



# Survey of imitation learning for robotic manipulation

Bin Fang<sup>1</sup> · Shidong Jia<sup>2</sup> · Di Guo<sup>1</sup> · Muhua Xu<sup>2</sup> · Shuhuan Wen<sup>3</sup> · Fuchun Sun<sup>1</sup>

Received: 29 May 2019 / Accepted: 9 September 2019  
© Springer Nature Singapore Pte Ltd. 2019

## Abstract

With the development of robotics, the application of robots has gradually evolved from industrial scenes to more intelligent service scenarios. For multitasking operations of robots in complex and uncertain environments, the traditional manual coding method is not only cumbersome but also unable to adapt to sudden changes in the environment. Imitation learning that avoids learning skills from scratch by using the expert demonstration has become the most effective way for robotic manipulation. The paper is intended to provide the survey of imitation learning of robotic manipulation and explore the future research trend. The review of the art of imitation learning for robotic manipulation involves three aspects that are demonstration, representation and learning algorithms. Towards the end of the paper, we highlight areas of future research potential.

**Keywords** Imitation learning · Robotic manipulation · Demonstration · Representation

## 1 Introduction

Traditionally, the manipulation capabilities of robots are acquired by hard coding. Different programs are required to adjust to different tasks and environments. It is difficult for the robot to handle dynamically changing events. This directly affects the flexibility and efficiency of the robot, and thus more intelligent methods are required to improve the manipulation capabilities of robots.

With the rapid development of artificial intelligence, machine learning methods are used to make robot learn and acquire manipulation skills in dynamic environment, which makes up for the defects of traditional programming methods. Nowadays, imitation learning, which is also known as learning by demonstration, apprenticeship learning, has

become the most effective way for robots to acquire skills (Yang et al. 2016). Compared with traditional methods, the imitation learning algorithm avoids tedious codes designed for specific scenes and specific tasks (Hussein et al. 2017). The advantages of imitation learning mechanisms are summarized as follows.

- *Enhancing adaptability* If an individual robot has the ability to imitate by observing the movements of others, it is possible to quickly learn useful actions and then adapt to the new environment.
- *Improving communication efficiency* Imitation provides individuals with an effective means of non-verbal communication so that individuals can learn from other individuals of different types or different hardware. Since a large amount of important information is sent during each action, communication through imitation is also efficient.
- *Improving learning efficiency* The high efficiency of the learning process is the most important advantage of imitation learning. When an individual acquires a new behavior, it will spread quickly in the population. Imitation combines the learning abilities of all individuals, allowing useful behavior to spread quickly, thereby increasing the adaptability and viability of the entire group.
- *Compatible with other learning mechanisms* Imitation learning can be combined with other machine learning

✉ Bin Fang  
fangbin@mail.tsinghua.edu.cn

<sup>1</sup> Department of Computer Science and Technology, Beijing National Research Center for Information Science and Technology, The Institute for Artificial Intelligence, State Key Lab of Intelligent Technology and Systems, Tsinghua University, Beijing 100084, People's Republic of China

<sup>2</sup> Key Lab of Industrial Computer Control Engineering of Hebei Province, Yanshan University, Qinhuangdao 066004, People's Republic of China

<sup>3</sup> Princeton International School of Mathematics and Science, 19 Lambert Drive, Princeton, NJ 08540, USA

mechanisms, such as reinforcement learning that can improve the speed and accuracy of imitation learning.

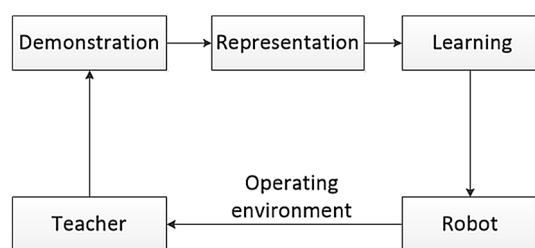
Benefiting from these advantages, imitation learning has developed rapidly in recent years and has become a hot research topic in the area of robotic learning. The process of the imitation learning of robotic system generally consists of three parts: demonstration, representation and imitation learning algorithm, as shown in Fig. 1. We will analyze the current research status of these three parts in the following section.

