

# Learning Tactile Insertion in the Real World

Daniel Palenicek,<sup>\*†</sup> Theo Gruner,<sup>\*†</sup> Tim Schneider,<sup>\*†</sup> Alina Böhm,<sup>\*†</sup> Janis Lenz,<sup>\*†</sup> Inga Pfenning,<sup>\*†</sup>  
Eric Krämer<sup>†</sup> and Jan Peters<sup>†</sup>

**Abstract**—Humans have exceptional tactile sensing capabilities, which they can leverage to solve challenging, partially observable tasks that cannot be solved from visual observation alone. Research in tactile sensing attempts to unlock this new input modality for robots. Lately, these sensors have become cheaper and, thus, widely available. At the same time, the question of how to integrate them into control loops is still an active area of research, with central challenges being partial observability and the contact-rich nature of manipulation tasks. In this study, we propose to use Reinforcement Learning to learn an end-to-end policy, mapping directly from tactile sensor readings to actions. Specifically, we use Dreamer-v3 on a challenging, partially observable robotic insertion task with a Franka Research 3, both in simulation and on a real system. For the real setup, we built a robotic platform capable of resetting itself fully autonomously, allowing for extensive training runs without human supervision. Our preliminary results indicate that Dreamer is capable of utilizing tactile inputs to solve robotic manipulation tasks in simulation and reality. Furthermore, we find that providing the robot with tactile feedback generally improves task performance, though, in our setup, we do not yet include other sensing modalities. In the future, we plan to utilize our platform to evaluate a wide range of other Reinforcement Learning algorithms on tactile tasks.

## I. INTRODUCTION

Humans heavily rely on their tactile sense to solve dexterous manipulation tasks [1, 2]. Haptic feedback allows humans to achieve high levels of accuracy in tasks ranging from classical assembly tasks to critical medical surgeries. At the same time, the dexterous manipulation capabilities of even the most sophisticated robots lag behind those of even a small child if the environment is not strictly controlled. One way of closing this gap is to equip robots with tactile sensors. Tactile sensors provide robots with crucial feedback at the points of contact, which their end-effectors often occlude from vision.

Recently, the development of vision-based tactile sensors [3–5] has sparked interest in applying machine learning to various tactile-based perception tasks, such as object classification [6], texture recognition [7, 8], and shape reconstruction [9]. In contrast to conventional sensors that estimate physical quantities like contact forces or torques, vision-based tactile sensors provide an image of the deformation of the contact surface. This shift in paradigm necessitates pre-processing the sensor data before utilization in a control loop. However, advances in computer vision in the last decades

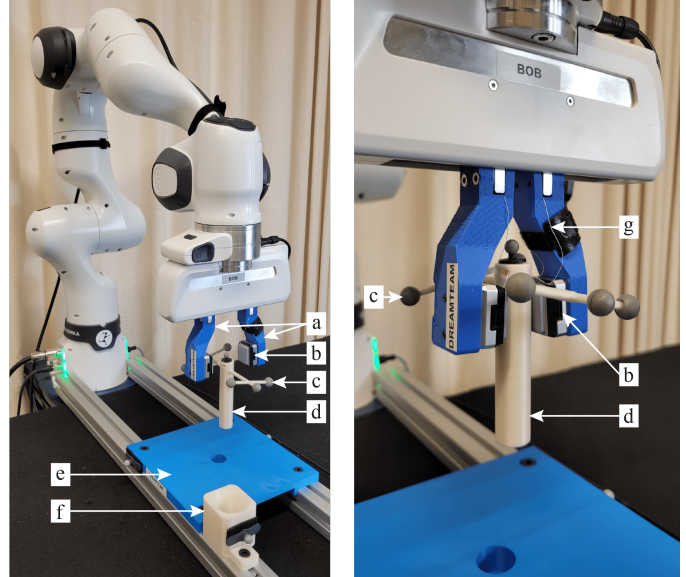


Fig. 1: **Tactile Insertion Platform Setup.** We label the gripper (a), GelSight Mini sensor (b), OptiTrack marker (c), cylinder (d), base plate (e), reset box (f), and thin thread (g). Note, that only one of the two GelSight Mini sensors is used in this work.

have made machine learning methods effective in extracting information from visual representations [10, 11].

