Tactile Sensing and Deep Reinforcement Learning for In-Hand Manipulation Tasks

Andrew Melnik[1], Luca Lach[1], Matthias Plappert[2], Timo Korthals[1], Robert Haschke[1], and Helge Ritter[1]

***Abstract*— Deep Reinforcement Learning techniques demonstrate advances in the domain of robotics. One of the limiting factors is the large number of interaction samples usually required for training in simulated and real-world environments. In this work, we demonstrate that tactile information substantially increases sample efficiency for training (by 97% on average), and simultaneously increases the performance in dexterous in-hand manipulation of objects tasks (by 21% on average). To examine the role of tactile-sensor parameters in these improvements, we conducted experiments with varied sensor-measurement accuracy (Boolean vs. float values), and varied spatial resolution of the tactile sensors (92 sensors vs. 16 sensors on the hand). We conclude that ground-truth touch-sensor readings as well as dense tactile resolution do not further improve performance and sample efficiency in the tasks. We make available these touch-sensors extensions as a part of OpenAI-Gym robotics Shadow-Dexterous-Hand environments.**

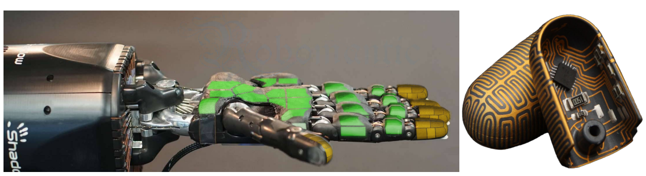

Fig. 1. Shadow Dexterous Hand equipped with fabrics-based tactile sensors in the palm and finger phalanges (indicated green) and fingertip sensors realized by Molded-Interconnect-Devices (indicated yellow) [1, 2]

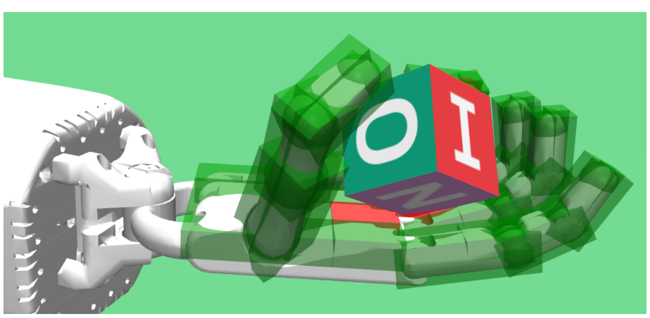

Fig. 2. 92 touch sensors covering the Shadow Dexterous Hand model. This is a technical visualization to represent the essence of our model. Red sites represent activated touch sensors, where a block is pressing against the touch sensitive area. Green sites represent inactive touch sensors. A video demonstration of the extended environments can be found at https://rebrand.ly/TouchSensors.

This work was supported by the Cluster of Excellence Cognitive Interaction Technology (CITEC) which is funded by DFG (EXC 277)

[1]CITEC, Bielefeld University, 33619 Bielefeld, Germany. andrew.melnik.papers@gmail.com

[2]OpenAI, San Francisco, USA

I. INTRODUCTION

Deep Reinforcement Learning techniques demonstrate advances in the domain of robotics. For example, dexterous in-hand manipulation of objects with an anthropomorphic robotic hand [3, 4]. An agent with a model-free policy was able to learn complex in-hand manipulation tasks using just proprioceptive feedback and visual information about the manipulated object. Despite the good performance of the agent, absence of tactile sensing imposes certain limitations on these approaches. In cases, like insufficient lighting conditions, or partial visibility of the manipulated object due to occlusion (e.g. by the manipulator itself), the tactile sensing may be necessary to perform an in-hand manipulation task [5]. Continuous haptic feedback can improve grasping acquisition in terms of robustness under uncertainty [6]. Different to these works with static objects, we present empirical results in simulation that show that including tactile information in the state improves the sample efficiency and performance for dynamic in-hand manipulation of objects.

