Scholar Publishing Group
Kinetic Mechanical Engineering
https://doi.org/10.38007/KME.2021.020206
ISSN 2790-2676 Vol. 2, Issue 2: 45-53

Improvement and Implementation of Machining and Positioning Method of Intelligent Construction Machinery Components Relying on Machine Vision

Pushpita Kumar*

Taif University, Saudi Arabia

**corresponding author*

Keywords: Machine Vision, Construction Machinery, Device Processing, Positioning Methods

Abstract: At present, industrial robots cannot achieve complete automatic assembly, for some specific occasions still need auxiliary facilities to cooperate. The identification and positioning of workpiece is the most important technical difficulty in automatic grasping of workpiece. This paper mainly studies the improvement and realization of the processing and positioning method of intelligent construction machinery components relying on machine vision. The overall scheme design of target workpiece positioning is presented. This paper introduces the technical background and significance of workpiece recognition and location based on machine vision. According to the actual engineering requirements of the workpiece, a monocular vision method of workpiece identification and positioning is designed. A localization algorithm based on image deep learning is designed. The deep learning network is used as the image feature extractor, and the extracted features are used to predict the corresponding Angle of the image by linear regression method. The experimental results show that the workpiece object detection method based on the improved deep learning model reduces the missing detection rate of small objects and improves the accuracy of the overall workpiece detection.

1. Introduction

In the current relatively mature industrial applications, most of the robot are fixed trajectory, according to the preset approach to work, to the environment change, the state of the target, and the end executor of the actual position of the lack of independent awareness, once the object in the working environment has changed more than expected, can affect the end of the robot positioning precision and work efficiency, As a result, current robots are mainly applied to relatively standard and repeatable tasks [1-2]. With the rapid development of intelligent manufacturing, the combination of robot and vision to achieve accurate recognition and positioning of target objects has become a hot spot of scholars at home and abroad. In the system of robot and vision fusion, the vision sensor will integrate and process the collected target information and convert it to the robot base coordinate system, so as to control the robot to realize the end positioning work. The positioning accuracy is an important index to evaluate its performance. However, geometric parameter errors are inevitably introduced in the process of robot machining, assembly and measurement. Meanwhile, the influence of non-geometric parameter errors such as robot joint flexibility, structural stiffness deformation and friction heat will further increase the pose errors of the end-effector, seriously affecting the positioning accuracy of its end [3-4]. In addition, the detection errors of the workpiece and other operating objects, and the error of the conversion matrix between the vision system and the robot control system will further reduce the positioning accuracy of the robot end, thus limiting the wide application of the system integrated with the robot and vision in industrial production [5-6].

With the advance of industrial automation, the technology of visual identification and positioning has gradually become mature. The traditional visual recognition and positioning method identifies the features of the target in the image and makes selection and determination by artificial restriction rules. However, this method can only be applied to simple pattern features, such as the circular Mark point is used to locate the PCB board, and the color domain difference is used to classify the workpiece with different colors. The area difference is used to identify the workpiece of different sizes [7-8]. Due to the variety of product designs, the features of target images are becoming more and more complex, and simple feature description can no longer adapt to the diversity of artifacts [9]. Image recognition technology uses template matching to adapt to the diversity of the workpiece, and feature extraction algorithms such as HOG, SIFT, SURF, ORB and LBP are applied to image processing. These algorithms use histogram features and descriptors to describe image features more accurately [10]. This kind of method shows good invariance in brightness, rotation, scale, affine and other aspects, and has good noise resistance and robustness, so that the image processing algorithm can adapt to more working conditions. Many scholars use this technology to solve industrial problems [11].

The combination of vision technology and robot technology is of great significance to increase the flexibility of robots in the production process, improve the work efficiency and the development of robot automation and intelligence. The acquisition, processing and analysis of the target information is the key to provide visual positioning and guidance for the robot.

2. Improvement of Machine Vision Processing Localization Method Based on Deep Learning

2.1. System Design

(1) Image acquisition system

In this paper, industrial robots are used to drill holes at specific positions of the target workpiece. In order to enable industrial robots to drill effectively on the workpiece, the most important thing is for industrial robots to get the position of the processing point [12]. The traditional industrial robot operating system often uses off-line Teaching method to obtain the position of the point to be processed, that is, teaching-then-work. The operating system needs to measure the position information of the point to be processed with the Teaching device in advance, and then the industrial robot completes the processing repeatedly according to the Teaching method. This kind of operating mode system has poor flexibility, low operating efficiency, and the accuracy cannot be guaranteed [13]. In order to improve the flexibility of the industrial robot and adapt to the change of the working environment, the processing system designed in this paper controls the industrial robot to carry out relevant operations by collecting workpiece information through machine vision, which has better flexibility and accuracy. This paper needs to rely on the vision system to complete the positioning of feature points on the workpiece, and use the image acquisition system to get the image of the workpiece to be processed. According to the image acquisition system, it can be divided into several categories as shown in Table 1 [14] according to the installation method, the number of industrial cameras and the control type.