## 2 Imitation learning for robotic manipulation

### 2.1 Demonstration

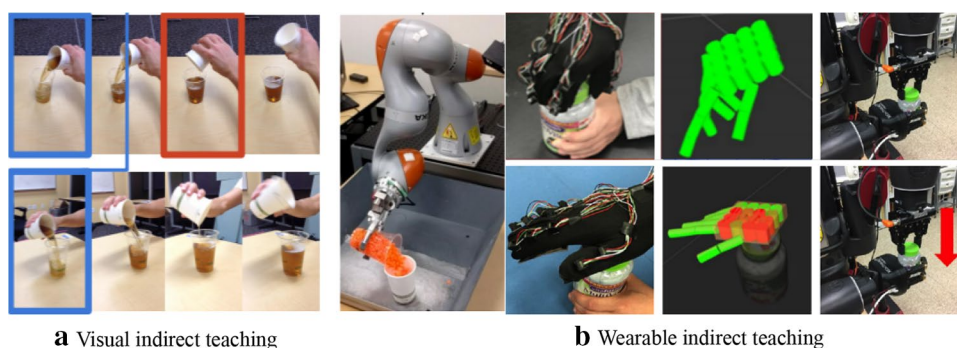
In the robotic manipulation learning process, operation information acquisition is the process of the impersonator obtaining the teacher's operation information through "observation", which is the basis of the imitation learning process (Attia and Dayan 2018). Currently, the demonstration methods of imitation learning are roughly divided into two categories: indirect demonstration and direct demonstration (Kumar et al. 2016).

Indirect demonstration does not involve contact with the robot, instead, the demonstration is built in a separate environment and manipulation information is collected during the teaching process (Wan et al. 2017). Indirect teaching



**Fig. 1** Framework of imitation learning for robotic manipulation

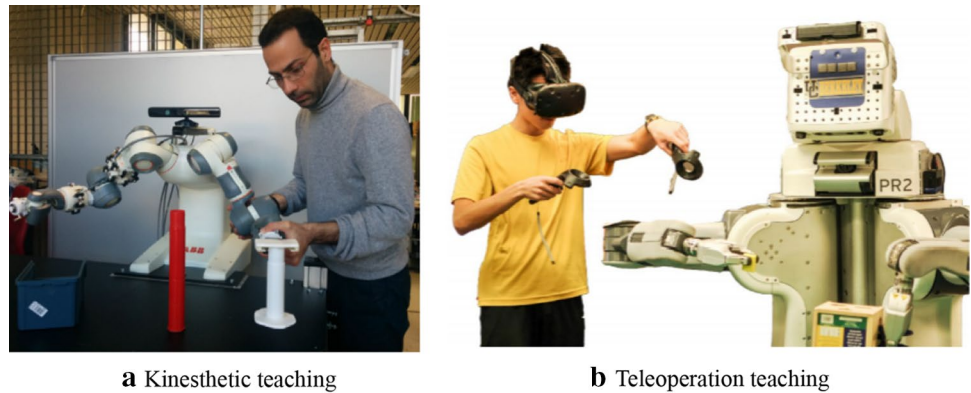
**Fig. 2** Example of indirect teaching



often uses visual systems (Sermanet et al. 2018) and wearable devices (Edmonds et al. 2017) to directly capture human motion information for robots to generate anthropomorphic operations, as shown in Fig. 2. Visual indirect teaching involves observing the teacher's image through machine learning, the high learning speed of this method has made it a hot research topic. However, such teaching samples lack some information, especially tactile information that is crucial for robotic manipulation. Wearable indirect teaching collects samples through wearable sensors, often obtaining more accurate and rich information (Fang et al. 2017a). However, since indirect teaching is separated from the robot and the operating characteristics of the robot itself are not considered, the quality of demonstration cannot be guaranteed.

Direct demonstration obtains teaching samples directly from the robot, making the process fast and simple. And the robot's action is more precise. Direct teaching can be divided into kinesthetic teaching (Amir and Matteo 2018) and teleoperation teaching (Zhang et al. 2018), as shown in Fig. 3. Kinesthetic teaching means that the operator directly contacts and guides the robot to complete a certain action, and the robot collects information by itself (Gaspar et al. 2018). This teaching method does not need to consider the different kinematic parameters between the robot and the human body, and the collected training data has less noise, reflecting a control process that is intuitive. Imitation learning based on kinesthetic teaching methods has become very popular. However, robots suitable for this type of interaction must be passively controllable and require direct contact. Therefore, it is not suitable for robots with multiple degrees of freedom, such as multiple degrees of freedom robotic arm or dual-arm robots.

