

Deep Learning for Robotics

Radouan Ait Mouha

Anhui Polytechnic University, Wuhu, China

Email: red1.aitmouha@gmail.com

How to cite this paper: Mouha, R.A. (2021) Deep Learning for Robotics. *Journal of Data Analysis and Information Processing*, 9, 63-76.

<https://doi.org/10.4236/jdaip.2021.92005>

Received: February 2, 2021

Accepted: April 18, 2021

Published: April 21, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The application of deep learning to robotics over the past decade has led to a wave of research into deep artificial neural networks and to a very specific problems and questions that are not usually addressed by the computer vision and machine learning communities. Robots have always faced many unique challenges as the robotic platforms move from the lab to the real world. Minutely, the sheer amount of diversity we encounter in real-world environments is a huge challenge to deal with today's robotic control algorithms and this necessitates the use of machine learning algorithms that are able to learn the controls of a given data. However, deep learning algorithms are general non-linear models capable of learning features directly from data making them an excellent choice for such robotic applications. Indeed, robotics and artificial intelligence (AI) are increasing and amplifying human potential, enhancing productivity and moving from simple thinking towards human-like cognitive abilities. In this paper, lots of learning, thinking and incarnation challenges of deep learning robots were discussed. The problem addressed was robotic grasping and tracking motion planning for robots which was the most fundamental and formidable challenge of designing autonomous robots. This paper hope to provide the reader an overview of DL and robotic grasping, also the problem of tracking and motion planning. The system is tested on simulated data and real experiments with success.

Keywords

Deep Learning, Neural Networks, Artificial Intelligence, Robotic Grasping, Motion Planning

1. Introduction

Some researchers may debate a specific definition of robots, or argue whether the definition is relative or dependent on the context of the situation, such as the concept of privacy. This may be a better approach as more and more rules and

regulations are created around its use in various contexts.

A robot is a machine especially programmable by a computer that is able to automatically perform a complex series of actions. Robots may be directed by an external control device or the control may be embedded within it. Robots can be created along the lines of a human being, but most robots are designed to perform a task without regard to their aesthetics [1]. Robots can be autonomous or semi-autonomous and range from humanoids.

The field of technology that deals with the design, construction, operation and application of robots, as long as computer systems for their control, sensory feedback and information processing is robotics [2]. These technologies deal with automated machines that can take the place of humans in dangerous environments or manufacturing processes, or resemble humans in appearance, behavior or cognition. For the time being, robots are inspired by nature contributing to the field of bio-inspired robotics. These robots have also created a newer branch of robotics named soft robotics.

Actually, robotics has already moved from the lab to the real world. A robot perceives the world with its different sensors, builds a coherent model of the world and it updates that model over time. However, the robot eventually has to take decisions, plan actions and implement these actions to accomplish a useful task.

Traditionally, the robot scientist or a team of robots design consoles for every task we want the robot to perform. Even for tasks, users can perform these controls intuitively, like holding or cutting food (Domestic robot) [3]. They are difficult to design because we are unable to easily translate this natural intuition into the code (also called analytical approaches). It can also be very strenuous to scale these methods up to the sheer amount of diversity that our robots have to deal with in the real world, picking up everything in your home, chopping any food and so on. These control algorithms are designed based on expert human knowledge of the robot and its environment in the specific task. The result of this approach illustrates the kinematic relationship between the parameters of the robot and the structured environment around it. The kinematic model helps moreover improve control strategies. Nevertheless, direct planning of the results from the kinetic model to the control device of the robot joint is open-loop in nature and is set to cause task space distractions, this model suggested by Ju and Yang [4]. Although this hand-coded instruction is known for effective task performance, such an approach has limitations; In particular, the program is limited to situations predicted by the programmer, but in cases that require frequent changes in the robot's programming, due to the changes in the environment or other factors, this approach becomes impractical. Thus, an unorganized environment stays a real challenge for intelligent robots that require complexity analytical approach to solution formation. For situations like this, Konidaris *et al.* [5] in their work consider experimental methods will provide augmentation of the cognitive and adaptive capacity of robots while reducing or eliminating

the need to manually design an automated solution. These early works explore the adaptive and cognitive ability of robots to learn tasks, but the repetition of tasks is limited to the described tasks.

To overcome this problem, machine learning is used as it can deal with a wide range of situations. This paradigm is growing in popularity especially after successful applications of deep learning. These early works [6] [7] show the researchers' progress in research methods in accommodating robots. These studies discuss early attempts to catch new things using experimental methods.

