Design of workpiece recognition and sorting system based on deep learning

Zhoulin Chang1,*, Zixie Lu1, Fu Mo1, Linzhao Hao1, Gaofeng Yu2

1College of Electrical and Mechanical Engineering, Guangdong University of science and technology, Dongguan, China
2Huahangweishi Robot Technology Co., Ltd, Beijing, China

*Corresponding author e-mail: zhoulinchang@gdust.edu.cn

Abstract. Aiming at the actual needs of industrial robots in the handling process of sorting and picking according to the types of objects, a design scheme of visual inspection and workpiece recognition based on deep learning is proposed. The solution uses deep learning to model the object recognition system, builds the mapping relationship between the image and the object category, and realizes the effective recognition of the object image information. After the analysis is completed, the industrial robot classifies the object according to the given information. Experiments show that the use of deep learning image detection mode can realize the classification and picking operation of industrial robots, which can meet the needs of some actual scene operations.

1. Introduction
In the design of traditional robot handling workstations, the workstation equipment mostly uses a number of detection sensors plus a motor pneumatic actuator to detect and classify objects. This method uses a single scenario and is limited by the characteristics of the sensor itself. The designed system is expanded. If more sensors are installed in the system, there are still problems of difficulty in on-site installation of the handling station and high system construction cost. In order to overcome the above shortcomings, an object recognition and sorting system based on deep learning is designed. By applying the deep learning visual detection method to the industrial robot handling station, it overcomes the detection limitation caused by using a single sensor to detect objects to the greatest extent. It also solves the problem of poor information matching caused by multiple sensor detection. This design greatly improves the accuracy of the robot sorting system to perform sorting tasks while increasing the flexibility of the workstation

2. System structure design
The industrial robot material identification and sorting system is mainly divided into two parts, hardware implementation and software design. The hardware part is mainly composed of ABB industrial robot body, robot control cabinet, PC, camera, pneumatic gripper, suction cup and vacuum generator [1], as shown in Figure 1. The industrial robot control cabinet is the core component. The control cabinet is connected to the industrial robot body through the control cable and communicates with the PC through the network cable. The communication between the control cabinet and the PC
uses the Ethernet communication protocol based on TCP/IP; The camera for detecting objects is fixed on a dedicated bracket near the industrial robot and connected to the PC through a USB interface; the execution tool used for object handling is composed of pneumatic grippers and suction cups, both of which are connected to the vacuum generator and are used in industrial Under the control of the robot control cabinet, the objects are clamped and transported.

![System hardware structure block diagram](image1)

**Figure 1.** System hardware structure block diagram

The software part of the system is mainly composed of four parts, object detection module, image capture module, image recognition module, and industrial robot sorting and marking module. The workflow is shown in Figure 2. Specifically, when the control cabinet receives the detection signal of the object in place, the PC starts the camera to take pictures of the objects below it. After this step, the camera transfers the captured pictures to the PC through the USB protocol and passes the trained neural network model recognizes and analyzes the object; after the recognition is completed, the PC sends the type of the object to the control cabinet through the TCP/IP protocol; the industrial robot checks the relevant information after receiving the analysis signal from the PC and executes the corresponding Sorting action, and then complete the handling of items and item sorting operations.

![System software workflow diagram](image2)

**Figure 2.** System software workflow diagram

3. Object image recognition

The image recognition part of the object is mainly realized by the cooperation of the camera and the photoelectric sensor. When the photoelectric sensor detects that the object has reached the designated area, the sensor informs the industrial robot control cabinet through the I/O signal that the object has reached the designated position. After receiving the in-position detection signal, the control cabinet sends the PC through the Ethernet to turn on the camera. Then the PC will analyze and process the image transmitted by the camera through the deep learning network of the image feature recognition module.

3.1. Image feature recognition

Image feature recognition is the key to realize object recognition sorting workstation. This design adopts the method of retraining the existing deep learning model to reconstruct the image recognition model. The parameters of the image recognition model are re-adjusted to achieve the purpose of classifying and identifying the given object. The basic learning model used in this design is GoogLeNet. GoogLeNet is a deep convolutional neural network model based on deep learning
proposed by the Google team in the ILSVRC challenge in 2014 [2]. The model is composed of 22-layer network structure, which adopts a modular structure (Inception structure) for easy addition and modification; the final stage of the network uses average pooling (average pooling) to replace the fully connected layer, which has proven to improve accuracy; In order to avoid the disappearance of the gradient, the network adds an auxiliary Softmax to forward the gradient (auxiliary classifier), which is very helpful for the training of the entire network [3].

The GoogLeNet-V4 model uses the RMSProp optimizer and borrows the idea of Factorization into small convolutions to solve the 7x7 volume integral into two one-dimensional convolutions (1x7, 7x1), and also solve the 3x3 volume integral into two one-dimensional volumes Product (1x3, 3x1), which can speed up the calculation, and can split 1 conv into 2 convs, which further increases the depth of the network and increases the nonlinearity of the network; It is worth noting that the network input data size is 299x299, and 35x35/17x17/8x8 modules are more refined. The specific network model structure is shown in Figure 3.