Teleoperation teaching can be done by joystick (Yang et al. 2017), tactile sensor (Argall and Billard 2010), control panel (Schreiber et al. 2010), infrared sensor (Jin et al. 2016), wearable device (Fang et al. 2017b) and other remote control devices to teach. The demonstrator is not bounded by the robot itself during the teaching process, and safety is guaranteed since there is no direct contact between the teacher and the robot (Liu and Wang 2018). Thus, teleoperation

**Fig. 3** Direct teaching example

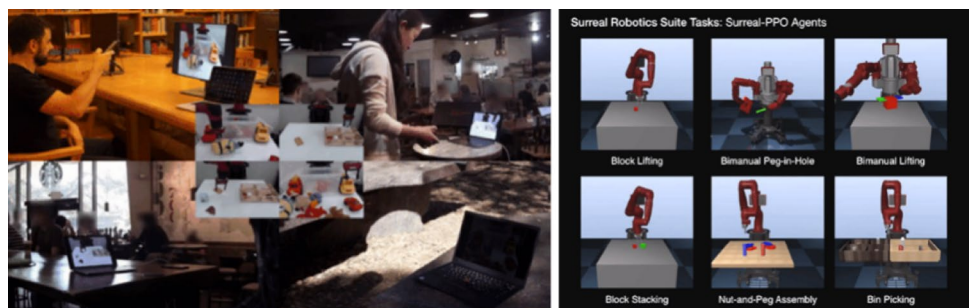
demonstration shows obvious advantages such as high safety and wide application range (Gong et al. 2018), and it can obtain high-quality teaching samples through teleoperation. For this purpose, Li Feifei's team built the RoboTurk network platform for demonstrations collection (Ajay et al. 2018). The demonstrator can use a mobile phone or mouse to freely remotely operate the robot arm, thus conveniently obtaining a large dataset, as shown in Fig. 4. Despite these advantages, most of the current teleoperations are merely teaching based on posture or trajectory, and lack of information on the actual operational force, making it difficult for the robot to perform the task of fine operation.

## 2.2 Representation

After the demonstrator completes the manipulation, the sample may contain many features, such as its environment, operational objects, etc. However, in most demonstrations, the operating environment is often so complex that it contains a large amount of irrelevant and redundant features that don't exactly match the robot. This forbids the robot from completing subsequent imitation operations. Therefore, it is important to characterize the demonstration sufficiently and efficiently into a form that can be both recognized and applied by the robot (Alibeigi et al. 2017). In imitation learning, manipulation demonstration can be divided into three types: symbolic characterization, trajectory characterization, and motion-state spatial characterization (Osa et al. 2017).

In symbolic characterization, the strategy that a robot generates through learning is a series of options, each of which contains a subsequence of actions sorted by time. Different tasks are accomplished by selecting different options in the list. For the complex task it is difficult to be accomplished by the single action, the symbolic representation strategy characterizes the complex task by sorting the simple action sequence in the options. Symbolic representation is a high-level behavioral representation of robots in imitation learning. The advantages are: (1) The same basic action can be reused to accomplish different tasks. (2) Convenient to replace the actions in the option's action sequence to adjust the relevant motion plan. Imitation learning with symbolic representation is a convenient and feasible way to improve the intelligence of multi-modal integrative robots, and it also provides a practical solution for solving complex, multi-step task learning problems. However, when facing complex and fine-grained transformations, it is difficult to achieve accurate segmentation and ordering of motions by means of symbolic characterization, which makes it difficult to achieve effective operational characterization.

