Grasping 3D Deformable Objects via Reinforcement Learning: A Benchmark and Evaluation

Melvin Laux
Robotics Research Group
University of Bremen
Bremen, Germany
melvin.laux@uni-bremen.de

Chandandeep Singh
Robotics Innovation Center
DFKI GmbH
Bremen, Germany
chandandeep.singh@dfki.de

Alexander Fabisch
Robotics Innovation Center
DFKI GmbH
Bremen, Germany
alexander.fabisch@dfki.de

***Abstract*—Robotic manipulation of deformable objects is a challenging task that has been tackled with a variety of approaches. However, due to the highly difficult task of modeling the dynamics of deformable objects in a fast and accurate way, many real-world use cases remain unsolved. Recent advances in data-driven approaches like reinforcement learning (RL) promise that these methods push forward the envelope of feasibility in the field of deformable object manipulation. Despite the growing interest in this field, data-driven approaches mainly focus on the manipulation of 1D and 2D deformable objects like ropes and cloth. In this work, we present the benchmark *DeformableGym* to facilitate the evaluation of RL methods for grasping 3D deformable objects. We use a set of simulated benchmark environments to evaluate existing model-free state-of-the-art algorithms and investigate the main challenges and potential pitfalls of applying them in this challenging setting.**

***Index Terms*—Reinforcement Learning, Robotic Grasping, Volumetric Deformable Objects**

I. Introduction

Manipulation of deformable objects is an important skill for robots in a large variety of tasks. Examples include human interaction, assistive robotics, medical use cases, or robotic surgery in non-industrial settings. Examples in industrial settings include fruit harvesting, food processing, or packaging of deformable objects. While recent advances in the field of robotic control have shown impressive results [24], these were mostly limited to entirely rigid environments. Despite a growing interest of the research community in the manipulation of deformable objects (DO), these efforts are largely focused on 1D (linear) or 2D (planar) DOs. When it comes to the manipulation of 3D (volumetric) DOs, the majority of works aim to solve the task of controlling the shape of an object, while the problem of grasping 3D DOs remains largely unexplored, especially when it comes to data-driven approaches such as reinforcement learning (RL). This is likely due to the lack of accurate models for 3D DOs.

We propose to use a simulation of 3D DOs and rigid robotic hands in combination with model-free RL to obtain a solution to manipulation problems in form of a policy. The advantage of this approach is that the computationally expensive DO model only has to be computed during training time and is not required during test time because the solution is encoded in a learned policy, which allows real-time application as it is easier to compute. The RL framework allows us to easily obtain closed-loop policies that take into account contact force measurements, so that generalization over unknown objects and other variations of the problem is possible.

This work was supported by the European Commission under the Horizon 2020 framework program for Research and Innovation (project acronym: APRIL, project number: 870142).

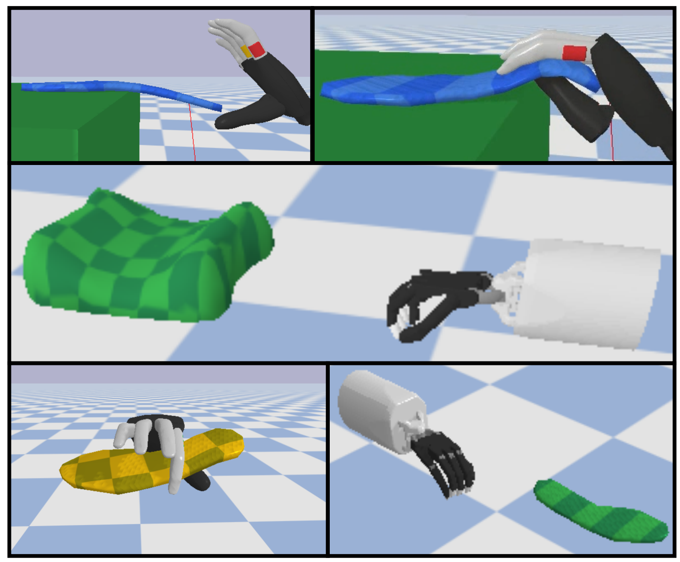

Fig. 1. Selection of available grasping environments in DeformableGym. Top: grasping process in *MiaInsoleOnConveyor* environment. Middle: *ShadowFloatingPillow* environment. Bottom: *MiaFloatingInsole* and *ShadowFloatingInsole* environments.

