Publications

- **CSGP: Closed-loop Safe Grasp Planning via Attention-based Deep Reinforcement Learning from Demonstrations**. ***Tang Z***, *Shi Y, Xu X.* ***IEEE Robotics and Automation Letters (RA-L)****, vol.8, no. 6, pp. 3158-3165, Jun 2023*

- **SymmetryGrasp: Symmetry-Aware Antipodal Grasp Detection From Single-View RGB-D Images**. *Shi Y,* ***Tang Z****, Cai X, Zhang H, Hu D, Xu X.* ***IEEE Robotics and Automation Letters (RA-L)****, vol. 7, no. 4, pp. 12235-12242, Oct 2022. (also selected by* ***ICRA 2023 for oral presentation****)*

- **A deep Koopman Operator-based Modeling Approach for Long-term Prediction of Dynamics with Pixel-level Measurements**. *Xiao Y,* ***Tang Z****, Xu X, Zhang X, Shi Y.* ***CAAI Transactions on Intelligence Technology****, Feb 2023.*

- **Efficient Reinforcement Learning with Least-squares Soft Bellman Residual for Robotic Grasping**. *Lan Y, Ren J, Tang T, Xu X, Shi Y,* ***Tang Z****.* ***Robotics and Autonomous Systems (RAS)****, vol. 164, 104385, Apr 2023.*

- **Grasp Planning Based on Deep Reinforcement Learning: A Brief Survey**. **Tang Z**, Xu X, Shi Y. **China Automation Congress (CAC)**, pp. 7293-7299, Oct 2021.

CSGP: Closed-Loop Safe Grasp Planning via Attention-Based Deep Reinforcement Learning From Demonstrations

Zixin Tang, Yifei Shi, *Member, IEEE*, and Xin Xu, *Senior Member, IEEE*

Abstract—**Grasping is at the core of many robotic manipulation tasks. Despite the recent progress, closed-loop grasp planning in stacked scenes is still unsatisfactory, in terms of efficiency, stability, and most importantly, safety. In this letter, we present CSGP, a closed-loop safe grasp planning approach via attention-based deep reinforcement learning (DRL) from demonstrations, which is capable of learning safe grasping policies that make surrounding objects less disturbed or damaged during manipulation. In CSGP, a 6-DoF safe grasping policy network with a Next-Best-Region attention module is proposed to intrinsically identify the safe regions in the view, facilitating the learning of safe grasping actions. Moreover, we design a fully automatic pipeline in the simulator to collect safe grasping demonstrations, which are utilized to pre-train the policy with behavior cloning and fine-tune it with DRL. To effectively and stably improve the policy during fine-tuning, a DRL from demonstrations algorithm named Safe-DDPGfD is presented in CSGP with a truncated height-anneal exploration mechanism for safe exploration. Moreover, we provide a benchmark that contains scenes with multiple levels of stack layers for method evaluation. Simulation results demonstrate the state-of-the-art performance of our method, achieving the Overall score of 88% in our benchmark. Also, real-world robot grasping experiments also show the effectiveness of our method.**

Index Terms—**Deep learning in grasping and manipulation, deep learning for visual perception, learning from demonstration.**

I. Introduction

As one of the most fundamental skills in robotic manipulation, grasping has been extensively studied in the past decades and has benefited a wide range of applications, such as bin picking [1] and object rearrangement [2]. Unlike open-loop grasp which requires synthesizing target grasp poses before moving, closed-loop grasp aims at finding a collision-free trajectory of the robot that first approaches and then grasps an object via real-time feedback (e.g., visual servoing) [3]. This property makes grasping more robust to perceptual errors, kinematic noises, and perturbations in scenes [4].