Humans perform better in dexterous in-hand manipulation tasks than robotic systems. One of the reasons is the rich tactile perception available to humans which allows to recognize and manipulate an object even without vision [7], instead just rely on tactile information. In such cases, tactile perception is one of the key abilities for in-hand manipulation of objects and tool usage. The importance of tactile sensing for object recognition was demonstrated on a multifingered robotic hand [8] as well as for successful grasping on a two-fingered gripper with high-resolution tactile sensors [9].

Multisensory fusion [10] techniques may help at the level of geometry, contact and force physics. For example, an autoencoder can translate high-dimensional sensor representations in a compact space, and a reinforcement learner can learn stable, non-linear policies [11]. Learning curves for tactile representations learned with different types of autoencoders can be found here [11].

Recent works describe approaches to bring the tactile sensing to anthropomorphic hands like the Shadow Dexterous Hand, by providing integrated tactile fingertips [1] as shown in Fig. 1 and constructing a flexible tactile skin [2]. The tactile skin comprises stretchable and flexible, fabric-based tactile sensors capable of capturing typical human interaction forces within the palm and proximal and distal phalanges of the hand. This enables the hand to exploit tactile information, e.g. for contact or slip detection [12, 13]. The distribution of tactile sensors in these works resembles our segmentation of the simulated Shadow Dexterous Hand into 92 tactile-sensitive areas. [14] proposed a deep tactile model-predictive control framework for non-prehensile manipulation to perform tactile servoing and to reposition an object to user-specified configurations, indicated by a goal tactile reading, using the learned tactile predictive model. [15] studied how tactile feedback can be exploited to adapt to unknown rollable objects located on a table and demonstrated the possibility of learning feedback controllers for in-hand manipulation using reinforcement learning on an underactuated, compliant platform. The feedback controller was hand-tuned to successfully complete the tasks.

Another challenge is the sample efficiency in learning of model free DRL methods. Assuming the importance of tactile information for in-hand manipulation of objects, it is necessary to measure how sensory information influences learning and overall performance for in-hand manipulation tasks. We selected well established and benchmarked OpenAI Gym simulated environments for robotics which include in-hand manipulation of objects with anthropomorphic-hand tasks. We extended these environments with touch sensors and compared the steepness of the learning curves with and without tactile information. We demonstrate that the sample efficiency and performance can be increased when tactile information is available to the agent. To fully examine the roll of tactile sensing, we are providing analysis of sensor measurement accuracy - a highly important aspect in using tactile sensing on physical robots, and examining the role of the spatial resolution of the tactile sensors on the hand. These experiments provide beneficial knowledge for those looking to build robots with tactile sensors for manipulation.

A logical next step is to employ such hands to study the impact of tactile information on manual skill learning. With the current sample efficiency of model free reinforcement learning this will require a combined strategy that connects simulation and real-world learning stages. As a step towards such an approach and to generate insights about how the availability of touch sensing impacts on manual skill learning with model free DRL approaches, we here present results of a simulation study, particularly with respect to the factors sample efficiency and performance level. To this end, we (i) extended the OpenAI Gym robotics environments with touch sensors designed to closely mimic the touch sensing of the robot hand [1][2], (ii) compared learning results for OpenAI Gym robotics environments with and without touch sensing. We find for all four learning tasks a significantly increase in sampling efficiency along with improved performance of the trained agents.

II. METHODS

OpenAI Gym [16] contains several simulated robotics environments with the Shadow Dexterous Hand. These environments use the MuJoCo [17] physics engine for fast and accurate simulation. These are open source environments designed for experimental work and research with Deep Reinforcement Learning. The anthropomorphic Shadow Dexterous Hand model, comprising 24 degrees of freedom (20 actuated and 4 coupled), has to manipulate an object (block, egg, or pen) so that it matches a given goal orientation, position, or both position and orientation.