II. Related Work

Enabling robotic systems to manipulate deformable objects promises new applications in the industrial, service, and healthcare sectors. However, in comparison to rigid objects, deformable object manipulation (DOM) offers challenges in a multitude of domains such as gripper design, sensing, modeling, planning, and control [1], [5], [20], [39]. These challenges arise primarily due to the high dimensional state representation and complex dynamics of deformable objects [20], [22], [23].

Deformable objects can be categorized as uniparameteric (linear), biparametric (planar or cloth-like), and triparametric (volumetric) [28]. While significant work has been done towards manipulation of uniparameteric and biparametric objects [19], [21], [27], [34], [35], triparametric objects are the least researched, primarily due to their high computation costs for simulation. Recent advances in computing facilitate real-time simulation of realistic deformations [28].

There exist several approaches for 2D DOM based on imitation learning [27], model-based approaches [14], planning-based strategies to grasp elastic foam objects with a Shadow Dexterous hand using tactile feedback [4], supervised learning and planning for a two-armed robot using haptic data [9], or RL for cloth manipulation [36]. In the domain of 2D DOM, SoftGym [18] is a set of benchmarks for RL. There exist several approaches for 2D DOM based on imitation learning [27], model-based approaches [14], planning-based strategies to grasp elastic foam objects with a Shadow Dexterous hand using tactile feedback [4], supervised learning and planning for a two-armed robot using haptic data [9], or RL for cloth manipulation [36]. In the domain of 2D DOM, SoftGym [18] is a set of benchmarks for RL.

Methods for 3D DOM include the following. [16], [17] propose a deformation-aware, data-driven grasp synthesis method by adding information about object stiffness to state-of-the-art (SotA) grasp planner based on depth images. [7] present a FEM-based control approach for dexterous manipulation of 3D deformable objects. They use a multi-fingered hand for deformation control by applying closed-loop inverse kinematics (CLIK) in Cartesian space. The planner combines a contact interaction model with non-linear isotropic mass-spring system to guarantee stable grasps. However, this work is restricted to open-loop control. [12] learn a deformation model which is used in a visual servoing feedback controller to actively manipulate objects to match a given target shape. [38] present a grasp planner for a BarrettHand mounted on an industrial arm for grasping deformable objects. This approach, however, requires an accurate estimate of the object's location. [13] present DefGraspSim, a simulation that can be used to evaluate the quality of grasps of 3D DOs and a set of grasp features that can be used to evaluate grasps. [31] present a simulation framework for manipulating complex volumetric DOs in realistic scenes and use it for shape control through planning based on feedback from camera images. [37] provide a model-based strategy for grasping 3D DOs with multi-fingered hands that uses tactile sensors of the hand.

In previous work [6] we focused on generating initial open-loop reaching motions from human demonstrations. Object pose estimation errors and deformations might lead to unsuccessful grasps that we want to refine with closed-loop policies obtained through RL.

Our contribution is a closed-loop, model-free approach to grasp 3D DOs that relies on force measurements in the fingers of the hand to compensate for errors of the pose estimation of the object based on RL. We use multi-fingered, anthropomorphic hands. In addition, we provide a benchmark that can be used to evaluate RL algorithms for grasping of DOs, and test existing algorithms on it to establish baselines.

III. DeformableGym

In order to facilitate progress and research in the field of robotic grasping of 3D DOs with RL, we propose DeformableGym[1] (see Figure 1). DeformableGym uses PyBullet [3] with the stable Neo-Hookean model [32] to simulate 3D DOs, as its underlying physics engine and implements the OpenAI Gym interface [2], allowing easy set-up and testing with commonly used, standard libraries of SotA algorithms, e.g., stable-baselines3 [26].

A. Configuration Options

DeformableGym contains a variety of highly configurable learning environments that allow an easy evaluation of different control mechanisms, i.e., position control or velocity control, closed-loop servoing or open-loop viapoint control. Furthermore, the modular implementation allows to use different types of robotic setups and deformable objects in each grasping task. For instance, currently supported robotic hands are the Prensilia Mia Hand [25] and the Shadow Dexterous Hand [30]. Each gripper can be used in a *floating* scenario or a complete scenario, in which the gripper is mounted on a UR robotic arm [33].