Manuscript received 7 November 2022; accepted 13 February 2023. Date of publication 6 March 2023; date of current version 17 April 2023. This letter was recommended for publication by Associate Editor G. Averta and Editor M. Vincze upon evaluation of the reviewers' comments. This work was supported by National Natural Science Foundation of China (NSFC) under Grants 61825305 and 62002379. *(Corresponding author: Yifei Shi and Xin Xu.)*

The authors are with the College of Intelligence Science and Technology, National University of Defense Technology, Changsha 410073, China (e-mail: tangzixin6135@outlook.com; yifei.j.shi@gmail.com; xinxu@nudt.edu.cn).

This letter has supplementary downloadable material available at https://doi.org/10.1109/LRA.2023.3253023, provided by the authors.

Digital Object Identifier 10.1109/LRA.2023.3253023

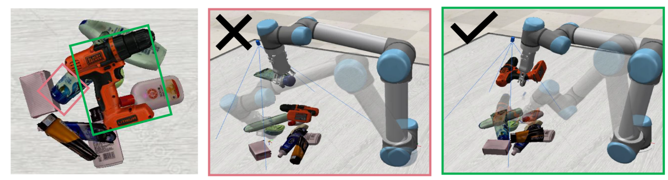

Fig. 1. An example of grasping in the stacked scene. **Left:** stacked scenario with multiple objects, where the object framed in pink is buried and the one framed in green is a safe object. **Middle:** grasping a buried object leads to severe damage to surrounding objects. **Right:** grasping a safe object makes surrounding objects intact.

Plenty of works on closed-loop grasp planning have focused on adopting advanced learning schemes [5] or leveraging large-scale datasets [3], [1], [4], [6] to improve grasping performance in various experimental settings. Despite the above progress, closed-loop grasp is still far from mature, especially in stacked scenes with fragile objects (see Fig. 1). In such circumstances, each grasping attempt should perform a *safe grasping policy* that makes surrounding objects less disturbed or damaged during manipulation. However, learning the 6-DoF safe grasping policy in a closed-loop manner is little explored in previous research. Many existing methods, which either do not consider safety [3], [6] or consider safe grasping in an open-loop manner [7], [8], cannot be directly applied.

In general, two main challenges hinder the learning of the safe grasping policy. First, the diversity of object geometries and complex stacked layouts make safe grasps hard to be determined, which requires careful analyses of both the objects [9], [10] and their stacked relations [8], [11]. Moreover, learning such policies with deep reinforcement learning (DRL) is challenging, in terms of efficiency and stability [1], [12].

To address the above challenges, we present CSGP, a closed-loop safe grasp planning method via attention-based DRL from demonstrations. In CSGP, a 6-DoF safe grasping policy network with a *Next-Best-Region (NBR) attention module* is proposed to intrinsically identify the safe regions. At its core, a cross-attention mechanism with a learnable descriptor is adopted in the module. To learn the network efficiently, in CSGP, we make sufficient use of demonstrations in a two-stage training scheme that pre-trains the network via behavior cloning (BC) and fine-tunes it with DRL algorithms. To this end, we first design a fully automatic pipeline in the simulator to collect diverse demonstrations satisfying our demand for safe grasping.

SymmetryGrasp: Symmetry-Aware Antipodal Grasp Detection From Single-View RGB-D Images

Yifei Shi, Zixin Tang, Xiangting Cai, Hongjia Zhang, Dewen Hu, *Senior Member, IEEE*, and Xin Xu, *Senior Member, IEEE*

***Abstract*—Symmetry is ubiquitous in everyday objects. Humans tend to grasp objects by recognizing the symmetric regions. In this letter, we investigate how symmetry could boost robotic grasp detection. To this end, we present a learning-based method for detecting grasp from single-view RGB-D images. The key insight is to explicitly incorporate symmetry estimation into grasp detection, improving the quality of the detected grasps. Specifically, we first introduce a new grasp parameterization in grasp detection for parallel grippers based on symmetry. Based on this representation, a symmetry-aware grasp detection network method is present to simultaneously estimate object symmetry and detect grasp. We find that the learning of grasp detection greatly benefits from symmetry estimation, improving the training efficiency and the grasp quality. Besides, to facilitate the cross-instance generality of grasping unseen objects, we propose Principal-directional scale-Invariant Feature Transformer (PIFT), a plug-and-play module, that allows spatial deformation of points during the feature aggregation. The module essentially learns feature invariance to anisotropic scaling along the shape principal directions. Extensive experiments demonstrate the effectiveness of the proposed method. In particular, it outperforms previous methods, achieving state-of-the-art performance in terms of grasp quality on GraspNet-1-Billion and success rate on a real robot grasping experiment.**