TABLE I
NEW OPENAI GYM ROBOTICS ENVIRONMENTS WITH TOUCH SENSORS: -V0 (BOOLEAN), -V1 (FLOAT-VALUE)
HTTPS://GITHUB.COM/OPENAI/GYM

HandManipulateBlockRotateZTouchSensors
HandManipulateBlockRotateParallelTouchSensors
HandManipulateBlockRotateXYZTouchSensors
HandManipulateBlockTouchSensors
HandManipulateEggRotateTouchSensors
HandManipulateEggTouchSensors
HandManipulatePenRotateTouchSensors
HandManipulatePenTouchSensors

TABLE II
THE 92 AND 16 TOUCH-SENSOR ENVIRONMENTS.

lower phalanx of the fingers (4x)	7 sensors x 4	1 sensor x 4
middle phalanxes of the fingers (4x)	5 sensors x 4	1 sensor x 4
tip phalanxes of the fingers (4x)	5 sensors x 4	1 sensor x 4
thumb phalanxes (3x)	5 sensors x 3	1 sensor x 3
palm (1x)	9 sensors x 1	1 sensor x 1
All touch sensors	92 sensors	16 sensors

As a step towards touch-augmented RL we extended the Shadow Dexterous Hand model with touch sensors available as new environments (Table I) in the OpenAI Gym package [16]. We covered all five fingers and the palm of the Shadow Dexterous Hand model with 92 touch sensors (Fig. 2; Table II). For the agent, the only difference between the robotics environments with and without touch sensors is the length of the state vector that the agent gets as an input at each time step. For the original OpenAI Gym simulated environments for robotics (Table I) without touch information, the state vector is 68-dimensional (Table III) [4]. In the environments with 92 touch sensors the state vector is 160-dimensional (68+92). As an additional experiment, we grouped 92 sensors into 16 sub-groups (Table II) to reduce the tactile sensory resolution ("16 Sensors-v0"). If any of sensors in a group has a greater than zero value, then the sub-group returns True, otherwise False. The grouping was done per phalanx (3 phalanx x 5 digits) plus a palm resulting in 16 sub-groups (Table II). In the environments with 16 touch sensors sub-groups the state vector is 84-dimensional

TABLE III
NEURAL NETWORK INPUT VECTOR IN THE HAND ENVIRONMENTS.

Type	Length
Joint angles	24
Joint angles‘ velocity	24
Object‘s position XYZ	3
Object‘s velocity XYZ	3
Object‘s orientation (quaternion)	4
Object‘s angular velocities	3
Target position of the object XYZ	3
Target orientation of the object XYZ (quaternion)	4
All values without touch sensors	68
Touch sensors	92
All values with touch sensors	160

(68+16). For a given state of the environment, a trained policy outputs an action vector of 20 float numbers used for position-control (actuation_center + action * actuation_range) of the 20 actuated degrees of freedom.

The MuJoCo [17] physics engine provides methods to mimic touch sensing at specified locations. This is based on specifying the tactile sensors' active zones by so-called sites. Each site can be represented as either ellipsoid or box. In Fig. 2, the sites are visualized as red and green transparent shapes attached to the hand model. If a body's contact point falls within a site's volume, and involves a geometry attached to the same body as the site, the corresponding contact force is included in the sensor reading. The output of this sensor is non-negative scalar. At runtime, the contact normal forces of all included contact points of a single zone are added and returned as the simulated output of the corresponding sensor [18][19].

The learning agent in the environments is an Actor and Critic network: 3 layers with 256 units each and ReLU non-linearities. We evaluate the learning using Deep Deterministic Policy Gradients (DDPG) [20] and Hindsight Experience Replay (HER) [21] techniques. All hyperparameters and the training procedure are described in detail in [21, 4], and are available in a code implementation as a part of the OpenAI Baselines (HER) repository (https://github.com/openai/baselines)[22].

For all environments, we train on a single machine with 19 CPU cores. Each core generates experience using two parallel rollouts and uses MPI for synchronization. We train for 200 epochs, which amounts to a total of 38x106 timesteps. We evaluate the performance after each epoch by performing 10 deterministic test rollouts per MPI worker and then compute the test success rate by averaging across rollouts and MPI workers. A test is counted as successful when, at the end of the test rollout, the manipulated object occurs in a rewarded state determined by the reward function. The training environments - the original and the newly contributed touch-augmented ones as listed in Table I - are available as part of OpenAI-Gym (https://github.com/openai/gym) [22]. In all cases, we repeat an experiment with five different random seeds (4, 20, 1299, 272032, 8112502) and report results by computing the median test success rate as well as the interquartile range.