Table 1. Image acquisition system classification

Classification	Category
Installation	Eye-in-Hand
	Eye-to-Hand
Number of camera	Monocular
	Binocular and multi-eyed
Whether the end is closed	End of the open loop
	End of the closed loop

This system mainly realizes the use of industrial robot driven pneumatic rotor operation, industrial camera to obtain workpiece position information, industrial robot according to the image processing information will be pneumatic tool to the specified position. It is mainly for the work pieces on the two-dimensional plane, so the image acquisition system adopts the open-loop vision scheme of the eye-to-hand end of monocular vision, that is, the monocular industrial camera is fixed at the top of the processing area.

(2) Visual system hardware selection

According to the overall scheme design of the system, the vision system in this paper includes eye-to-hand monocular vision system. In the application of machine vision system, industrial cameras are usually used for image acquisition. When selecting the industrial camera and lens, the resolution of the industrial camera, the focal length of the lens, the interface of the device and other factors should be considered. For the eye-to-hand monocular vision system in the system, the industrial camera is fixed on the workbench [15-16]. In the process of operation, the robot should be given sufficient space for action. The installation height of the industrial camera designed in this system is at least 1000mm above the horizontal plane of the table, and the visual field of the industrial camera on the robot operation plane can not be less than 200mm$\times$200mm. This system uses the industrial camera in machine vision to complete the recognition and positioning of the feature points on the workpiece. In this part, the selection of industrial camera is a very key link, its selection will affect the quality of the collected image, the quality of preprocessing results, image processing rate and drilling efficiency. According to the actual situation of the system in the processing operation, the field of view of the acquired image is 200mm$\times$200mm, and the accuracy requirement is 0.08mm/Pixel. According to the calculation formula (1) of the resolution of industrial cameras, the required minimum resolution can be calculated.

$$\lambda = l / d \tag{1}$$

λ- resolution; L - the width or length of the visual field; d - Precision size.

For industry lens selection, the related parameters of the camera have already decided on the selected the size of the lens and the kinds of connectors, this paper chooses industrial camera sensor target surface size of 1.1 ", according to the regulation of industrial camera choice face reading not less than 1.1 ", if there are less than the situation of industrial camera target surface size, In this case, the maximum aperture of industrial lenses cannot meet the needs of the camera, resulting in black circles at the edges and affecting image quality [17]. In terms of interface, the selected camera satisfies the C and CS interfaces. The distance from the industrial lens to the plane of the workpiece to be processed is called the working distance, and the size of this distance will affect the field of view of the image collected by the industrial camera. The working distance selected by this system is 700mm-750mm [18]. The focal length of the lens can be expressed as the distance between the main point and the focus after lens optics. The formula for calculating the focal length is shown in Equation (2) :

$$f = v \times D / V \tag{2}$$

Where: F -- focal length of lens; D - the distance between the lens and the target; V - longitudinal dimension of visual field; V - COMS vertical dimension.

2.2. Workpiece Location Algorithm Based on Deep Learning

The workpiece positioning algorithm based on deep learning first needs to collect enough images for the holding pose of a specific workpiece and the deviation Angle of the workpiece from the calibration position under each image. The above images are used as the training set for learning, and the corresponding Angle value of each image is used as the label, and then input to the network model for supervised learning, to obtain the prediction model for each type of workpiece. In the industrial production line, the workpiece deviation Angle can be obtained by inputting the collected pictures online into the prediction model.

Regression prediction is to take all factors affecting the target as independent variables, the target value as the dependent variable, through the analysis of historical data, using mathematical methods to express the function relationship between the independent variable and the dependent variable, and using this relationship to calculate the value of the new independent variable, the result is the target value to be predicted. In this project, the collected image can be taken as the independent variable and the corresponding Angle value as the dependent variable, and the relationship between image and Angle can be analyzed by the method of regression prediction. Common models include linear regression, logistic regression, ridge regression, lasso regression and elastic regression.

Since the whole regression positioning model requires large prediction measurement range (20 °) and high prediction accuracy (±0.01 °), it is difficult to take into account the requirements of large range and high precision prediction at the same time. Therefore, the whole prediction model is refined into two sub-models when designing the model: coarse correction model; Fine correction model.

