

ПРИМЕНЕНИЕ КОНВОЛЮЦИОННЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ОБУЧЕНИЯ МАНИПУЛЯЦИОННОГО РОБОТА ОПЕРАЦИИ ЗАХВАТА ОБЪЕКТОВ

Маис Жамал

zhamalm@student.bmstu.ru

SPIN-код: 3799-7355

МГТУ им. Н.Э. Баумана, Москва, Российская Федерация

Аннотация

Захватывание различных объектов является основной задачей, выполняемой роботом при исследовании и манипулировании в окружающей среде. Учитывая значительные трудности, связанные с программированием необходимого положения захвата для каждого из объектов, предлагается применять конволюционную нейронную сеть для обучения робота операции захвата различных объектов с учетом их пространственного положения. Сеть была обучена с помощью выборки, состоящей из 800 изображений 20 объектов. Экспериментальное исследование показало, что сеть обеспечивает 53,04 % успеха захвата новых объектов, которые отсутствовали в обучающей выборке. Это позволяет предположить, что при увеличении объема обучающей выборки робот сможет успешно захватывать объекты, удовлетворяющие определенным условиям и отсутствующие в обучающей выборке.

Ключевые слова

Робот, манипулятор, операция захвата, нейронная сеть, конволюционная нейросеть, машинное обучение, обучение нейросети, искусственный интеллект

Поступила в редакцию 31.10.2019

© МГТУ им. Н.Э. Баумана, 2019

Введение. Захват объектов является важной способностью манипуляционного робота. Он позволяет роботам манипулировать объектами в окружающей среде или изменять состояние окружающей среды, если это необходимо. Основная задача промышленных манипуляторов также заключается в захватывании и перемещении некоторых объектов.

Роботизированный захват был активным предметом исследований в течение десятилетий [1–5], и много усилий было потрачено на разработку алгоритмов захватывания. Опираясь на данные обзорных статей [6, 7], методы процесса захвата можно подразделить на аналитические методы, зависящие от анализа механики захвата и контактных взаимодействий «захват – объект», и эмпирические методы управления данными, которые опираются на выборку возможных объектов захвата из визуального источника и ранжирование их по определенному критерию.

Е.В. Селиверстова [8] использовала аналитический метод захватывания, рассматривая задачу захвата как задачу минимизации сил контакта между объек-

том и губками захвата. Однако в ее работе было выдвинуто предположение, что точные геометрические и физические модели объекта доступны роботу, что не всегда возможно. Кроме того, могут быть неизвестны свойства поверхности, а также коэффициенты трения, вес, центр масс и распределение веса объекта.

S. Kumra и др. [9] разработали процесс захвата объектов манипулятором, используя глубокое обучение с учетом 3D-изображения и RGB-изображения (Red, Green, Blue) объектов. Они рассматривали эту задачу как задачу классификации и применяли конволюционную нейронную сеть. Предлагаемая в этой работе нейросеть, которая состоит из 50 слоев, обеспечила 89,21 % успешных результатов захвата. Вместе с тем размер сети является очень большим, входные сети существенно зависят от изображения RGB и глубины изображения, что сопряжено со значительными вычислительными затратами и не позволяет работать в реальном масштабе времени.

Настоящее исследование основано на методах управления данными: получив достаточно большую выборку обучающих данных и используя конволюционную нейронную сеть, можно организовать обучение робота, позволяющее найти правильное положение захвата с учетом 3D-изображение (изображение глубины) объекта. Структура сети максимально упростится, когда сеть будет работать в реальном масштабе времени.

Обучающая выборка. В качестве обучающей выборки в поставленной выше задаче можно использовать стандартный набор данных для захвата, разработанный в Корнуэльском университете [10]. Он содержит множество объектов, каждый из которых имеет несколько заранее определенных способов захвата. Способы захвата задаются как ориентированные прямоугольники на плоскости (рис. 1). Подготовка обучающей выборки предложена в работе D. Morrison и др. [11]; эта процедура иллюстрируется на рис. 1 на примере подготовки обучающего набора для одного объекта. Из базы данных были взяты RGB-изображение объекта и 3D-изображение, учитывающее глубину изображения (Depth). Эти изображения определяют три способа захвата, отмеченных прямоугольниками на левой части рис. 1. 3D-изображения были нормализованы, чтобы обеспечить инвариантность по глубине.