Following [4], in all tasks we use rewards that are sparse and binary: The agent obtains a reward of 0 if the goal has been achieved (within some task-specific tolerance) and 1 otherwise. Actions are 20-dimensional: we use absolute position control for all non-coupled joints of the hand. We apply the same action in 20 subsequent simulator steps (with t = 0.002 s each) before returning control to the agent, i.e. the agent's action frequency is f = 25 Hz. Observations include the 24 positions and velocities of the robot's joints. In case of an object that is being manipulated, we also include its Cartesian position and rotation represented by a quaternion (hence 7-dimensional) as well as its linear and angular velocities. In case of a "...TouchSensors-v0" environment with touch-sensors, we also include a vector (92 Boolean values) representing tactile information. In case of a "...TouchSensors-v1" environment with touch-sensors, we include a vector (92 float values) representing tactile information.

Manipulate Block: A block is placed on the palm of the hand. The task is to manipulate the block such that a target pose is achieved. The goal is 7-dimensional and includes the target position (in Cartesian coordinates) and target rotation (in quaternions). We include multiple variants:

HandManipulateBlockFullTouchSensors: Random target rotation for all axes of the block. Random target position.

HandManipulateBlockRotateXYZTouchSensors: Random target rotation for all axes of the block. No target position. A goal is considered achieved if the distance between the manipulated objects position and its desired position is less than 1 cm (applicable only in the Full variant) and the difference in rotation is less than 0.1 rad.

Manipulate Egg: An egg-shaped object is placed on the palm of the hand. The goal is 7-dimensional and includes the target position (in Cartesian coordinates) and target rotation (in quaternions).

HandManipulateEggFullTouchSensors: Random target rotation for all axes of the egg. Random target position. A goal is considered achieved if the distance between the manipulated objects position and its desired position is less than 1 cm and the difference in rotation is less than 0.1 rad.

Manipulate Pen: A pen-shaped object is placed on the palm of the hand. The goal is 7-dimensional and includes the target position (in Cartesian coordinates) and target rotation (in quaternions).

HandManipulatePenRotateTouchSensors: Random target rotation x and y axes of the pen and no target rotation around the z axis. No target position. A goal is considered achieved if the difference in rotation, ignoring the z axis, is less than 0.1 rad.

For the sake of brevity, further details about training procedure, reward function, goal-aware observation space, and neural network parameters are available in [4], since our main contribution focuses on the extension of the existing Shadow Dexterous Hand model by tactile sensors.

III. EXPERIMENTAL RESULTS

We tested how tactile information helps to increase performance and samples efficiency of training on the object manipulation tasks. To this end, we have reproduced the experiments where no touch sensors were used [4], and conducted the same experiments with additional touch-sensor readings. To provide insights about how different aspects of tactile information (accuracy, tactile resolution) influence learning and performance we conducted three experiments. In the first experiment we added float-value readings from 92 sensors to the state (red curves in Fig. 3). This experiment can be reproduced in the OpenAI-gym-robotics environments ending at "...TouchSensors-v1". In the second experiment we added Boolean-value reading from the same 92 sensors to the state (black curves in Fig. 3). The experiment can be reproduced in the OpenAI-gym-robotics environments

TABLE IV

AVERAGE PERFORMANCE IN THE RANGE OR 150-200 EPOCHS (FIG. 3) WITH AND WITHOUT TACTILE INFORMATION (LEFT FOUR COLUMNS). PERFORMANCE-INCREASE RATIO OF LEARNING WITH TACTILE INFORMATION IN COMPARISON TO LEARNING WITHOUT TACTILE INFORMATION (RIGHT THREE COLUMNS).