The prediction range of the coarse correction model above is one period of the workpiece, namely (-10 °, 10 °), so the training samples are collected by means of a step size of 0.1 °, 200 images per cycle, 18 cycles, 3600 images in total. The prediction range of the precision correction model is (-1 °, 1 °), the sample sampling step is 0.01 °, 18 cycles, a total of 3600 images. The measurement range of this step is far greater than the accuracy of the coarse correction model, which is more conducive to ensure that the workpiece must be within the correction range of each step.

As an image feature extractor, the model of convolutional neural network mainly refers to the classical image classification network ResNet, which is different from the original network to perform 1000 class image classification tasks on ImageNet competition. The size of the input image in the experiment is 2000 × 60, and the final network value outputs 1 value as the result of the regression, and the loss function is modified to the absolute value of the error between the predicted value and the label. On the root processing module, the size of the convolution kernel is changed to 3×3, and the step size is changed to 1. The size of the residual learning module remains the same, and the number of convolution kernels is reduced.

The convolutional neural network for extracting the features of the convolution image is mainly the superposition of multiple residual learning units, using the continuous series of smaller convolution kernels, and the later feature map represents the more complex features in the image. Finally, the feature is input into the fully connected layer, and the prediction result is obtained by linear regression. In order to reduce overfitting, optimization methods such as drop-out and L2 regularization are used.

TensorFlow has been recognized by the majority of users due to its excellent performance, and it is also relatively outstanding in model design, interface, deployment, performance and architecture design. The framework is updated and iterated in real time rapidly, and there is an active community to learn from each other. Moreover, it can simultaneously meet the needs of industrial production and academic research, support distributed computing, has strong portability, support CPU, GPU and other platforms, so that all industries begin to use it for production practice. In view of this, this paper takes TensorFlow as the framework to build a deep learning network model for accurate positioning of artifacts.

3. Experimental Platform Building and Network Training

In this experiment, TensorFlow framework was used to build an improved deep learning neural network. The computer configuration of the experiment is shown in Table 2.

Table 2. Computer configuration

Configuration items	Configuration information
CPU	Intel Xeon Gold5115
GPU	Nvidia Quadro P4000
Memory	32GB
Disk	1T

The input image size is 224x224, the activation function is Relu, the batch size is set to 32, the regularization coefficient is set to 0.0003, the IOU of the prediction box is set to 0.6, the NMS is set to 0.3, and the total number of iterations is 20,000.

4. Analysis of Experimental Results

4.1. Comparison Before and After Improvement

As shown in Figure 1, before the model is iterated 5000 times, the accuracy of the machine learning model and the improved model in the artifact dataset is roughly equal, and when the number of model iterations exceeds 10000, the accuracy of the improved model in the dataset is slightly higher than that of the machine learning model in the dataset.

