International Journal of Data Science

ISSN online: 2053-082X - ISSN print: 2053-0811
https://www.inderscience.com/ijds

Structure design and system implementation of a supermarket shopping robot based on deep learning

Jinyun Jiang, Shiyi Ying, Wanxin Fu, Xiaoliang Jiang

DOI: 10.1504/IJDS.2023.10053383

Article History:

Received:	12 April 2022
Accepted:	05 September 2022
Published online:	09 March 2023

Structure design and system implementation of a supermarket shopping robot based on deep learning

Jinyun Jiang, Shiyi Ying*, Wanxin Fu and Xiaoliang Jiang

College of Mechanical Engineering,
Quzhou University,
Quzhou, 324000, China
Email: jjy_qzu@163.com
Email: 2218713668@qq.com
Email: 2635005306@qq.com
Email: jxl_swjtu@163.com
*Corresponding author

Abstract: To enhance the operating efficiency of supermarkets, reduce their labour costs, and satisfy people's shopping experiences, we present the design and implementation of a supermarket shopping robot based on deep learning. Firstly, the robot adopts high-performance STM32F407 as its main control chip and is powered by 4 DC motors. It relies on 12 grey-scale sensors, gyroscopes and other devices for path planning. Through the infrared detection module, detect whether there are goods in the cargo window, and use the manipulator to accurately grasp and place the goods. Secondly, design a shopping cart to replace the manual cart to maintain the robot's control of the shopping cart during the shopping process. Finally, the AlexNet network is used as the feature extractor to realise the rapid identification of the target cargos. The experimental results show that under the simulation of the real supermarket environment, the designed robot runs flexibly, stably and reliably, and can well complete the purchase and supply of commodities, which is in line with the development trend of artificial intelligence.

Keywords: supermarket shopping robot; deep learning; single-chip microcomputer; recognition; AlexNet.

Reference to this paper should be made as follows: Jiang, J., Ying, S., Fu, W. and Jiang, X. (2023) 'Structure design and system implementation of a supermarket shopping robot based on deep learning', *Int. J. Data Science*, Vol. 8, No. 1, pp.1–15.

Biographical notes: Jinyun Jiang received her MS in Mechanical Design and Theory from Beijing University of Chemical Technology, Beijing, China. Her research interests include material processing and image segmentation.

Shiyi Ying is a junior majoring in mechanical and electronic engineering at Quzhou University in China. His research interests include mechanical and electronic engineering.

WanXin Fu is an undergraduate student of Quzhou University. Her research interests include machine vision and digital testing technology.

Xiaoliang Jiang received his BS in Mechanical Design, the PhD in the College of Mechanical Engineering from Southwest Jiaotong University, Chengdu, China, in 2010, and 2016. He is currently an Associate Professor at the College of Mechanical Engineering, Quzhou University. His research interests include image processing and image recognition.

1 Introduction

In recent years, service robots on the market have been impressive in various occasions, and their increasingly intelligent functions have improved the quality of life to a certain extent (Kivrak et al., 2021). For example, although there are some shopping robots in supermarkets, they cannot meet the shopping needs of special groups. Therefore, we hope to develop a service robot that can realise autonomous shopping. The shopping cart can be controlled by the robot, so as to truly free hands and facilitate the handling of goods (Sales et al., 2016). In large supermarkets, customers only need to wait a moment for their shopping robots to complete their shopping lists. This work can be realised through the combination of artificial intelligence and mobile robot, which provides a powerful solution for the realisation of some specific functions of our supermarket shopping robot (Ryumin et al., 2020; Cebollada et al., 2021).

The experimental work of our supermarket shopping robot is carried out in the simulated supermarket shopping site (Chen et al., 2020). The site is equipped with four shelves in the outer ring and the central area is the warehouse. The ground of the site is sprayed with black matte paint and divided into several 400mm long square areas with 30mm wide white lines. The robot we designed needs to start from the specified "starting point area", find the "target goods" within the specified time, and use the shopping cart to return the goods to the starting point area. On this basis, we carry out the structural design and system implementation of supermarket shopping robot. For the position planning of the sensor, we need to ensure the parking position of the trolley in the outer ring shelf and the inner ring warehouse area at the same time, so that the self-developed manipulator can detect in real time and complete the task of catching targets (Zhang et al., 2020). In addition, due to the many commodity placement angles and complex and changeable sizes in the image, the accuracy of supermarket shopping robot in commodity detection and recognition will be difficult to control (Lin et al., 2021).