	No Sensors (ave. perform.)	**92 Sensors-v1** (ave. perform.)	**92 Sensors-v0** (ave. perform.)	**16 Sensors-v0** (ave. perform.)	**92 Sensors-v1** (perform. incr.)	**92 Sensors-v0** (perform. incr.)	**16 Sensors-v0** (perform. incr.)
HandManipulate Egg	0.74	0.83	0.85	0.85	1.12	1.15	1.15
HandManipulate BlockRoatateXYZ	0.92	0.94	0.94	0.93	1.02	1.02	1.01
HandManipulate PenRotate	0.28	0.32	0.31	0.30	1.14	1.11	1.07
HandManipulate Block	0.19	0.30	0.30	0.29	1.58	1.58	1.53
Mean					**1.22**	**1.22**	**1.19**

TABLE V

LEFT FOUR COLUMNS: FIRST EPOCH WHEN A CONVERGENCE CURVE (FIG. 3) REACHES THE AVERAGE PERFORMANCE (EPOCHS 150-200) WITHOUT TACTILE INFORMATION. RIGHT THREE COLUMNS: SAMPLE EFFICIENCY OF LEARNING WITH TACTILE INFORMATION IN COMPARISON TO LEARNING WITHOUT TACTILE INFORMATION.

	No Sensors (first epoch)	**92 Sensors-v1** (first epoch)	**92 Sensors-v0** (first epoch)	**16 Sensors-v0** (first epoch)	**92 Sensors-v1** (sample eff.)	**92 Sensors-v0** (sample eff.)	**16 Sensors-v0** (sample eff.)
HandManipulate Egg	136	77	59	58	1.77	2.31	2.35
HandManipulate BlockRoatateXYZ	85	56	60	75	1.52	1.42	1.13
HandManipulate PenRotate	108	34	40	56	3.18	2.70	1.93
HandManipulate Block	113	62	60	69	1.82	1.88	1.64
Mean					**2.07**	**2.08**	**1.76**

ending at "...TouchSensors-v0". In the third experiment we grouped 92 sensors into 16 sub-groups (Table II) to reduce the tactile sensory resolution. If any of sensors in a group has a Boolen-True value, then the sub-group returns True, otherwise False. The grouping was done per phalanx (3 phalanx x 5 digits) plus a palm resulting in 16 sub-groups (Table II). Thus in the third experiment we added Boolean-value readings from the 16 sub-groups to the state (green curves in Fig. 3).

Fig. 3 and results in Table IV demonstrate that tactile information increases performance of the agent in the tasks. We define 100% performance level as the average over the last 50 epochs (epochs 150-200) of median test success rate in the original "NoSensors" environments (blue curves in Fig. 3). To compare performance-increase ratio of learning with tactile information in comparison to learning without tactile information we divide the average performance over the last 50 epochs (epochs 150-200) of the three experiments ("92 Sensors-v1", "92 Sensors-v0", "16 Sensors-v0") by the average performance in the original experiment ("NoSensors"). In each experiment, we observe on average more than 1.19 times better performance (Table IV) when tactile information is available.

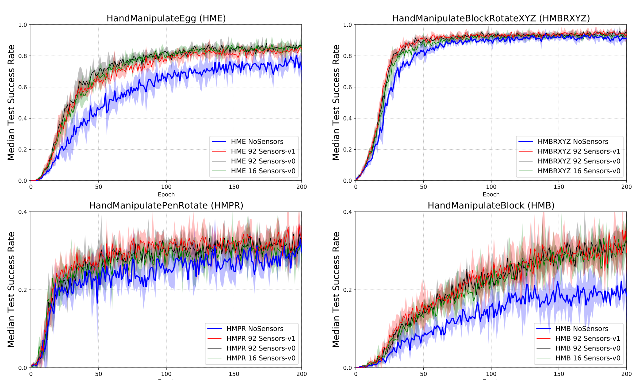

Fig. 3. Curves - median test success rate. Shaded areas - interquartile range [five random seeds]. Blue curves - learning without tactile information (NoSensors), red curves - learning with float-value tactile readings (Sensors-v1) from 92 sensors, black curves - learning with Boolean-value tactile readings (Sensors-v0) from 92 sensors, green curves - learning with Boolean-value tactile readings (Sensors-v0) from 16 sensor sub-groups.