Trajectory characterization can be understood as the mapping of task-related conditions and trajectories, which is a lower-level representation comparing to symbolic representation. The task-related conditions here may be the initial state and the target state of the system, such as the initial position of the gripper in a grabbing task and the target position of the object being grasped. And the trajectories can

**Fig. 4** Roboturk remote learning data collection platform

be abstracted into time series of system states or inputs. In symbolic representation, individual actions in the sequence of actions can be characterized by means of trajectory characterization. Trajectory characterization is like the dynamic motion primitive method in behavioral cloning (Pastor et al. 2009), which introduces an additional forcing term in the critically damped spring system to normalize its characterization of complex individual actions. Another example is the probabilistic motion primitive method used in Dermý's paper (Dermý et al. 2017), which describes the trajectory as a motion primitive distributed by probability. In practical applications, Yang's paper (Yang et al. 2017) extracted the impedance of the human body and applied it to the robot, which extended the dynamic motion primitive method. However, the method of trajectory characterization requires that the teaching sample contains as many dynamic features as possible, which makes it difficult to be performed on samples containing only non-dynamic features such as images and touch.

The action-state space is characterized differently from symbolic characterization and trajectory characterization. Instead of generating a series of complete executable strategy options and action sequences in advance, it generates a series of action-state decisions that are determined when a certain state occurs. The corresponding control actions establish the mapping relationship between the task-related conditions and the system state to the control actions. The method of motion-state space representation is used in dynamic system stability observation method of behavioral cloning. This method obtains a nonlinear autonomous dynamic system capable of generating action-state decision (Khansari-Zadeh and Billard 2011). Action-state space characterization also includes the inverse reinforcement learning method used in Faha et al.'s (2018) and Piot et al.'s (2016) paper, which assumes the optimality of expert teaching and learns the reward function and uses the reward function to generate an action-state decision. However, because the action-state space is characterized by a series of decisions describing short-term or instantaneous behavior, it is easy to accumulate the characterization error in the process of long-term representation.

Demonstrations may contain a wide variety of features such as dynamics, vision, and touch. Yang indicates the dynamic characteristics of the joint angle, joint velocity and end pose of the robot arm are generated by teaching, and the operation is characterized by the trajectory characterization (Yang et al. 2018). The humanoid robot made by Hwang et al. (2016) simulates human motion through visual observation, creating a three-dimensional image sequence using a stereoscopic vision capture system, and the extracted visual features are then used to estimate the trajectory of the teacher. In addition to dynamics and visual features, tactile information is also of great importance. The introduction

of tactile sequences in operational teaching and imitation has also become a recent hot research topic. In the past few years, many machine learning methods, such as nearest neighbor, support vector machine, Gaussian process, and nonparametric Bayesian learning, have been successfully used for tactile material identification. The tactile features are firstly introduced in Liu et al. (2019), which improves the success rate of robotic cap opening operations. At present, most of the research still focused on the operational characterization of dynamics and visual features. However, multimodal information combined with tactile features can extend the operational characterization to obtain better characterization results. Therefore, the study of multimodal operational information characterization has great significance in imitation learning.

### 2.3 Imitation learning algorithm

In the complete process of robotic imitation learning, demonstration teaching provides a teaching sample containing rich features, operation characterization characterizes the features in the teaching sample as valid forms that the robot can recognize. The ultimate goal of imitative learning is to make the robot "master" behavior, which means that the robot has to reproduce the behavior and generalize the behavior into other unknown scenes. The process that robots use the teaching information to "master" behavior can be called operation imitation. The current mainstream methods of operational imitation are roughly divided into three categories: behavioral cloning, inverse reinforcement learning, and adversarial imitation learning.

Behavioral cloning method uses the teaching information to establish a direct mapping of state and task-related conditions with trajectories and actions, similar to the feature-tag in supervised learning (Torabi et al. 2018). According to whether it depends on the model, it can be divided into model-based behavioral cloning and non-model-based behavioral cloning. Based on operational characterization, it can be divided into behavioral cloning of trajectory characterization, behavioral cloning of state-action space characterization and symbolic representation behavior clone. There are many types of behavioral cloning methods that are freely combined according to the above two classification methods. Examples are the dynamic motion primitives in Ijspeert et al. (2013) and the mixed motion primitive methods used in Gams et al. (2014) and Amor et al. (2014), such methods can generate continuous, smooth and generalizable representations in terms of trajectories. In Andrew (2015), picture information is taken as the state, and steering wheel angle is taken as the action, thus the state-action space decision is learned. The above operation imitation method does not need to consider the model. This is because based on the characteristics of the model, the actions or trajectories



generated by learning the teaching information are prone to be unreachable and not executable during the execution of the robot (Wu et al. 2019). Therefore, in Finn et al. (2016), a dynamic model of robot with high-degree of freedom is established using a deep network. However, in the case of a limited number of samples, the strategy obtained by behavioral cloning is not very adaptable.