A central challenge of many manipulation tasks is partial observability, as object tracking is often inaccurate, especially due to occlusion caused by the robot’s end-effector. While tactile sensors can provide crucial information about the points of contact, the information obtained from them will usually still be incomplete, rendering the decision-making problem a *Partially Observable Markov Decision Process (POMDP)*. Yet, reinforcement learning (RL) algorithms commonly rely on the assumption that the state-space  $\mathcal{S}$  is Markovian and fully observable [12–16]. Recently, a number of RL algorithms have been proposed that tackle POMDPs by learning a latent state-space model of the environment [17–26].

In this paper, we utilize *Dreamer-v3* [27] to solve a challenging, partially observable robotic insertion task purely from tactile feedback, both in simulation and the real world. Dreamer is a model-based RL approach that learns a latent state-space model and an actor-critic model simultaneously, which has shown great success in challenging Atari games [24] and even on a physical quadruped task [26]. We show that Dreamer is capable of utilizing tactile feedback from a *Gel-sight Mini* [3] sensor effectively while simultaneously being sample efficient enough to train in the real world from scratch.

<sup>\*</sup>Equal contribution. <sup>†</sup>Technical University of Darmstadt, Germany.

This project has received funding from the Aristotle BMBF project, BMWSB ZukunftBau (no. 10.08.18.7-21.34.) and was supported by the Research Clusters “The Adaptive Mind” and “Third Wave of AI”, funded by the Excellence Program of the Hessian Ministry of Higher Education, Science, Research and the Arts, hessian.AI.

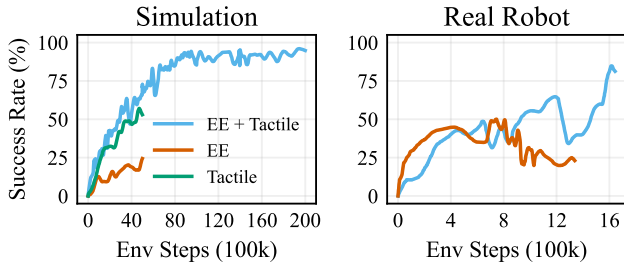


Fig. 2: **Evaluation Results.** Insertion success rates during training in simulation (left) and on the real system (right).

Our contributions are (i) the development of a tactile insertion simulation setup based on *Taxim* [28], allowing for rapid development and testing, (ii) the design of a real robot platform with fully autonomous robust resetting of the task, allowing for extensive training runs without human supervision, and (iii) a preliminary evaluation the influence of touch on learning insertion policies in simulation and on the physical system.

## II. ROBOTIC PLATFORM

The robotic setup, consisting of a 7 DoF Franka Research 3, a circular peg, and a base plate with an insertion hole, is presented in Fig. 1. We mount a vision-based GelSight Mini tactile sensor [3] sensor in the fingertips of the robot’s gripper. A key challenge is the development of a reset procedure that places the agent back in its initial state, with the peg properly placed in its gripper. To prevent the loss of the cylinder if the peg slips from the gripper, we suspend the peg with a thin thread connected to the gripper. Upon losing contact, the peg hangs from the thread. The robot then moves to a reset box, lowers the peg inside, re-grasps, and returns to its initial position above the insertion hole. This resetting procedure enables the agent to train consecutively for multiple days without requiring human intervention.

The agent’s observation comprises a  $64 \times 64 \times 3$  RGB image from the Gelsight sensor and the end-effector position  $\mathbf{p}_e \in \mathbb{R}^3$  in Cartesian coordinates at 25 Hz. The policy returns the relative new positions of the end-effector at 20 Hz. These target positions are then fed into the *franky* [29] control library, which internally uses *ruckig* [30] for smooth motion planning and Franka’s internal cartesian impedance controller for execution. We ensure safe exploration of the environment by restricting the workspace  $\mathcal{W}$  to be in the vicinity of the hole. For continuous tracking of the cylinder pose, we utilize OptiTrack [31]. This information is only used for evaluation of the robot’s performance and not as an input to the policy.

To simulate variation of the hole position, we add a random, per episode constant, offset of up to 5 mm to the observed end-effector position. This way, even though the hole position never changes, for the agent it seems as if it did since its base-frame is slightly different in every episode. Furthermore, we randomly vary the starting position of the robot’s end-effector.

As shown in Table I, the reward comprises of four components: (i) proximity to the goal, (ii) a terminal reward upon reaching the goal  $\mathcal{G} = \{\mathbf{x} \in \mathbb{R}^3 : |\mathbf{p}_g - \mathbf{x}| < (5, 5, 10)^T [\text{mm}]\}$ , (iii) a terminal penalty for leaving the workspace, and (iv)

an action penalty to encourage smooth motions. The policy training includes several neural network models: encoder and decoder networks that project between the observation space and the 4160-dimensional latent space, a recurrent state-space model, a reward model in the latent space, and the policy.

