Proceedings of the 2nd International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/93/20240928

The application of Deep Learning in the field of Robotics

Haochen Liu

China university of Mining & Technology, Beijing, No.11 Ding, Xueyuan Road, Haidian District, Beijing, 10083, China

2817300802@qq.com

Abstract. As a new field with rapid development in the past two decades, deep learning has received more and more attention from researchers. Neural network adopts widely interconnected structure and effective learning mechanisms to simulate the process of information processing in the human brain, which is an important method in the development of artificial intelligence. Compared with shallow models, deep learning can achieve stronger unsupervised autonomous learning capabilities by increasing the number of layers of the network. In the past, the robots needed to be controlled by the humans; they controlled the robot to finish different tasks. In recent years, the researchers have tried to apply deep learning in the field of robots. Deep neural networks can simulate the complex cognitive laws of the human brain, so many researchers apply deep learning to robots to make robots have the thinking ability of humans. This article will review the cases of combining robot and in-depth learning in recent decades, Summarizing the achievements and problems faced in the three fields of robot vision, trajectory planning and motion control.

Keywords. robot, deep learning, robot vision, motion control, trajectory planning.

1. Introduction

Since the 21st century, the robotic domain has drawn more and more attention from different national governments and institutions because of its great development prospects. The United States, China, Japan, South Korea, Europe and other governments have put forward the development of a robot plan: in China, the ministry released in December 2021, the focus of the robot industry development planning is to encourage innovation, to make China become the leader of the global robot industry development by making unremitting efforts. The International Federation of Robotics (IFR) has released the World Robotics Report 2023. The report records 553,052 industrial robots installed in factories worldwide, an increase of 5% from a year earlier. By region, 73 percent of the newly deployed robots are installed in Asia, 15 percent in Europe, and 10 percent in the Americas. The US government has also proposed the US National Robot Program (NRI), seeking to conduct research on robot systems, as well as Japan's "new robotics strategy, Germany's High-tech Strategy 2025 (HTS), all reflect the importance governments attach to the development of robotics.

The original robots were industrial robots, which people relied on to handle repetitive work on the production line, such as assembly and welding. With the continuous improvement of technology, the commercialization degree of industrial robots is gradually improved, and gradually toward industrialization, and then promoted in various process links in the large-scale production represented

DOI: 10.54254/2755-2721/93/20240928

by the automobile manufacturing industry, such as handling, painting, arc welding, etc. These advances have eased the lack of labor force in various countries after World War II. This kind of robot not only has higher production efficiency than normal workers, but also has higher accuracy. With deepening research in the field of robots, the robot began to be employed in service, agriculture, health care, education, military, fire, and other fields, such as medical robots can help medical staff to complete the operations, agricultural robots can help farmers complete picking, irrigation, rescue robot can deep normal people cannot reach the area, quickly found trapped people. As the demand continues to rise, Robots are also growing, and breakthroughs are needed in areas such as robot vision, trajectory planning and motion control. Traditional robots are no longer able to meet the growing demand of people. Robots are starting to become intelligent, and researchers hope that robots can think like people and take different measures in different situations.

With the improvement of computing power and the advent of the era of big data, more and more researchers have begun to pay attention to the field of deep learning. The core idea of deep learning is to imitate the operation of the human brain. By increasing the number of layers of the network, the machine can automatically learn from data an understand abstract concepts, Deep learning (Deep Learning, DL) is a kind of machine learning. Its algorithm is based on data representation learning. It can discover complex structures in big data, and use back propagation to guide the machine on how to represent the previous layer network, so as to change the internal parameters of each layer[1]. The purpose of deep learning is to give robots (Robotics) the same analytical learning ability as humans, because they can enhance their perception, decision-making and control ability[2].

This paper takes the application of deep learning in the robot field as the main line, reviews the classic cases of combining deep learning and robot in recent decades, focuses on the application of deep learning in robot vision, trajectory planning and motion control, and discusses the future development.

