# Stick Roller: A Case Study on Efficiently Learning Precise Tactile Dynamics for In-hand Manipulation

**Yipai Du**
Hong Kong University of Science and Technology
**Michael Yu Wang**
Monash University
**Wenzhao Lian**
University of Chinese Academy of Sciences
**Yu She** ( shey@purdue.edu )
Purdue University West Lafayette

**Research Article**

**Keywords:** Vision-based Tactile sensing, Tactile Control, In-hand Manipulation

**Posted Date:** August 21st, 2023

**DOI:** https://doi.org/10.21203/rs.3.rs-3252500/v1



**Additional Declarations:** No competing interests reported.

# Stick Roller: A Case Study on Efficiently Learning Precise Tactile Dynamics for In-hand Manipulation

Yipai Du[1], Michael Yu Wang[2], Wenzhao Lian[3], Yu She[4*]

[1]Hong Kong University of Science and Technology, Hong Kong, China.
[2]Monash University, Melbourne, Australia.
[3]University of Chinese Academy of Sciences, China.
[4*]Purdue University, West Lafayette, IN, USA.

*Corresponding author. E-mail: shey@purdue.edu;
Contributing authors: yipai.du@connect.ust.hk; michael.y.wang@monash.edu;
lianwenzhao@gmail.com;

**Abstract**

In-hand manipulation is challenging in robotics due to the intricate contact dynamics and high degrees of control freedom. Precise manipulation with high accuracy often requires tactile perception, which adds further complexity to the system. This paper addresses these challenges in perception and control and proposes a framework for learning high-resolution tactile dynamics on real hardware that is feasible and scalable. Specifically, we use a case study on manipulating a small stick using the Allegro hand equipped with the Digit vision-based tactile sensor to demonstrate the framework's effectiveness. The framework includes an action space reduction module, tactile perception module, and learning with uncertainty module, all designed to operate in low data regimes. With minimal manually collected data, our learned contact dynamics model achieves grasp stability, sub-millimeter precision, and zero-shot generalizability across novel objects. The proposed framework demonstrates promising results for enabling precise in-hand manipulation with tactile feedback on real hardware.

**Keywords:** Vision-based Tactile sensing, Tactile Control, In-hand Manipulation

## 1 Introduction

In-hand manipulation is a crucial skill for humans to manipulate objects, allowing people to utilize the power of various tools. However, robots have not yet achieved that level of dexterity. Robotic in-hand manipulation will play an essential role in extending the deployment of robots for accomplishing more complex industrial tasks. The parallel-jaw grippers can only reply on extrinsic contact (Dafle et al., 2014) to regrasp an object to change its relative pose in hand, but a robotic hand can achieve it with more degrees of freedom of the fingers (Fig. 1). The motion primitive of rolling a stick-like object between two fingers can be used to tighten or loosen the screws or change the pose of the grasped object after the initial grasp. It can greatly improve the flexibility of robotic systems. For example, the manipulator has to move back and forth to adjust the pose of the in-hand cable (She et al., 2021), but it is not suitable for confined space industrial manipulation. In that case, the in-hand adjustment becomes the optimal motion to adopt. Although

this task seems to be simple, it is non-trivial to accomplish. Robotic hands' high degrees of freedom increases their flexibility but also induces challenges in designing control policies. The precision requirement poses difficulty to sim-to-real approaches (Handa et al., 2023) and makes high-resolution tactile sensing indispensable. Among various types of tactile sensors, vision-based tactile sensors can provide superior sensing resolution at a low design and fabrication cost (Shimonomura, 2019). Although many prior works have been trying to reduce the learning complexity of in-hand manipulation, either with tactile feedback (Nagabandi, Konolige, Levine, & Kumar, 2020; Yin, Huang, Qin, Chen, & Wang, 2023) or without (Qi, Kumar, Calandra, Ma, & Malik, 2023), they cannot capture the fine contact dynamics, such as the contact position and contact strength, which are crucial for tasks requiring precise contact control within the fingertip. To explore the potential of learning tactile in-hand manipulation on real hardware, we try to safely and precisely achieve the rolling stick motion with a small amount of real data. The primary contributions of this paper include:

- A simple method to accurately estimate the grasping strength and pose of stick-like objects in hand from tactile feedback.
- A framework to learn from a small amount of data while achieving high precision in learning tactile dynamics
- A real-hardware case study to show the proposed method can enable rolling object pose control with submillimeter accuracy.