To solve these problems, we first plan the installation location of various sensors (Yang and Liu, 2021). Sensor detection device is an important part of robot. The robot must recognise the target goods and complete the capture on the basis of accurate parking of the shopping cart. The infrared sensor and camera we choose are installed together, which can accurately locate the car in front of the goods, then take photos, identify and grab them, and put them into the shopping cart. In addition, we use a 12-channel grey-scale sensor and apply the PID algorithm (Hu et al., 2019) to ensure that the shopping cart does not deviate from the white line. This just ensures the tracking of the outer ring. When the trolley completes the capture of the 'target goods" in the outer ring, we need to expand the designed rotary 12 channel grey-scale to make the robot track closer to the inner ring warehouse area, so as to facilitate the capture of goods.

For commodity detection and recognition, we use deep learning for image recognition (Qi, 2020; Li et al., 2019; Zhao et al., 2021), and use the AlexNet network as the feature extractor (Chen and Dong, 2021; Gao et al., 2021; Yu et al., 2013). The feature extraction of traditional machine learning (Sahil and Vijay, 2021) mainly depends on manpower, which is limited by the number of trainable parameters, resulting in the limited features it can learn. Deep learning has good learning performance and is very helpful for product identification. However, this method relies heavily on the amount of data, and it is difficult to get ideal results when the amount of data is small. Therefore, we increase the number of photos of goods, train the images of each angle and each commodity (Ollah and Karim, 2021), and maintain good accuracy. Then set the set value to 0.85 and compare the similarity with the training photos. When the similarity reaches the threshold (0.85), it means that the object is the target object, so as to distinguish distracters.

PID control algorithm is a linear control algorithm that adjusts the deviation between expectation and current situation to make it reach the expectation. We expect the supermarket shopping robot to keep driving in a straight line on the field, so using PID control algorithm to control the motor speed (Garcia-Martinez et al., 2020) is a good choice.

In addition, there is the design of the shopping cart clamping device. We use the steering gear to control the gear meshing, thus achieving the purpose of clamping the shopping cart. To facilitate the supermarket shopping robot combination to walk normally in the field, the shopping cart is clamped on the right side of the supermarket shopping robot in this design. The shopping cart cannot be connected with the powered equipment. Therefore, the connection structure is designed and installed on the right side of the third floor of robot.

The contribution of this study can be summarised as:

1 The competition theme of this design is supermarket shopping robot. We propose a design scheme of supermarket shopping robot based on deep learning, and develop an intelligent supermarket shopping robot that simulates instead of hand pushing shopping cart. In this robot, we designed some new mechanical structures to improve the performance of the robot. To improve the stability of the robot, the virtual environment of the supermarket is simulated, and the feasibility of our supermarket shopping robot structure and the scheme of carrying shopping cart by the robot is verified.

2 We use deep learning algorithm for image recognition, and use the Alexnet network as the feature extractor. Experiments show the effectiveness of this method.

2 Our approach

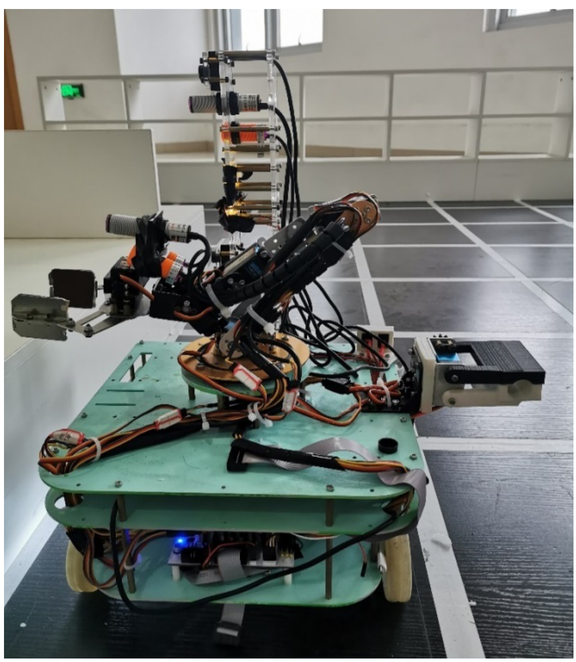
The design of supermarket shopping robot mainly includes two aspects: robot hardware design and robot software design. Since this design is mainly aimed at the innovation of supermarket shopping robot structure, the description of robot hardware design will be described in more detail below.

2.1 Robot hardware design

2.1.1 Robot structure

Supermarket shopping robot appearance is mainly divided into the upper middle and lower three layers. As shown in Figure 1, four 24v 900 speed reduction motors, two 12 channel grey sensors and five digital single channel grey sensors are fixed under the bottom plate of the first floor, The angle correction and steering of the robot are realised by the speed difference between the left and right motors. To reduce the centre of gravity of the robot as much as possible, in the design process, we reasonably arrange the main control board, DC motor drive module, voltage stabilising module, adjustable voltage stabilising power supply module, manipulator steering gear control board and other electronic modules on the base plate. The 24V lithium battery is placed in the centre of the base plate to concentrate the weight of the robot and improve the stability. The second layer is mainly used to place computers, process image information, and transmit data with single chip microcomputer through serial port line. The third layer is mainly used to fix the camera and photoelectric sensor combination device, mechanical arm and shopping cart handle connection device.