2. Application of deep learning in three fields of robotics

2.1. Robot vision

Over the past several decades, robot vision has evolved significantly as a critical component of robotic systems, enabling the extraction and processing of information from the external environment. Since 1960, researchers have focused on advancing this technology; however, initial progress was impeded by the substantial hardware requirements and processing times needed to handle vast amounts of visual data.

Beginning in the 1980s, renewed interest in robot vision spurred global research efforts, leading to advancements in methods, theories, and applications such as Optical Character Recognition (OCR) and intelligent camera systems. As a result, the 1990s witnessed the establishment of dedicated vision companies and the development of image processing products.

Subsequently, increased investment in robot vision technologies for manufacturing processes catalyzed rapid expansion in the field. Hundreds of enterprises began marketing machine vision systems, ultimately fostering the formation of a comprehensive robot vision industry. Recent developments in LED lighting sensors, and control structures have further accelerated the advancement of robot vision technology, reducing production costs and driving ongoing growth in the sector. Robot vision is to extract the information obtained from the outside world, and then transfer it back to the robot through special technology processing. Since 1960, researchers have been to on robot vision. Due to the huge amount of information of the visual system, the huge hardware system for processing these information takes a long time, so robot vision has been unable to be applied to practice. Since the 1980s, robot vision has set off a global research boom, updated methods and theories, OCR and intelligent camera in this stage, and gradually triggered the wider dissemination and application of robot vision related technology. In the early 1990s, the vision company was founded and developed image processing products. Later, the robot vision related technology has been continuously invested in the manufacturing process, which makes the robot vision field expand rapidly.

Hundreds of enterprises began to sell the machine vision system in large numbers, and a complete robot vision industry has gradually formed. At this stage, the rapid development of LED lights, sensors and control structures has further accelerated the progress of the robot vision industry, and gradually reduced the production cost of the industry.

In recent years, the deep learning has been wider and wider used in the field of robot vision. The robot vision means that robot acquires external environment information through visual sensors and converts it into digital signals for processing and analysis, so as to realize the understanding and cognitive ability of the environment. The applications of the deep learning used in robot vision are as fellows below:

Target recognition: the Deep learning has a big advantage in the area of it by running model training, such as deep convolutional neural network.

Visual SLAM: It could relax more accurate and robust autonomous localization and mapping tasks by using deep learning.

Object grasping: It could realize more efficient and accurate object grasping by training deep learning model.

2.1.1. Picking robot target recognition

In [2], it describes the current most advanced technology in regards to the application of deep learning methods to generalised robotic grasping and discusses how to improve the overall performance of the robot grasp detection by every element of the deep learning. the system obtains an RGB-D image from a Kinect mounted on the robot, and searches over a large space of possible grasps, for which some candidates are shown. For each of these, it extracts a set of raw features corresponding to the color and depth images and surface normals, then uses these as inputs to a deep network which scores each rectangle. Finally, the top-ranked rectangle is selected and the corresponding grasp is executed using the parameters of the detected rectangle and the surface normal at its center. Red and green lines correspond to gripper plates, blue in RGB-D features indicates masked-out pixels. [2]

Modern fruit cultivation urgently needs an automatic device capable of automatic picking to improve the production efficiency of picking and sorting. Therefore, a picking robot target recognition technology based on machine vision and deep learning,which can realize the automatic recognition of tangerine. [3]

In his study, the deep learning framework is used to process and analyze the collected images, mainly based on deep learning. The author mainly used the SSD algorithm to detect the fruit target. Different from other target detection algorithms, Its prediction real box does not include the candidate target area and domain. Different sizes can be used to identify and measure the target object. SSD uses the small convolution algorithm to detect and identify the target feature border offset and classification, and repeatedly screens the repetition boxes by using the non-maximum suppression algorithm. The convolutional neural network feature graph is relatively large, which is suitable for identifying small targets.