The inverse reinforcement learning method restores the reward function representing the expert's intention by assuming the optimality of the expert teaching information and uses the reinforcement learning method to obtain the final control strategy based on the reward function (Montaser et al. 2019; Rajeswaran et al. 2018). The advantage of inverse reinforcement learning is that in the case of insufficient teaching samples, the reward function can be obtained in reverse, resulting in a generalized strategy (Justin et al. 2018). Similar to the classification of behavioral cloning, inverse reinforcement learning can also be classified according to whether it depends on the model and the three characterizations. Ratliff et al. (2006) find a reward function that can make the optimal strategy the most different from other strategies (Zucker et al. 2011). Based on this, it can be generalized to a reward function that can be used for nonlinearity. Ziebart et al. (2008) used the probabilistic model to propose the inverse reinforcement learning of the maximum entropy model, which overcame the random deviation of the reward function caused by the expert teaching preference. In terms of application, Finn et al. (2017) obtained a nonlinear reward function through inverse reinforcement learning, which can guide the robotic arm to complete complex housework tasks. Inverse reinforcement learning restores the reward function based on the optimality of the expert teaching information, making the strategy generated less effective in an environment that is very different from the teaching environment.

Behavioral cloning and inverse reinforcement learning methods only learn strategies from expert teaching information and do not interact with expert teaching information to further optimize learning strategies (Cai et al. 2019). The generation of adversarial imitation learning is a method of completing imitation learning in combination with the generative adversarial nets (Goodfellow et al. 2014), which confronts the generated trajectory with the expert teaching trajectory, distinguishes the expert trajectory or the imitator trajectory by the classifier, and then performs iterative confrontation training to make the distribution between the two as close as possible to complete the operational imitation (Kuefler and Morton 2017; Ho 2016). Baram et al. proposed a method based on forward model to make the stochastic strategy completely differentiable to generate imitation learning (Baram et al. 2017). Henderson et al. (2018) proposed an imitation learning method for the optional framework of hierarchical strategy. It all implements an extension

to the generation of adversarial learning methods. Confronting the imitation generated trajectory to the expert teaching trajectory and simulating whether the generated trajectory is reachable and executable are essential to the process. Therefore, the system with insufficient model information will seriously affect the generation of adversarial imitation learning.

### 3 Discussion and conclusions

In summary, imitation learning has become an important key technology in the field of robot operation. However, the current research work in this area still faces many challenges, which are summarized as follows:

1. *Demonstration* Although great progress has been made in teaching via wearable devices, most of teleoperation teachings only consider the position and posture, and most of them are the mechanical arms of the end jaws, lacking information about the overall operation of the hand-arm system. For a multi-degree-of-freedom humanoid manipulator, it is necessary to consider the operational configuration, position, attitude, and the dexterous hand's operating force, that is, the tactile teaching. But how to integrate multimodal teaching methods to achieve high-quality teaching samples remains a challenging issue.
2. *Representation* Using the teaching operation sample to characterize the learning state and operational intent of the teacher is an important step in imitation learning. Most of the current research focuses on the trajectory or visual representation. Although visual and tactile representations in our teaching samples can provide more information to imitation learning, how to exploit the correlation between visual and tactile information with robotic operations and how to learn the characterization of multimodal information are very important issues in practical applications. Research in this area has just begun. And it is not only the cornerstone of operational characterization but also an important direction for future multimodal teaching information characterization.
3. *Learning* The existing imitative operation learning has a low utilization rate of teaching samples, and cannot achieve efficient strategy learning. At the same time, the operation of the imitation learning algorithm is more sensible for multi-modal characteristics, operational space locality, and small samples which pose a great challenge to the generalization of imitation operations. How to design an efficient robotic imitation learning framework is still at the forefront of robot learning.

In general, the multi-modal imitation learning of robotic operation technology provides a more efficient and high-quality way for the robot to better grasp the operational skills, which is of great significance for improving the operational intelligence of robots. There are still many challenging academic problems in this field, and it is necessary to carry out in-depth exploration and analysis from the perspectives of signal processing, machine learning, and robot operation theory.