DeformableGym contains the robots MiaHand (i.e., floating Mia hand), ShadowHand, URMia (i.e., Mia hand mounted on UR arm), URShadow; and the objects FloatingInsole, FloatingPillow, ConveyorInsole (i.e., insole on conveyor belt). Figure 1 shows a selection of the available environments.

B. Object Meshes

Volumetric object simulation requires tetrahedral meshes. We obtained surface meshes of real DOs by measuring their dimensions and modeling them in blender. Then we applied TetWild [11] to convert these surface meshes to tetrahedral volume meshes. PyBullet requires Lamé parameters, which we compute from estimates of Poisson's ratio and Young's modulus of the real DOs.

C. Observations, Actions and Rewards

We define the environment's action space $\mathcal{A}$ as the gripper's pose offset and finger velocity. The observation space $\mathcal{S}$ contains the current end-effector pose, the force-torque sensor information, and the current position of the object. In the case of the Mia hand, this leads to a 10-dimensional action space and a 16-dimensional observation space. In all grasping environments, we do not rely on complicated, hand-crafted reward functions, but instead use a sparse binary reward signal in combination with a simple grasp success condition. We consider a grasp to be successful if the grasp object does not fall below a certain height within a given time frame at the end of an episode. Successful grasps result in a reward of 1, unsuccessful grasps in a reward of -1. All intermediate steps receive an immediate reward of 0.

[1] Source code available at https://github.com/dfki-ric/deformable_gym

IV. EXPERIMENTS

We use DeformableGym to benchmark existing SotA model-free RL algorithms in order to identify specific challenges when trying to solve 3D DO grasping using RL. We intend to answer the following research questions:

- Can current SotA RL algorithms solve the problem of 3D DO grasping using a sparse reward formulation?
- Are learned policies able to generalize to task variations?
- Can adversarial exploration speed up the learning of robust policies?
- How should we model states to solve the grasping task sample-efficiently?

A. Experimental Setup

To answer these questions, we trained policies with SAC [10], TD3 [8], and PPO [29] using a variety of training regimens and observation formulations in the MiaInsoleOnConveyor environment. In this setting, the agent needs to grasp an insole that is placed on a conveyor belt (see Figure 1). To facilitate the running of experiments and to ensure comparable and high-quality implementations of the used algorithms, we use stable-baselines3 [26] for our evaluation. We test three different training regimens: fixed, randomized, and adversarial exploration, which we denote using the suffixes -F, -R, and -A, respectively. In the fixed setup, the initial hand position is identical throughout the entire training process. In the randomized setup, the initial position is sampled uniformly from a 6 cm x 6 cm area around the fixed initial position in the x-y plane. The adversarial exploration setup uses the fixed initial position in combination with an adversarial framework [15] to set the hand's initial position based on the grasping policy's current knowledge. We also evaluate the effect of using a history of observations, i.e., compare the performance when using only the current observation and when using a stack of the past four observations. The hyperparameters of the algorithms were tuned in a previous set of experiments.

B. Evaluation

We evaluate the policies' robustness by testing their ability to grasp the same object but from different initial positions at the beginning of each grasping attempt. These evaluations are done by systematically varying the initial pose in form of a grid of 25 initial end-effector positions in a 6 cm x 6 cm area in the x-y plane.

C. Results

Figure 2 shows exemplary learning curves of the PPO and SAC algorithms using the fixed and randomized training regimen in the *MiaInsoleOnConveyor* environment with a known object position and using a truncated history. It can be observed that PPO is able to learn a successful grasping behavior for a single fixed initial position. However, both PPO and SAC are unable to learn generalizing grasping policies.

Table I shows the evaluation of generalization per training procedure, algorithm, and state representation. It becomes evident that including a truncated history of observations in the state space leads to a better performance over observing the current state only. For all state representations the best training procedure is adversarial training with SAC. We evaluated the adversarial training regimen exemplary using SAC. However, this regimen may be applied to any value-function based RL algorithm, e.g., TD3.

