Секция 3. **Радиотехника, связь, информационные технологии**

УДК 004.896

Ali Rawan, master student
E-mail: eng.rawan.sy@gmail.com
Al Akkad M. Aiman, CSc in engineering, associate professor
E-mail: aimanakkad@yandex.ru
Department of Computer Science
Kalashnikov Izhevsk State Technical University, Izhevsk, Russian Federation[1]

Machine Learning in Robotic Grasping Tasks: A Survey

Recently machine learning techniques, including deep learning and reinforcement learning, have been considered as the milestone in the field of computer vision and vision-based robot tasks, such as grasping. This survey presents a set of recent approaches in the field of object pick-and-place tasks and object grasping tasks. These approaches are categorized into two groups, deep-learning-based approaches, and reinforcement-learning-based approaches. Task-oriented grasping decisions for humans, are made intuitively, while it is a big challenge for robots to achieve the grasping tasks as proficient as humans. Several conditions affect the performance of robot grasping such as changes in environment and illumination, existence of a huge number of objects with different properties, complex backgrounds, and occlusion between objects Machine learning techniques are implemented in robotic systems to improve the capability of these robots to handle these conditions and guarantee high performance.

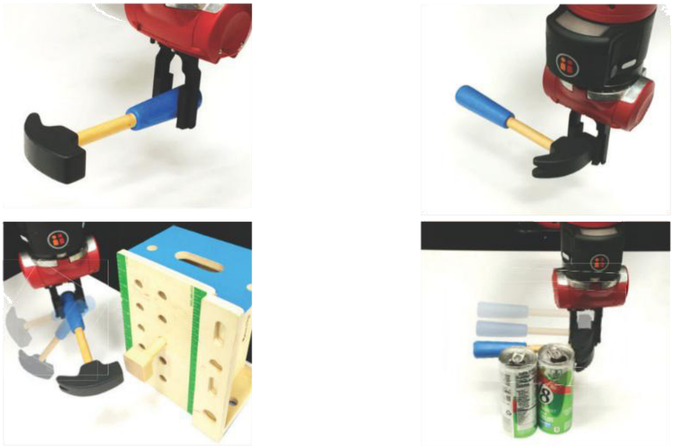
Keywords: Robotics, Machine learning, Deep learning, Reinforcement learning, Grasping tasks, Object detection.

1. Introduction

In the last few years, the use of robots in grasping and pick-and-place tasks is rapidly increased, either in the industrial fields [1] or the other life activities such as assistive robots [2].

Grasping task requires a set of decisions depend on the aim of the manipulation task. Fig. 1 [4] shows that a robot grasps the same object in two different ways according to the manipulation task; wherein Fig. 1a the task is hammering, while in image Fig. 1b it is sweeping. This type of grasping known as task-oriented grasping [3, 4].

For human, these decisions are made intuitively, in contrast, it is a big challenge for robots to achieve the grasping tasks as proficient as humans. Several conditions affect the performance of robot grasping such as changes in environment and illumination, a huge number of objects with different properties, complex backgrounds, occlusion between objects, etc. [5]



a b

Fig. 1. Example of task-oriented grasping

Therefore, machine learning techniques are implemented in robotic systems to improve the capability of these robots to handle these conditions and guarantee high performance. Variety approaches have emerged with great breakthroughs in machine learning algorithms.

Deep learning (DL) is part of machine learning techniques. Based on algorithms for learning multiple levels of representation in order to model complex relationships among data. Inspired by the biological nervous system, deep learning model consists of a network of parallel and simultaneous mathematical operations are performed directly on the available data to obtain a set of representational heuristics between the input and output data. These heuristics are then used in decision making [6, 7].

Reinforcement learning (RL) is an area of machine learning inspired by behaviorist psychology. RL is a type of dynamic programming that trains algorithms using a system of reward and punishment. In another expression, a robot with a reinforcement learning algorithm learns by interacting with its environment. It receives rewards by performing correctly and penalties for performing incorrectly. The robot, over time, makes decisions to maximize its reward and minimize its penalty using the dynamic programming [8, 9].

In this survey, a set of most recent researches related to the implementation of DL and RL in the field of robotic grasping tasks are discussed.