Figure 1 Appearance of supermarket shopping robot (see online version for colours)



To improve the strength of robot, we use m3 copper column to connect 3mm custom aluminium plate. In the design process of customised aluminium plate, the installation holes of electronic modules such as main control board, DC motor drive module, voltage stabilising module, adjustable voltage stabilising power supply module and mechanical arm steering gear control board are fully considered. All electronic modules are overhead with the help of plastic copper column, separated from aluminium plate, and covered with acrylic plate to reduce the damage of electronic modules.

2.1.2 Sensor detection

This design involves a variety of sensors (Zhang, 2021), mainly including gyroscope, grey sensor, photoelectric sensor, industrial camera and so on. Firstly, to control the steering angle of robot, we use the 9-axis hi229 attitude module to detect the attitude of the robot. The steering is realised by controlling the speed difference between the left and right motors, and the steering angle is controlled with the help of gyroscope. Since the gyroscope will be disturbed by the magnetic field of the motor, resulting in the misalignment of the detection angle, we will carry out the elevation treatment when installing the gyroscope.

Our supermarket shopping robot needs to have the ability of autonomous action (Bavelos et al., 2021). The test site is pasted with a 30mm wide white line. We decided to use a multi-directional grey-scale sensor to achieve the purpose of accurate line inspection. The photosensitive resistance in the 12 channels grey-scale sensor can be identified according to the principle that the resistance value is also different according to the strength of the reflected light on the ground. The detection signal is a digital quantity. When a white line is detected, the signal received by single chip microcomputer is 0, and when a white line is not detected, the signal is 1. Supplemented by PID control algorithm, the precise line patrol of the robot can be realised. For the patrol inspection of the outer ring, we use 12 channels grey sensor fixed in the middle of the front and rear ends for detection. Considering the line patrol of the inner ring, we have improved the 12 channels grey sensor at the front end. Through the control of the steering gear, the rotation of the grey sensor is realised. Without changing the robot route, the line patrol range of the robot is expanded to complete the line patrol of the inner ring. The five 1-way grey sensors installed on the left can realise the two positioning of the camera for the cargo window and the manipulator for the cargo window. It is convenient for the camera to collect commodity images and for the manipulator to grasp.

Photoelectric sensors are needed to judge whether there are goods in the shelf and warehouse area. There are many kinds of photoelectric sensors. From the application principle, it can be divided into slot photoelectric sensor, opposite photoelectric sensor and reflective photoelectric sensor. The application of reflective photoelectric sensor is a better choice in this design. During the design and manufacture of the emission photoelectric sensor, the light emitter and the light receiver are placed in the same position. The detection principle of papers is that when papers pass through the optical path emitted by the illuminator, the light will be blocked, and the light receiver can't receive light. In this way, the photoelectric switch will act and generate a switch control signal. We have designed a combined device of photoelectric sensor and camera. The height of photoelectric sensor is consistent with that of cargo window to realise accurate detection. Adjust the working distance of the lens to collect clear and undistorted images for efficient image processing.

2.1.3 single-chip microcomputer control system

The supermarket shopping robot we designed adopts STM32F407ZGT6 microcomputer as the main control chip. It is mainly used to:

1. control the motor to keep the robot moving correctly
2. Communicate with computer

3 Connect with the control board of the manipulator steering gear

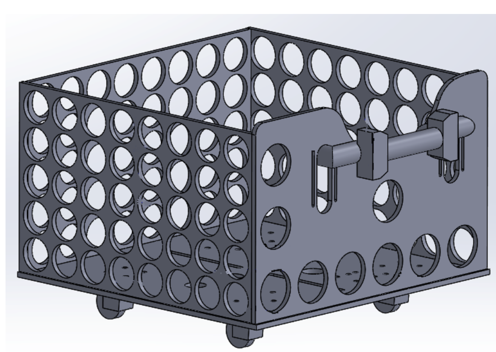
4 The photoelectric sensor and grey sensor are controlled to detect the cargo window and patrol the site.

The 24V lithium battery outputs 23v voltage through the voltage stabilising module, and then the 5V voltage output after being stabilised by the voltage stabilising power supply module supplies power to the single chip microcomputer. Because the robot takes a long time to perform tasks and the manipulator consumes a lot of power, the circuit design of the supermarket shopping robot adopts the mode of independent power supply. The robot walking mechanism, sensor power supply, single chip microcomputer steering gear port power supply and mechanical arm steering gear control board power supply are separated to ensure that the mechanical arm performs accurate actions and the robot performs tasks with a long endurance time. The sensor socket power supply on the main controller is taken from the 10V voltage output of the regulated power supply module. In addition, due to the need for steering gear control, we have set multiple steering gear sockets on the main controller, and the 24V lithium battery outputs 6.5V voltage to supply power to the steering gear socket through the voltage stabilising module.