The proposed learning framework includes action space reduction, model-based tactile signal extraction, and uncertainty-aware dynamics learning. In addition to manipulating the rolling object, the authors hope the framework will also inspire researchers to solve industrial in-hand manipulation problems on real hardware in a more data-efficient manner.

This paper is structured as follows. Section 2 reviews related work on vision-based tactile signal processing and in-hand manipulation. In Section 3, we introduce the three essential modules of the framework, each specifically designed to facilitate efficient learning. In Section 4, quantitative experiments are conducted to validate the accuracy of the learned contact dynamics model, using both test data and hardware tests. We also extend the experiments to include novel object testing and pose control testing, demonstrating the usefulness of the learned model for manipulating unseen objects. Section 5 and 6 give the limitation and conclusion of the work.

## 2 Related Works

Robotic in-hand manipulation has been considered important due to its flexibility for various daily tasks. Early studies rely on task-specific hardware to achieve fast and accurate object manipulation (Ishihara, Namiki, Ishikawa, & Shimojo, 2006). In recent years stronger and faster robotic hands have been developed (Kumar, Xu, & Todorov, 2013), allowing the exploration of more generic human-like tasks. To effectively utilize the high degrees of freedom in the place of complex contact dynamics, reinforcement learning approaches become more suitable (Van Hoof, Hermans, Neumann, & Peters, 2015). For example, Andrychowicz et al. (2020) showed for the first time using end-to-end deep-RL based method for controlling the contact-rich in-hand manipulation, but the training and sim-to-real transfer are difficult to reproduce. Therefore, the core question in multi-finger in-hand manipulation has been improving the sample efficiency of learning in high dimensionality (Kober, Bagnell, & Peters, 2013) to make the methods more scalable and manageable. Chen Chen et al. (2013) studied the human hand kinematics and dynamics constraints, which inspires novel robotic hand hardware and algorithm design to reduce the action complexity. Solak and Jamone (2019) attempted to use kinesthetic human demonstrations and learned dynamical movement primitives to translate and rotate the objects in the hand while maintaining a stable grasp. Sub-optimal controllers using domain knowledge (Khandate, Mehlman, Wei, & Ciocarlie, 2023) and demonstration through human hand teleoperation (Arunachalam, Silwal, Evans, & Pinto, 2023) were also proven to be effective in improving the sample efficiency in reinforcement learning for in-hand manipulation. Nagabandi et al. (2020) combined deep dynamics model and online planning for closed-loop control, which is promising for efficiently learning complex dexterous manipulation skills with a limited amount of data. Shaw, Bahl, and Pathak (2023) extracted

**Fig. 1** The hardware setup for this work. Two fingers of the Allegro robotic hand are used to manipulate a piece of stick. The Digit vision-based tactile sensor on the thumb finger reads the contact state and enables manipulation without visual feedback

basic manipulation primitives from internet videos as visual, action, and physical priors. Many techniques (Q. Chen et al., 2023; T. Chen et al., 2023; Qi et al., 2023) on sim-to-real transfer were also emerging that allow the learned models in simulation to be deployed on real hardware. A problem with these systems above is that they are designed for palm-scale in-hand manipulation rather than fingertip-scale and cannot capture the fine contact dynamics within the fingertip, hence unsuitable for precise manipulation.