***Index Terms*—RGB-D perception, deep learning in grasping and manipulation, deep learning for visual perception.**

I. Introduction

As ONE of the most fundamental skills for intelligent robots, grasping has a wide range of applications from bin-picking for industrial robots to general object grasping for home service robots. With the recent progress of RGB-D sensing and 3D learning techniques, detecting feasible grasp from cluttered scenes has attracted a surge of research attention from robotics and computer vision communities.

Recent advances in grasp detection are dominated by learning-based approaches. According to whether the object geometry is known beforehand or not, these approaches can be divided into two categories: the model-based methods [36] and the model-free methods [9]. While the former is not always applicable to novel object grasping, the latter requires a large number of data with careful annotations to be trained with. In particular, those methods are prone to fall short in grasping novel objects located in cluttered scenes, due to the inferior generality caused by insufficient network training. To solve this problem, recent research has been focused on improving the quality of the train data [9], [38] or involving the advanced learning models [18], [24]. In this paper, we study the problem of grasp detection from a new perspective. The key insight is simple: *symmetry is prevalent in everyday objects, and humans tend to grasp objects with fingers placed at symmetric regions*. Theoretically, symmetry detection and grasp detection are highly relevant tasks in both psychology [8], [16] and robotics [1], [2]. Besides, from the prospect of geometry learning, symmetry is a lightweight yet informative representation of 3D geometry, which can potentially narrow the searching space of grasp detection. As such, understanding object symmetry seems to be an intermediate and complementary stage of grasp detection. It is therefore interesting to study whether symmetry detection could be incorporated into the grasp detection pipelines and boost the performance.

We present SymmetryGrasp, a learning-based method for detecting antipodal grasp from single-view RGB-D images with the guidance of symmetry estimation (Fig. 1). Our contributions are two-fold. First, we introduce a new parameterization of grasp detection for parallel grippers based on symmetry. The parameterization is designed to explicitly exploit symmetry clues and is expected to improve the robustness and generality of general grasp detection. Based on this parameterization, a symmetry-aware grasp detection network method is presented to simultaneously estimate symmetry and detect grasp. The detected grasps are located in local symmetric regions and regularized by the estimated symmetry. With a dedicated network design,

Fig. 1. SymmetryGrasp couples the detection of symmetry and grasp into a unified framework. Given a single-view RGB-D image (left), it takes advantage of the estimated symmetries (middle) to determine the feasible antipodal grasps (right), greatly increasing grasp quality and generality.

Manuscript received 1 June 2022; accepted 26 September 2022. Date of publication 14 October 2022; date of current version 25 October 2022. This letter was recommended for publication by Associate Editor D. Jha and Editor C. C. Lerma upon evaluation of the reviewers' comments. This work was supported in part by NSFC under Grants 61825305 and 62002379 and in part by the National Key Research and Development Program of China under Grants 2018YFB1305105. *(Corresponding author: Xin Xu.)*

The authors are with the College of Intelligence Science and Technology, National University of Defense Technology, Changsha 410073, China (e-mail: yifei.j.shi@gmail.com; tangzixin6135@outlook.com; 649112160@qq.com; h_zhang00@163.com; dwhu@nudt.edu.cn; xinxu@nudt.edu.cn).

