TEST Engineering & Management

July - August 2020
ISSN: 0193-4120 Page No. 5587 - 5594

Research on Robot Path Planning and Design Based on Deep Learning Image Recognition Technology

HuazhenChen[1], GuoqingXia[1,*], Yanbiao Huang[2]and Wei Yang[1]
[1]South China Institute of Software Engineering of Guangzhou University, Guangzhou, Guangdong, China, 510990
[2]Guangdong Industry Polytechnic, Guangzhou, Guangdong, China, 510300

Article Info
Volume 83
Page Number: 5587 – 5594
Publication Issue:
July - August 2020

Article History
Article Received: *25 April 2020*
Revised: *29 May 2020*
Accepted: *20 June 2020*
Publication: *28 August 2020*

Abstract

With the gradual popularization of robot applications, the complexity and diversity of the task requirements of mobile robots are also increasing, and robots are required to have higher environmental adaptability and mobility in the unstructured complex environment. Based on this, this paper first analyses the concept of deep learning and its application in image recognition and path planning. Secondly, it studies the analysis of image recognition technology based on deep learning. Finally, it gives the robot path planning and design strategy based on deep learning image recognition technology, and verifies the adaptability of deep learning in robot autonomous path planning and design.

Keywords: *Robot Path Planning, Deep Learning, Image Recognition Technology ;*

1. Introduction

With the rapid growth of artificial intelligence technology, the intelligent and autonomous level of robot is constantly improved. The mobile robot which can independently carry out path planning has obtained more research and progress, and has been gradually applied and popularized in many fields. Robot with autonomous path planning based on deep learning technology plays an important run in many fields, especially in the consumer market. Autonomous path planning robot is mainly based on image recognition and radar positioning to achieve the mission requirements, and it has been more in-depth development in the domain of human-computer environment interaction, service and military fields.

The robot based on deep learning image recognition technology help realize the recognition and reconstruction of terrain environment information, and plan and process the robot's path and action through the real-time control system. The path planning and design of this kind of robot mainly has different types, such as remote control, relying on external information or completely autonomous path control strategy. Obviously, the higher the degree of automation, the greater the difficulty and complexity. The traditional mobile robot based on remote control signal and external information input is easy to control, with relatively simple mechanical structure and low work energy consumption, which has strong advantages of cost and technical feasibility. However, with the increasing complexity and diversity of the task requirements of mobile robots, the requirements for autonomous path planning ability of mobile robots are constantly strengthened, so that robots can use more complex unstructured use environment and conditions.

In the unstructured complex environment, the application of traditional robot is greatly limited, and it is urgent to further improve its environmental adaptability and mobility. With the in-depth

Published by: The Mattingley Publishing Co., Inc.

July - August 2020
ISSN: 0193-4120 Page No. 5587 - 5594

development of mobile robots in several fields as shown in Figure 1, the intelligent autonomous path planning and control of mobile robots have brought new development. Based on these technologies, the adaptability and mobility ability of mobile robots in complex terrain environment can be greatly improved.

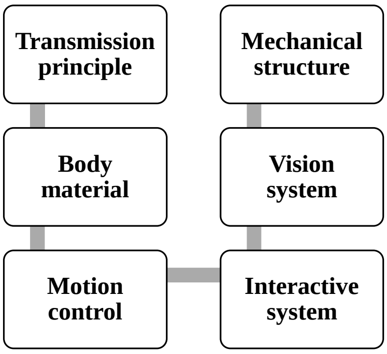

Figure 1. The support fields of mobile robots.

As a mobile robot that can realize autonomous perception, thinking and action, the application of deep learning, artificial intelligence and image recognition technology makes it have the conditions and functions of autonomous and flexible action similar to human brain. In recent years, intelligent mobile robots with the ability of perception, planning and control can independently collect environmental information through perceptron, make planning through internal procedures, and control its own actions. Therefore, it has been widely used in many industrial fields and civil consumption markets.