To control the subtle contact within the fingertips, high-resolution tactile sensing is desired as feedback. Vision-based tactile sensors using digital cameras to capture the sensing surface deformation are promising as they have high sensing resolution, easy manufacturing methods, and robustness in harsh settings. It also synergizes tactile signal processing with visual signal processing tools (Dong et al., 2021). Various kinds of vision-based tactile sensors (Du, Zhang, & Wang, 2022; Sferrazza & D'Andrea, 2019; Taylor, Dong, & Rodriguez, 2022; Ward-Cherrier et al., 2018; Yuan, Dong, & Adelson, 2017) are being developed that satisfy different needs on the tactile signal perception. Among them, the Digit tactile sensor (Lambeta et al., 2020) achieved enhanced reliability and compact design that can fit into the Allegro hand fingertip. Lambeta et al. (2020) adopted an autoencoder structure to extract the contact center and strength for Digit tactile sensor in a self-supervised manner. They achieved the manipulation of a marble ball within the two fingertips but consumed more than 10 hours of real data. We hope to solve the rolling stick problem, which is more difficult than marble manipulation, in a data-efficient manner to make learning precise contact dynamics more scalable on real hardware.

# 3 Method Description

The proposed method for learning precise tactile dynamics relies on three key components: the action space proposal, the tactile perception, and learning the dynamics model with uncertainty, each specifically designed to enhance learning efficiency. The whole pipeline of the learned model is shown in Fig. 2.

## 3.1 Action Space Reduction

This work aims to focus the effort on the part of the action space that is more conducive to maintaining a stable grasp. To achieve this, we reduced the 8 DOF (degrees of freedom) two-finger system to 7 DOF by fixing the base joint of the index finger. This is because simultaneously moving the base joints of the index finger and the thumb is unnecessary for this task. Then a uniformly sampled discrete action space was formed

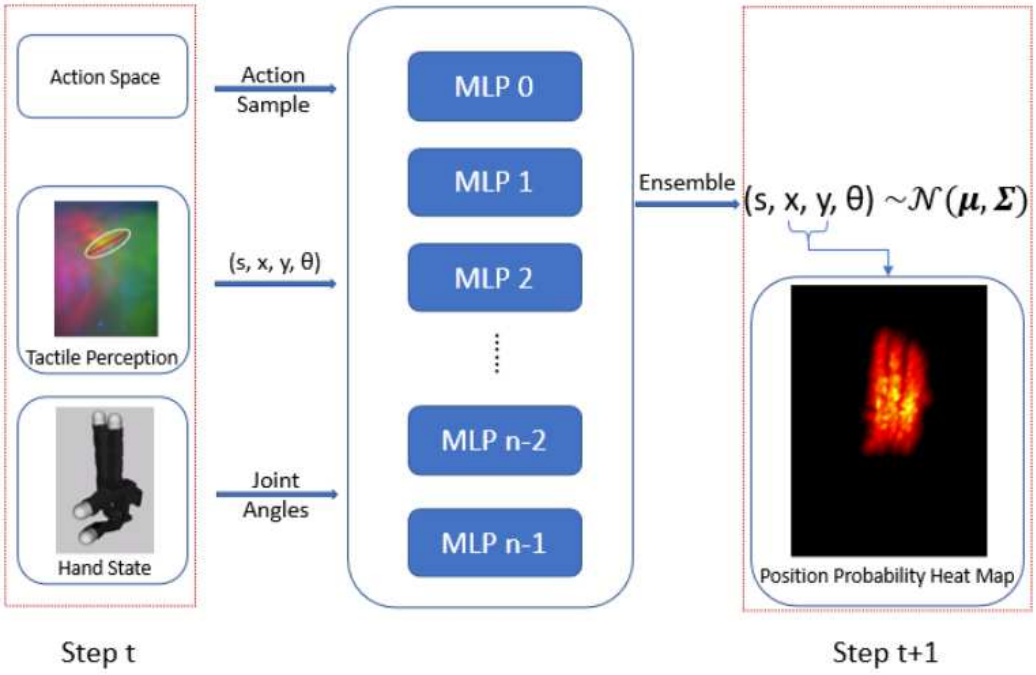


**Fig. 2** The pipeline of the learned tactile dynamics model. The predictions from multiple multilayer perceptrons are ensembled to produce the future tactile state distribution (parameterized by a multivariate normal distribution) given the current tactile state, hand state, and action. By selecting the maximum value of the contact position marginal distribution across all actions, a heat map representation can be generated to visualize the outcomes of the actions