The main drive behind the use of deep learning in robotics is that it is more general than any other learning algorithm. It has been proven deep networks are capable of thinking and abstraction at a high level. Therefore, it makes it an ideal choice for robots in an unregulated environment. In addition, thanks to advances in parallel processing and sophisticated, deep numerical libraries networks have become very effective. Time-critical robotics tasks need to be addressed high-frequency response unit for motion control.

This paper reviews deep learning approaches for the detection of robotic grasping poses for a given object captured in an image or video in real time. The paper is organized as follows: Section 2 provides an overview about deep learning and neural network; Section 3 identifies several challenges which faced the deep learning in robotics; Section 4 introduces the perception and learning-based approaches for detecting robotic grasps; Section 5 handle the problem of tracking and motion planning in real time and finally a conclusion can be found in Section 6.

2. Deep Learning

The reality that deep learning is a type of machine learning is very important. Thus it must make sure about the type of relation between artificial intelligence, machine learning and deep learning. Actually, deep learning has been in the spotlight because it has adroitly solved some problems that face artificial intelligence as real challenges. Its performance surely is exceptional in many branches. However, it faces limitations as well.

The limitation of deep learning [8] stems from its fundamental concepts that have been inherited from its progenitor which is machine learning. Deep learning as a type of machine learning cannot avoid the primary problems that machine learning faces. That is why before going deeply into the concept of deep learning, you can review machine learning and its algorithms [9] [10] [11] [12].

The term deep learning is loosely related to a wide variety of neural networks architectures [13]. That is why deep learning models are often referred to as deep neural networks. The term "deep" generally refers to the number of hidden layers in a neural network, and traditional one have only two or three hidden layers, whereas deep learning networks have up to 150 layers. Deep learning models are trained using large sets of disaggregated (not rated) data and neural network architectures that learn features directly from the data without the need

to manually extract features.

Neural Network is a series of algorithms that copyists the operations of the human brain to recognize relationships between large amounts of data. In this meaning, neural networks refer to systems of neurons, either organic or artificial in nature. Each neuron has a mathematical model for determining its outputs from its input. Neural networks can acclimate to changing input. So the network generates the best possible result without needing to redesign the output criteria. **Figure 1** illustrates a simple example of a neural network with two hidden layers.

A graph structure where each node generally referred to as neurons is connected to either input signal or other nodes with weighted edges. The output of node is the linear combination of edge weights and the inputs and probably followed by a non-linear function.

An artificial neuron takes an input vector and outputs a scalar value. The parameters of the neuron are determined by a set of weights. Each weight is used as a multiple of a numerical entry. The neuron output is the result of applying the nonlinear activation function on the total weighted input. Thus, a neuron with weights w , input x , output y , and nonlinear activation function are represented as follows:

$$y = \varnothing \left(\sum_{i=0}^n W_i \cdot x_i \right)$$

with n the size of x , and W_0 representing a bias for the neuron input if X_0 is set to 1.

The nonlinear activation functions $\varphi(x)$ are used to improve the expressive strength of the network. They are valid as long as they are continuous, limited, and monotonously increasing. Additional requirement for learning algorithms above $\varphi(x)$ is the differential. Hence, the activation function used often is the sigmoid logistic function:

$$\varnothing(x) = \frac{1}{1 + \exp(x)}$$

DNN, and generally any ANN, is determined by both parameters and hyper-parameters. Hyper-parameters are related to network architecture (number of

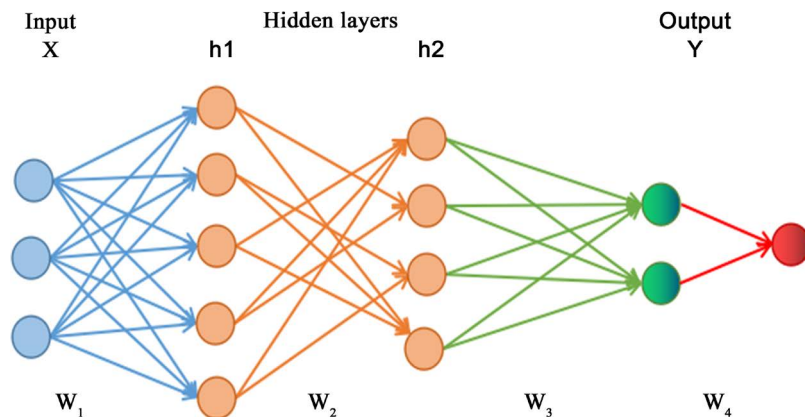


Figure 1. Simple example of a neural network.