2.1.4 Shopping cart design

In order to give customers a more convenient and comfortable shopping experience, this design also includes a supermarket shopping cart controlled by a robot. As shown in Figure 2, considering that the supermarket shopping robot needs to carry the shopping cart to run on the simulated supermarket site, and the shopping cart needs to load goods, we have made the structural design of the shopping cart as light and flexible as possible. The shopping cart carries a lot of goods and needs to meet certain strength conditions. The shopping cart is made of acrylic plate with a thickness of 5 mm, and four corner wheels are installed at the bottom. The splicing process refers to the tenon and tenon structure in the structural style of ancient Chinese wooden buildings. After splicing, it is fixed with special acrylic glue. We have adopted hollow out design at the bottom plate and surrounding fence, which reduces our own weight on the premise of ensuring the strength requirements. Finally, the robot's mechanical gripper controls the shopping cart by clamping the handle connected to the shopping cart body.

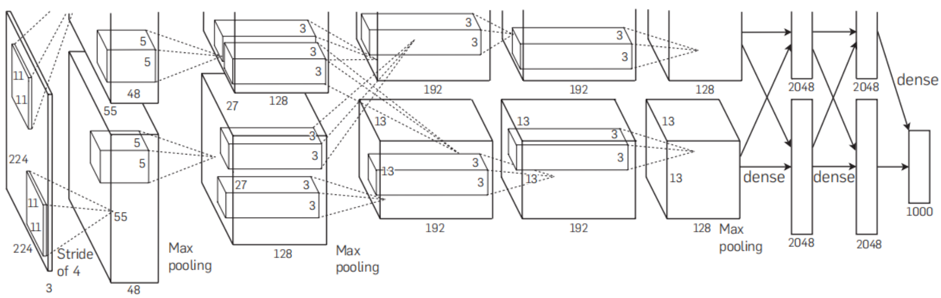
Figure 2 Structural design of shopping cart



2.1.5 Mechanical arm structure design

The supermarket shopping robot is installed and fixed with a mechanical arm on the third floor (Du et al., 2021), as shown in Figure 3, to grasp the target goods. The 30 kg steering gear fixed on the mechanical arm connector is connected with the mechanical arm steering gear control board and single chip microcomputer steering gear port through 3P steering gear extension line. The manipulator can rotate 360 degrees through the steering gear fixed on the third layer disc, so as to select, clamp and place the target goods. Each joint of the manipulator is equipped with a 30 kg 360 degree adjustable metal gear steering gear. The whole arm has three movable joints, which can occupy a small space during recovery and a large range of motion during grasping.

Figure 3 AlexNet network structure



The mechanical claw is controlled by a 30 kg 360 degree adjustable metal gear steering gear, which is meshed through gear transmission. The structural design has moderate opening distance and simple and ingenious clamping mechanism. The maximum opening width of the mechanical claw is about 98mm, which meets the size requirements for commodities. The material of the mechanical claw is thin, which can prevent the mechanical claw from touching the goods on both sides when clamping and picking up the goods in the middle of the cargo window.

2.2 Robot software design

2.2.1 Image recognition algorithm and application

The supermarket shopping robot we designed uses the AlexNet network based on Tensorflow learning system (Joseba et al., 2021) in image recognition algorithm. AlexNet realises the deep convolution network structure in large scale images dataset, which accelerates the promotion of deep learning theory. Alexnet network consists of 8 weight layer, containing 5 convolution and 3 full connection layers, as listed in Figure 3. The feature information of each convolutional layer is connected by multiple feature maps calculated in the higher layer. The main use of convolution layer is feature extraction (Zhu et al., 2018), which is defined as:

$$x_n^l = \sum_{i \in M_n} x_i^{l-1} * k_{in}^l + b_n^l \tag{1}$$

where x_n^l is the n characteristic diagram of layer l, M_n shows a set of feature maps; k_{in}^l represents the i elements in n convolution kernel in the layer l, "*" is the convolution operation.