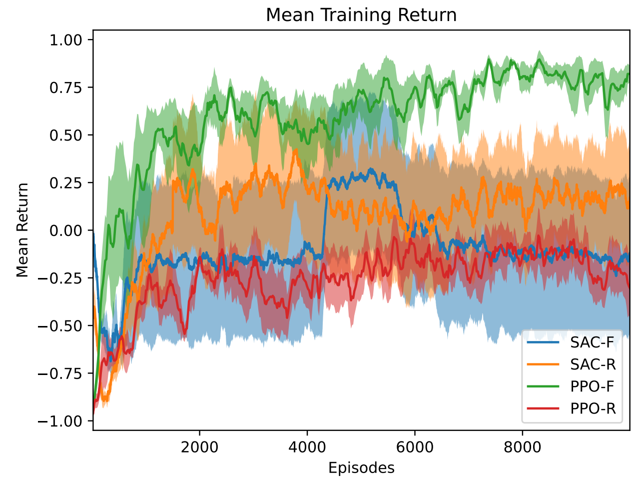

Fig. 2. Mean training return over five runs of SAC and PPO in *MiaInsoleOnConveyor*. Shaded areas represent the standard error of five runs. The suffixes -F and -R represent the fixed and randomized training regimen, respectively.

TABLE I
RESULTS. PERFORMANCE IS EVALUATED AFTER 10,000 EPISODES FOR EACH COMBINATION OF ALGORITHM, STATE REPRESENTATION, AND TRAINING REGIMEN.

	Mean Success Rate ± Standard Error			
Algorithm	K-N	K-H	U-N	U-H
PPO-F	.464 ± .077	.488 ± .086	.410 ± .091	.432 ± .045
PPO-R	.344 ± .086	.536 ± .064	.376 ± .131	.392 ± .077
TD3-F	.456 ± .126	.296 ± .098	.264 ± .084	.296 ± .152
TD3-R	-	-	-	.368 ± .118
SAC-F	.408 ± .170	.408 ± .141	.312 ± .135	.152 ± .095
SAC-R	.216 ± .130	.690 ± .160	.264 ± .154	.520 ± .121
SAC-A	**.640** ± .137	**.744** ± .037	**.544** ± .158	**.800** ± .041

Abbreviations: K-N – known position, no history; K-H – known position, with history; U-N – unknown position, no history; U-H – unknown position, with history.

V. CONCLUSION AND FUTURE WORK

We present a benchmark for manipulation of 3D DOs using robotic hands through RL. RL is suitable for these tasks as it does not need a model of object deformations to generate closed-loop grasping behavior after training in simulation. We use the benchmark to evaluate SotA deep RL algorithms under various training regimens and state representations. We found that using adversarial training performs well over a range of different problem definitions.

In the future, we intend to perform a more thorough evaluation of SotA RL algorithms on this novel benchmark including a detailed analysis of the impact of different state and action representations. Furthermore, we plan to examine the generalization capabilities of the tested algorithms not only with respect to initial positions, but also to initial orientation and grasp object parameters like friction, size, and stiffness.

REFERENCES

[1] Mario Bollini, Stefanie Tellex, Tyler Thompson, Nicholas Roy, and Daniela Rus. Interpreting and executing recipes with a cooking robot. In *Experimental Robotics: The 13th International Symposium on Experimental Robotics*, pages 481–495. Springer, 2013.

[2] Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. OpenAI Gym. Technical report, June 2016. arXiv:1606.01540 [cs] type: article.

[3] Erwin Coumans and Yunfei Bai. Pybullet, a python module for physics simulation for games, robotics and machine learning. http://pybullet.org, 2016–2021.

[4] A. Delgado, C. A. Jara, D. Mira, and F. Torres. A tactile-based grasping strategy for deformable objects' manipulation and deformability estimation. In *2015 12th International Conference on Informatics in Control, Automation and Robotics (ICINCO)*, volume 02, pages 369–374, July 2015.

[5] Zackory Erickson, Vamsee Gangaram, Ariel Kapusta, C Karen Liu, and Charles C Kemp. Assistive gym: A physics simulation framework for assistive robotics. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 10169–10176. IEEE, 2020.

[6] Alexander Fabisch, Manuela Uliano, Dennis Marschner, Melvin Laux, Johannes Brust, and Marco Controzzi. A modular approach to the embodiment of hand motions from human demonstrations. In *2022 IEEE-RAS 21st International Conference on Humanoid Robots (Humanoids)*, pages 801–808, 2022.