**Acknowledgements** This work is jointly supported by Foshan-Tsinghua industry-university-research cooperation collaborative innovation special fund no. 2018THFS04, Tsinghua University Initiative Scientific Research Program no. 2019Z08QCX15, National Natural Science Foundation of China under with Grant nos. 91848206, U1613212 and 61703284.

## References

- Ajay, M., Zhu, Y., Li, F.: Roburk: a crowdsourcing platform for robotic skill learning through imitation. In: Conference on Robot Learning, pp. 1–15 (2018)
- Alibeigi, M., Ahmadabadi, M.N., Araabi, B.N.: A fast, robust, and incremental model for learning high-level concepts from human motions by imitation. *IEEE Trans. Robot.* **33**(1), 153–168 (2017)
- Amir, M., Matteo, R.: Robot learning from demonstrations: emulation learning in environments with moving obstacles. *Robot. Auton. Syst.* **101**, 45–56 (2018)
- Amor, H., Neumann, G., Kamthe, S., Kroemer, O., Peters, J.: Interaction primitives for human–robot cooperation tasks. In: IEEE International Conference on Robotics and Automation, pp. 2831–2837 (2014)
- Andrew, J.: An invitation to imitation. Technical report, Robotics Institute, Carnegie Mellon University (2015)
- Argall, B., Billard, A.: A survey of tactile human–robot interactions. *Robot. Auton. Syst.* **58**, 1159–1176 (2010)
- Attia, A., Dayan, S.: Global overview of imitation learning. <https://arxiv.org/abs/1801.06503> (2018)
- Baram, N., Anschel, O., Caspi, I., Mannor, S.: End-to-end differentiable adversarial imitation learning. In: International Conference on Machine Learning (ICML), pp. 390–399 (2017)
- Cai, Q., Hong, M., Chen, Y., Wang, Z.: On the global convergence of imitation learning: a case for linear quadratic regulator. <https://arxiv.org/abs/1901.03674> (2019)
- Dermi, O., Charpillet, F., Ivaldi, S.: Multi-modal intention prediction with probabilistic movement primitives. In: International Workshop on Human-Friendly Robotics, pp. 181–196 (2017)
- Edmonds, M., Gao, F., Xie, X., Liu, H., Qi, S., Zhu, Y., Rothrock, B., Zhu, S.: Feeling the force: integrating force and pose for fluent discovery through imitation learning to open medicine bottles. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 3530–3537 (2017)
- Fahad, M., Chen, Z., Guo, Y.: Learning how to pedestrians navigate: a deep inverse reinforcement learning approach. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (2018)
- Fang, B., Sun, F., Liu, H., Guo, D.: Development of a wearable device for motion capturing based on magnetic and inertial measurement units. *Sci. Program.* (2017a). <https://doi.org/10.1155/2017/7594763>
- Fang, B., Sun, F., Liu, H., Guo, D.: Robotic teleoperation systems using a wearable multi-modal fusion device. *Int. J. Adv. Robot. Syst.* 1–11 (2017b)
- Finn, C., Abbeel, P., Levine, S.: Model-agnostic meta-learning for fast adaptation of deep networks. In: International Conference on Machine Learning (ICML), pp. 1126–1135 (2017)
- Finn, C., Levine, S., Abbeel, P.: Guided cost learning: deep inverse optimal control via policy optimization. In: International Conference on Machine Learning, pp. 49–58 (2016)
- Gams, A., Nemec, B., Ijspeert, A.J., Ude, A.: Coupling movement primitives: interaction with the environment and bimanual tasks. *IEEE Trans. Robot.* **30**(4), 816–830 (2014)
- Gaspar, T., Nemec, B., Morimoto, J., Ude, A.: Skill learning and action recognition by arc-length dynamic movement primitives. *Robot. Auton. Syst.* **100**, 225–235 (2018)
- Gong, D., Zhao, J., Yu, J., Zuo, G.: Motion mapping of the heterogeneous master–slave system for intuitive telemanipulation. *Int. J. Adv. Rob. Syst.* **15**, 1–9 (2018)
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. In: International Conference on Neural Information Processing Systems, pp. 2672–2680 (2014)
- Henderson, P., Chang, W., Bacon, P., Meger, D., Pineau, J., Precup, D.: OptionGan: learning joint reward-policy options using generative adversarial inverse reinforcement learning. In: National conference on Artificial Intelligence (2018)
- Ho, J., Ermon, S.: Generative adversarial imitation learning. *Adv. Neural Inf. Process. Syst.* (2016). <https://arxiv.org/abs/1606.03476>
- Hussein, A., Gaber, M.M., Elyan, E., Jayne, C.: Imitation learning: a survey of learning methods. *ACM Comput. Surv.* (2017). <https://doi.org/10.1145/3054912>
- Hwang, C., Chen, B., Syu, H., Wang, C., Karkoub, M.: Humanoid robot's visual imitation of 3-D motion of a human subject using neural network based inverse kinematics. *IEEE Syst. J.* **10**(2), 685–696 (2016)
- Ijspeert, A., Nakanishi, J., Hoffmann, H., Pastor, P., Schaal, S.: Dynamical movement primitives: learning attractor models for motor behaviors. *Neural Comput.* **25**(2), 328–373 (2013)
- Jin, H., Chen, Q., Chen, Z., Hu, Y., Zhang, J.: Multi-leap motion sensor based demonstration for robotic refine tabletop object manipulation task. *Trans. Intell. Technol.* **1**, 104–113 (2016)
- Justin, F., Luo, K., Levine, S.: Learning robust rewards with adversarial inverse reinforcement learning. In: International Conference on Learning Representations, pp. 1–15 (2018)
- Kuefler, A., Morton, J., Wheeler, T.: Imitating driver behavior with generative adversarial networks. In: IEEE Intelligent Vehicles Symposium, pp. 204–211 (2017)
- Khansari-Zadeh, S.M., Billard, A.: Learning stable nonlinear dynamical systems. *IEEE Trans. Robot.* **27**, 943–957 (2011)
- Kumar, V., Gupta, A., Todorov, E., Levine, S.: Learning dexterous manipulation policies from experience and imitation. <https://arxiv.org/abs/1611.05095> (2016)
- Liu, H., Wang, L.: Gesture recognition for human–robot collaboration: a review. *Int. J. Ind. Ergon.* **68**, 355–367 (2018)
- Liu, H., Zhang, C., Zhu, Y., Jiang, C., Zhu, S.: Mirroring without overimitation: learning functionally equivalent manipulation actions. In: Proceedings of the AAAI Conference on Artificial