by taking a step size of 0.1 radius in the 7 DOF joint space. It is because it is easier to perform kinematics filtering on discrete actions than on continuous actions since for discrete actions we can do offline filtering and only keep the remaining ones, but for continuous actions we have to filter online as well, which may introduce extra overhead. Then the actions were filtered by calculating the forward kinematics of the fingers. Although for in-hand manipulation, the forward kinematics of robotic hands are usually not as precise as robotics arms, mainly due to less powerful hardware, it is still a helpful tool to remove most of the irrelevant robot finger configurations. The filtering was based on a box constraint to make the two fingertips close to each other and an orientation constraint to make the two fingertips roughly facing each other. The parameters for these two constraints were naively adjusted by hand. Still, the method effectively reduced the action space as most configurations did not satisfy these two simple constraints (see Section 4.1). This idea can be extended to use learned classifiers (filters) from demonstrations (Arunachalam et al., 2023) for space reduction in more complex tasks and is left as future work.

## 3.2 Tactile Perception

To achieve efficient tactile perception, we considered a strategy that exploits the contact geometry model. The pipeline (shown in Fig. 3) was inspired by the method for cable manipulation by She et al. (2021) but used the color difference of the tactile image instead of the contact depth image for contact detection. The observation is that thresholding on the absolute value of the difference image is often as effective as thresholding on the contact depth for PCA (principal component analysis). Requiring no contact depth reconstruction also implies that the tactile sensor needs not to be calibrated, which is not the case in general for GelSight tactile sensors (Yuan et al., 2017). This relaxation saves a lot of effort since the calibration is harder for the Digit sensor than GelSight due to non-uniform illumination in different colors across sensors. The tactile perception produces four numbers $(s, x, y, \theta)$. $s$ is the mean value of the absolute difference, averaged across the locations and the color channels, and serves as a contact strength

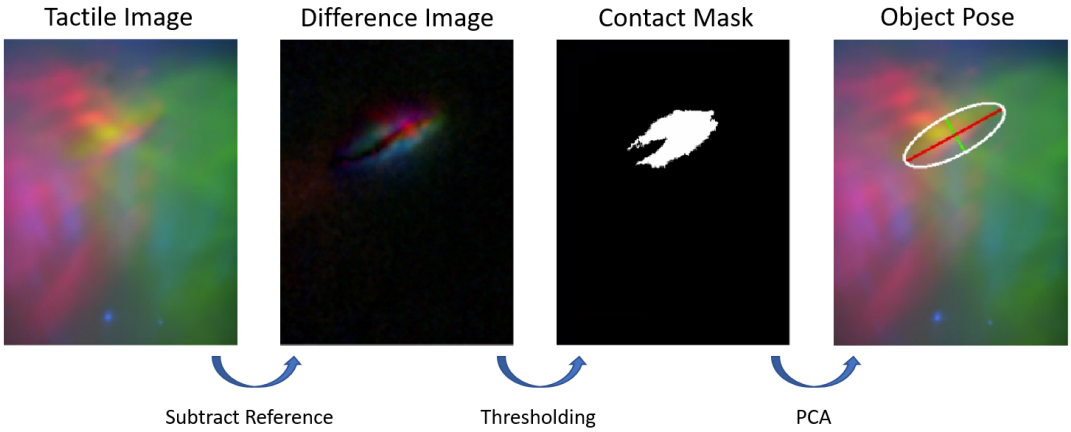


**Fig. 3** The procedure to get the contact center position and orientation

indicator, i.e., how strong the current contact is. By inspection, we set a contact strength threshold of 1.8 to distinguish a successful grasp from a failed one. $x, y$ and $\theta$ are the contact center and object angle calculated with PCA.

