch of the garbage can, detection of human proximity and garbage overflow warning.
2. Overall design
The intelligent classification trash can designed in this paper is mainly composed of image acquisition module, microprocessor module, classification module, motor module, sensor module, power module and communication module, as shown in Figure 1. The image acquisition module is installed at the entrance of the garbage can. When the infrared sensor detects the garbage dumping behavior, then take a picture of the garbage using camera and uploaded the garbage images to the microprocessor STM32. Then STM32 uploads the garbage images to the cloud server which is running convolutional neural network and migration learning through the NB-IOT module. According to the constructed network model, the cloud server classifies the garbage and sends the identification results back to STM32, thus to control the motor module. The motor is installed at the bottom of the trash can to drive the rotation of the trash can lid. The dustbin lid has a groove and a valve at the bottom. When the result of garbage sorting is battery, the control motor rotates the groove which with the batteries to the corresponding trash can according to the garbage sorting rules. The sensor module is composed of infrared sensor and ultrasonic sensor, which can detect the approach of human body and obtain the information stored inside the trash can. The power module can provide stable power for the whole system to ensure the system running steadily.

![Figure 1. Overall design frame diagram](image)

3. Hardware Design

3.1. STM32 microprocessor
The smart trash can designed in this paper adopts STM32F103RCT6 as the microprocessor of the system. The microcontroller based on ARM core can reach the highest frequency of 72 MHz when it is running normally. It has a highly centralize for timer, counter, AD converter and many other components inside and outside the chip. Compared with the traditional 51 microcontroller, it has advantages of high performance, low power consumption, and the price is similar to the previous 8-bit microprocessor. The STM32 minimum system designed in this paper mainly includes two circuits: clock and reset. In order to obtain a more accurate master clock, an external crystal oscillator circuit is used to select the clock source. It consists of an 8 MHz crystal oscillator and two 22pF capacitors. The crystal oscillator and two capacitors on the circuit are usually connected close to the STM32 interface, in order to reduce the distortion of the external clock output and effectively reduce the startup waiting time. Reset circuit is the guarantee of running the microcontroller steadily. The working frequency of the processor is generally high, and it is easily disturbed by external factors such as voltage fluctuation or electromagnetic interference in the outside area.

3.2. Image acquisition module
The image acquisition module of the intelligent classification trash can designed in this paper mainly includes camera and FIFO frame buffer. The camera uses OV7670 product by Omni Vision, which has
its own AL422B FIFO frame buffer. It is used to improve the stability of data transmission. The communication between STM32 F103 single chip microcomputer and OV7725 image acquisition module is completed through the I2C interface. OV7725 leads out 20 pins in total, and the pins mainly include :D0–D9 pixel data ports, XCLK and PCLK are the input clock and pixel synchronization clock of the system. HREF and VSYNC are horizontal and vertical synchronization signals.

3.3. The motor module
The motor module of intelligent garbage can classification designed in this paper selects a stepping motor and controls the rotation angle of the motor by STM32 and achieved the purpose of driving the rotation of the garbage can cover finally. because the output pulse of STM32 is too weak to drive the stepping motor up, so a driving circuit needs to be selected to start the motor. The ring distributor obtains the control signal from STM32 and then control and change the state of the motor, amplifies the signal into a signal capable to driving the motor through the signal processing module, the push stage and the power amplifier. The role of the protection circuit is to prevent the electric current of the stepping motor from being too large, and cut off the main circuit in time when it appears. In this design, RD0218 driver is selected to control the motor of the intelligent trash can. It has the characteristics of small size, low noise, high precision and so on, which meet the requirements of the intelligent trash can.

3.4. The NB-IOT communication module
Since the working time of the garbage can is determined by the times of throw garbage, and the transmitting data size is not large, and the demands for garbage cans are huge, therefore the communication module of the intelligent garbage can designed in this paper adopts NB-IOT. NB-IOT communication is a new technology in the IOT field proposed in recent years. It has the characteristics of strong link, low cost, low power consumption and high coverage. This design uses BC26 chipset from Shanghai Yiyuan, with 44 LCC pins and 14 LGA pins.