![Inception Architecture of GoogLeNet (Inception v4)](image)

**Figure 3.** The inception architecture of GoogLeNet (Inception v4)

The Input layer in Figure 3 represents the data input layer. The image data that this layer can receive is an RGB three-channel picture with a resolution of 299x299x3 pixels. The design of the sorting system uses the same input layer as GoogLeNet. In order to meet the input conditions of the network, the system will segment and correct the size of the captured image before inputting the model, and adjust the pixel resolution of the image captured by the camera to a size of 299×299 pixels. In this design, the above-mentioned network hierarchy is inherited, and the subsequent dropout layer is also included in the network model to reduce the impact of over-fitting.

3.2. Image recognition model training

In order to make the vision module more suitable and to make up for the shortcomings of insufficient data collection in reality, the training and testing part of the convolutional neural network uses a mixed picture set composed of pictures collected by the network and pictures taken in the experimental scene. Conduct training and testing. The training model retains all the previous layers of GoogleNet, and only re trainable the latter three layers: the fully connected layer, the Softmax classifier, and the output layer.
The sample data used in the training is 5500 pictures of workpieces. There are six kinds of workpieces which 4,000 pictures are used for training and 1,500 pictures are used for testing. The training uses a powered stochastic gradient descent algorithm (SGDM) for training. The initial learning rate is 0.001, the minimum batch is 32, and the maximum number of training times is 2260. The training process is shown in Figure 4, and the accuracy of the training result is 95.20%.

![Figure 4. The classification model training process](image)

4. **Industrial robot sorting operation process**

After the deep learning network completes the analysis of the image data, the PC sends the name and category information of the object to the program data receiving end of the industrial robot through the TCP/IP protocol [5]. By monitoring and querying the information on the PC, the industrial robot can verify the received object image information after the PC sends the information and load the corresponding sorting strategy after confirming that the information is correct. If the robot is in an idle state, it performs the corresponding sorting action and transports the items to the corresponding material box. If the robot is busy, the industrial robot loads the corresponding strategy after completing the current task, recognizes the object, and completes the sorting operation. After the robot completes the sorting of the objects according to the pre-taught path, the robot sends the sorting completion instruction signal to the PC through the corresponding program module [6]. The specific design process of the robot sorting part is shown in Figure 5.
5. Conclusion

This article uses industrial robots as a carrier, combined with a camera and a PC to realize the design of a workpiece recognition and sorting system, and completes the purpose of sorting different kinds of items. On the object visual recognition algorithm, through the migration learning of the original deep learning network, the accurate recognition and sorting of objects is realized. The visual recognition system can be adjusted according to the actual situation. By changing the input data of the deep learning model and combining the corresponding industrial robot handling fixture, the workstation can realize the handling of different objects. From the actual test operation situation, the method has high recognition accuracy, clear and controllable sorting process, and the system has the advantages of reliable action, good stability, high recognition accuracy, and strong scalability, which can meet object sorting in many scenarios.

However, due to the impact of various types and complex environments, the design still has certain limitations. Sometimes it is difficult to express material characteristics well to complete accurate identification. It is necessary to further improve the deep learning method and expand the generalization of visual identification materials in order to coping with more factors including various types, blurred imaging, light differences, occlusion and overlap, etc. and adapting better to the requirements of flexible manufacturing.

Acknowledgments

This work was financially supported by the follow projects. The special innovation projects of ordinary universities in Guangdong Province. Project name: Research on visual recognition and sorting robots based on deep learning; Project number: 2018KTSCX260. Project of Guangdong University of science and technology. Project name: Research on sorting robot based on visual recognition technology; Project number: GKY-2019KYYB-7. Project of cooperation education with production China Ministry of education. Project name: Application Research of intelligent robot vision recognition technology; Project number: 201901019029.
References

[1] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 1–9, 2015.

[2] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. Rethinking the inception architecture for computer vision. arXiv preprint arXiv:1512.00567, 2015.

[3] C. Dong, C. C. Loy, K. He, and X. Tang. Learning a deep convolutional network for image super-resolution. In Computer Vision–ECCV 2014, pages 184–199. Springer, 2014.

[4] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 3431–3440, 2015.

[5] Ji Dengchen Ji. Research on scattered multi-objective sorting system based on machine vision [J]. Industrial control computer, 2018 (12): 34-35.

[6] Wenshan Hui. Robot tactile recognition and adaptive grasping control based on CNN-LSTM[J]. Journal of Instrumentation, 2019(1): 211-218.

[7] Xuqing Wu. Design of Intelligent Sorting System for Parallel Robots[J]. Mechanical and Electrical Engineering. 2019 (2): 224-228.

[8] Yuqiang Zhou. Part recognition and sorting system based on convolutional neural network[J]. Coal Mine Machinery. 2018 (6): 151-153.

[9] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei. Large-scale video classification with convolutional neural networks. In Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on, pages 1725–1732. IEEE, 2014.

[10] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems, pages 1097–1105, 2012.

[11] A. Lavin. Fast algorithms for convolutional neural networks. arXiv preprint arXiv:1509.09308, 2015.

[12] T. Tieleman and G. Hinton. Divide the gradient by a running average of its recent magnitude. COURSERA: Neural Networks for Machine Learning, 4, 2012. Accessed: 2015-11-05.