layers, number of neurons per layer, activation function, neuron connection, etc) and to the parameters of the learning process. The network parameters correspond to the set of weights of all neurons. The network can be trained to select the optimum weights (parameters) using numerical optimization methods. The optimum hyper-parameters cannot be learned directly from the data; so many network architectures must be trained and tested to select the best hyper-parameters. Thus, to be able to adapt network parameters, select hyper-parameters, and characterize their performance, the available data must be divided into three non-overlapping datasets: training, validation, and testing. A training dataset is used to learn model parameters (weights) of the network. A validation dataset is used for selection Hyper-parameters by evaluating network performance under different hyper-parameter configurations. The test dataset is used to characterize the trained network by estimating the generalization error [14].

In Feed forward neural network (FFNN) more specifically called multi-layered perception (MLP); perceptions are coordinated into interconnected layers. The input layer combines the input patterns. The output layer contains ratings or output signals to which the input modes may be assigned. One of the most common types of deep neural network is called folding neural networks, called convolutional neural networks (CNN or ConvNet) [15]. Convolutional neural network (CNN) is remarkably a special type of FFNN, reduces the number of parameters in a deep neural network with many units without losing model quality. CNN has found apps in images and text treatment where they beat many pre-set criteria. **Figure 2** shows its main algorithm.

The hidden layer usually contains two distinct layers: the first stage is the result of local wrapping of the previous layer (the core contains trainable weights), and the second stage is the maximum assembly stage, in which the number of units is greatly reduced by keeping only the maximum response of several units of the stage First. After several hidden layers, the final layer is usually a fully joined layer. It contains a unit for every class the network expects, and each of those units receives inputs from all the units of the previous layer.

Recurrent Neural Networks (RNNs) are used to label, classify or create sequences. A sequence is a matrix, each row of which is a feature vector and the

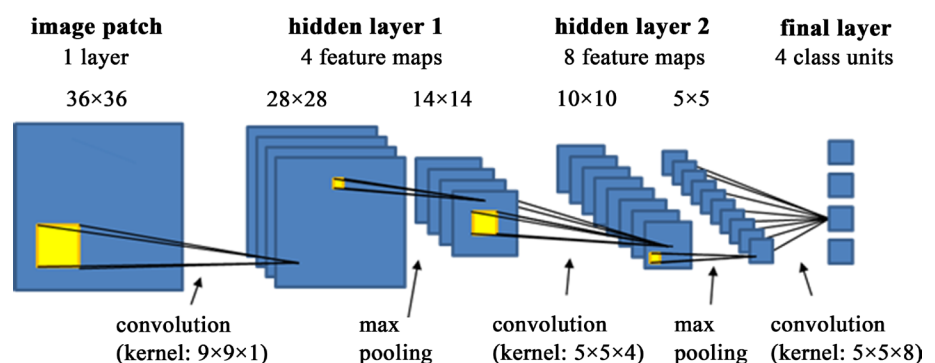


Figure 2. Convolutional neural network algorithm.

order of rows is important. Naming a sequence is a class prediction for each feature vector in a sequence. To classify a sequence is the category prediction for an entire sequence. To generate a sequence is the output of another sequence, it is somewhat related to the input sequence [16]. RNNs are often used in word processing marks or character sequences [17]. For the same reason, repetitive neural networks are also used in speech processing [18].

3. Deep Learning in Robotics

The success of deep learning in computer vision has inspired some applications in robotics. One of the main challenges is using robots in an unrestricted environment. The lack of algorithms was strong and general. Deep learning methods demonstrated that accurate and robust performance can now be achieved in many applications. Especially, image classification and natural language processing have benefited greatly from this new technology.

Deep learning models can be used as a preprocessing unit, often on local images and measurements, which converts raw sensor data into a feature space with fewer dimensions can be used to control. Automated constipation detection [19] uses raw images for detection grab point for various objects that can be used later to grab. DL can also be used in a comprehensive manner where the network is responsible for processing the primary input along the way to generate the control signal for the robot's transporters.