### 3.3 The Dynamics Model

An uncertainty-aware dynamics model was employed following the idea by Lakshminarayanan, Pritzel, and Blundell (2017) to adapt to a low data regime and ensure a safe grasp outcome. We chose a simple MLP (multilayer perceptron) structure because it can model complex contact dynamics without prior knowledge of the actual physics, making the method more generic. 16 MLPs were trained separately under negative log-likelihood loss with independent random initialization and shuffling. Then their outputs were ensembled to obtain a single prediction and its uncertainty (covariance). The idea is to ensure the actions to be executed have high confidence in a contact strength greater than 1.8 for the grasp to be solid. The contact location and orientation predictions also come with uncertainty, making it possible to plan and control with uncertainty for more accurate and generalizable results. Note that the full loss (in $s$, $x$, $y$ and $\theta$) is incurred only when the next grasp is a successful (stable) grasp. When the next grasp fails, only the loss in $s$ is backpropagated. This strategy helps to preserve the high uncertainty in the prediction of $x$, $y$ and $\theta$ when the grasp fails and increases the robustness of the system. We set the multilayer perceptrons to have five layers with 32 hidden units for each and use ReLU activation functions. The dropout rate is set at 0.1. We also use 2 numbers to represent a single rotation to facilitate learning angle continuity, as discussed by Zhou, Barnes, Lu, Yang, and Li (2019).

## 4 Experiments and Results

### 4.1 Hardware Setup and Data Collection

With 0.1 radius sampling step size, the original 8 degree of freedom joint angle action space is a vast space with 2244806784 discrete actions. However, action filtering retained only 424 actions, which is only around 0.000019% of the original space. This indicates that action filtering through kinematics is a simple yet effective strategy to reduce redundant actions and make the data collection process more efficient. A wooden stick with a 2.5 mm radius and 12.5 cm length (shown as object 1 in Fig 4) was used to collect the training data. Its radius is much smaller than the radius of the ball by Lambeta et al. (2020), requiring more estimation and control accuracy to manipulate. Note that only object 1 in Fig. 4 was used to collect the data that trained the dynamics model. The other 3 objects were used to test the model's generalization ability (see Section 4.4). The object was manipulated by randomly selecting actions from the 424 proposed actions, and the process continued indefinitely if the grasp of the object was maintained in the hand. When the grasp failed, the robotic hand was reset to a random state, and the object was manually put back into the hand.