[7] F. Ficuciello, A. Migliozzi, E. Coevoet, A. Petit, and C. Duriez. FEM-Based Deformation Control for Dexterous Manipulation of 3D Soft Objects. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4007–4013, October 2018. ISSN: 2153-0866.

[8] Scott Fujimoto, Herke van Hoof, and David Meger. Addressing Function Approximation Error in Actor-Critic Methods. In Jennifer G. Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018*, volume 80 of *Proceedings of Machine Learning Research*, pages 1582–1591. PMLR, 2018.

[9] Mevlana C. Gemici and Ashutosh Saxena. Learning haptic representation for manipulating deformable food objects. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 638–645, September 2014. ISSN: 2153-0866.

[10] Tuomas Haarnoja, Aurick Zhou, Pieter Abbeel, and Sergey Levine. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In Jennifer Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 1861–1870. PMLR, 10–15 Jul 2018.

[11] Yixin Hu, Qingnan Zhou, Xifeng Gao, Alec Jacobson, Denis Zorin, and Daniele Panozzo. Tetrahedral meshing in the wild. *ACM Trans. Graph.*, 37(4):60:1–60:14, July 2018.

[12] Zhe Hu, Tao Han, Peigen Sun, Jia Pan, and Dinesh Manocha. 3-D Deformable Object Manipulation Using Deep Neural Networks. *IEEE Robotics and Automation Letters*, 4(4):4255–4261, October 2019.

[13] Isabella Huang, Yashraj Narang, Clemens Eppner, Balakumar Sundaralingam, Miles Macklin, Tucker Hermans, and Dieter Fox. DefGraspSim: Simulation-based grasping of 3D deformable objects. Technical report, July 2021. arXiv:2107.05778 [cs] type: article.

[14] P. Jiménez. Survey on model-based manipulation planning of deformable objects. *Robotics and Computer-Integrated Manufacturing*, 28(2):154–163, April 2012.

[15] Melvin Laux, Oleg Arenz, Jan Peters, and Joni Pajarinen. Deep adversarial reinforcement learning for object disentangling. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2020, Las Vegas, NV, USA, October 24, 2020 - January 24, 2021*, pages 5504–5510. IEEE, 2020.

[16] Tran Nguyen Le, Jens Lundell, Fares J. Abu-Dakka, and Ville Kyrki. Towards synthesizing grasps for 3D deformable objects with physics-based simulation. Technical report, July 2021. arXiv:2107.08898 [cs] type: article.

[17] Tran Nguyen Le, Jens Lundell, Fares J. Abu-Dakka, and Ville Kyrki. Deformation-Aware Data-Driven Grasp Synthesis. *IEEE Robotics and Automation Letters*, 7(2):3038–3045, April 2022.

[18] Xingyu Lin, Yufei Wang, Jake Olkin, and David Held. SoftGym: Benchmarking Deep Reinforcement Learning for Deformable Object Manipulation. In Jens Kober, Fabio Ramos, and Claire J. Tomlin, editors, *4th Conference on Robot Learning, CoRL 2020, 16-18 November 2020, Virtual Event / Cambridge, MA, USA*, volume 155 of *Proceedings of Machine Learning Research*, pages 432–448. PMLR, 2020.

[19] Xiao Ma, David Hsu, and Wee Sun Lee. Learning latent graph dynamics for deformable object manipulation. *arXiv preprint arXiv:2104.12149*, 2, 2021.

[20] Jeremy Maitin-Shepard, Marco Cusumano-Towner, Jinna Lei, and Pieter Abbeel. Cloth grasp point detection based on multiple-view geometric cues with application to robotic towel folding. In *2010 IEEE International Conference on Robotics and Automation*, pages 2308–2315. IEEE, 2010.

[21] Jan Matas, Stephen James, and Andrew J. Davison. Sim-to-real reinforcement learning for deformable object manipulation. *CoRR*, abs/1806.07851, 2018.

[22] Stephen Miller, Mario Fritz, Trevor Darrell, and Pieter Abbeel. Parametrized shape models for clothing. In *2011 IEEE International Conference on Robotics and Automation*, pages 4861–4868. IEEE, 2011.

[23] Stephen Miller, Jur Van Den Berg, Mario Fritz, Trevor Darrell, Ken Goldberg, and Pieter Abbeel. A geometric approach to robotic laundry folding. *The International Journal of Robotics Research*, 31(2):249–267, 2012.