Robotics presents many unique challenges for learning algorithms. First, robots must perform a wide range of tasks and it often takes a long or even not easy time coding entirely new learning algorithms and features for each task. Second, robots have to deal with a great deal of diversity in the real world, which is difficult for many learning algorithms to deal with it. Lastly, most of the time is a premium at most of robotic applications, so learning algorithms must be rapidly inferred to be useful for robotic applications.

Walking and running like a human, teaching through demonstration, mobile navigation in pedestrian environments and collaborative automation, automatic combat recovery, automatic container, automatic shelving picking, automation aircraft inspection and maintenance and automated disaster mitigation and recovery [20]. All these numerous targets and goals were set by the robotic community for the near future, but for reaching these targets, the reality faced several challenges which deep neural network technology (DNN) has high potential impact.

The first challenge: Learn new complex high-dimensional dynamics: Analytical deriving complex dynamics requires anthropologists and is time consuming and subtractive a tract off between state dimensionality and tractability. Make strong models like this to uncertainty is difficult and complete case information is often unknown. Systems than can, there is a need for rapid and independent adaptation to the new dynamics to solve such problems, like grasping new things and traveling over surfaces with unknown or uncertain characteristics, manag-

ing or adapting interactions between a new tool and or environment, degradation and or failure of robot subsystems [21] [22] [23].

The second challenge: Learn the politics of controlling dynamics environments: As with dynamics, control systems that accommodate high degrees of freedom for applications such as mobile manipulation with multiple arms, anthropomorphic hands and swarm robots need. Such systems will be called to operate reliably and safely environments with a high degree of uncertainty and limited state information [24] [25].

The third challenge: Advanced manipulation: In spite of the progress made over the last decades, active searches and robust, general solutions to tasks such as absorbing distortion and/or complex geometric shapes, using tools and operating systems in the environment. Remain out of reach especially in new situations. This challenge includes motor planning, locomotion and understanding inherent in tasks like this [26] [27] [28].

The fourth challenge: Advanced object recognition: Deep neural networks have already proven to be high adept at identifying and classifying objects. An advanced application examples include identifying distorted objects, estimating their condition and morphology understanding semantic task and route specification and learn about properties of objects and surfaces such as wet/slippery floors or sharp objects which can be dangerous to human collaborators [29].

The fifth challenge: Interpreting and anticipation human action: This challenge is crucial, robots must work with or between people in applications such as collaborative robots to manufacture, care for the elderly and autonomous vehicles that operate on public rounds or move around in pedestrian environments. It will enable teaching by demonstration, which in turn facilitate assignment of tasks by individuals who have no experience with robots or programming. This challenge may also extend to the awareness of human needs, anticipate when an automated intervention will be appropriate [30] [31] [32].

The sixth challenge: Sensor fusion and dimensional reduction: Low cost spread sensing technologies have been a boon to robots as they offer a plethora of possibilities rich, high-dimensional and multi-media data. This challenge refers to methods build meaningful and informative representations of the state from this data [28] [33] [34].

The seventh challenge: High-level task planning: Robots will need reliable high-level execution orders that merge the previous challenges to achieve a new level of utility, particularly if it will benefit all the general public [35].

4. Robotic Grasping

In this part, we will focus on the perception and learning-based approaches for robotic grasping which is one of the main problems addressed in our work. It's very important to note that most of researches define "grasp" as an end-effector configuration that is achieved partial or complete form or forced closure of a specific object. This is the challenging problem, because it depends on the place-

ment and configuration of the automatic clutch. As well as the shape and physical properties of the object to be understood and typically requires research into a large number of possible gripper configuration [36] [37] [38].

In order for robots to gain a more general benefit, cognition is a necessary skill to master. These general purpose, robots may use their cognitive capabilities to visually recognize certain grasps for a given object. **Figure 3** shows a real example. The grasps describes how a robotic end-effector can be arranged to safely grasp an object successfully raised it without slipping. Conventionally, grasp detection required specialized human knowledge to formulate the algorithm for the task analytically, but this is a laborious and time-consuming approach [39].

Things grasping is a difficult challenge because of a wide range of factors like the different shapes of objects and unlimited object poses. Successful robotic gripping systems must be able to defeat this challenge to achieve beneficial results. Counter to robots, humans can determine how to ingest a file specific object. Robot grasping performance is much lower than the grasping standards for human things, but it is constantly improving due to the high demand. Actually it has some following sub-systems which presented as the following: a) Perception detection sub-system which is to discover gripping positions from images of objects in their image plane coordinates. b) Grasp planning sub-system, which is to map the detected image plane coordinates to the world coordinates. c) Control sub-system which is to define the reverse kinematic solution of the previous subsystem.