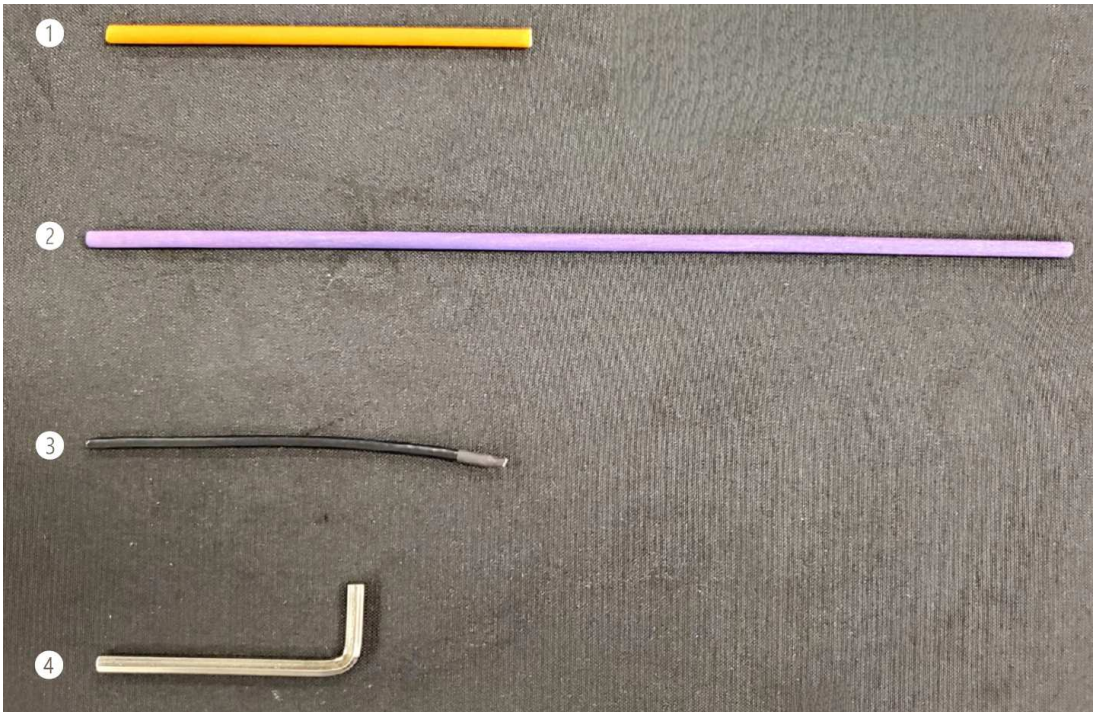


**Fig. 4** Different objects used in the experiments. Only object 1 was used to collect the dataset for training the dynamics model. The other objects were used to test the generalization ability of the learned model

The tactile state $(s, x, y, \theta)$ and hand state (joint angle and joint torque) were recorded before and after the action. A total of 2250 state transitions were recorded during the data collection process, and the grasp failed in 361 of them after the sampled action was applied, resulting in a failure rate of 16%. This failure rate is significantly lower than that of random sampling in the original entire action space, as the robotic hand can easily go into an unreasonable state. Compared with Lambeta et al. (2020), which used 96000 state transitions to train, the data size in this work is only 2.3% of that. However, the approach we adopted collected the data and trained the model more wisely, thus making the method data efficient and scale better.

## 4.2 Training the Dynamics Model

The original 2250 sampled data was split into a training set of 1800 samples and a test set of 450 samples. Two different cases were compared when training the dynamics model of manipulating the object. One used the full hand state, including the joint angle feedback from the robotic hand and the joint torque. The other was only using joint angle feedback as the hand state. The idea was to see if the hand torque proprioception helps accomplish the task. The evaluation of the neural network performance on test data is summarized in Table 1. The results show that predicting the contact strength is easier when the next state is a successful grasp, as indicated by the lower prediction error. However, both successful and failed grasps achieve reasonably small prediction errors. The position errors of both $x$ and $y$ axis are less than millimeter level, which indicates that the neural network can predict the future object locations given the current object pose and the hand state. Interestingly, the performances with and without joint torque data do not show a significant distinction. The possible explanation is that the tactile perception (the object pose and contact strength) already encodes the interaction information between the object and the robotic fingers, which is lower-dimensional and more informative than the hand torque data. Therefore, introducing the hand torque state to the input will provide little additional information to the system. Hence, the joint torque data will not be fed into the system in the remainder of this paper to reduce the neural network complexity.