Meanwhile, compared with the classical learning algorithm, it utilise ReLu as activation function to improve the training speed by about 6 times, it is given as:

$$\mathrm{f}(\mathrm{x}) = \max(0, x) \tag{2}$$

Alexnet presented a local normalisation layer, which is added after the activation of ReLu to create a competition mechanism for local neuronal activity and increase the generalization ability of the model. It is given by:

$$b_{x,y}^i = a_{x,y}^i \Bigg/ \left(k + \alpha \sum_{j=\max(0,i-n/2)}^{\min(N-1,i+n/2)} (a_{x,y}^j)^2 \right)^\beta \tag{3}$$

In addition, dropout is introduced into AlexNet to handle the problem of over-fitting in the process of model training (Srivastava et al., 2014). In the training phase, a probabilistic process is added to each unit of the training network. The network calculation formula without dropout is:

$$z_i^{(l+1)} = w_i^{(l+1)} y^l + b_i^{l+1} \tag{4}$$

$$y_i^{(l+1)} = f(z_i^{(l+1)}) \tag{5}$$

2.2.2 PID tracking control algorithm

PID regulation control is a simple linear control algorithm, which is composed of proportional P, integral I and differential D. It calculates the deviation e according to the object value R and the true value y, that is $e = r - y$, then performs proportional, integral and differential operations according to the obtained deviation, and finally obtains the output value U. Here we mainly use the position PID algorithm, and the adjustment algorithm formula is:

$$u = K_p * e(k) + K_i * \sum e(k) + K_d[e(k) - e(k-1)] \tag{6}$$

Including: K_p is the scale item parameter, K_p is the integral term parameter, K_d is the parameter of differential term; ek is the current deviation, $e(k-1)$ is the last deviation, and $\sum e(k)$ is the cumulative sum of $e(k)$ deviations. In the formula, the deviation e is the difference between the output value of the grey sensor and the target value. The car position is adjusted by PID algorithm to track accurately. Figure 4 shows the C language implementation of PID control algorithm.

Figure 4 C language implementation of PID algorithm (see online version for colours)

```
int Position_PID (int Encoder,float Position_KP,float Position_KI,float Position_KD)
{
   static float Bias,Pwm,Integral_bias,Last_Bias;
   Bias=Encoder;
   Integral_bias+=Bias;
   Pwm=Position_KP*Bias+Position_KI*Integral_bias+Position_KD*(Bias-Last_Bias);
   Last_Bias=Bias;
   return Pwm;
}
```

3 Test

3.1 Test device

To prove the effectiveness of the designed shopping robot, we propose a test scheme: first, set up a "shopping list", as shown in Table 1. In addition, we will also place some goods not in the "shopping list" as distractors. The robot is required to carry the shopping cart to find the target goods required in the 'shopping list" on the simulated site, grab them from the shelf, place them in the shopping cart and return them to the starting area. In the whole shopping process, the robot needs to maintain control over the shopping cart, and the goods captured from the shelf cannot change significantly. In addition, the robot is not allowed to disturb the shopping environment.

Table 1 Shopping list

No.	*Target goods to be purchased*	*Shelf area*
1	Red square brick (50×50 cube)	A, B shelf
2	Disturbed standard third-order Rubik's Cube	A, B shelf
3	Blue square brick (50×50 cube)	A, B shelf
4	Snow beer (empty can)	C, Dshelf
5	Red Bull (empty tank)	C, Dshelf
6	Yellow tennis	C, Dshelf
7	Wahaha AD calcium milk (empty bottle)	Warehouse E
8	Terentsu pure milk (empty box)	Warehouse E

The shopping environment we constructed is shown in Figure 5, which is a square structure composed of four supermarket shelf areas A, B, C and D and E warehouse area. Each shelf in the shelf area is segmented into two layers, containing 12 independent cargo windows. A maximum of 3 goods of the same kind should be placed at each cargo window. One of the non-empty cargo windows will be placed in the middle of the shelf cargo window. The details of goods placement in the cargo window are shown in Figure 6. The distribution of goods in each cargo window is mainly divided into the following six situations.

Figure 5 Diagram of simulated supermarket (see online version for colours)

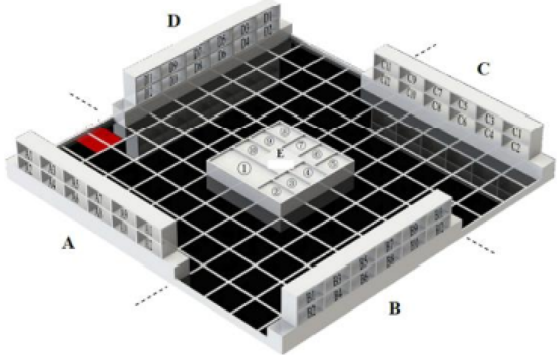
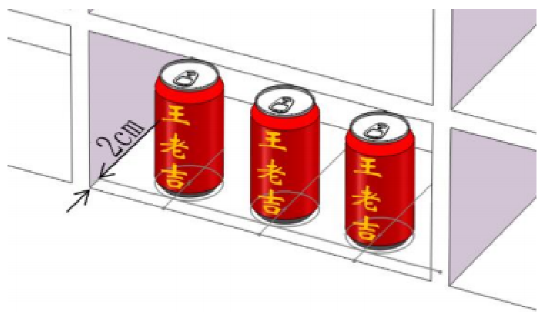


Figure 6 Diagram of cargo placement details in cargo window (see online version for colours)