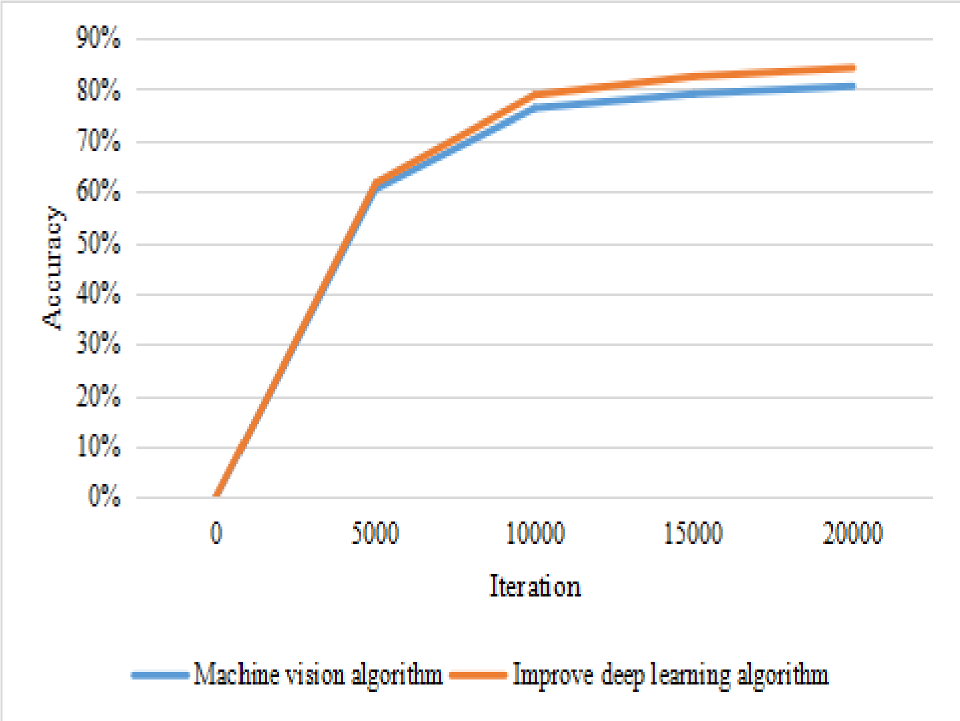

Figure 1. Relationship between iteration times and detection accuracy

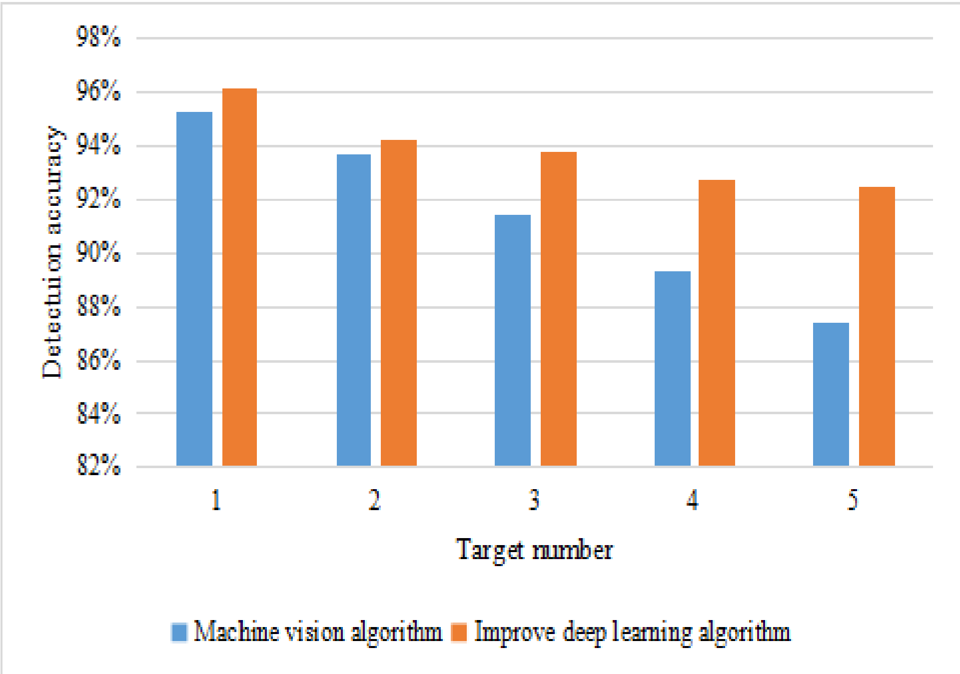

Figure 2. The relationship between the number of targets and the detection accuracy

As shown in Figure 2, the more the number of objects in a single image, the higher the probability of missed detection and the lower the accuracy of object detection. Therefore, the small number of targets has a great impact on the accuracy of target detection. However, the improved model is of great help to improve the probability of missing detection. When the number of objects is more than 2, the recognition accuracy of the improved model is higher than that of the SSD model. When the number of objects is 5, the detection accuracy of the improved model is still higher than 90%.

4.2. Test Results of Different Models

Table 3. Test results of each model

Model	mAP/%	FPS
CNN	93.2%	15
Machine vision algorithm	91.7%	29
Improve deep learning algorithm	94.6%	24

As can be seen from Table 3, after the algorithm is trained, the mAP value based on the improved model reaches 94.6%, and the accuracy rate is 2.9% higher than that of the machine vision algorithm. In terms of speed, the improved model is not as fast as machine vision, but higher than CNN. This is because the improved SSD speed is reduced by increasing the number of layers for feature extraction.

5. Conclusion

This paper mainly solves the problem of workpiece identification and positioning in the process of automated assembly, conducts an in-depth study on the detection and positioning of the target workpiece, establishes a monocular vision system, and detects the target workpiece through deep learning methods. This paper implements the identification and location of scattered workpiece, and the feasibility of the algorithm are verified through the experiment, but still there are some problems need to be further solved: identification of a target artifact result is not very stable, there may be some shade, overlapping artifacts cannot be accurately detected, some of the volume is too small target artifact may be ignored. Improved deep learning model workpiece target detection reduces the missed detection rate of small targets, but the increase of computation affects the speed of target detection. In the later stage, the basic network will be trimmed and optimized to a certain extent to further achieve the accuracy and real-time performance of target detection.

Funding

This article is not supported by any foundation.

Data Availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Conflict of Interest

The author states that this article has no conflict of interest.