A successful grasp describes how an automated responder can be pointed over an object to safely hold the object between its clutch and catch the thing which its the goal. As humans, we use our eyesight to visually recognize objects in our area around and know how to approach them to catch them. Likewise, optical perception sensors can be used on an automated system to produce information about the environment, it can be interpreted in a useful form. The mapping method is essential to classify each pixel scene on the basis of belonging or lack of



Figure 3. Right hand robotics grasping a cup.

affiliation to a successful understanding.

In most of the earlier works, the knobs were represented as points on real scenes or images from 3D network models based on simulations. Saxena *et al.* [40] have used the supervised learning approach, investigated the regression learning method for inferring the 3-Dimensional location of a point of assimilation in cartesian coordinate system. They used a probabilistic model about potential grip points while thinking uncertainty in camera placement. To expand their investigation, they estimated 3D workspace for finding the constipation point g , given by $g = (x, y, z)$. Whereas Zhang *et al.* [41] by using the reinforcement learning approach for grasp point detection have defined a grasp as a point in the plan 2D image. The main drawback of this specific point was the fists. However, it only determines where to grasp an object and doesn't specify how wide the clutch had to be opened or the desired direction of the handle to successfully grasp the object.

Our problem belongs to many vision problems, for example; camera calibration, stereo matching, structure from motion, motion tracking and object recognition. To solve our problem, a number of solutions have been proposed, in this work we suggested angles detectors which also were suggested, such as counter-based approach, counter-based approach, colour-based approach, model-based approach and machine learning-based approach. The result of our approach presented in **Figure 4**.

Sunghu kim [42] has proposed corner detection system consists of a spatial filtering part and a detection part. Filter part perform direct bending estimation through an application flexural filter after directional filter. The corner detection part performs the local maximum on bending the final field and angles are extracted by the application from threshold. Density based angle detection strong when used in textured images, given its visuals filter system, but weak when used for detection angles with structural meaning, like obtuse angles, it has a low identification accuracy. In contrast, the contour based corner detection is powerful when used for detection structured objects, because an appreciation of their curvature strategic, but weak when used to detect fabric images, due to Canny's fragile edge detection. **Figure 5** describes the hysteresis thresholding



Figure 4. Grasping detection using corner detection. The left image is the original image which is the data, and the right one shows the result of corner detection using python.

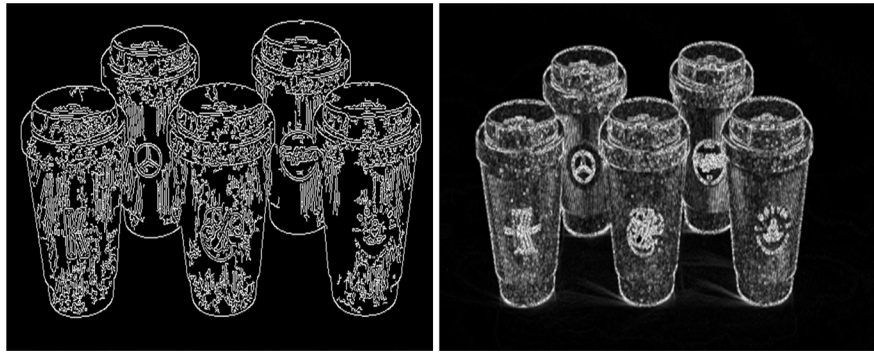


Figure 5. The left image: Hysteresis thresholding stage. The right image: The edge image.

stage and last result.

Due to the availability of inexpensive depth sensors, many robotic systems use RGB-D supported by open source software, does not require sophisticated hardware and availability unique sensing capabilities, for more reading [43].

5. Tracking and Motion Planning

Modern robots have come a long way from their predecessors in the past, which are traditionally been in structured environment. Therefore the robot's interactions are limited and the behavior can be direct determined by human action. But with the era of the fourth industrial revolution, it is considerably known as Industry 4.0.

Industrial robots movement that is part of the larger production process is usually programmed in an inflexible manner and requires precise control of conditions from the movement mission. For example, simple pick and place movement requires accurate knowledge of the situation from the object to be captured and about which container the object is placed in it. Small deviations for either the object or container may cause the operation to fail, such as motion inflexible programming is unable to handle it, small differences in the environment. This is a general problem, the question raises how can robots be enabled to act with these differences in the environment and independently adapt to them to plan their movement in a flexible way [44]. The real experiment result from a real video has shown in **Figure 6**.