2. Deep Learning in Object Grasping

A lot of advancements were made in the latest year in the field of vision-based techniques in robotic grasping using deep learning. S Wang et al. proposed a model for robotic grasping based on fully Convolutional Neural Network (CNN) using high-resolution RGB-D images (400×400) for each pixel. The proposed method uses one encoder extracts the features from the original image, such as color image, depth image, grasp position map, grasp angle map, and grasp width map; and one decoder which outputs these pixel-wise parameters. According to this implementation, the results show a high accuracy of about 94.42 % for image-wise and 91.02 % for object-wise and fast prediction time about 8 ms. Also, this model offers the ability to generate a robotic grasping for various objects without training directly [10].

Y. Xu, et al. present a GraspCNN approach using a single CNN. The proposed algorithm is a grasp pose localization algorithm to detect oriented diameter circle back to 6D grasp pose in a point cloud directly from the RGB image. In basic this method combined in design between the standard CNN and YOLO model to achieve the best object detecting performance in a clustered environment. AS a result of demonstrations this model shows a 96.5 % accuracy, high speed, and stability [11].

Z. Zhao, et al. developed a grasp network approach that was a mix between two types of networks. The first one Grasp Prediction Networks (GPNs) depends on the standard CCNs and Mixture Density Networks (MDNs). This GPNs uses for prediction samples based on depth image mapping to a set of features for the Gaussian Mixture Model (GMM) from groups of grasp points can decide which candidate group can be a sample. The second one Grasp Evaluation Networks (GENs) which integrates the GPNs work by evaluating the candidate groups and select the optimal one. The experimental results show a high quality of grasping with GPNs and a high resolution of evaluating with GENs. However, the cost of these designs and implementation is high compared with other methods also there are lots of limits in dealing with depth image mapping [12].

P. Sharma, et al. address the design approach in how to improve 3D object detection for a robot arm by utilizing 2D machine learning methods. The design methods consider the 3D parameters of the images into 2D sets of information, and process the object detection to obtain high accuracy performance of identification and distance information for the Navigation and grasp of the robot arm. Deep Convolutional Neural Network (DCNN) is proposed with the usage of RGB-D images, which includes one 2D-image in RGB, and the other in-depth image form. Two methods are proposed, one will be using two parallel DCNN model and another method will be using three parallel DCNN model. Afterwards, the parallel models need to be concatenated to get 3D-object detection [13].

Y. Song, et al. introduced an effective single-state grasp detection network based on region proposal networks from Faster R-CNN. The proposed approach consists of two steps; firstly, multiple oriented reference anchors are generated. Then, the grasp rectangles are regressed and classified based on these anchors. The performance of this approach is evaluated based on the Cornell grasp dataset and the Jacquard dataset, and the experiment results show high grasp detection accuracies [14].

3. Reinforcement Learning

While collecting robotic grasping dataset via human labeling can be quite challenging, as there may be multiple grasping regions in the object and human notions of grasping are biased by semantics. Therefore, some researchers try to train robotic grasping models through trial-and-error experiments which can auto-collect a huge amount of training data.

D. Kalashnikov, et al. suggest a closed-loop vision-based grasping method using a scalable self-supervised deep reinforcement learning algorithm. The introduced method can exploit over 580k real-world grasp attempts to train a deep neural network Q-function with over 1.2M parameters to perform closed-loop, real-world grasping that generalizes to 96 % grasp success on previously unknown objects. Unlike the static learning behaviors that choose a grasp point and then execute the desired grasp, this method enables closed-loop vision-based control, whereby the robot continuously updates its grasp strategy based on the most recent observations to optimize long-horizon grasp success. Experiment results prove that the proposed method can be generalized effectively for complex real-world grasping tasks [15].

A. Zeng, et al. introduced a new approach to achieve a synergy between pushing and grasping. This approach is based on a pixel-wise version of deep networks that combines deep reinforcement learning with affordance-based manipulation. Experiment outcomes prove that the system learns to perform complex sequences of pushing and grasping on a real robot intractable training times [16].

A. Rajeswaran, et al., considering a set of 4 object manipulation tasks of human-like five-fingered hand, shown in Fig. 2 [17], developed a model-free reinforcement learning model augmented with human demonstrations collected in virtual reality [17].