[24] OpenAI, Marcin Andrychowicz, Bowen Baker, Maciek Chociej, Rafał Józefowicz, Bob McGrew, Jakub Pachocki, Arthur Petron, Matthias Plappert, Glenn Powell, Alex Ray, Jonas Schneider, Szymon Sidor, Josh Tobin, Peter Welinder, Lilian Weng, and Wojciech Zaremba. Learning dexterous in-hand manipulation. *CoRR*, 2018.

[25] Prensilia. Prensilia mia hand. https://www.prensilia.com/portfolio/mia/, 2023.

[26] Antonin Raffin, Ashley Hill, Adam Gleave, Anssi Kanervisto, Maximilian Ernestus, and Noah Dormann. Stable-baselines3: reliable reinforcement learning implementations. page 12348–12355, 1 2021.

[27] Gautam Salhotra, I-Chun Arthur Liu, Marcus Dominguez-Kuhne, and Gaurav S Sukhatme. Learning deformable object manipulation from expert demonstrations. *IEEE Robotics and Automation Letters*, 7(4):8775–8782, 2022.

[28] Jose Sanchez, Juan-Antonio Corrales, Belhassen-Chedli Bouzgarrou, and Youcef Mezouar. Robotic manipulation and sensing of deformable objects in domestic and industrial applications: a survey. *The International Journal of Robotics Research*, 37(7):688–716, June 2018.

[29] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal Policy Optimization Algorithms. *CoRR*, abs/1707.06347, 2017. arXiv: 1707.06347.

[30] ShadowRobot. Shadow dexterous hand. https://www.shadowrobot.com/dexterous-hand-series/, 2005.

[31] Bokui Shen, Zhenyu Jiang, Christopher Choy, Silvio Savarese, Leonidas J. Guibas, Anima Anandkumar, and Yuke Zhu. ACID: Action-Conditional Implicit Visual Dynamics for Deformable Object Manipulation. volume 18, June 2022.

[32] Breannan Smith, Fernando De Goes, and Theodore Kim. Stable neo-hookean flesh simulation. 37(2), mar 2018.

[33] Universal Robots A/S. Collaborative robots from universal robots. https://www.universal-robots.com/products/, 2022. Accessed: 2023-03-23.

[34] Angelina Wang, Thanard Kurutach, Kara Liu, Pieter Abbeel, and Aviv Tamar. Learning robotic manipulation through visual planning and acting. *arXiv preprint arXiv:1905.04411*, 2019.

[35] Yilin Wu, Wilson Yan, Thanard Kurutach, Lerrel Pinto, and Pieter Abbeel. Learning to manipulate deformable objects without demonstrations. *arXiv preprint arXiv:1910.13439*, 2019.

[36] Yilin Wu, Wilson Yan, Thanard Kurutach, Lerrel Pinto, and Pieter Abbeel. Learning to Manipulate Deformable Objects without Demonstrations. In Marc Toussaint, Antonio Bicchi, and Tucker Hermans, editors, *Robotics: Science and Systems XVI, Virtual Event / Corvalis, Oregon, USA, July 12-16, 2020*, 2020.

[37] Lazher Zaidi, Juan Antonio Corrales, Belhassen Chedli Bouzgarrou, Youcef Mezouar, and Laurent Sabourin. Model-based strategy for grasping 3D deformable objects using a multi-fingered robotic hand. *Robotics and Autonomous Systems*, 95:196–206, September 2017.

[38] Lazher Zaidi, Juan Antonio Corrales Ramon, Laurent Sabourin, Belhassen Chedli Bouzgarrou, and Youcef Mezouar. Grasp Planning Pipeline for Robust Manipulation of 3D Deformable Objects with Industrial Robotic Hand + Arm Systems. *Applied Sciences*, 10(23):8736, January 2020.

[39] Jihong Zhu, Andrea Cherubini, Claire Dune, David Navarro-Alarcon, Farshid Alambeigi, Dmitry Berenson, Fanny Ficuciello, Kensuke Harada, Jens Kober, Xiang Li, Jia Pan, Wenzhen Yuan, and Michael Gienger. Challenges and Outlook in Robotic Manipulation of Deformable Objects. *IEEE Robotics & Automation Magazine*, 29(3):67–77, September 2022.