Далее вычисляют параметры, определяющие качество захвата, угол ориентации захвата и его ширину (средняя часть рис. 1). Для определения качества захвата Q^T (Grasp Quality) всем пикселям, соответствующим области пересечения прямоугольников захвата с объектом, присваивается значение единица, а остальным пикселям — значение ноль. Угол ориентации захвата φ^T (Grasp Angle) определяется как угол поворота прямоугольника захвата относительно горизонтальной оси и принимает значения в диапазоне $[-\pi/2; +\pi/2]$. Для устранения неоднозначности, возникающей при изменении угла на $\pm\pi/2$, каждый угол определяется двумя значениями: $\sin(2\varphi^T)$, $\cos(2\varphi^T)$ (см. правую часть рис. 1). При вычислении ширины захвата W^T (Grasp Width) ширине присваивается значение длины

перпендикулярна к стороне соответствующего прямоугольника. Поскольку набор данных [10], который снимают с реальной камеры, уже содержит реальный шум датчика, добавление шума не является необходимым.

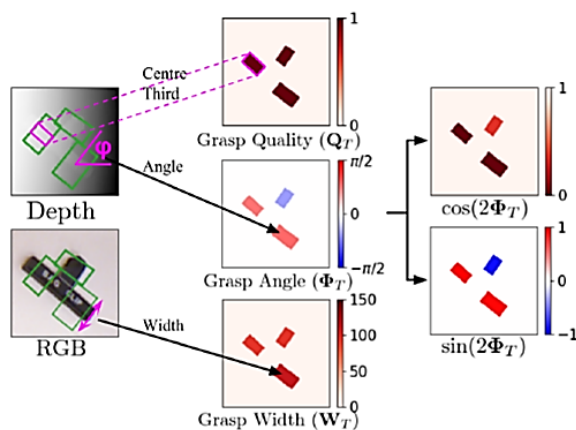


Рис. 1. Подготовка обучающего набора

Далее проводится обучение нейросети методом обратного распространения ошибки. Поскольку положение каждого объекта заранее неизвестно, при обучении нейросети каждое изображение было повернуто на 10 разных углов. Таким образом, обучающая выборка увеличилась в 10 раз по сравнению с рекомендациями [10] и составила 800 единиц для каждого объекта.

Структура нейросети. Выбранная для проведения эксперимента структура нейросети (рис. 2) была основана на исследовании D. Morrison и др. [11]. Сеть состоит из шести слоев, из которых три первые слоя являются конволюционными, а три последующих — транспонированными по отношению к конволюционным. На схеме указаны размеры каждого слоя, количество фильтров, а также число шагов, т. е. количество пикселей, которые фильтр пропускает после каждого применения.

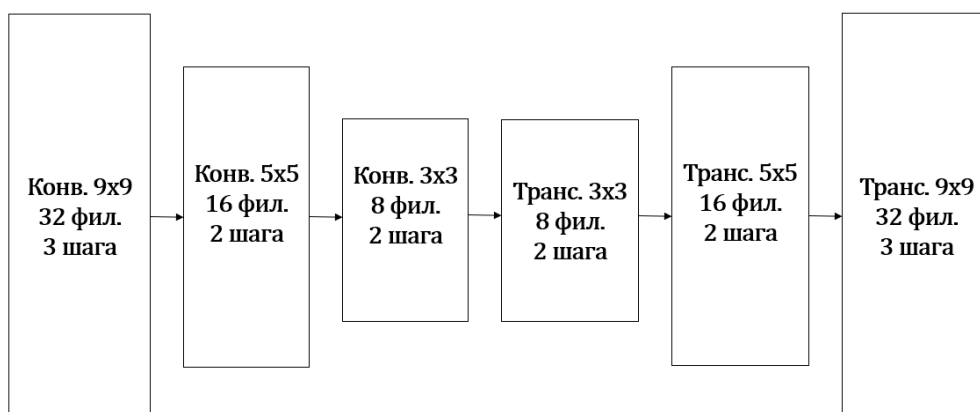


Рис. 2. Конволюционная нейронная сеть

Экспериментальное исследование. Задача исследования состояла в том, чтобы определить надежность захвата роботом объектов, которые отсутствовали в обучающей выборке.

Вначале сеть была обучена по 800 изображениям 20 объектов, и после проверки сеть достигла 53,04 % успеха захвата на наборе новых объектов. Ниже приведены примеры полученных экспериментальных результатов. На рис. 3 даны два примера успешных захватов. Для каждого из объектов на первом изображении показаны возможные захваты (зеленые прямоугольники), из которых обученная сеть выбирает наиболее подходящий для захвата. Он показан также красным цветом на втором изображении, которое определяет выбранную сеть ширину и ориентацию захвата. Определяются отдельно качество захвата и угол его ориентации (третье и четвертое изображения). Качество захвата определяется интенсивностью красного цвета. Лучшее место захвата определяется по наиболее высокому значению интенсивности, имеющему 256 градаций.