2.1.2. Visual SLAM based on deep learning

Deep learning is widely used in visual SLAM, including: combining deep learning with visual SLAM front-end to improve the accuracy of image feature extraction; integrating object recognition, target detection, speech segmentation and other technologies, and then increasing the perception and understanding of the surrounding environment information to estimate monocular visual depth through a deep neural network.

In[4] Godard proposed a convolutional neural network instead of training directly with depth map data, using the easily accessible angle of binocular stereo vision to estimate the depth of individual images, so as to perform end-to-end unsupervised monocular depth estimation. It improves its own robustness.

DOI: 10.54254/2755-2721/93/20240928

In[5], the author proposes the Semanticfusion method, which performs semantic segmentation using CNN, plus the conditional random field, and designs the semantic mapping system, which is a fusion scheme for CNN and state-of-the-art dense SLAM.

In [6], Zhou proposes an unsupervised learning network for monocular depth and pose estimation that features single-view depth estimation and multi-view pose estimation from unstructured video sequences in the case of an unsupervised neural network. Optimal grasping position detection method based on deep learning.

2.1.3. Optimal grasping position detection method based on deep learning

In recent years, the visual robot image processing technology has developed from simple binary image processing to large data volumes of high resolution image processing, and microelectronics, artificial intelligence, artificial neural networks, the rapid development of the discipline, visual information interaction technology mature, visual robot grasping application laid a good foundation for it. [7] Deep learning-based image processing methods have the ability of continuous autonomous learning and efficient problem solving, deep learning has been widely used in position detection, and image segmentation based on deep learning And the recognition method has gradually become the current mainstream image processing method.

In 2014, Ross Girshick et al. proposed the R-CNN model, which located the target for image recognition and image segmentation with the advantages of simple operation and fast [8].

In [9], Chinese Academy of Sciences proposed an optimal grasping position detection method based on deep learning, and conducted experimental verification on UR 5 robot.

2.2. Motion control

Motion control has evolved with science and technology. It began in the mechanical era and progressed to electrical and compute control eras. Currently, we are in the era of intelligent control, where artificial intelligence and machine learning have made motion control more intelligent and human-like.

Artificial intelligence is in a good era of innovation and development, which will inject new momentum into the convergence and development of a new round of technological and industrial revolution, and at the same time give more possibilities for the new generation of intelligent robots to realize intelligent perception and automatic control. Nowadays, robot motion control using deep reinforcement learning, deep reinforcement learning (DRL) method is attracting the attention of more and more researchers as an emerging research field.

The application of DRL used in motion control could be divided into five parts: Autonomous navigation, object grasping, gait control, man-machine collaboration, and group collaboration[10].

2.2.1. Autonomous navigation

Compared with traditional ways, DRL can provide a low- cost scheme based on learning for robot autonomous navigation in a dynamic unknown environment through the effective use of environmental information. In the case of not relying on map information, Li put forward a mobile robot path planning method based on DQN and visual servo, with the initial environment image and target image for input, through the training to establish the corresponding relationship between each other and control strategy, to complete the indoor dominant, navigation task, compared with the traditional visual servo method, it has strong robustness and generalization ability[11].

2.2.2. Gait control

The use of Deep Reinforcement Learning (DRL) presents a viable solution to address the challenges of insufficient accuracy in complex dynamics modeling and motion trajectory tracking in robot gait control. Training data is gathered by introducing random discrete actions, which are then used to establish an offline estimator for pose identification.

Compared to traditional wheeled robots, foot robots offer greater flexibility in walking, maneuvering through barriers, and working in broad spaces, such as jungles and mountains. Gait control is a critical technology for bionic robots, and traditional gait control methods typically require point selection, trajectory optimization, and operating space control modules. These methods also need to be adjusted for different movement tasks. DRL can solve the challenges of complex dynamic modeling and trajectory tracking accuracy in robot gait control, providing an alternative solution. The DRL could provides another feasible solution to solve the problems such as the insufficient accuracy of complex dynamics modeling and motion trajectory tracking in robot gait control. The training data is collected by randomly introducing discrete actions, and a pose identification model is initially established as an offline estimator.