In addition, we develop a digital twin of our setup in PyBullet [32], tactile sensors are simulated with *Taxim* [28].

## III. EMPIRICAL EVALUATIONS

We train Dreamer from scratch, both in simulation and reality. In either environment, we conduct experiments with and without tactile sensors to evaluate whether tactile sensing yields a benefit for our task.

### a) Simulation results:

Fig. 2 shows that Dreamer achieves a success rate up to 90% given end-effector and tactile information. We investigate the importance of proprioception and tactile sensing by removing the end-effector and tactile observations in two experiments. The results clearly show that using tactile images greatly improves learning performance, while observing the end-effector position yields no significant benefit.

b) **Real Robot results:** On the real setup, our study comprises an experiment using only the end-effector position and one including the tactile feedback. Although the difference in performance between these two experiments is less pronounced than in simulation, especially in the beginning, tactile sensing still seems to provide a substantial advantage in performance. We hypothesize that the inherent softness of the gripper in the real system might simplify the task compared to the simulated setup, where we use rigid body physics. However, the results are still preliminary, and more experiments will have to be conducted in the future.

## IV. CONCLUSION & FUTURE WORK

In this work, we developed a robot platform that facilitates autonomous training of an RL agent on a Franka Research 3 robot to investigate whether reinforcement learning is a viable option for incorporating tactile feedback in control loops. To ensure reliable and safe training, we implement a robust resetting routine and restrict the action space, enabling autonomous learning of an insertion policy without human supervision. Leveraging tactile feedback from a vision-based tactile sensor attached to the gripper, our policy autonomously learns to insert a peg into a hole using the model-based RL approach Dreamer. Our simulation results suggest that the inclusion of tactile information significantly enhances learning outcomes. With our platform, we intend to benchmark other RL algorithms on tactile insertion tasks to get a clear understanding of which methods are capable of dealing with the inherent complexity of tactile manipulation tasks in the real world. Furthermore, we plan to increase task complexity in future experiments, enabling a thorough exploration of touch’s role in dexterous manipulation.

TABLE I: **Reward function,** with multiple components.

$$\begin{aligned}
 r &= r_d + r_g + r_p + r_a \\
 r_d &= -5 \cdot (|\mathbf{p}_g - \mathbf{p}_e|) \\
 r_g &= 100 \cdot \mathbb{1}_{\{\mathcal{G}\}}(\mathbf{p}_g) \\
 r_p &= -100 \cdot \mathbb{1}_{\{\mathbb{R}^3/\mathcal{W}\}}(\mathbf{p}_e) \\
 r_a &= 10^{-3} \cdot |\mathbf{a}|
 \end{aligned}$$