Results in Table V demonstrate sample-efficiency increase while training when tactile information is available. To compare the sample efficiency for learning with and without tactile information, we measured how many training epochs were necessary to reach the performance level of an agent trained without tactile information in an environment. For example, in the HandManipulateBlock environment, the "NoSensors" performance level equals 0.19 (Table IV) and is reached after 113 epochs (Table V) when training without tactile information. When training with float-value tactile readings (92 Sensors-v1), the agent reaches the "NoSensors"-performance level of 0.19 first time already after 62 epochs (Table V), which results in the 1.82 times increase (Table V) in sample efficiency for training. When training with Boolean-values tactile readings (92 Sensors-v0), the agent reaches the "NoSensors"-performance level of 0.19 first time already after 60 epochs, which results in the 1.88 times increase in sample efficiency for training. When training with Boolean-values tactile readings with lower resolution (16 Sensors-v0)(Table II), the agent reaches the "NoSensors"-performance level of 0.19 first time already after 69 epochs, which results in the 1.64 times increase in sample efficiency for training. In each experiment, we observe on average more than 1.76 times faster convergence (Table V) when tactile information is available (Table V).

IV. DISCUSSION

In this work, we study the inclusion of tactile information into model-free deep reinforcement learning. We consider tasks with dynamic in-hand manipulation of known objects by the Shadow-Dexterous-Hand model. We observe that feedback from tactile sensors improves performance and sample efficiency for training. Next to vision, tactile sensing is a crucial source of information for manipulation of objects and tool usage for humans and robots [23, 24, 25, 26]. Visual reconstruction of contact areas suffers from occlusions by the manipulator at the contact areas [27]. Better inferring the geometry, mass, contact and force physics of manipulated objects is possible when touch information is available. Tactile sensing thus can provide essential information for manual interaction and grasping, making it one of key components for better generalization. Tactile information also allows training of gentle manipulation of fragile objects. It is a challenging objective to tap the potential of touch for dexterous robot hands. This requires to combine hardware and software developments for robot hand touch sensing.

In this work, we concatenated tactile, proprioceptive, and visual information at the input level of a neural network. A possible further extension of this work is a multi-modal sensor fusion. The multi-modal sensor fusion [10] allows end-to-end training of Bayesian information fusion on raw data for all subsets of a sensor setup. It can potentially deliver better performance and more sample efficient training with model-free deep reinforcement learning approaches.

V. CONCLUSIONS

In this work, we introduce the touch-sensors extensions to OpenAI-Gym [16] robotics Shadow-Dexterous-Hand environments [4] modeled after our touch sensor developments [1, 2]. We find that adding tactile information substantially increases sample efficiency for training (by 97% on average, Table V) and performance (by 21% on average, Table IV) in the environments, when training with deep reinforcement learning techniques [21]. To examine the role of tactile-sensor parameters in these improvements, we conducted experiments (Fig. 3) with varied sensor-measurement accuracy (Boolean vs. float values), and varied spatial resolution of the tactile sensors (92 sensors vs. 16 sensors on the hand). We conclude that accurate sensory readings as well as dense tactile resolution do not substantially improve performance and sample efficiency when training with deep reinforcement learning techniques, in comparison to Boolean sensor readings and sparse sensor localization (one sensor per phalanx). The performance and sample efficiency for training are similar in these case. These experiments provide beneficial knowledge for those looking to build robots with tactile sensors for manipulation.

REFERENCES

[1] Risto Koiva et al. "A highly sensitive 3D-shaped tactile sensor". In: *2013 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*. IEEE. 2013, pp. 1084–1089.

[2] Gereon H Büscher et al. "Flexible and stretchable fabric-based tactile sensor". In: *Robotics and Autonomous Systems* 63 (2015), pp. 244–252.

[3] Marcin Andrychowicz et al. "Learning dexterous in-hand manipulation". In: *arXiv preprint arXiv:1808.00177* (2018).

[4] Matthias Plappert et al. "Multi-goal reinforcement learning: Challenging robotics environments and request for research". In: *arXiv preprint arXiv:1802.09464* (2018).