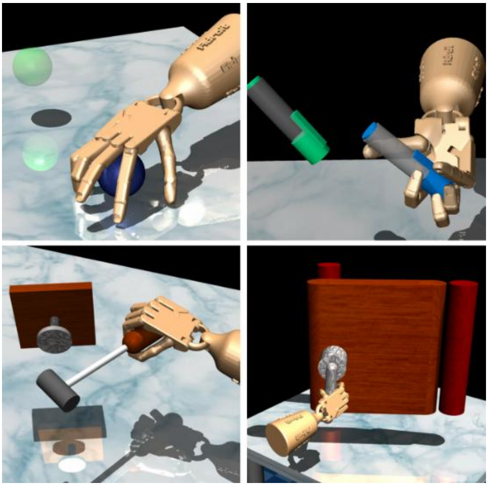


Fig. 2. Grasping Tasks that robot supposed to achieve

4. Conclusion and Future Work

In this survey, a set of most recent researches in the field of robot grasping task were investigated. Several conditions affect the performance of robot grasping such as changes in environment and illumination, existence of a huge number of objects with different properties, complex backgrounds, and occlusion between objects. Therefore, machine learning techniques are implemented to improve the capability of these robots to handle these conditions and guarantee high performance. Two main categories are considered, which are deep learning and reinforcement learning. It can be concluded that deep learning techniques are more common compared with reinforcement learning. However, the latter one is more suitable for grasping of unknown objects or if it is somehow difficult to collect a sufficient dataset. As a future work, 3D object grasping according to the matching between the object's splines and the robot palm splines will be studied. Both DL and RL techniques will be investigated to figure out which of them is the suitable option to obtain an accurate and efficient grasping performance.

References

1. *Solowjow, E.* Industrial robot grasping with deep learning using a programmable logic controller (PLC) / E. Solowjow, I. Ugalde, Y. Shahapurkar, J. Aparicio, J. Mahler, V. Satish, K. Goldberg, H. Claussen. 2020. arXiv: 2004.10251.

2. *Rakhimkul, S.* Autonomous object detection and grasping using deep learning for design of an intelligent assistive robot manipulation system / S. Rakhimkul,

A. Kim, A. Pazylbekov, A. Shintemirov // 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC) (6-9 Oct. 2019). Bari, Italy : IEEE, 2019. Pp. 3962-3968. DOI: 10.1109/SMC.2019.8914465.

3. *Al Akkad, M. A.* Complexity reduction for two human-like ambidextrous robotic hands control // 2019 International Siberian Conference on Control and Communications (SIBCON) (18-20 April 2019). Tomsk, Russia : IEEE, 2019. Pp. 1-7. DOI: 10.1109/SIBCON.2019.8729631.

4. *Fang, K.* Learning task-oriented grasping for tool manipulation from simulated self-supervision / K. Fang, Y. Zhu, A. Garg, A. Kurenkov, V. Mehta, L. Fei-Fei, S. Savarese. 2018. arXiv: 1806.09266.

5. *Li, Y.* A review: machine learning on robotic grasping / Y. Li, Q. Lei, C. P. Cheng, G. Zhang, W. Wang, Zh. Xu // Proceedings SPIE. Vol. 11041. Eleventh International Conference on Machine Vision (ICMV 2018) (15 March 2019). 2019. 110412U. International Society for Optics and Photonics. DOI: 10.1117/12.2522945.

6. *Krizhevsky, A., Sutskever, I., Hinton, G. E.* ImageNet classification with deep convolutional neural networks // Advances in neural information processing systems. 2012. Vol. 25. No. 2. DOI: 10.1145/3065386.

7. *Druzhkov, P., Kustikova, V. D.* A survey of deep learning methods and software tools for image classification and object detection // Pattern Recognition and Image Analysis. 2016. Vol. 26. No. 1. Pp. 9–15. DOI: 10.1134/S1054661816010065.

8. *Mathe, S., Pirinen, A., Sminchisescu, C.* Reinforcement learning for visual object detection // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (27-30 June 2016). Las Vegas : USA : IEEE, 2016. Pp. 2894-2902. DOI: 10.1109/CVPR.2016.316.