## REFERENCES

- [1] Roberta L Klatzky and Susan J Lederman. “Identifying objects from a haptic glance”. In: *Perception & psychophysics* (1995).
- [2] Roland S Johansson and J Randall Flanagan. “Coding and use of tactile signals from the fingertips in object manipulation tasks”. In: *Nature reviews neuroscience* (2009).
- [3] Wenzhen Yuan, Siyuan Dong, and Edward H Adelson. “Gelsight: High-resolution robot tactile sensors for estimating geometry and force”. In: *Sensors* (2017).
- [4] Benjamin Ward-Cherrier, Nicholas Pestell, Luke Cramphorn, et al. “The tactip family: Soft optical tactile sensors with 3d-printed biomimetic morphologies”. In: *Soft robotics* (2018).
- [5] Mike Lambeta, Po-Wei Chou, Stephen Tian, et al. “Digit: A novel design for a low-cost compact high-resolution tactile sensor with application to in-hand manipulation”. In: *Robotics and automation letters* (2020).
- [6] Tadeo Corradi, Peter Hall, and Pejman Irvani. “Bayesian tactile object recognition: Learning and recognising objects using a new inexpensive tactile sensor”. In: *International conference on robotics and automation*. 2015.
- [7] Alina Böhm, Tim Schneider, Boris Belousov, et al. “What Matters for Active Texture Recognition With Vision-Based Tactile Sensors”. In: *International conference on robotics and automation*. 2024.
- [8] Wenzhen Yuan, Yuchen Mo, Shaoxiong Wang, et al. “Active Clothing Material Perception Using Tactile Sensing and Deep Learning”. In: *International conference on robotics and automation*. 2018.
- [9] Sudharshan Suresh, Zilin Si, Joshua G Mangelson, et al. “Shapemap 3-d: Efficient shape mapping through dense touch and vision”. In: *International conference on robotics and automation*. 2022.
- [10] Yann LeCun, Bernhard Boser, John Denker, et al. “Handwritten digit recognition with a back-propagation network”. In: *Advances in neural information processing systems*. 1989.
- [11] Kaiming He, Xiangyu Zhang, Shaoqing Ren, et al. “Deep Residual Learning for Image Recognition”. In: *arXiv:1512.03385*. 2015.
- [12] John Schulman, Filip Wolski, Prafulla Dhariwal, et al. “Proximal policy optimization algorithms”. In: *arXiv:1707.06347*. 2017.
- [13] Tuomas Haarnoja, Aurick Zhou, Pieter Abbeel, et al. “Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor”. In: *International conference on machine learning*. 2018.
- [14] Scott Fujimoto, Herke Hoof, and David Meger. “Addressing function approximation error in actor-critic methods”. In: *International conference on machine learning*. 2018.
- [15] Xinyue Chen, Che Wang, Zijian Zhou, et al. “Randomized ensembled double Q-learning: Learning fast without a model”. In: *International conference on learning representations*. 2021.
- [16] Aditya Bhatt, Daniel Palenicek, Boris Belousov, et al. “CrossQ: Batch Normalization in Deep Reinforcement Learning for Greater Sample Efficiency and Simplicity”. In: *International conference on learning representations*. 2023.
- [17] Danijar Hafner, Timothy Lillicrap, Ian Fischer, et al. “Learning Latent Dynamics for Planning from Pixels”. In: *International conference on machine learning*. 2019.
- [18] Alex X Lee, Anusha Nagabandi, Pieter Abbeel, et al. “Stochastic latent actor-critic: Deep reinforcement learning with a latent variable model”. In: *Advances in neural information processing systems*. 2020.
- [19] Aaron J. Havens, Yi Ouyang, Prabhat Nagarajan, et al. “Learning Latent State Spaces for Planning through Reward Prediction”. In: *arXiv:1912.04201*. 2019.
- [20] Oleh Rybkin, Chuning Zhu, Anusha Nagabandi, et al. “Model-Based Reinforcement Learning via Latent-Space Collocation”. In: *arXiv:2106.13229*. 2021.
- [21] Tung D. Nguyen, Rui Shu, Tuan Pham, et al. “Temporal Predictive Coding For Model-Based Planning In Latent Space”. In: *arXiv:2106.07156*. 2021.
- [22] Carles Gelada, Saurabh Kumar, Jacob Buckman, et al. “DeepMDP: Learning Continuous Latent Space Models for Representation Learning”. In: *arXiv:1906.02736*. 2019.
- [23] Danijar Hafner, T. Lillicrap, Jimmy Ba, et al. “Dream to Control: Learning Behaviors by Latent Imagination”. In: *arXiv:1912.01603*. 2020.
- [24] Danijar Hafner, T. Lillicrap, Mohammad Norouzi, et al. “Mastering Atari with Discrete World Models”. In: *arXiv:2010.02193*. 2020.
- [25] F. Deng, Ingook Jang, and Sungjin Ahn. “Dreamer-Pro: Reconstruction-Free Model-Based Reinforcement Learning with Prototypical Representations”. In: *International conference on machine learning*. 2021.
- [26] Philipp Wu, Alejandro Escontrela, Danijar Hafner, et al. “DayDreamer: World Models for Physical Robot Learning”. In: *Conference on robot learning*. 2022.
- [27] Danijar Hafner, Jurgis Pasukonis, Jimmy Ba, et al. “Mastering diverse domains through world models”. In: *arXiv:2301.04104*. 2023.
- [28] Tim Schneider. *Taxim-GPU*. URL: <https://git.ias.informatik.tu-darmstadt.de/tactile-sensing/taxim-torch>.
- [29] *franky*. 2024. URL: <https://github.com/TimSchneider42/franky>.
- [30] Lars Berscheid and Torsten Kröger. “Jerk-limited real-time trajectory generation with arbitrary target states”. In: *arXiv:2105.04830* (2021).
- [31] *Motion Capture Systems*. 2024. URL: [optitrack.com](https://optitrack.com).
- [32] Erwin Coumans and Yunfei Bai. *Pybullet, a python module for physics simulation for games, robotics and machine learning*. 2016.