Path planning technology is to control the robot itself according to the given task or changing environment[1]. As an important factor of mobile robot's activity ability, image recognition technology can significantly affect its autonomous ability of environment perception, environment modeling and path planning. At present, there are many optimization algorithms in the research of robot path planning, but most of the research methods have limitations. The image recognition technology of deep learning can combine the kinematics and dynamics characteristics and the environment characteristics of the robot to select the path efficiently and in real time. Image recognition technology is based on the ability of automatically processing image information to realize the image analysis and recognition of the environment in which the robot is located.

In short, machine learning has brought a lot of intelligent path planning methods. From machine vision to natural image processing, deep learning technology has a broad application in the field of mobile robot path planning and design, which provides a new development direction for mobile robot path planning, so it has important research value.

2. Path planning based on deep learning

2.1. Deep learning algorithm

Deep learning combines low-level features to form more abstract high-level representation attribute categories or features to discover distributed feature representation of data. Deep neural network contains neural networks with multiple hidden layers. Deep learning can express a much larger set of functions than shallow networks in a more compact and concise way. Deep learning network emphasizes the depth of the model structure, usually has multi-layer hidden layer nodes, which clearly highlights the functions of feature learning.

Compared with the method of constructing features by artificial rules, deep learning algorithm uses big data to learn features, which can better describe the rich internal information of data. By learning a deep nonlinear network structure, it can realize complex function approximation and represent the distributed representation of input data.

2.2. The similarities and differences between deep learning and neural network

Deep learning and neural network both adopt hierarchical structure, as shown in Figure 2 below, that is, deep learning and neural network systems both contain multi-layer networks, and nodes of adjacent layers are connected, while nodes of other layers are not connected with each other, and each layer can be regarded as a logistic regression model.

Published by: The Mattingley Publishing Co., Inc.

July - August 2020
ISSN: 0193-4120 Page No. 5587 - 5594

The differences between them are shown in Table 1 below.

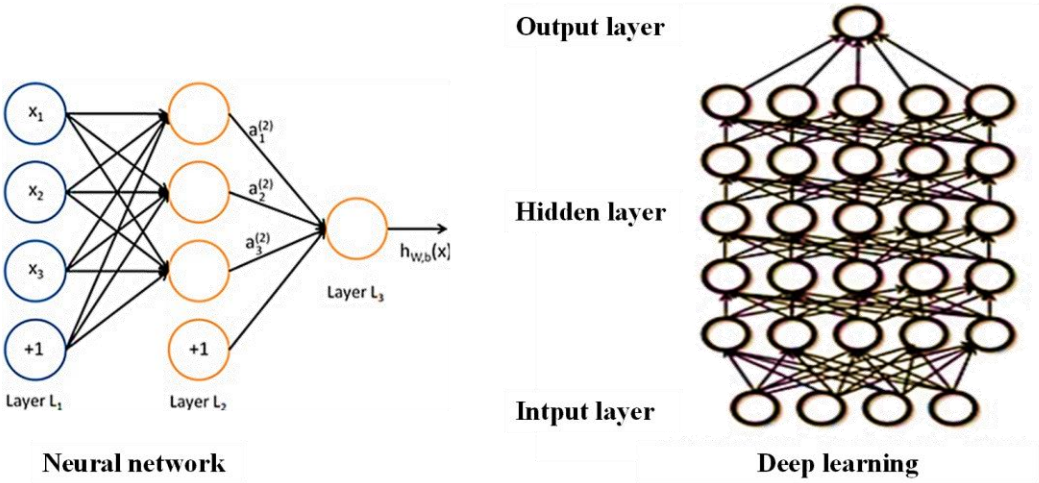

Figure 2. Hierarchical structure of neural networks and deep learning.

It is difficult to adjust the parameters of neural network because it is easy to over fit and needs many skills. The training speed is relatively slow, and the effect is not better than other methods in the case of less levels[2]. In addition, because the BP algorithm needs label data to train, but most of the data is unlabeled, and the gradient becomes more and more sparse and the convergence is easy to reach the local minimum when the feedback is adjusted, so the deep learning training method does not use BP algorithm.