Digital Object Identifier 10.1109/LRA.2022.3214785

Received: 10 March 2022 | Revised: 23 June 2022 | Accepted: 23 November 2022

CAAI Transactions on Intelligence Technology

DOI: 10.1049/cit2.12149

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH

A deep Koopman operator-based modelling approach for long-term prediction of dynamics with pixel-level measurements

Yongqian Xiao | Zixin Tang | Xin Xu | Xinglong Zhang | Yifei Shi

College of Intelligence Science and Technology, National University of Defense Technology, Changsha, China

Correspondence

Xin Xu and Xinglong Zhang, College of Intelligence Science and Technology, National University of Defense Technology, Changsha 410073, China.
Email: xuxin_mail@263.net and zhangxinglong18@nudt.edu.cn

Funding information

National Natural Science Foundation of China, Grant/Award Numbers: 61825305, 62003361, U21A20518; China Postdoctoral Science Foundation, Grant/Award Number: 47680

Abstract

Although previous studies have made some clear leap in learning latent dynamics from high-dimensional representations, the performances in terms of accuracy and inference time of long-term model prediction still need to be improved. In this study, a deep convolutional network based on the Koopman operator (CKNet) is proposed to model non-linear systems with pixel-level measurements for long-term prediction. CKNet adopts an autoencoder network architecture, consisting of an encoder to generate latent states and a linear dynamical model (i.e., the Koopman operator) which evolves in the latent state space spanned by the encoder. The decoder is used to recover images from latent states. According to a multi-step ahead prediction loss function, the system matrices for approximating the Koopman operator are trained synchronously with the autoencoder in a mini-batch manner. In this manner, gradients can be synchronously transmitted to both the system matrices and the autoencoder to help the encoder self-adaptively tune the latent state space in the training process, and the resulting model is time-invariant in the latent space. Therefore, the proposed CKNet has the advantages of less inference time and high accuracy for long-term prediction. Experiments are performed on OpenAI Gym and Mujoco environments, including two and four non-linear forced dynamical systems with continuous action spaces. The experimental results show that CKNet has strong long-term prediction capabilities with sufficient precision.

KEYWORDS

deep neural networks, image motion analysis, image sequences, sequential estimation

1 | INTRODUCTION

For many physical systems, accurate models play important roles in model-based planning, control, and trajectory forecasting [1-4]. Therefore, it is meaningful to develop data-driven modelling approaches for systems with unknown dynamics [5, 6]. Prior works were developed for modelling complex non-linear systems, including vehicle dynamics [7], fluid dynamics [8] etc. In some cases without sensor support for obtaining state measurements, pixel-level measurements usually can be captured with cameras [9], such as the position of a basketball in the air, and traffic lane lines. Besides, end-to-end control with pixel-level measurements is also a popular research direction, such as autonomous driving [10]. In these cases, the control problem can be significantly simplified if the associated prediction models can be achieved. A natural solution is to utilise deep neural network architectures with high-dimensional measurements as the inputs for modelling dynamics [11-14]. Nevertheless, this kind of modelling method has shortcomings. Firstly, the model prediction process is a non-linear evolution, so it has to go throughout the whole network architecture repeatedly for every prediction step. This results in large accumulative errors from networks and plenty of inference time for long-term predictions. Secondly, it could be difficult to design a model-based optimal control controller with the resulting non-linear model.

The Koopman operator has received considerable attention as a tool for the identification and analysis of non-linear

Robotics and Autonomous Systems 164 (2023) 104385

Contents lists available at ScienceDirect

Robotics and Autonomous Systems

journal homepage: www.elsevier.com/locate/robot

Efficient reinforcement learning with least-squares soft Bellman residual for robotic grasping☆

Yixing Lan[1], Junkai Ren*[,1], Tao Tang, Xin Xu*, Yifei Shi, Zixin Tang

College of Intelligence Science and Technology, National University of Defense Technology, Changsha, China

ARTICLE INFO

Article history:
Available online 23 February 2023

Keywords:
Reinforcement learning
Policy evaluation
Robotic grasping
Sparse kernel