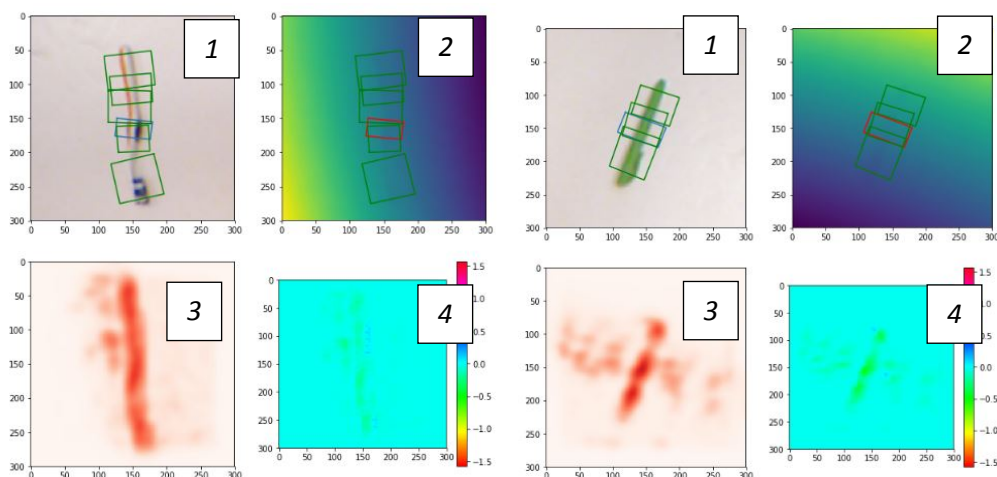


Рис. 3. Успешные решения захвата объектов. Для каждого объекта показано четыре изображения:

- 1 — изображение объекта с зелеными прямоугольниками для возможных захватов и с синим прямоугольником для захвата, полученного нейросетью; 2 — изображение ширины захвата объекта; 3 — изображение, определяющее качество захвата; 4 — изображение, определяющее угол поворота захвата

Примеры неудачных захватов приведены на рис. 4. Здесь приняты те же обозначения, что и на рис. 3. Можно предположить, что неудачные решения нейросети обусловлены либо сложной конфигурацией самих объектов (левое изображение), либо их неудачной ориентацией по отношению к оптической системе наблюдения (правое изображение).

Надежность получаемых нейросетью решений может быть существенно повышена при увеличении объема обучающей выборки.

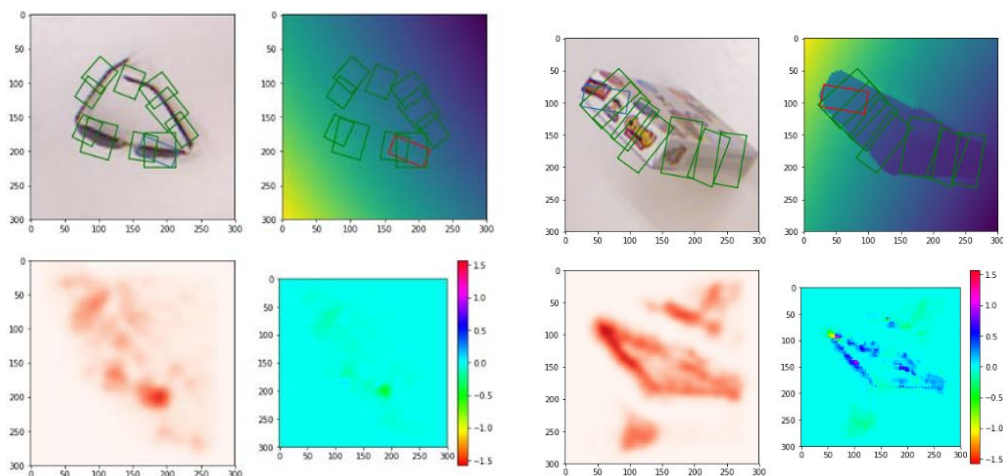


Рис. 4. Примеры неудачных решений захвата объектов, полученных сетью

Заключение. По результатам проведенного исследования можно заключить, что построенная нейросеть достаточно хорошо обучалась правильным положениям захвата. В меньшей мере процесс обучения был удачным по отношению к углам ориентации и к ширине захвата. Трудности возникают и в особом случае, когда угол равен $+\pi/2$ или $-\pi/2$ и возникает ошибка в определении знака угла. Отмеченные недостатки можно объяснить недостаточно большим объемом обучающей выборки.