Table 1. The differences betweendeep learning and neural network.

Aspects	Deep learning	Neural network
Parameters adjusting	Wake-sleep algorithm	BP algorithm
Trainingmethod	Layer by layer training	Iterative algorithm
Initial value setting	Feature learning	Random setting

2.3. Deep learning training process

The deep learning training process is based on bottom-up unsupervised learning, and a single layer of neurons is constructed layer by layer based on wake sleep algorithm[3]. The step flow of wake sleep algorithm is shown in Figure 3 below, in which the cognitive process of wake stage realizes the reconstruction of weight information, while the generation process of sleep stage and the creation of abstract scene.

In addition, based on the top-down supervised learning, the parameters of each layer can be improved and the parameters of the whole network can be adjusted. The deep learning model obtains the network parameters by unsupervised learning the framework of input data, so that the initial value comes close to the global optimum, so as to achieve initial targets.

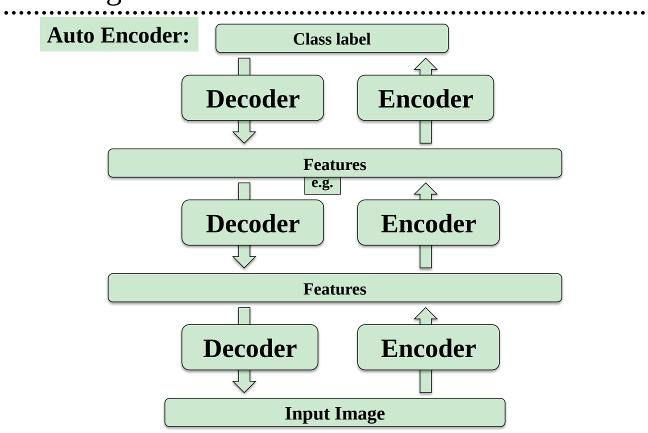

Figure 3. The step flow of wake sleep algorithm.

2.4. The concrete model of deep learning

The specific models of deep learning include automatic encoder, sparse automatic encoder, and

Published by: The Mattingley Publishing Co., Inc.

noise reduction automatic encoder. The automatic encoder generates a top-down path through output characteristics, realizes image/function input through decoding, and realizes decoding based on feed-forward bottom-up path to complete the whole cycle. The process architecture of the automatic encoder is shown in Figure 4.

Automatic encoder. In the training stage of sparse automatic encoder, a set of bases is obtained based on a given series of sample images, and the K-SVD method is used to adjust alternately until convergence, so as to obtain a set of dictionaries which can well represent the series. Secondly, in the coding stage, based on the given new image and the dictionary obtained, OMP algorithm is used to solve a lasso problem to obtain sparse vector[4].

As for the noise reduction automatic encoder, the noise of training data is increased, so that the automatic encoder must learn to remove this noise and obtain the input of filtering noise. This requires the encoder to learn the robust representation of the input signal. It can be seen that noise reduction automatic encoder usually has stronger generalization ability.

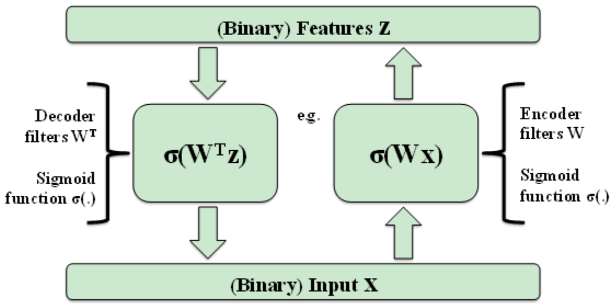

Figure 4. The process architecture of the automatic encoder convolution neural network (CNN).