4. Software Design

4.1. Convolutional neural network algorithm
By introducing local connection, weight sharing, maximum pooling, nonlinear activation and other methods, the convolutional neural network (CNN) allows the divine network to learn features from images automatically, and avoids the complicated pre-processing on the image such as edge detection and threshold segmentation. Compared with traditional machine learning methods, it has stronger characteristic learning and expressive ability, so it has been widely used in image classification, target detection and other aspects. CNN is mainly composed with input layer, convolution layer, pooling layer, full connection layer and output layer. There has a clear division of labor and close connection between each layer. The input layer is responsible for pre-processing and standardized the data. The convolution layer is responsible for extracting features and consists of convolution units. Pooling layer is in the middle or behind of the convolution layer, is the output of the convolution layer, and the main work of pooling layer is to retain the main features and reduce the operations. The full connection layer is responsible for comparing the analyzed features with samples. There are many layers in the full connection layer, each layer has a large number of neurons, and each neuron is connected with all the neurons in the previous layer. The full connection layer is responsible for integrating the important information processed by the convolution layer or pooling layer. Among them, the model structure can be extended by extending the convolution layer, pooling layer and full connection layer. The basic working principle of CNN model is shown in Figure 2. Firstly, the garbage images taken by the camera are extracted by the sliding operation from convolution kernel. Then, the features are extracted twice results from the alternative actions of two convolution layers and pooling layers. Finally, the features are integrated and input into the fully connected layer for feature fusion then output the results after classification.
After comprehensive considered of the advantages and disadvantages of each network models, the model construction operation cost and other factors, this design uses Inceptionv3 network as the feature extractor. On the basis of maintaining its original weight, build neural network and retraining. The network structure of InceptionV3 is shown in Figure 3.

There are 47 layers in the Inceptionv3 network. Two different convolution kernels with step size of 1 are using in 33 layer and 11 layer which designed for the convolution layer. By weighting each small pixel region on the image of the previous layer, the feature information of the pixel region is extracted. The small pixel region is called the local feature value, and the weight of the region is the convolution kernel. After the convolution operation, the image is added with the bias value to form the feature graph. The transformation process is shown in Formula 1. Where $X^l_j$ is the eigenvalue obtained by the convolution operation, $X^{l-1}_{i}$ is the eigeninformation of the image calculation at the previous layer, $K^l_{ij}$ is the convolution kernel participating in the operation, and $b^l_j$ is the bias value.

$$X^l_j = \sum_{i \in M_1} \sum_{j \in M_f} X^{l-1}_i K^l_{ij} + b^l_j \quad (1)$$

The size of the pooling layer is mainly 33 and 88. The maximum pooling method is adopted to retain the maximum pixel value every time, in order to achieve the effect of feature reduction and data compression. The process is shown in Formula 2.

$$X^l_j = \text{down}(X^{l-1}_j) \quad (2)$$

The subtlety of this network also lies in the design of multiple Inception-Module groups. By using multiple 1*1 convolution kernels and multi-branch convolution aggregation, the Module groups greatly

---

**Figure 2.** The basic working principle diagram of convolutional neural network

**Figure 3.** InceptionV3 Network model structure
reduced the parameters and layers of the network, and reduced the calculated quantity of convolutional neural network. This system use a user-defined data set as the input layer of the network, uses the shared weight value of Inceptionv3 network and fine-tuning the output, at the same time redefines the full connection layer. Transform the formula through the full connection layer, as shown in Formula 3.

\[
\begin{bmatrix}
  y_1 \\
  y_2 \\
  \vdots \\
  y_n
\end{bmatrix} = \begin{bmatrix}
  W_{1,1} & W_{1,2} & \cdots & W_{1,n} \\
  W_{2,1} & W_{2,2} & \cdots & W_{2,n} \\
  \vdots & \vdots & \ddots & \vdots \\
  W_{n,1} & W_{n,2} & \cdots & W_{n,n}
\end{bmatrix} \begin{bmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_n
\end{bmatrix} + \begin{bmatrix}
  b_1 \\
  b_2 \\
  \vdots \\
  b_n
\end{bmatrix}
\] (3)

Successfully integrated into a one-dimensional vector. The softmax function is selected as the output classifier, as shown in Formula 4.

\[
\text{softmax}(y_i) = \frac{e^{y_i}}{\sum_{j=1}^{n} e^{y_j}}
\] (4)

In formula 4, i refers to a certain item category, so as to output the probability distribution and achieve multi-classification goals[5].