References

[1] Jesse R, Polina D, Josh S, et al. 7 Characterization of Feeder Cattle Behavior Using an Integrated Machine Vision Learning System. Journal of Animal Science, 2021(Supplement_3):Supplement_3.

[2] T Yasuda, Y Adachi, T Uchida. Data Processing Program And Data Processing Device. Kyushu Institute of Technology, 2018, 3(30):27-32.

[3] Moshabaki A, Razi N, Kordnaeij A, et al. A model for B2B salesperson performance with service ecosystems perspective: a grounded theory. Journal of Business & Industrial Marketing, 2021, 37(6):1314-1337. https://doi.org/10.1108/JBIM-07-2020-0316

[4] Maconachie T, Leary M, Zhang J, et al. Effect of build orientation on the quasi-static and dynamic response of SLM AlSi10Mg. Materials Science & Engineering, A. Structural Materials: Properties, Misrostructure and Processing, 2020(788-):788. https://doi.org/10.1016/j.msea.2020.139445

[5] Pang M, Ma L, Zhai S, et al. Liquid-phase assisted laser preparation of 304 stainless steel surface texture and tribological properties evaluated. Industrial Lubrication and Tribology, 2021, 74(6):654-662. https://doi.org/10.1108/ILT-12-2021-0499

[6] Niu C, Wei C, Tan Y, et al. Dry-sliding tribological properties of a tin-bronze alloy produced by multiaxial forging and annealing. Industrial Lubrication and Tribology, 2021, 74(6):692-697.

[7] Gha B, Yang S A, Ygb C. Tailoring metallic surface properties induced by laser surface processing for industrial applications - ScienceDirect. Nanotechnology and Precision Engineering, 2019, 2(1):29-34. https://doi.org/10.1016/j.npe.2019.03.003

[8] Singh R P, Kumar N, Gupta A K, et al. Investigation into rotary mode ultrasonic drilling of bioceramic: an experimental study with PSO-TLBO based evolutionary optimization. World Journal of Engineering, 2021, 19(3):274-290. https://doi.org/10.1108/WJE-03-2021-0179

[9] Rana S, Hampson R, Dobie G. Breast Cancer: Model Reconstruction and Image Registration from Segmented Deformed Image using Visual and Force based Analysis. IEEE Transactions on Medical Imaging, 2019, PP(99):1-1.

[10] Zhang S, Shan J, Sun F, et al. Multimode fusion perception for transparent glass recognition. Industrial Robot: the international journal of robotics research and application, 2021, 49(4):625-633. https://doi.org/10.1108/IR-12-2021-0295

[11] Bianchi V, Ciampolini P, Munari I D. RSSI-Based Indoor Localization and Identification for ZigBee Wireless Sensor Networks in Smart Homes. Instrumentation & Measurement IEEE Transactions on, 2019, 68(2):566-575. https://doi.org/10.1109/TIM.2018.2851675

[12] Moongu S, Kwanghee K. Learning-based essential matrix estimation for visual localization. Journal of Computational Design and Engineering, 2021(3):3.

[13] Barrington A R, Chung D, Del Barrio J A, et al. Punctiform and Polychromatic Pre-Descemet Corneal Dystrophy: report of two families and the identification of a segregating intronic variant in PDZD8using whole-exome sequencing. Investigative ophthalmology & visual science, 2019(9):60.

[14] Liao H, Mesfin A, Luo J. Joint vertebrae identification and localization in spinal CT images by combining short-and long-range contextual information. IEEE transactions on medical imaging, 2018, 37(5): 1266-1275. https://doi.org/10.1109/TMI.2018.2798293

[15] Ozimek M, Juenemann A G M, Choragiewicz T, et al. The immunohistochemical identification and localization of homocysteine in the human retina with the features of retinal neurodegenerative diseases. Investigative Ophthalmology & Visual Science, 2019, 60(9):

1218-1218.
[16] Gajjar R, Gajjar N, Thakor V J, et al. Real-time detection and identification of plant leaf diseases using convolutional neural networks on an embedded platform. The Visual Computer, 2021, 38(8): 2923-2938. https://doi.org/10.1007/s00371-021-02164-9
[17] Bianchi V, Ciampolini P, De Munari I. RSSI-based indoor localization and identification for ZigBee wireless sensor networks in smart homes. IEEE Transactions on Instrumentation and Measurement, 2018, 68(2): 566-575. https://doi.org/10.1109/TIM.2018.2851675
[18] Idoudi M, Bourennane E B, Grayaa K. Wireless visual sensor network platform for indoor localization and tracking of a patient for rehabilitation task. IEEE Sensors Journal, 2018, 18(14): 5915-5928. https://doi.org/10.1109/JSEN.2018.2838676