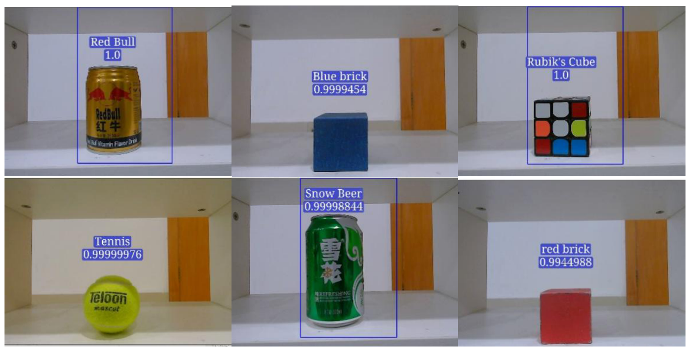
Robot path planning is mainly divided into inner and outer circles. Outer ring path: when the robot executes the outer ring path program, it needs to store the red square bricks (50) stored in four shelves A, B, C and D (50×50 cube), disturbed standard third-order magic cube, blue square brick (50×50 cube), snowflake beer, red bull, yellow tennis ball and other 6 categories of 18 goods are clamped and placed in the shopping cart and carried back to the starting area. Inner circle path: when the robot executes the inner circle path program, it needs to clip and place the class 2 and 6 goods such as Wahaha AD calcium milk and terentsu pure milk stored in the four platform areas of warehouse e into the shopping cart and carry them back to the starting area.

3.2 Experimental results and analysis

In the experiment, the precision of commodity recognition is the key factor for the success of shopping robot design. This work utilises AlexNet neural network to classify and identify goods (Frank et al., 2021). Therefore, this design constructs a small-scale database in the test process, which includes all the ‘shopping list” goods used in the experiment. The database contains 8 kinds of goods and several interferences. There are about 1500 images of each kind of goods from 0 to 7. Through the two real-time cameras of the photoelectric sensor and camera combination module of the supermarket shopping robot, the cargo window image information of the supermarket shopping robot driving and stopping at the corresponding white line position in the middle of each cargo window is photographed respectively. In view of the different number of goods in the cargo window, when designing the image recognition program, the image clipping function of open CV is used to set the coordinates of the fixed image clipping frame for the three windows that have determined the cargo placement position on the left, middle and right.

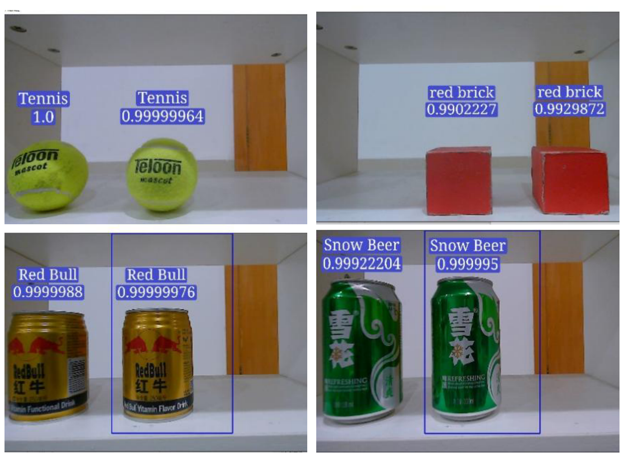
The image recognition program aims at the recognition results of 6 categories of single goods that may be placed on the four shelves a, B, C and D in the cargo window, as shown in Figure 7. Through the analysis and comparison of the real-time collected image with the trained database, the cargo type information of the image and the accuracy compared with the classified single cargo photo database in the database can be obtained.

Figure 7 Single cargo type identification of outer ring (see online version for colours)



The recognition results of the image recognition program for the two cases of placing the target goods in the outer circle are shown in Figure 8.

Figure 8 Identification of two cargo types and quantities in the outer ring (see online version for colours)



The recognition results of three target goods are shown in Figure 9. Through the recognition results of the above four figures, it is found that the accuracy of the neural network and image recognition program designed and used in this paper is often more than 99% when identifying the target goods in the “shopping list”, rather than the target goods in the “shopping list”. For interferents, the accuracy is often lower than 85%

(Aaisha and Neeraj, 2021; HyeokSoo and Jongpil, 2021). Therefore, by comparing the accuracy in Python, we can divide the final identified goods output into two categories:

1. Goods with accuracy higher than 85%, i.e., the output of the target goods identification result in the "shopping list" is 0, which means the goods that need to be clamped by the manipulator.
2. For goods with accuracy lower than 85%, i.e., the output of interference identification result is 1, which means goods that do not need to be clamped by mechanical arm.