2.5. CNN has become the main topic in the field of image recognition

The weight sharing network structure of CNN reduces the complexity of the network model and the network parameters by unsupervised learning the framework of input date[5]. The input image is convoluted with three trainable convolution kernels and additive bias, as shown in Figure 5 below. Then, after sub sampling, the feature map of C1 layer is weighted and biased, and then three feature maps of S2 layer are obtained through a sigmoid function.

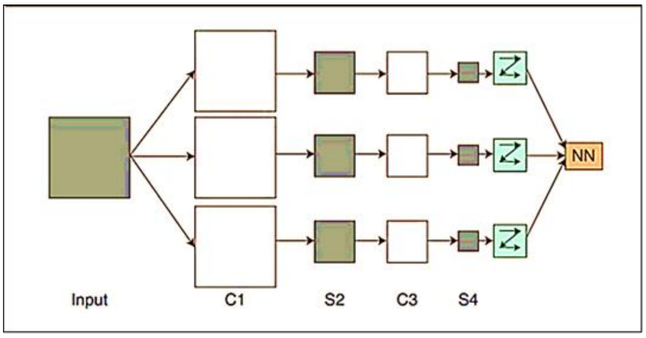

Figure5. Schematic diagram of convolution neural network.

Convolution neural network can reduce parameters and share weights, and avoid explicit feature sampling[6]. Based on structure reorganization and weight reduction, feature extraction function is integrated into multi-layer perceptron, so it can directly process gray images for image-based classification. In addition, the input image based on convolutional neural network can well match the topology of the network, and the weight sharing makes it more adaptive.The learning process of convolutional neural network mainly uses two learning rules of forward-propagation and back-propagation to optimize the weights and learn an optimal filter to extract features. The output of the forward propagation network layer is as follows:

$$x^l = f(u^l), \ in\ which, u^l = W^l x^{l-1} + b^l \quad (1)$$

In which, $f(.)$ is the output activation role of the network.The back-propagation algorithm will give a label to each training sample, and train and change the weight by adjusting the error between the network output and the given label. The cost function of training samples is as follows:

$$E^N = \frac{1}{2}\sum_{n=1}^{N}\sum_{k=1}^{2}(t_k^n - y_k^n)^2 \quad (2)$$

For the nth training sample among N training samples, its cost function is expressed as follows:

$$E^n = \frac{1}{2}\sum_{k=1}^{2}(t_k^n - y_k^n)^2 = \frac{1}{2}\|t^n - y^n\|_2^2 \quad (3)$$

The sensitivity of the error to the bias basis b and the derivative of the error to the input u of a node

July - August 2020
ISSN: 0193-4120 Page No. 5587 - 5594

are equal[7]. For the non-output layer, the sensitivity of the first layer can be expressed as:

$$\delta^l = (W^{l+1})^T \delta^{l+1} \bullet f'(u^l) \quad (4)$$

The partial derivative of the cost function corresponding to each weight in the convolutional neural network is calculated:

$$\frac{\partial E}{\partial b} = \frac{\partial E}{\partial u}\frac{\partial u}{\partial b} = \delta \quad (5)$$

In the pooling layer, each feature map is subsampled, so the output feature map is as follows:

$$x_j^l = f(\beta_j^l \, down(x_j^{l-1}) + b_j^l) \quad (6)$$

After calculating the sensitivity of each neuron in the pool layer, adding all the elements in the sensitivity map, it could get the following results:

$$\frac{\partial E}{\partial b_j} = \sum_{u,v} (\delta_j^l)_{uv} \quad (7)$$

If the corresponding feature map is saved in the forward propagation, there is no need to calculate in the back-propagation process.

3. Image recognition technology of deep learning

Image pattern recognition is to analyze the image content and find out the typical elements and features in the image. Each object is detected by image segmentation, and their image is separated from the rest of the scene. Then the object is measured, and some important characteristics of the object are quantified based on the calculation. Finally, the category of each object is determined. Image recognition is mainly based on the pattern recognition of the image, through the input image information to establish an image recognition model, analyze and extract the characteristics of the image, and then establish a classifier, and finally complete the classification and recognition according to the characteristics of the image.