В целом результаты проведенных исследований подтверждают возможность применения искусственных нейронных сетей конволюционного типа для обучения манипуляционного робота процессу захвата различных объектов, включая объекты, отсутствующие в обучающей выборке. В качестве дальнейшего развития рассмотренного метода прежде всего необходимо увеличить объем обучающей выборки, что позволит увеличить и число распознаваемых объектов.

Рассмотренная нейросеть может быть наиболее успешно использована в тех случаях, когда распознаваемые объекты все же имеют некоторое сходство, различаясь по размерам, ориентации и конфигурации. Например, в роботизированной аптеке, где робот должен подбирать заказанные покупателем медикаменты на стеллажах. Аналогично можно применять подобную нейросеть для управления роботом на складах роботизированных магазинов для отбора изделий согласно заказам, поступившим по электронной почте.

Литература

- [1] Saxena A., Driemeyer J., Ng A.Y. Robotic grasping of novel objects using vision. *The Int. J. Robot. Res.*, 2008, vol. 27, no. 2, pp. 157–173. DOI: <https://doi.org/10.1177%2F0278364907087172>
- [2] Jiang Y., Moseson S., Saxena A. Efficient grasping from RGBD images: learning using a new rectangle representation. *IEEE ICRA*, 2011, pp. 3304–3311. DOI: <https://doi.org/10.1109/ICRA.2011.5980145>

-
- [3] Ciocarlie M., Hsiao K., Jones E.G., et al. Towards reliable grasping and manipulation in household environments. In: *Experimental robotics*. Springer, 2014. pp. 241–252.
- [4] Lenz I., Lee H., Saxena A. Deep learning for detecting robotic grasps. *The Int. J. Robot. Res.*, 2015 vol. 34, no. 4-5, pp. 705–724. DOI: <https://doi.org/10.1177%2F0278364914549607>
- [5] Redmon J., Angelova A. Real-time grasp detection using convolutional neural networks. *IEEE ICRA*, 2015, pp. 1316–1322. DOI: <https://doi.org/10.1109/ICRA.2015.7139361>
- [6] Sahbani A., El-Khoury S., Bidaud P. An overview of 3-D object grasp synthesis algorithms. *Rob. Auton. Syst.*, 2012, vol. 60, no. 3, pp. 326–336. DOI: <https://doi.org/10.1016/j.robot.2011.07.016>
- [7] Caldera S., Rassau A., Chai D. Review of deep learning methods in robotic grasp detection. *Multimodal Technologies Interact.*, 2018, vol. 2, no. 3, pp. 57. DOI: <https://doi.org/10.3390/mti2030057>
- [8] Селиверстова Е. Управление многопалым захватным устройством автономного робота при захвате и манипулировании деформируемыми объектами. Дисс. ... канд. тех. наук. М., МГТУ им. Н.Э. Баумана, 2018.
- [9] Kumra S., Kanan C. Robotic grasp detection using deep convolutional neural networks. *IEEE/RSJ IROS*, pp. 769–776. DOI: <https://doi.org/10.1109/IROS.2017.8202237>
- [10] Barclay J.R. Stream discharge from Harford, NY. *ecommons.cornell.edu: веб-сайт*. URL: <http://hdl.handle.net/1813/34425> (дата обращения: 15.11.2019).
- [11] Morrison D., Leitner J., Corke P. Closing the loop for robotic grasping: a real-time, generative grasp synthesis approach. *RSS*, 2018. DOI: <https://doi.org/10.15607/RSS.2018.XIV.021>

Маис Жамал — студентка магистратуры, кафедры «Робототехнические системы и мехатроника», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

Научный руководитель — Ющенко Аркадий Семенович, доктор технических наук, профессор кафедры «Робототехнические системы и мехатроника», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

Ссылку на эту статью просим оформлять следующим образом:

Маис Жамал. Применение конволюционных нейронных сетей для обучения манипуляционного робота операции захвата объектов. *Политехнический молодежный журнал*, 2020, № 01(42). <http://dx.doi.org/10.18698/2541-8009-2020-01-571>