Figure 9 Identification of three cargo types and quantities in the outer ring (see online version for colours)

4 Discussion

In many logistics warehouses or factory inventory systems, code scanning is mostly used to determine whether it is the target object. In the more complex working environment of supermarket, the working challenges that robots need to face are more severe. Recently, deep learning, which has attracted extensive attention, has promoted the success of our work. The experimental results prove the effectiveness of our shopping robot.

We constructed a simulated supermarket shopping environment using 8 kinds of 24 goods and some interferents with similar shapes and colours. The goods we selected include square bricks of different colours, disordered third-order Rubik's cube and some canned drinks. There is little difference in appearance, which makes it difficult to detect and identify these items. In the test, our supermarket shopping robot performs very well in the face of similar goods. It is believed that our structural design and method can work in large supermarkets with richer types of goods and higher similarity of goods after improvement and upgrading.

At present, our robot has the functions of autonomous line patrol, image recognition, automatic capture and so on. Different from the general supermarket service robot, the robot also has the ability to carry the shopping cart independently, and can play the role of replacing the manual cart in the supermarket. For the structural improvement, we stacked three photoelectric sensors on the mechanical claw. The height required for detecting goods can be adjusted by controlling the mechanical arm, and the top photoelectric sensor is used to detect whether there are goods. After image recognition, if the other two photoelectric sensors detect the item, the robotic arm will grab the item. Such a structural design is proved to be effective after testing.

However, our supermarket shopping robot is in the testing stage. During the movement of the robot, it can track accurately, the sensor detection is sensitive and the cargo image recognition is accurate, but these are all run in the testing environment. In real life, large supermarkets have large customer flow, dense shelves, and multiple supermarket shopping robots work together. Therefore, how to make our robots successfully complete their work in such an environment and realise the cooperation of multiple robots has become the next challenge.

5 Conclusion

This paper proposes a supermarket shopping robot based on deep learning. We propose the structural design of some new supermarket shopping robots. The supermarket shopping robot uses stm32f407zgt6 as MCU, and the Alexnet neural network structure plays a great role in the goods classification in the "shopping list". To verify the reliability of overall structure design, line patrol algorithm and image recognition algorithm of supermarket shopping robot car, we established a simulated supermarket shopping environment and collected the images of goods in the "shopping list" for experiments. Experimental simulation proves the feasibility of the supermarket shopping robot.

Acknowledgements

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions to improve this paper. Besides, this work is supported by the National Natural Science Foundation of China (No. 62102227, 51805124, 62101206), Zhejiang Basic Public Welfare Research Project (No. LZY22E050001, LZY22D010001, LGG19E050013), Science and Technology Major Projects of Quzhou (2021K29).

Conflict of interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

Aaisha, M. and Neeraj, K. (2021) 'PROTECTOR: an optimized deep learning-based framework for image spam detection and prevention', *Future Generation Computer Systems*, Vol. 125, pp.41–58.

Bavelos, A.C., Kousi, N., Gkournelos, C., Lotsaris, K., Aivaliotis, S., Michalos, G. and Makris, S. (2021) 'Enabling flexibility in manufacturing by integrating shopfloor and process perception for mobile robot workers', *Applied Sciences*, Vol. 11, No. 9, p.3985.

Cebollada, S., Paya, L., Flores, M., Peidro, A. and Reinoso, O. (2021) 'A state-of-the-art review on mobile robotics tasks using artificial intelligence and visual data', *Expert Systems with Applications*, Vol. 167, pp.114195.

Chen, A., Yang, B., Cui, Y., Chen, Y., Zhang, S. and Zhao, X. (2020) 'Designing a supermarket service robot based on deep convolutional neural networks', *Symmetry*, Vol. 12, No. 3, p.360.

Chen, Z. and Dong, R. (2021) 'Research on fast recognition method of complex sorting images based on deep learning', *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 35, No. 6, pp.2152005.

Du, Y.C., Taryudi, T., Tsai, C.T. and Wang, M.S. (2021) 'Eye-to-hand robotic tracking and grabbing based on binocular vision', Microsystem Technologies, Vol. 27, No. 4, pp.1699–1710.

Frank, H.D.M., Shimaalsadat, K.G., Karl, and Bernd, M. (2021) 'Identification of acoustic emission sources for structural health monitoring applications based on convolutional neural networks and deep transfer learning', *Neurocomputing*, Vol. 453, pp.1–12.

Gao, H., Wang, Z., Cai, L. and Ji, S. (2021) 'Channelnets: compact and efficient convolutional neural networks via channel-wise convolutions', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 43, No. 8, pp.2570–2581.

Garcia-Martinez, J.R., Cruz-Miguel, E.E., Carrillo-Serrano, R.V., Mendoza-Mondragon, F., Toledano-Ayala, M. and Rodriguez-Resendiz, J. (2020) 'A PID-type fuzzy logic controller-based approach for motion control applications', *Sensors*, Vol. 20, No. 18, p.5323.