- Intelligence, vol 33 (2019). <https://doi.org/10.1609/aaai.v33i01.33018025>
- Montaser, M., Waleed, D., Benjamin, R.: Transfer learning for prosthetics using imitation learning. <https://arxiv.org/abs/1901.04772> (2019)
- Osa, T., Pajarinen, J., Neumann, G., Bagnell, J.A., Abbeel, P., Peters, J.: An algorithmic perspective on imitation learning. *Found. Trends Robot.* **7**(1), 1–179 (2017)
- Pastor, P., Hoffmann, H., Asfour, T., Schaal, S.: Learning and generalization of motor skills by learning from demonstration. In: *IEEE International Conference on Robotics and Automation* (2009)
- Piot, B., Geist, M., Pietquin, O.: Bridging the gap between imitation learning and inverse reinforcement learning. *IEEE Trans. Neural Netw. Learn. Syst.* **28**, 1814–1826 (2016)
- Rajeswaran, A., Kumar, V., Gupta, A., Vezzani, G., Schulman, J., Todorov, E., Levine, S.: Learning complex dexterous manipulation with deep reinforcement learning and demonstrations. *Robot. Sci. Syst.* (2018). <https://doi.org/10.15607/RSS.2018.XIV.049>
- Ratliff, N., Bagnell, J., Zinkevich, M.: Maximum margin planning. In: *International Conference on Machine learning (ICML)*, pp. 729–736 (2006)
- Schreiber, G., Stemmer, A., Bischoff, R.: The fast research interface for the KUKA lightweight robot. In: *The Workshop on IEEE ICRA 2010 Workshop on Innovative Robot Control Architectures for Demanding*, pp. 15–21 (2010)
- Sermanet, P., Lynch, C., Hsu, J., Levine, S.: Time-contrastive networks: self-supervised learning from video. In: *IEEE International Conference on Robotics and Automation*, pp. 1134–1141 (2018)
- Torabi, F., Warnell, G., Stone, P.: Behavioral cloning from observation. In: *International Joint Conference on Artificial Intelligence*, pp. 4950–4957 (2018)
- Wan, W., Lu, F., Wu, Z., Harada, K.: Teaching robots to do object assembly using multi-modal 3D vision. *Neurocomputing* **259**, 85–93 (2017)
- Wu, Y., Charoenphakdee, N., Bao, H., Tangkaratt, V., Sugiyama, M.: Imitation learning from imperfect demonstration. <https://arxiv.org/abs/1901.09387> (2019)
- Yang, C., Lu, F., Wu, Z., Harada, K.: Development of a robotic teaching interface for human to human skill transfer. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 710–716 (2016)
- Yang, C., Zeng, C., Liang, P., Li, Z., Li, R., Sun, C.: Interface design of a physical human–robot interaction system for human impedance adaptive skill transfer. In: *IEEE Robotics and Automation Society*, pp. 329–340 (2017)
- Yang, C., Wang, X., Cheng, L., Ma, H.: Neural-learning-based telero-bot control with guaranteed performance. *IEEE Trans. Cybern.* **47**, 3148–3159 (2017)
- Yang, C., Zeng, C., Fang, C., He, W., Li, Z.: A DMPs-based framework for robot learning and generalization of humanlike variable impedance skills. *IEEE/ASME Trans. Mechatron.* **23**, 1193–1203 (2018)
- Zhang, T., McCarthy, Z., Jow, O., Lee, D., Goldberg, K., Abbeel, P.: Deep imitation learning for complex manipulation tasks from virtual reality teleoperation. In: *IEEE International Conference on Robotics and Automation*, pp. 5628–5635 (2018)
- Ziebart, B., Maas, A., Bagnell, J.: Maximum entropy inverse reinforcement learning. In: *AAAI Conference on Artificial Intelligence*, pp. 1433–1438 (2008)
- Zucker, M., Ratliff, N., Stolle, M., Chestnutt, J., Andrew Bagnell, J., Atkeson, C., Kuffner, J.: Optimization and learning for rough terrain legged locomotion. *Int. J. Robot. Res.* **30**(2), 175–191 (2011)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Bin Fang** received the Doctor degree in Mechanical Engineering from Beihang University in 2014, and now he is the research assistant in the Department of Computer Science and Technology in Tsinghua University. His research interests include sensor fusion, wearable device, robotics and human-robot interaction.