9. *Sutton, R. S., Barto, A. G.* Reinforcement learning: An introduction. 2nd ed. Cambridge, USA : MIT Press, 2018. 552 pp.

10. *Wang, S.* Efficient fully convolution neural network for generating pixel wise robotic grasps with high resolution images / S. Wang, X. Jiang, J. Zhao, X. Wang, W. Zhou, Y. Liu // 2019 IEEE International Conference on Robotics and Biomimetics (ROBIO). Dali, China : IEEE, 2019. DOI: 10.1109/ROBIO49542.2019.8961711.

11. *Xu, Y.* GraspCNN: Real-Time Grasp Detection Using a New Oriented Diameter Circle Representation / Y. Xu, L. Wang, A. Yang, L. Chen // IEEE Access. 2019. Vol. 7. Pp. 159322-159331. DOI: 10.1109/ACCESS.2019.2950535.

12. *Zhao, Z.* Grasp prediction and evaluation of multi-fingered dexterous hands using deep learning / Z. Zhao, W. Shang, H. He, Z. Li // Robotics and Autonomous Systems. 2020. Vol. 129. P. 103550. DOI: 10.1016/j.robot.2020.103550.

13. *Sharma, P., Valles, D.* Deep convolutional neural network design approach for 3D object detection for robotic grasping // 2020 10th Annual Computing and Communication Workshop and Conference (CCWC) (6-8 Jan. 2020). Las Vegas, USA : IEEE, 2020. Pp. 0311-0316. DOI: 10.1109/CCWC47524.2020.9031186.

14. *Song, Y.* A novel robotic grasp detection method based on region proposal networks / Y. Song, L. Gao, X. Li, W. Shen // Robotics and Computer-Integrated Manufacturing. 2020. Vol. 65. P. 101963. DOI: 10.1016/j.rcim.2020.101963.

15. *Kalashnikov, D.* QT-Opt: scalable deep reinforcement learning for vision-based robotic manipulation / D. Kalashnikov, A. Irpan, P. Pastor, J. Ibarz, A. Herzog,

E. Jang, D. Quillen, E. Holly, M. Kalakrishnan, V. Vanhoucke, S. Levine. 2018. arXiv: 1806.10293.

16. *Zeng, A.* Learning synergies between pushing and grasping with self-supervised deep reinforcement learning / A. Zeng, S. Song, S. Welker, J. Lee, A. Rodriguez, T. Funkhouser // 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (1-5 Oct. 2018). Madrid, Spain : IEEE, 2018. Pp. 4238-4245. DOI: 10.1109/IROS.2018.8593986.

17. Rajeswaran, A. Learning complex dexterous manipulation with deep reinforcement learning and demonstrations / A. Rajeswaran, V. Kumar, A. Gupta, G. Vezzani, J. Schulman, E. Todorov, S. Levine. 2017. arXiv: 1709.10087.

Р. Али, магистрант
E-mail: eng.rawan.sy@gmail.com
М. А. Аль Аккад, канд. техн. наук, доц.
E-mail: aimanakkad@yandex.ru
Кафедра «Программное обеспечение»
Ижевский государственный технический университет имени М. Т. Калашникова

Машинное обучение в роботизированных задачах захвата: обзор

В последнее время методы машинного обучения, включая глубокое обучение и обучение с подкреплением, были признаны важной вехой в области компьютерного зрения и основанных на зрении задач робототехники, таких как захватывание. В этом обзоре представлены новейшие подходы в области задач выбора и размещения объектов и задач захвата объектов. Эти подходы подразделяются на две группы: подходы, основанные на глубоком обучении, и подходы, основанные на подкреплении. Задачи захвата людьми решаются интуитивно, в то время как для роботов является большой проблемой решить этих задач так же хорошо, как это делают люди. На эффективность захвата робота влияют несколько условий, такие как изменение окружающей среды и освещенности, наличие огромного количества объектов с различными свойствами, сложные фоны и окклюзия между объектами. Методы машинного обучения применяют в робототехнических системах для улучшения способности роботов справляться с этими условиями и гарантировать высокую производительность.

Keywords: робототехника, машинное обучение, глубокое обучение, обучение с подкреплением, задачи захвата, обнаружение объекта.