**Table 1** Absolute Error on Test Data

| | With Joint Torque | Without Joint Torque |
|---|---|---|
| Failure Strength | $\mathbf{0.334 \pm 0.017}$ | $0.364 \pm 0.028$ |
| Success Strength | $\mathbf{0.122} \pm 0.006$ | $0.130 \pm 0.006$ |
| X Position (mm) | $\mathbf{0.512} \pm 0.0252$ | $0.517 \pm 0.0230$ |
| Y Position (mm) | $\mathbf{0.570} \pm 0.0474$ | $0.629 \pm 0.0409$ |
| Angle (Radius) | $0.148 \pm 0.008$ | $\mathbf{0.147} \pm 0.008$ |

## 4.3 Model Prediction Accuracy

The prediction accuracy of the learned dynamics model is also tested on real hardware. Specifically, we hope to test the model's prediction ability on actions with different output confidence. To achieve this, the finger was initialized to grasp the object that the model was trained on with a random pose. All the discrete actions from Section 3.1 were fed into the learned dynamics model to obtain the predicted contact strength, object pose, and their corresponding confidence. To only preserve actions that would result in a stable grasp, all the actions that had contact strengths less than 1.8 (the threshold between a solid contact and a loose contact) with greater than 5% probability were filtered. Among the remaining actions, their resulting contact center location can be plotted as a probability heat map (as shown in Fig. 2) to visualize the positional distribution of the actions better. The probability heat map was generated by taking the maximal probability among all the safe actions at each location in the tactile image. Mathematically, for a point $M(x, y)$ on the map $M$,

$$M(x, y) = \max_{a \in \mathcal{A}} \Pr(x, y \mid a) \tag{1}$$

where $\mathcal{A}$ is the set of the remaining actions. The maximum point in the heat map (brightest point in Fig. 2) corresponds to an action with the most confident resulting location. By selecting locations with different probability levels on the heat map and executing the actions behind these locations, we can test the dynamics model's prediction behavior on actions with varying confidence levels. The values on the probability heat map are normalized to the range between 0 to 1, which is referred to as the relative probability ratio in Table 2 for visualization and testing purposes. In each run, the hand chose the action with the corresponding relative probability ratio, and the resulting contact location was observed. The experiment was repeated for 50 runs for each probability ratio to get the average performance. It can be seen on Table 2 that the object did not fall out of the hand for all the cases. It is evident that filtering the actions by the contact strengths effectively produces a next state with safe and stable contact. The correlation between the magnitude of the errors and the actions' confidence level is not very strong. This is because although the relative probability ratio categorizes different actions for showing the statistics, they have different absolute probabilities in each run. Hence the relative probability ratio only serves as a way to organize and present the experiment. In practice, the absolute probability will be used to estimate the action errors.

For the prediction accuracy on the object's angle, because only two fingers are employed during the manipulation, the predicted distribution of the change of the object orientation is unimodal and narrowly spread. Therefore, only the one action with the least standard deviation for output angle is executed during testing. The prediction error averaged over 50 runs is $0.078 \pm 0.090$ radius.

## 4.4 Generalization to Novel Objects

To test the dynamics model's ability to generalize to other novel objects, in addition to the trained object ( 1# as shown in Fig. 4), three other objects were used to conduct the same experiment as in Section 4.3. Object 2# is the same as the trained object but 2 times longer. Object 3# is a small cable piece representing a different material. Object 4# is a hex key of even more complicated inertial property. The results are summarized in Table 3. The results indicate that object 3# is

**Table 2** Error Statistics on Real Hardware

| Probability Ratio | X Position Error (mm) | Y Position Error (mm) | Object Fall Frequency |
|---|---|---|---|
| 1.0 | $\mathbf{0.302 \pm 0.190}$ | $\mathbf{0.346 \pm 0.234}$ | **0/50** |
| 0.9 | $0.507 \pm 0.229$ | $0.414 \pm 0.370$ | **0/50** |
| 0.8 | $0.439 \pm 0.292$ | $0.521 \pm 0.366$ | **0/50** |
| 0.7 | $0.307 \pm 0.258$ | $0.385 \pm 0.283$ | **0/50** |
| 0.6 | $0.575 \pm 0.439$ | $0.531 \pm 0.375$ | **0/50** |
| 0.5 | $0.595 \pm 0.439$ | $0.404 \pm 0.322$ | **0/50** |
| 0.4 | $0.712 \pm 0.429$ | $0.482 \pm 0.317$ | **0/50** |