THE USE OF CONVOLUTIONAL NEURAL NETWORKS FOR TRAINING THE MANIPULATION ROBOT TO CAPTURE OBJECTS

Mais Jamal

zhamalm@student.bmstu.ru

SPIN-code: 3799-7355

Bauman Moscow State Technical University, Moscow, Russian Federation

Abstract

The capture of various objects is the main task performed by the robot in the study and manipulation in the environment. Given the significant difficulties associated with programming the necessary capture position for each of the objects, it is proposed to use a convolutional neural network to train the robot to capture various objects, taking into account their spatial position. The network was trained using a sample of 800 images of 20 objects. An experimental study showed that the network provides 53.04% success in capturing new objects that were not in the training set. This suggests that with an increase in the training sample, the robot will be able to successfully capture objects that meet certain conditions and are not in the training sample.

Keywords

Robot, manipulator, capture operation, neural network, convolutional neural network, machine learning, neural network training, artificial intelligence

Received 31.10.2019

© Bauman Moscow State Technical University, 2019

References

- [1] Saxena A., Driemeyer J., Ng A.Y. Robotic grasping of novel objects using vision. *The Int. J. Robot. Res.*, 2008, vol. 27, no. 2, pp. 157–173. DOI: <https://doi.org/10.1177%2F0278364907087172>
- [2] Jiang Y., Moseson S., Saxena A. Efficient grasping from RGBD images: learning using a new rectangle representation. *IEEE ICRA*, 2011, pp. 3304–3311. DOI: <https://doi.org/10.1109/ICRA.2011.5980145>
- [3] Ciocarlie M., Hsiao K., Jones E.G., et al. Towards reliable grasping and manipulation in household environments. In: *Experimental robotics*. Springer, 2014, pp. 241–252.
- [4] Lenz I., Lee H., Saxena A. Deep learning for detecting robotic grasps. *The Int. J. Robot. Res.*, 2015 vol. 34, no. 4-5, pp. 705–724. DOI: <https://doi.org/10.1177%2F0278364914549607>
- [5] Redmon J., Angelova A. Real-time grasp detection using convolutional neural networks. *IEEE ICRA*, 2015, pp. 1316–1322. DOI: <https://doi.org/10.1109/ICRA.2015.7139361>
- [6] Sahbani A., El-Khoury S., Bidaud P. An overview of 3-D object grasp synthesis algorithms. *Rob. Auton. Syst.*, 2012, vol. 60, no. 3, pp. 326–336. DOI: <https://doi.org/10.1016/j.robot.2011.07.016>
- [7] Caldera S., Rassau A., Chai D. Review of deep learning methods in robotic grasp detection. *Multimodal Technologies Interact.*, 2018, vol. 2, no. 3, pp. 57. DOI: <https://doi.org/10.3390/mti2030057>
- [8] Seliverstova E. Upravlenie mnogopalym zakhvatnym ustroystvom avtonomnogo robota pri zakhvate i manipulirovaniy deformiruemyimi ob"ektami. Diss. ... kand. tekhn. nauk [Control on polydactyl robotic grasp of autonomous robot at grasping and manipulating

- of deformable object. Kand. tech. sci. diss.]. Moscow, Bauman MSTU Publ., 2018 (in Russ.).
- [9] Kumra S., Kanan C. Robotic grasp detection using deep convolutional neural networks. *IEEE/RSJ IROS*, pp. 769–776. DOI: <https://doi.org/10.1109/IROS.2017.8202237>
- [10] Barclay J.R. Stream discharge from Harford, NY. *ecommons.cornell.edu: веб-сайт*. URL: <http://hdl.handle.net/1813/34425> (дата обращения: 15.11.2019).
- [11] Morrison D., Leitner J., Corke P. Closing the loop for robotic grasping: a real-time, generative grasp synthesis approach. *RSS*, 2018. DOI: <https://doi.org/10.15607/RSS.2018.XIV.021>

Mais Jamal — Master’s Degree Student, Department of Robotics and Mechatronics, Bauman Moscow State Technical University, Moscow, Russian Federation.

Scientific advisor — Yushchenko A.S., Dr. Sc. (Eng.), Professor, Department of Robotics and Mechatronics, Bauman Moscow State Technical University, Moscow, Russian Federation.

Please cite this article in English as:

Mais Jamal. The use of convolutional neural networks for training the manipulation robot to capture objects. *Politekhnicheskii molodezhnyy zhurnal* [Politechnical student journal], 2020, no. 01(42). <http://dx.doi.org/10.18698/2541-8009-2020-01-571.html> (in Russ.).