3.1. Image recognition pre-processing and recognition classification

Image recognition pre-processing will train or test the image to a certain extent, making it clearer or more suitable for algorithm requirements[8]. Common pre-processing methods include de-noising, image enhancement and normalization. The classifier cannot recognize the pixel level shallow signal; only recognize the high-level signal, so it could extract useful high-level features from the training data. The recognition and classification of image recognition is based on the process that the algorithm forms a classification standard after certain training, which can classify the images to be identified in the test set into a certain category. The general process of image recognition is shown in Figure 6.

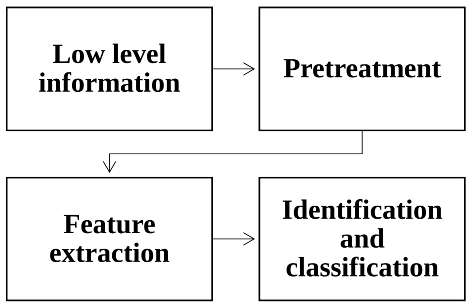

Figure6. The general process of image recognition.

3.2. Image recognition method based on deep learning

The spatial connection of the image is local. Based on the local receptive field, each neuron can only feel the local image area, and then at a higher level, these neurons that feel different parts are integrated to obtain the global information of the image. The perceptual region of neural unit in image recognition based on deep learning comes from some upper neural units.

3.3. Local receptive field.

The receptive field is the size of the visual perception area. In the convolution neural network, the definition of receptive field is the area size of the original image mapped by the pixels on the feature map of each layer of the convolution neural network, as shown in Figure 7 below. The filter size and step size of all previous network layers has great influence on the size of receptive field in the deep convolution layer. In the calculation, the size of image padding is ignored. In the calculation, the size of image padding is ignored.

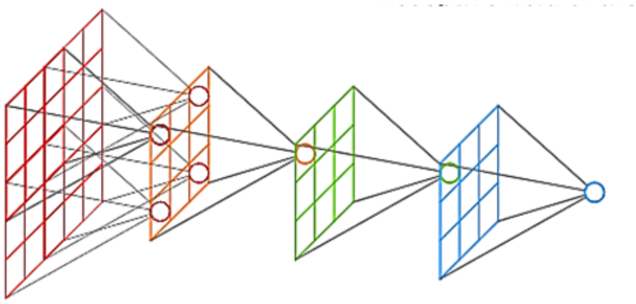

Figure 7. Characteristic graph of output of each layer of convolutional neural network.

3.4. Shared weights

The number of parameters in the hidden layer with shared weights related to the size of the filter and the number of filter types[9]. Multiple filters are needed to extract different features. The parameters of each filter are different, which means that it proposes different features of the input image. Feature map is to de-convolute the image with each filter to get the reflection of different characteristics of the image. Generally speaking, there are as many convolution kernels as there are feature maps to form a layer of neurons[10].

3.5. Pooling

According to the principle of image local correlation, only one pixel in a neighbourhood of an image can express the information of the whole region. The image has static property, and can take the average value or maximum value of the features in a certain region of the image. Pooling is the process of realizing this aggregation. If the continuous range in the image is selected as the pooled region, and only the features produced by the repeated hidden units are pooled, then these pooling units have translation invariance. Even after a small translation, the pooled features will still be generated. A typical example of image pooling is shown in Figure 8.

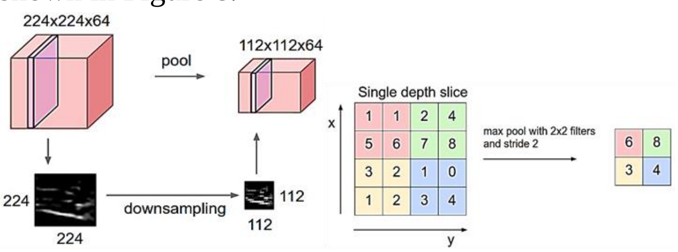

Figure 8. Image pooling - sampling process.