Foot robot compared with the traditional wheeled robot, has more flexible the ability to walk, barrier and broad working space in the jungle, mountain and other special environment, gait control is capable of the action of bionic robot has one of the key technology, the traditional gait control method usually need to contact point selection, trajectory optimization and operating space control work module, and in the face of different movement tasks need to adjust the design, and DRL also to solve the complex dynamic modeling and robot gait control trajectory tracking accuracy, provide another feasible solution.

2.2.3. Man-machine collaboration

By predicting the possible actions that humans perform at a certain state, the DRL method establishes a more intuitive cognitive model for the robot movement control of human- machine cooperation, so that the human-machine cooperation is closer to the interaction between human beings.

In rehabilitation robot human-computer interaction, neural networks can be used to analyze brain-computer signals. The structure of artificial neural network mainly includes network input layer, network hidden layer and network output layer. In the early stage, artificial neural networks had few hidden layers, so the structure is generally relatively simple. With the increase of hidden layers and the proposal of various network structure models, many types of neural network models have gradually emerged, among which the convolutional neural network model has been widely used in the analysis field of EEG signals.[12]

2.2.4. Group collaboration

Cooperative organization refers to the phenomenon of bird colony and ant colony clusters in the biosphere, so researchers have carried out in-depth research on cluster control. The DRL method finds a new way in the development path of group collaborative control among multiple robots by reducing human analysis and analysis in traditional collaborative control and deriving modeling process.

In the study of cluster cooperative control, in [13], Peng propose a dynamic surface control method of neural network in terms of single trajectory guidance, and designed an unmanned boat distributed formation controller, which improves the transient performance of neural network learning and control input signals. In[14], In terms of multi-trajectory guidance, Peng proposes a distributed state feedback and output feedback inclusion controller based on an iterative learning neural network for one-directional network connectivity with model uncertainty and Marine environment disturbance

2.3. Trajectory planning

Trajectory planning can be regarded as a solution frominput to output. The robot system inputs are the kinematics, kinetic parameters and the desired trajectory, and the output is the robot end-effector including the time series of displacement, speed and acceleration and the amount of motion of each joint. A trajectory planning is a curve obtained when the position of a point in space changes continuously over time.

At present, the most promising application is the AI trajectory planning method of deep learning. Deep learning is one of the fields of artificial intelligence. It is a technology such as human learning ability through computers. It has been mainly used in the field of signal processing and image processing[15].

The below are some examples that deep learning used in trajectory planning.

BAE proposed a multi-robot multi-trajectory planning algorithm combining Deep q reinforcement learning with convolutional neural network algorithm in [16]

In [4], It introduces an intelligent vehicle model transfer trajectory planing method to address the problem of the model error and tracking dependence in the process of intelligent vehicle motion planning. It introduce a network structure .For complex continuous state space Σ(Sensor) and continuous action space $\Sigma(\delta)$, it is necessary to train the deep reinforcement learning model M θ in a virtual environment M using the DDPG algorithm. The DDPG algorithm consists of an Actor policy network and a Critic evaluation network. [17]

Grigorescu proposed NeuroTrajectory based on the neural network, which is able to learn the vehicle local state trajectory from the occupancy grids and the corresponding driving commands, in [18].

3. Conclusions

Deep learning has significantly impacted various aspects of robotics, including robot vision, motion control, and trajectory planning. In the realm of robot vision, deep learning algorithms have been primarily employed for object detection, visual Simultaneous Localization and Mapping (SLAM), and object grasping. By autonomously processing image data from the external environment, these algorithms enhance the robot's ability to perceive and comprehend its surroundings, leading to improved accuracy in target detection. Within the domain of motion control, Deep Reinforcement Learning (DRL) has been applied to autonomous navigation, gait control, human-robot collaboration, and multi-robot collaboration. The application of DRL enables robots to achieve more human-like movements, enhance control performance, and streamline movement processes. In trajectory planning, deep learning techniques have been predominantly utilized in signal processing and image processing. A promising direction in this field is end-to-end path planning based on sensor data, target identification, and mapping. This approach allows robots to autonomously determine the most efficient path, resulting in significant cost reductions. Overall, deep learning has the potential to revolutionize various facets of robotics, ultimately leading to enhanced performance and more versatile applications.