ABSTRACT

Grasping control of intelligent robots has to deal with the difficulties of model uncertainties and nonlinearities. In this paper, we propose the Kernel-based Least-Squares Soft Bellman residual Actor–Critic (KLSAC) algorithm for robotic grasping. In the proposed approach, a novel linear temporal-difference learning algorithm using the least-squares soft Bellman residual (LS^2BR) method is designed for policy evaluation. In addition, KLSAC adopts a sparse-kernel feature representation method based on approximate linear dependency (ALD) analysis to construct features for continuous state–action space. Compared with typical deep reinforcement learning algorithms, KLSAC has two main advantages: firstly, the critic module has the capacity for rapid convergence by computing the fixed point of the linear soft Bellman equation via the least-squares optimization method. Secondly, the kernel-based features construction approach only requires predefining the basic kernel function and can improve the generalization ability of KLSAC. The simulation studies on robotic grasping control were conducted in the V-REP simulator. The results demonstrate that compared with other typical RL algorithms (e.g., SAC and BMPO), the proposed KLSAC algorithm can achieve better performance in terms of sample efficiency and asymptotic convergence property. Furthermore, experimental results on a real UR5 robot validated that KLSAC performed well in the real world.

1. Introduction

Robotic grasping control is an important research topic in the field of robotics, which has shown great potential in practical applications [1], such as express sorting [2], bin picking [3], human–robot interaction [4], cloth manipulation [5] and many sorts of dexterous manipulation [6–8]. The fast development of intelligent manufacturing has increasing demands on robots with precise manipulation and competent services. Although model-based control methods can solve robotic grasping tasks based on accurate models of robots and prior knowledge of motion patterns, it is still difficult to model complex and dynamic grasping tasks in practice. In addition, the performance of model-based methods will be unsatisfactory when the model is inaccurate or dynamically changing.

Reinforcement learning (RL), as an important class of machine learning approaches for solving sequential optimization decision-making issues, has played a critical role in improving the autonomous learning-control performance of robotic systems. RL agents learn an optimized control policy by interacting with the environment without accurate models, which significantly reduces the development cost and enables robots to be more adaptive to uncertain complex environments. Recently, RL has been applied to robotic grasping tasks to improve the grasping performance in various scenes [9–16]. For example, Kalashnikov et al. [10] proposed the QT-Opt algorithm to learn closed-loop vision-based manipulation skills with deep Q network. However, a large number of training samples are required to learn a near-optimal policy. Yuan et al. [13] applied DQN to object rearranging manipulation tasks with visual feedback and increased the operational success rate. Although these methods showed good grasping performance in simulation environments, additional efforts are required to enable the robots to adapt to the real world with different environmental background, system dynamics, and external noises. As discussed in [17], the training stability and generalization capability of current RL-based object grasping methods still need to be improved. As is well known, many existing deep reinforcement learning (DRL) algorithms require substantial interaction data from the environments to obtain good performance and still face challenges such as low sample efficiency, slow convergence speeds, and unstable training processes. Moreover, as collecting training data on physical robot systems is costly, time-consuming, and potentially unsafe, RL algorithms are usually trained in simulators and often may fail on

☆ This work was supported by the National Natural Science Foundation of China under Grants 61825305, 62002379.

* Corresponding authors.
E-mail addresses: jk.ren@nudt.edu.cn (J. Ren), xinxu@nudt.edu.cn (X. Xu).
[1] These authors have contributed equally to this work.

https://doi.org/10.1016/j.robot.2023.104385
0921-8890/© 2023 Published by Elsevier B.V.