4.2. Transfer Learning

In the field of image recognition, convolutional neural network has an excellent performance, but in order to pursue better training effect, it often needs large-scale data sets for support. Although there are many public data sets available, getting enough tagged training samples for some special areas will be a extremely time-consuming and expensive work [6]. If the sample data set is not enough to support the training of deep network model, transfer learning can be adopted. Transfer study is defined as: there has a given source domain \(DS\) and learning task \(TS\), target domain \(DT\) and learning task \(TT\). Through source domain \(DS\) and learning task \(TS\) to improve the prediction accuracy of the target decision function \(f(.)\) in the target domain \(DT\) [7]. It can be seen from the definition that the main idea of transfer learning is to apply the knowledge acquired in the source domain to the learning tasks which are included in the target domain, so as to improve the performance of the target task [8]. One of the most commonly used methods in transfer learning is fine-tuning. The idea is using a pre-trained model on a large scale data set to retrain the data set which on the target domain. By fine-tuning parameters of several layers, not only the speed of network training can be accelerated, but also better results can be obtained [9].

![Figure 4. Flow chart of Transfer Learning](image-url)
In this paper, the convolutional neural network Inception V3 model pre-trained on the ImageNet data set is used as the feature extractor to freeze all layers except the fully connected layer, which means the bottom layer is used to extract features, and does not participate in the training, but only fine-tuning the top-level parameters. The initial learning rate is set, and then trained the target data set fine-tuning. Finally, the convolutional neural network model for garbage recognition is obtained. The specific process is shown in the figure 4.

In the process of model training, Adam method is used to optimize the network which can make it converge quickly. Adam algorithm uses the gradient first moment estimated value and second moment estimated value to adjust the learning rate of each parameter dynamically [10]. Formula 5 is the gradient first order moment estimate value \( m_t \), and formula 6 is the calculation formula for the second order moment estimate value \( v_t \) which is expressed in the form of moving average. Of which, \( g_t \) is the first-step degree of step \( t \), \( \beta_1 \) and \( \beta_2 \) are respectively express the exponentially decaying rate of the gradient estimated first-order moment value and second-order moment value, which need to be manually configured to acquire satisfactory performance. In general, \( \beta_1 \) is set to 0.87 and \( \beta_2 \) to 0.98. The calculation formula of Adam algorithm is shown in Formula 7, of which \( n \) is the step size of Adam algorithm, \( \theta_t \) and \( \theta_{t+1} \) are the weight value of step \( t \) and step \( t+1 \) respectively, \( \epsilon \) is 10-8.

\[
\begin{align*}
    m_t &= \beta_1 m_{t-1} + (1 - \beta_1)g_t \\
    v_t &= \beta_2 v_{t-1} + (1 - \beta_2)g_t^2 \\
    \theta_{t+1} &= \theta_t - \frac{n}{\sqrt{v_{t+1}} + \epsilon} m_t
\end{align*}
\]

4.3. Full bucket detection alert

The intelligent trash can designed in this paper also equipped with ultrasonic sensors, which can realize ranging functions. The main working principle of the ultrasonic sensors is that using ultrasonic module to receives the specific ultrasound issued previously, then calculated the distance \( D \) between the garbage and ultrasonic module using the time difference values. The formula of the distance \( D \) is shown in Formula 8.

\[
D = \frac{340}{s} \times t
\]

Therefore, the distance can be monitored by real-time ultrasonic after each waste delivery. When the distance is less than 30cm, the wireless network communication will carry out through the NB-IOT module, which improves the working efficiency of the sanitation workers when they changing the garbage bags. The flow chart is shown in Figure 5.