[5] Timo Korthals et al. "Multisensory Assisted In-hand Manipulation of Objects with a Dexterous Hand". In: *juxi.net* (2019).

[6] Hamza Merzić et al. "Leveraging Contact Forces for Learning to Grasp". In: *2019 International Conference on Robotics and Automation (ICRA)*. IEEE. 2019, pp. 3615–3621.

[7] Susan J Lederman and Roberta L Klatzky. "Hand movements: A window into haptic object recognition". In: *Cognitive psychology* 19.3 (1987), pp. 342–368.

[8] Alexander Schmitz et al. "Tactile object recognition using deep learning and dropout". In: *2014 IEEE-RAS International Conference on Humanoid Robots*. IEEE. 2014, pp. 1044–1050.

[9] Roberto Calandra et al. "More than a feeling: Learning to grasp and regrasp using vision and touch". In: *IEEE Robotics and Automation Letters* 3.4 (2018), pp. 3300–3307.

[10] Timo Korthals et al. "Jointly Trained Variational Autoencoder for Multi-Modal Sensor Fusion". In: *22st International Conference on Information Fusion,(FUSION) 2019, Ottawa, CA, July 2-5, 2019*. 2019.

[11] Herke Van Hoof et al. "Stable reinforcement learning with autoencoders for tactile and visual data". In: *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2016, pp. 3928–3934.

[12] Martin Meier et al. "Distinguishing sliding from slipping during object pushing". In: *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2016, pp. 5579–5584.

[13] Guillaume Walck et al. "Robot self-protection by virtual actuator fatigue: Application to tendon-driven dexterous hands during grasping". In: *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2017, pp. 2200–2205.

[14] Stephen Tian et al. "Manipulation by feel: Touch-based control with deep predictive models". In: *arXiv preprint arXiv:1903.04128* (2019).

[15] Herke Van Hoof et al. "Learning robot in-hand manipulation with tactile features". In: *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*. IEEE. 2015, pp. 121–127.

[16] Greg Brockman et al. *OpenAI Gym*. 2016. eprint: `arXiv:1606.01540`.

[17] Emanuel Todorov, Tom Erez, and Yuval Tassa. "Mujoco: A physics engine for model-based control". In: *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 2012, pp. 5026–5033.

[18] Emanuel Todorov. *Sensor / touch*. `http://www.mujoco.org/book/XMLreference.html#sensor-touch`. 2019.

[19] Emanuel Todorov. *MPL sensors*. `http://www.mujoco.org/book/haptix.html#mpSensors.`. 2019.

[20] Timothy P Lillicrap et al. "Continuous control with deep reinforcement learning". In: *arXiv preprint arXiv:1509.02971* (2015).

[21] Marcin Andrychowicz et al. "Hindsight experience replay". In: *Advances in Neural Information Processing Systems*. 2017, pp. 5048–5058.

[22] Prafulla Dhariwal et al. "Openai baselines". In: *GitHub, GitHub repository* (2017).

[23] Jonathan Maycock et al. "Approaching manual intelligence". In: *KI-Künstliche Intelligenz* 24.4 (2010), pp. 287–294.

[24] Roland S Johansson and Goran Westling. "Roles of glabrous skin receptors and sensorimotor memory in automatic control of precision grip when lifting rougher or more slippery objects". In: *Experimental brain research* 56.3 (1984), pp. 550–564.

[25] Hao Dang, Jonathan Weisz, and Peter K Allen. "Blind grasping: Stable robotic grasping using tactile feedback and hand kinematics." In: *ICRA*. 2011, pp. 5917–5922.

[26] Markus Fritzsche, Norbert Elkmann, and Erik Schulenburg. "Tactile sensing: A key technology for safe physical human robot interaction". In: *2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE. 2011, pp. 139–140.

[27] Timo Korthals et al. "Multisensory Assisted In-hand Manipulation of Objects with a Dexterous Hand". In: *2019 IEEE International Conference on Robotics and Automation Workshop on Integrating Vision and Touch for Multimodal and Cross-modal Perception,(ViTac) 2019, Montreal, CA, May 20-25, 2019*. 2019.