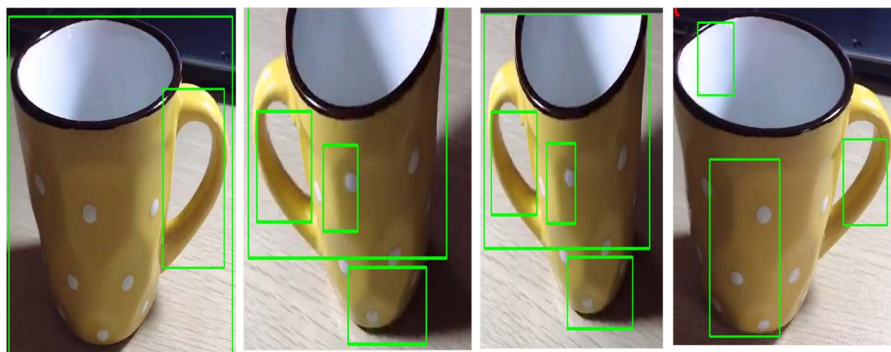


Figure 6. Experiment from video.mp4; status: movement.

6. Conclusion

Deep learning has shown promise in sensing, perception, action problems and even the ability to typically combine these separate functions in one system. Deep neural networks can run on raw sensor data and extract key features from its data without human assistance, which could drastically reduce pre-engineering time. They are also adept at integrating multi-media and high-dimensional data. The improvement has been demonstrated with experience, making it easier to adapt the unstructured dynamics in which robots operate. An important aspect of robots is their ability to manipulate their environment which has proven difficult to learn these skills. But to perceive things, a robot doesn't require knowing how to perform all the tasks of manipulation usefully as long as it can learn these skills easily when needed. However, robotic perception, robotic learning and robotic control tasks remain serious challenges on techniques usually applied. This work presented some of these current challenges which are relevant to deep learning in robotics.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] (2016) Definition of "Robot". Oxford English Dictionary.
- [2] Built In (2019) What Is Robotics ? Types of Robots. <https://builtin.com/robotics>
- [3] From Wikipedia, the Free Encyclopedia (2021) Definition of "Domestic Robot". https://en.wikipedia.org/wiki/Domestic_robot
- [4] Ju, Z.F., Yang, C.G. and Ma, H.B. (2014) Kinematics Modeling and Experimental Verification of Baxter Robot. 2014 33rd Chinese Control Conference (CCC), Nanjing, 28-30 July 2014, 8518-8523. <https://doi.org/10.1109/ChiCC.2014.6896430>
- [5] Konidaris, G., Kuindersma, S. and Grupen, R.A. (2011) Robot Learning from Demonstration by Constructing Skill Trees. *The International Journal of Robotic Research*, **31**, 360-375. <https://doi.org/10.1177/0278364911428653>
- [6] Saxena, A., Driemeyer, J., Kearns, J. and Ng, A.Y. (2006) Robotic Grasping of Novel Objects. In: *Advances in Neural Information Processing Systems 19: Proceedings of the 2006 Conference*, MIT Press, Cambridge, 1-15.
- [7] Saxena, A., Driemeyer, J. and Ng, A.Y. (2008) A Robotic Grasping of Novel Objects Using Vision. *The International Journal of Robotics Research*, **27**, 157-173. <https://doi.org/10.1177/0278364907087172>
- [8] Sünderhauf, N., *et al.* (2018) The Limits and Potential of Deep Learning for Robotics. *The International Journal of Robotics Research*, **37**, 405-420. <https://doi.org/10.1177/0278364918770733>
- [9] Brownlee, J. (2016) *Machine Learning Mastery with Python: Understand Your Data, Create Accurate Models, and Work Projects End-to-End*. Machine Learning Mastery, San Francisco.
- [10] Mohamed, A. (2005) Survey on Multiclass Classification Methods. Technical Report, Caltech, California Institute of Technology, Pasadena.