3.6. Rasterization

After the image is pooled and subsampled, a series of feature maps are obtained, and the input of MLP is a vector. Therefore, it is necessary to take out the pixels in these feature images and arrange them into a vector. The typical process of rasterization is shown in Figure 9.

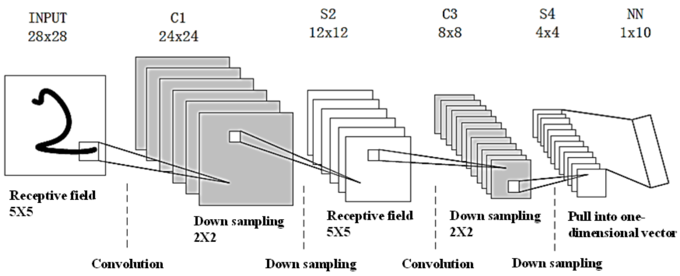

Figure 9. The typical process of rasterization.

3.7. Methods of training algorithm

Firstly, the network initialization is carried out, that is, in the forward propagation stage, a sample is taken from the sample set and input into the network to calculate the corresponding actual output. Secondly, in the backward propagation stage, the difference between the actual output and the corresponding ideal output is calculated. Finally, the weight matrix is adjusted by back propagation based on the method of minimizing error.

3.8. Gradient descent method

To search the maximum value of the function as soon as possible from any point, you should search the direction where the function changes the fastest[10]. If the function is a unitary function, the gradient is the derivative of the function:

$$\nabla f(x) = f'(x) \tag{8}$$

If it is a binary function, the gradient is defined as:

$$\nabla f(x_1, x_2) = \frac{\partial y}{\partial x_1} i + \frac{\partial y}{\partial x_2} j \tag{9}$$

If it needs to find the minimum point of the function, it should look in the direction of negative gradient. Gradient descent method has its own shortcomings.

3.9. Typical examples

Published by: The Mattingley Publishing Co., Inc.

July - August 2020
ISSN: 0193-4120 Page No. 5587 - 5594

For example, in order to calculate the gradient, it is necessary to calculate the gradient value for each training input separately, and then calculate the average value. When the number of training inputs is too large, it will take a long time, which will make learning very slow.

4. Robot path planning and design based on deep learning image recognition technology

4.1. Robot motion decision model

The construction of robot motion decision-making model is to make autonomous path planning decisions based on the environment information in structured and unstructured roads, so as to improve the robot's adaptive ability to the environment. Based on the local receptive field, weight sharing and pooling layer of convolutional neural network, the parameters such as displacement, scaling and distortion invariance in the process of robot surrounding environment image recognition are obtained respectively. Among them, the local receptive field output feature map of pixels on the original image mapping area. Weight sharing can effectively improve the efficiency and speed of training process. Pooling is used to reduce the resolution of features so as to keep the original features after image processing.

4.2. Design of robot autonomous path planning

Firstly, it is necessary to make decision on the basis of the complexity of the system, the category of the system planning, the complexity of the system, and so on. Secondly, based on convolution neural network, the robot can recognize the image of the surrounding environment, so as to help the robot to make its own path planning. In addition, based on the training of convolution neural network, a perfect robot path planning network model is obtained to predict the path data and control the robot motion based on the prediction results.

4.3. Dynamic path planning algorithm for robot

First of all, in the global path planning level of robot, neural network algorithm is used to meet the accuracy requirements of large-scale working environment because of its strong learning ability in path planning. Secondly, in the local path planning level of the robot, the image recognition information of the robot's surrounding environment is obtained of the deep learning algorithm, which lays the foundation for the analysis and calculation of the path planning information.

4.4. Implementation of dynamic programming algorithm.

The dynamic programming strategy is used to select the path and find the optimal solution. Firstly, the optimal solution of robot autonomous path planning problem is defined, and the typical characteristics of the optimal solution are described. Secondly, based on the deep learning image recognition technology, the optimal value of the problem is obtained from the bottom up, and the information in the calculation process is recorded. Then based on the obtained optimal value and its information, the optimal solution of the problem is obtained.