![Figure 5. Full bucket detection reminder flow chart](image-url)
5. Testing
In order to verify the recognition accuracy of the intelligent trash can designed in this paper when it used in actual application, the convolutional neural network data set was obtained through two approaches: First, 1000 pictures of different types of household garbage were collected through the Internet, such as plastic bottles, paper boxes, batteries, etc. Second, 200 pictures in different angles and different types of household garbage were taken through the camera on the garbage bin, and then cleaning, sorting, classifying the collected pictures. Finally, garbage images are placed in different folders which has different labels and the data pretreatment is complete. The environment required by the training model is built in the cloud server and then training the convolutional neural network Inceptionv3 network. The accuracy of the Convolutional neural network Inceptionv3 model was verified by the test set, and the results were shown in Table 1.

| Test garbage names | Average recognition rate |
|--------------------|--------------------------|
| A banana peel      | 83.15%                   |
| Cardboard          | 85.44%                   |
| Plastic bottles    | 82.64%                   |
| The battery        | 89.6%                    |
| Cans               | 86.08%                   |

The pre-trained Convolutional neural network Inception V3 model on the ImageNet data set is used as the feature extractor, all layers except the fully connected layer are frozen, and only the top-level parameters are fine-tuned. The initial learning rate is set to 0.001 and the batch size is set to 32. In other words, 32 images were randomly selected as input data in each iteration, and the default size of the original model is $299 \times 299 \times 3$ which was adopted on the input images in the experiment. The learning rate is adjusted after 20 times iterations. The experimental results are shown in Table 2. The accuracy of all kinds of garbage recognition has been generally improved, which indicates that transfer learning from the source domain which is similar to the target domain can not only accelerate the training speed, but also improve the accuracy.

| Test garbage names | Average recognition rate |
|--------------------|--------------------------|
| A banana peel      | 85.32%                   |
| Cardboard          | 87.38%                   |
| Plastic bottles    | 85.98%                   |
| The battery        | 92.2%                    |
| Cans               | 89.06%                   |

6. Summary
In this paper, the smart trash can which is based on the convolutional neural network and transfer learning technologies has realized core functions including the classification of garbage and waste storage, pailful detection and alert, and so on. Finally realizes the intelligent garbage classification which not only can effectively improve the rate of garbage sorting and raise the working efficiency of the sanitation workers, but also saving a huge number of human resources and reducing the waste of resources. This smart trash can improving the social benefit and economic benefit by a large margin and more suitable with modern intelligent standards.

Acknowledgments
This work is supported by the University-level quality engineering project in 2021-Online and Offline hybrid course Data structure, project number: 2021zlgc007.
References

[1] Zhao Biao, Zhao Ji, Liu Tianqi, Design and Implementation of Intelligent Garbage Bin Classification based on Arduino [J], Electronic World, 2020, 15.

[2] Huang Guowei. Research on Intelligent Classification of Urban Garbage Cans Based on Deep Learning [D], Anhui University of Science and Technology, 2019.5.

[3] Xing Zhixiang, Gu Fenglin, Qian Hui, Zhang Ying, Wang Lijin. Research progress of Pedestrian Detection Method based on Convolutional Neural Network [J], Safety and Environmental Engineering, 6 (2018).

[4] Wang Yuxin, Pu Linhua, Pu Ran, Li Bei, Ding Xia. Implementation of Garbage Classification System based on Convolutional Neural Network [J], Journal of Beijing Information Science and Technology University, October 2020.

[5] Wu Bicheng, Deng Xiangen, Zhang Zizi, Tang Xiaoyu. Intelligent Garbage Classification System based on convolutional Neural Network [J], Physics Experiment, November 2019.

[6] Tan Z, Li K. Differential evolution with mixed mutation strategy based on deep reinforcement learning[J]. Applied Soft Computing, 2021: 107678.

[7] Wang B B. Expression recognition of newborn pain based on transfer learning [D], Nanjing University of Posts and Communications, 2020.6.

[8] Lu Shuyi, WEI Shuang, Wan Siyuan. Intelligent Garbage Bin Classification Based on Deep Learning [J], Computer and Digital Engineering, 5th issue, 2021.

[9] Wu Ling, Wang Hao, Zhang Xiaochun, Zhou Jian, Duan Aihua. Design and Implementation of Garbage Classification System based on Deep Transfer Learning [J], Journal of Shenyang University, December 2020.

[10] Zhang Zhihe. Identification of Chinese orchid species based on convolutional Neural Network and Transfer Learning [D], Zhejiang A&F University, 2020. 12.