- [11] Hawkins, D.M. (2004) The Problem of Overfitting. *Journal of Chemical Information and Computer Sciences*, **44**, 1-12. <https://doi.org/10.1021/ci0342472>
- [12] Ghahramani, Z. (2003) Unsupervised Learning. In: *Summer School on Machine Learning*, Springer, Berlin, 72-112. https://doi.org/10.1007/978-3-540-28650-9_5
- [13] Burkov, A. (2019) The Hundred-Page Machine Learning Book. 159.
- [14] Guo, Y.M. and Liu, Y. (2015) Deep Learning for Visual Understanding: A Review. *Neurocomputing*, **187**, 27-48. <https://doi.org/10.1016/j.neucom.2015.09.116>
- [15] Ghorbani, M.A. (2016) A Comparative Study of Artificial Neural Network (MLP, RBF) and Support Vector Machine Models for River Flow Prediction. *Environmental Earth Sciences*, **75**, Article No. 476. <https://doi.org/10.1007/s12665-015-5096-x>
- [16] Vinyals, O., Ravuri, S.V. and Povey, D. (2012) Revisiting Recurrent Neural Networks for Robust ASR. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Kyoto, 25-30 March 2012, 115-133. <https://doi.org/10.1109/ICASSP.2012.6288816>
- [17] Cireşan, D.C., Meier, U., Gambardella, L.M. and Schmidhuber, J. (2010) Deep, Big, Simple Neural Nets for Handwritten Digit Recognition. *Neural Computation*, **22**, 3207-3220. https://doi.org/10.1162/NECO_a_00052
- [18] Graves, A., Mohamed, A.R. and Hinton, G. (2013) Speech Recognition with Deep Recurrent Neural Networks. *IEEE International Conference on Acoustics, Speech and Signal Processing*, Vancouver, 26-31 May 2013, 19-31. <https://doi.org/10.1109/ICASSP.2013.6638947>
- [19] Lenz, I. and Lee, H. (2015) Deep Learning for Detecting Robotic Grasps. *The International Journal of Robotics Research*, **34**, 705-724. <https://doi.org/10.1177/0278364914549607>
- [20] DARPA Robotics Challenge. 2012-2015. <https://www.darpa.mil/program/darpa-robotics-challenge>
- [21] Punjani, A. and Abbeel, P. (2015) Deep Learning Helicopter Dynamics Models. *IEEE International Conference on Robotics and Automation (ICRA)*, Seattle, 26-30 May 2015, 3-45. <https://doi.org/10.1109/ICRA.2015.7139643>
- [22] Mariolis, I., Peleka, G., Kargakos, A., *et al.* (2015) Pose and Category Recognition of Highly Deformable Objects Using Deep Learning. 2015 *International Conference on Advanced Robotics (ICAR)*, Istanbul, 27-31 July 2015, 2-25. <https://doi.org/10.1109/ICAR.2015.7251526>
- [23] Levine, S., Pastor, P., Krizhevsky, A., *et al.* (2016) Learning Hand-Eye Coordination for Robotic Grasping with Deep Learning and Large-Scale Data Collection. *The International Journal of Robotic Research*, **37**, 421-436. <https://doi.org/10.1177/0278364917710318>
- [24] Polydoros, A.S., Nalpantidis, L. and Kruger, V. (2015) Real-Time Deep Learning of Robotic Manipulator Inverse Dynamics. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Hamburg, 28 September-3 October 2015, 7.
- [25] Dong-ok, W., Klaus, R. and Lee, S. (2020) An Adaptive Deep Reinforcement Learning Framework Enables Curling Robots with Human-Like Performance in Real-World Conditions. *Science Robotics*, **5**, eabb9764. <https://doi.org/10.1126/scirobotics.abb9764>
- [26] Yang, Y., Li, Y., Fermüller, C., *et al.* (2015) Robot Learning Manipulation Action Plans by Watching Unconstrained Videos from the World Wide Web. *29th AAAI Conference on Artificial Intelligence (AAAI-15)*, Vol. 19, 3686-3692.
- [27] Yu, J., Weng, K., Liang, G., *et al.* (2013) A Vision-Based Robotic Grasping System