4.5. Error analysis

In the process of image recognition training of deep learning, it is easy to produce over fitting, so it is necessary to enhance the image in environment recognition. In order to ensure that the recognition image has rotation invariance, the classification layer and the previous convolution layer are not frozen in the training process, so that the model can adapt to the changes of images from different angles. The robot path planning based on deep learning image recognition technology can acquire the original images of the robot's surrounding environment for network training, which makes the robot's motion decision-making real-time.

5. Conclusion

In summary, in the unstructured complex environment, the application of robot is greatly limited, and it is necessary to further improve its environmental adaptability and mobility. With the usage and R&D of deep learning, artificial intelligence and image recognition technology, it has the conditions and functions of autonomous

Published by: The Mattingley Publishing Co., Inc.

recognition environment similar to human brain. The robot based on deep learning image recognition technology can realize the recognition and reconstruction of terrain environment information, and plan and process the robot's path and action through the real-time control system.

In this paper, through the studies of deep learning concepts and its application in image recognition and path planning, and analyzes the algorithm, model and training process of deep learning. Secondly, through the analysis of image recognition technology, the image recognition method based on deep learning is studied. Finally, the robot path planning and design strategy image recognition technology is studied, including robot motion decision-making model, autonomous path planning and design, dynamic planning algorithm and error analysis, which verifies the effectiveness of deep learning in robot autonomous path planning and design.

Acknowledgments

1. 2019/2020(NaturalScience)CharacteristicInnovation ProjectsofOrdinaryUniversitiesinGuangdongProvince"ResearchonAutonomousNavigationMechanismof MobileRobot"（2019KTSCX245）、"Research on Intelligent Control System of Multi-Posture Service Robot";

2. 2019 Scientific Research project of South China Institute of Software Engineering of Guangzhou University "Research on Intelligent Control System of Multi-Posture Robot" (KY201905);

3. 2018Thekeyscientificresearchplatformandscientificresearchprojectofordinaryuniversitiesin GuangdongProvince"ResearchonImageEnhancementTechnologyBasedonMachineLearning"(2018KQNCX395).

References

[1] K. He, G. Gkioxari, P. Dollar, Mask R-CNN. IEEE Transactions on pattern analysis & machine intelligence, 2017, (99):1-1..

[2] H. S. Li, Y. Zheng, Simulation of robot path planning optimization method under complex path. Computer simulation, 2014,31 (01): 407 - 410.

[3] Y. Zhang, E. C. Dai, Y. Luo, Path planning of mobile robot based on improved genetic algorithm. Computer measurement and control, 2016, 24 (1): 313-316.

[4] X. D. Li, M. Ye, T. Li, A review of target detection based on convolutional neural network. Computer application research, 2017, 34 (10): 2881-2886.

[5] Q. Cheng,G. Jiang, Three-dimensional mechanical analysis of electro-hydraulic hexapod robot. Mechanical transmission, 2015 (1): 43-47.

[6] X. Z. Hu, C. J. Dong, Overview of path planning methods for mobile robots based on Intelligent Technology. Western China Science and technology, 2015, 14 (10): 105-107.

[7] A. Rajkomar, S. Lingam, A. G. Taylor, High-throughput classification of radiographs using deep convolutional neural networks. Journal of Digital Imaging, 2017, 30(1):95-101.

[8] O, Bailo, F. Rameau, K. Joo, Efficient adaptive non-maximal suppression algorithms for homogeneous spatial keypoint distribution. Pattern Recognition Letters, 2018, 106:53-60.

[9] L. Bruzzone, P. Fanghella. Mantis: Hybrid leg-wheel ground mobile robot. Industrial Robot, 2014, 41(1).

[10] B. Zeng, Y. M. Yang, Real time path planning method based on Ant Colony Algorithm in dynamic environment. Computer application research, 2010, 27 (3): 860-86.