Han, X., Zhong, Y., Cao, L. and Zhang, L. (2017) 'Pre-trained alexnet architecture with pyramid pooling and supervision for high spatial resolution remote sensing image scene classification', *Remote Sensing*, Vol. 9, No. 8, p.848.

Hu, H., Wang, T., Zhao, S. and Wang, C. (2019) 'Speed control of brushless direct current motor using a genetic algorithm-optimized fuzzy proportional integral differential controller', *Advances in Mechanical Engineering*, Vol. 11, No. 11, p.1687814019890199.

HyeokSoo, L. and Jongpil, J. (2021) 'Mobile robot path optimization technique based on reinforcement learning algorithm in warehouse environment', *Applied Sciences*, Vol. 11, No. 3, p.1209.

Joseba, G.M., Rubén, A., Olatz, M., Javier, A. and Ángelet, M. (2021) 'A generalization performance study using deep learning networks in embedded systems', *Sensors*, Vol. 21, No. 4, p.1031.

Kivrak, H., Cakmak, F., Kose, H. and Yavuz, S. (2021) 'Social navigation framework for assistive robots in human inhabited unknown environments', *Engineering Science and Technology-An International Journal-Jestech*, Vol. 24, No. 2, pp.284–298.

Li, X., Tiong, A.M.H., Cao, L., Lai, W.J., Phan, P.T. and Phee, S.J. (2019) 'Deep learning for haptic feedback of flexible endoscopic robot without prior knowledge on sheath configuration', *International Journal of Mechanical Sciences*, Vol. 163, p.105129.

Lin, C.J., Lin, C.H. and Wang, S.H. (2021) 'Integrated image sensor and light convolutional neural network for image classification', *Mathematical Problems in Engineering*, Vol. 2021, pp.5573031.

Ollah, M.V. and Karim, F. (2021) 'Scaling human-object interaction recognition in the video through zero-shot learning', *Computational Intelligence and Neuroscience*, Vol. 2021, pp.9922697.

Qi, L. (2020) 'Automatic recognition and classification algorithm of medical images based on neural network machine learning', *Journal of Medical Imaging and Health Informatics*, Vol. 10, No. 1, pp.123–128.

Ryumin, D., Kagirov, I., Axyonov, A., Pavlyuk, N., Saveliev, A., Kipyatkova, I., Zelezny, M., Mporas, I. and Karpov, A. (2020) 'A multimodal user interface for an assistive robotic shopping cart', *Electronics*, Vol. 9, No. 12, pp.2093.

Sahil, S. and Vijay, K. (2021) 'Performance evaluation of machine learning based face recognition techniques', *Wireless Personal Communications*, Vol. 118, No. 4, pp.3403–3433.

Sales, J., Marti, J.V., Marin, R.E. and Sanz, P.J. (2016) 'CompaRob: the shopping cart assistance robot', *International Journal of Distributed Sensor Networks*, Vol. 12, No. 2, pp.4781280.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R. (2014) 'Dropout: A simple way to prevent neural networks from overfitting', *The Journal of Machine Learning Research*, Vol. 15, No. 1, pp.1929–1958.

Wu, Z. and He, S. (2021) 'Improvement of the alexNet networks for large-scale recognition applications', *Iranian Journal of Science and Technology-Transactions of Electrical Engineering*, Vol. 45, No. 2, pp.493–503.

Yang, B. and Liu, F. (2021) 'A dual-typed and omnidirectional infrared perceptual network for indoor human tracking', *Measurement Science and Technology*, Vol. 32, No. 11, pp.115124.

Yu, X., Lyu, Z., Hu, D. and Xu, J.D. (2013) 'Scale-invariant feature transform based on the frequency spectrum and the grid for remote sensing image registration', *Mapping Sciences and Remote Sensing*, Vol. 50, No. 5, pp.543–561.

Zhang, D. (2021) 'Interoperability technology of sports health monitoring equipment based on multi-sensor information fusion', *Eurasip Journal on Advances in Signal Processing,* Vol (2021) No. 1, pp.1–18.

Zhang, H., Tan, J., Zhao, C., Liang, Z., Liu, L., Zhong, H. and Fan, S. (2020) 'A fast detection and grasping method for mobile manipulator based on improved faster R-CNN', *The International Journal of Robotics Research and Application*, Vol. 47, No. 2, pp.167–175.

Zhao, H., Zhang, H.X., Cao, Q.J., Sun, S.J., Han, X.Z. and Palaoag, T.D. (2021) 'Design and development of image recognition toolkit based on deep learning', *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 35, No. 1, p.2159002.

Zhu, L., Li, Z., Li, C., Wu, J. and Yue, J. (2018) 'High performance vegetable classification from images based on Alexnet deep learning model', *International Journal of Agricultural and Biological Engineering*, Vol. 11, No. 4, pp.217–223.