References

[1] LECUN Y ,BENGIO Y ,HINTON G.Deep Learning[J].Nature ,2015 ,521(7553) :436-444
[2] Lenz I ,Lee H, Saxena A .Deep Learning for Detecting RoboticGrasps[J].The International Journal of Robotics Research, 2013, 34(4-5).DOI:10. 1177/02783 64914549607.
[3] Azkarate Saiz A .Deep learning review and its applications[J]. 2015.
[4] GODARD C , MAC AODHA O ,BROSTOW G J . Unsupervised monocular depth estimation with left -right consistency[C]/ / Proceedings of the IEEE conference on computer vision and pattern recognition . 2017: 270-279 .
[5] HOLMES S ,KLEIN G ,MURRAY D W . A square root unscented Kalman filter for visual monoSLAM [C]/ /2008 IEEE International Conference on Robotics and Automation .IEEE,2008: 3710-3716 .
[6] ZHOU T ,BROWN M ,SNAVELY N ,et al .Unsupervised learning of depth and ego - motion from video [C]/ /Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 1851 -1858 .
[7] Lu Rongsheng, Shi Yanqiong, Hu Haibing. Review of three-dimensional imaging techniques for robotic vision [J]. Progress in Laser and Optoelectronics, 2020,57 (04): 9-27.
[8] Girshick R , Donahue J , Darrell T , et al . Rich feature hierarchies for accurate object detection and semantic segmentation[C]. Computer Vision and Pattern Recognition ,2014: 580-587 .

[9] Du Xuedan, CAI Yinghao, Lu Tao, et al. A deep learning-based robotic arm grasping method [J]. Robotics, 2017,39 (06): 820-828 + 837
[10] Dong Hao, Yang Jing, Li Shaobo, etc. Progress in robot motion control based on deep reinforcement learning [J].Control and decision-making,2022,37(02):278- 292.DOI:10 . 13195/j.kzyjc.2020. 1382.
[11] Li Y, Jana K S. Learning view and target invariant visual servoing for navigation[C]. The IEEE International Conference on Robotics and Automation. Paris: IEEE, 2020: 658-664
[12] Dong Xian-Guang-Jianguang. Implementation of EEG signal detection and brain-computer interface based on convolutional neural network [D]. Jinan: Shandong University, 2016.
[13] PENG Z H, WANG D, LI T S. Predictor-based neural dynamic surface control for distributed formation tracking of multiple marine surface vehicles with improved transient performance[J]. Science China Information Sciences, 2016, 59(9): 92210.
[14] PENG Z H, WANG D, SHI Y, et al. Containment control of networked autonomous underwater vehicles with model uncertainty and ocean disturbances guided by multiple leaders[J]. Information Sciences, 2015, 316: 163–179.
[15] Niu Qichen, Zhang Gong, Zhang Gongxue, etc. Review and development trend of multi-robot trajectory planning research [J]. Machine Tool and Hydraulic Pressure, 2021,49 (12): 184-189.
[16] BAE H, KIM G, KIM J ,et al.Multi-robot path planning method using reinforcement learning[J]. Applied Sciences ,2019 ,9(15) : 3057 .
[17] Yu L , Shao X , Wei Y ,et al.Intelligent Land Vehicle Model Transfer Trajectory Planning Method Based on Deep Reinforcement Learning[J].Preprints, 2018 (9).DOI:10.20944/PREPRINTS201808.0049.V1.
[18] Grigorescu S, Trasnea B, Marina L, et al. Neurotrajectory: A neuroevolutionary approach to local state trajectory learning for autonomous vehicles[J]. IEEE Robotics and Automation Letters, 2019, 4: 3441– 3448.