Grasp Planning Based on Deep Reinforcement Learning: A Brief Survey

Zixin Tang, Xin Xu and Yifei Shi

College of Intelligence Science and Technology
National University of Defense Technology

2021 China Automation Congress (CAC) | 978-1-6654-2647-3/21/$31.00 ©2021 IEEE | DOI: 10.1109/CAC53003.2021.9727526

Abstract—**Grasping is a fundamental ability for robots to interact with the environment. Recent advances show great progress on grasping unstructured household objects by leveraging deep reinforcement learning. This paper presents a brief survey on grasp planning based on deep reinforcement learning. The goal of grasp planning is to determine the actions of the robotic manipulator moving from its initial position to the target one so that the object could be grasped and manipulated. However, objects could be stacked or occluded, making the target object ungraspable by applying a simple grasp action. To tackle this issue, auxiliary actions (such as pushing) should be conducted along with grasping. According to whether auxiliary actions are performed, we categorize the existing grasping planning methods into two classes: methods without auxiliary actions and methods with auxiliary actions. We review them in details and show the advantages and limitations of each type. Furthermore, we show several popular evaluation metrics that are widely used in grasp planning. Finally, difficulties and future research directions are discussed.**

Index Terms—**robot, grasp planning, deep reinforcement learning**

I. Introduction

Robotic grasping has a wide range of application prospects in industry and households. Grasping regular industrial parts has been relatively mature while grasping household objects puts higher demands on the flexibility and intelligence of robots due to various shapes and complex textures. Hence, grasping unstructured household objects is still an unsolved problem.

Grasp planning (GP) is a key component of robotic grasping, which forms the basis of some manipulation tasks [1], such as Pick-and-Place [2], Insertion [3], and Grasp the Invisible [4]. Formally, vision-based robotic grasping involves four key tasks, object localization, pose estimation, grasp detection, and motion planning, which can be accomplished independently or jointly [5]. Grasp planning (GP) is a combination of these four tasks. Namely, the robotic arm plans paths from the initial pose to the target grasp pose using scene information obtained from sensors. The path should satisfy the kinematic constraints of the robotic arm and avoid obstacles. The extra motion should also be prevented from the perspective of grasp efficiency. However, to grasp objects in cluttered scenes, some auxiliary actions, such as pushing, should be performed to create feasible grasping [6] [4] [7].

In recent years, with the powerful representation capabilities of neural networks [8], deep reinforcement learning (DRL) has shown promising sequential decision-making abilities in many tasks, such as video games [9] or robotic manipulation tasks [6] [2]. Since agent learns from interacting with the environment by trial-and-error, a large number of trials are required to train DRL algorithms, especially with sparse rewards [10] [11]. However, the training requirements of large amounts of data hinder the application of DRL algorithms to real-world robotic manipulation tasks.

As a result, many DRL based grasp planning methods only infer 3/4 DoF grasps, such as top-down grasps [12] [11]. Due to the limitation of DoF, some household objects like horizontal plates are hard to grasp. In order to alleviate the immense exploration search space faced by 6 DoF grasp planning, learning from demonstrations [13] and Sim2Real [7] for DRL based grasp planning are widely studied recently. Although there are some hurdles for applying DRL in robotic grasp planning, it is worth noting that DRL has been playing a more and more important role in it.

In this work, we present a brief survey on grasp planning based on deep reinforcement learning. According to whether auxiliary actions are performed, we divide them into two categories: GP without auxiliary actions [14] [3] [2] [15] and GP with auxiliary actions [4] [16] [17] [7].

For GP without auxiliary actions, according to whether the target grasp pose is detected before, grasp planning can be divided into open-loop methods [18] and closed-loop methods [11] [12] [7] [15]. For the open-loop method, the target grasp pose is known before planning, hence it only needs to plan the motion trajectory. Whereas, due to the tight coupling with the grasping detection algorithm [19] [20], there are two shortcomings. First, perceptual errors in the grasp detection algorithm will sensitively affect the results of planning. More importantly, since it is essentially a two-stage method, disturbances in the scene are hard to cope with, such as moving objects. On the contrary, the close-loop method directly plans grasp paths in an end-to-end manner without prior knowledge of grasp poses, which is more robust to perceptual noises, kinematic errors as well as robust in dynamic scenes [14]. Hence, we only exhaustively review the close-loop method for GP without auxiliary actions.

This work was supported by National Natural Science Foundation of China (61825305, 62002379) and the Zhejiang Lab's International Talent Fund for Young Professionals.