**Shidong Jia** received a bachelor's degree in Automation from Yan-shan University in 2016. Currently, she is studying for a master's degree in Control theory and Control Engineering at Yan-shan University. Her research interests include Gesture recognition and robotics.



**Di Guo** received the bachelors degree from the Department of Instrumentation Science and Optoelectronics Engineering, Beihang University, Beijing, China. She is currently pursuing the Ph.D. degree with the Department of Computer Science and Technology, Tsinghua University, Beijing. Her current research interests include robotic manipulation and sensor fusion.



**Muhua Xu** is a high-school student currently studying on the east coast of the US. Interested in applied math and computer science, she performs research in areas focusing on Neural Networks and robot intelligence.



**Shuhuan Wen** was received the PhD degree in control theory and control engineering from the Yanshan University, Qinhuangdao, China in 2005. She is currently a Professor of automatic control in the Department of Electric Engineering, Yanshan University. She has coauthored one book, about 40 papers. Her research interests include SLAM, computer vision, robotics, 3-D object recognition and reconstruction. Dr. Wen was a Visiting Scholar of the Ottawa University, Carleton University

and Simon Fraser University in Canada from 2011 to 2013.



**Fuchun Sun** is a Full Professor with the Department of Computer Science and Technology, Tsinghua University, Beijing, China. His current research interest includes robotic perception and cognition. Mr. Sun was a recipient of the National Science Fund for Distinguished Young Scholars. He serves as an Associate Editor of a series of international journals including IEEE transactions on systems, man and cybernetics: systems, the IEEE transactions on fuzzy systems, Mechatronics, Robotics

and Autonomous Systems.