- Using Deep Learning for 3D Object Recognition and Pose Estimation. 2013 *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Shenzhen, 12-14 December 2013, 14-37. <https://doi.org/10.1109/ROBIO.2013.6739623>
- [28] Redmon, J. and Angelova, A. (2015) Real-Time Grasp Detection Using Convolutional Neural Networks. 2015 *IEEE International Conference on Robotics and Automation (ICRA)*, Seattle, 26-30 May, 2015, 5-13. <https://doi.org/10.1109/ICRA.2015.7139361>
- [29] Schmitz, A., Bansho, Y., Noda, K., *et al.* (2015) Tactile Object Recognition Using Deep Learning and Dropout. 2014 *14th IEEE-RAS International Conference on Humanoid Robots, Humanoids 2014*, Madrid, 18-20 November 2014, 7. <https://doi.org/10.1109/HUMANOIDS.2014.7041493>
- [30] Neverova, N., Wolf, C., Taylor, G.W., *et al.* (2015) Multi-Scale Deep Learning for Gesture Detection and Localization. *European Conference on Computer Vision*, Zurich, 6-12 September 2015, 474-490. https://doi.org/10.1007/978-3-319-16178-5_33
- [31] Hwang, J., Jung, M., Madapana, N., *et al.* (2015) Achieving “Synergy” in Cognitive Behavior of Humanoids via Deep Learning of Dynamic Visual-Motor-Attentional Coordination. 2015 *IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*, Seoul, 3-5 November 2015, 1-19. <https://doi.org/10.1109/HUMANOIDS.2015.7363448>
- [32] Finn, C., Levine, S. and Abbeel, P. (2016) Guided Cost Learning: Deep Inverse Optimal Control via Policy Optimization. *International Conference on Machine Learning (ICML)*, New York, 19-24 June 2016, 49-58.
- [33] Lenz, I., Lee, H. and Saxena, A. (2013) Deep Learning for Detecting Robotic Grasps. *The International Journal of Robotics Research*, **34**, 705-724. <https://doi.org/10.15607/RSS.2013.IX.012>
- [34] Wu, J., Yildirim, I., Lim, J.J., Bill, F. and Josh, T. (2015) Galileo: Perceiving Physical Object Properties by Integrating a Physics Engine with Deep Learning. *Advances in Neural Information Processing Systems 28 (NIPS 2015)*, Montreal, 7-12 December 2015, 4-9.
- [35] Levine, S., Finn, C., Darrell, T. and Abbeel, P. (2016) End-to-End Training of Deep Visuomotor Policies. *Journal of Machine Learning Research*, **17**, 1-40.
- [36] Bohg, J., Morales, A., Asfour, T. and Kragic, D. (2014) Data-Driven Grasp Synthesis—A Survey. *IEEE Transactions on Robotics*, **30**, 289-309. <https://doi.org/10.1109/TRO.2013.2289018>
- [37] Bicchi, A. and Kumar, V. (2000) Robotic Grasping and Contact: A Review. *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings*, San Francisco, 24-28 April 2000, 11-27.
- [38] Kumra, S. and Kanan, C. (2016) Robotic Grasp Detection Using Deep Convolutional Neural Networks. 2017 *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2017)*, Vancouver, 24-28 September 2017, 8. <https://doi.org/10.1109/IROS.2017.8202237>
- [39] Caldera, S., Rassau, A. and Chai, D. (2018) Review of Deep Learning Methods in Robotic Grasp Detection. *Multi-Modal Technologies and Interaction*, **2**, 57. <https://doi.org/10.3390/mti2030057>
- [40] Saxena, A., Driemeyer, J., Kearns, J. and Ng, A.Y. (2006) Robotic Grasping of Novel Objects. *NIPS'06: Proceedings of the 19th International Conference on Neural Information Processing Systems*, Vancouver, 4-7 December 2006, 1209-1216.

- [41] Zhang, F., Leitner, J., Milford, M., Upcroft, B. and Corke, P. (2015) Towards Vision-Based Deep Reinforcement Learning for Robotic Motion Control. *Proceedings of the Australasian Conference on Robotics and Automation*, Canberra, 2-4 December 2015, 8.
- [42] Kahaki, S.M.M., Nordin, M.J. and Ashtari, A.H. (2014) Contour-Based Corner Detection and Classification by Using Mean Projection Transform. *Sensors*, **14**, 4126-4143. <https://doi.org/10.3390/s140304126>
- [43] Bo, L., Ren, X. and Fox, D. (2012) Unsupervised Feature Learning for RGB-D Based Object Recognition. *International Symposium on Experimental Robotics (ISER)*, Québec City, 18-21 June 2012, 387-402. https://doi.org/10.1007/978-3-319-00065-7_27
- [44] Meyes, R., Hassan, T., Thomas, T., *et al.* (2017) Motion Planning for Industrial Robots Using Reinforcement Learning. *Procedia CIRP*, **63**, 107-112. <https://doi.org/10.1016/j.procir.2017.03.095>