**Table 3** Error Statistics on Different Objects

| Probability Ratio | X Position Error (mm) | Y Position Error (mm) | Object Fall Frequency |
|---|---|---|---|
| **Object 2** | | | |
| 1.0 | $0.439 \pm 0.093$ | $0.302 \pm 0.253$ | **0/50** |
| 0.9 | $0.375 \pm 0.463$ | $1.062 \pm 0.370$ | 1/50 |
| 0.8 | $0.546 \pm 0.361$ | $0.517 \pm 0.346$ | 2/50 |
| 0.7 | $0.872 \pm 0.590$ | $0.453 \pm 0.322$ | **0/50** |
| 0.6 | $1.097 \pm 0.609$ | $0.653 \pm 0.585$ | 1/50 |
| 0.5 | $0.634 \pm 0.443$ | $0.487 \pm 0.395$ | 1/50 |
| 0.4 | $0.770 \pm 0.531$ | $0.336 \pm 0.424$ | 4/50 |
| **Object 3** | | | |
| 1.0 | $0.531 \pm 0.336$ | $0.439 \pm 0.249$ | **0/50** |
| 0.9 | $0.551 \pm 0.565$ | $0.453 \pm 0.322$ | **0/50** |
| 0.8 | $1.330 \pm 0.590$ | $0.551 \pm 0.370$ | **0/50** |
| 0.7 | $0.936 \pm 0.565$ | $0.595 \pm 0.482$ | 3/50 |
| 0.6 | $0.833 \pm 0.356$ | $0.453 \pm 0.366$ | **0/50** |
| 0.5 | $0.712 \pm 0.429$ | $0.434 \pm 0.317$ | 5/50 |
| 0.4 | $0.707 \pm 0.419$ | $0.653 \pm 0.419$ | 1/50 |
| **Object 4** | | | |
| 1.0 | $0.404 \pm 0.322$ | $0.707 \pm 0.414$ | **0/50** |
| 0.9 | $0.375 \pm 0.249$ | $0.750 \pm 0.668$ | 2/50 |
| 0.8 | $0.356 \pm 0.297$ | $0.838 \pm 0.658$ | 2/50 |
| 0.7 | $0.736 \pm 0.468$ | $0.575 \pm 0.453$ | 1/50 |
| 0.6 | $0.838 \pm 0.507$ | $0.614 \pm 0.453$ | 2/50 |
| 0.5 | $1.574 \pm 0.863$ | $0.936 \pm 0.755$ | 5/50 |
| 0.4 | $1.145 \pm 0.877$ | $1.330 \pm 1.019$ | 10/50 |

the easiest to manipulate compared to the other two, as it falls out of the hand less frequently than the others. Regarding the objects' physical properties, object 3# is indeed most similar to the trained object. Object 4# is the most difficult to deal with, as can be anticipated, because it is heavier, thicker, and with an imbalanced weight. The positional errors and fall frequencies did not deteriorate much for all three objects when the probability ratio was larger than 0.7. The positional errors were kept within the milliliter level for most cases. This reveals that the dynamics model learned some generalizable features and was, to some extent, robust to the variation of the object property.

### 4.5 Application: Pose Regulation Experiment

Using the learned dynamics model, we can control the contact position of the object with the finger. For example, the contact position can be adjusted to the exact center of the finger, given any initial configurations. In this experiment, we randomly placed object 1# approximately 6 mm away from the center of the finger. All the valid actions from Section 3.1 were fed into the dynamic model to obtain the corresponding future predictions. Actions that resulted in unstable contact were first removed (same as in Section 4.3), and among all the remaining actions, the action that gave a mean location prediction closest to the center of the finger was selected and executed. This process was then repeated for 5 steps to make the contact iteratively closer to the center. Another experiment was performed to examine that neural network learned a smooth landscape in action space by using additional gradient-based optimization on the original proposed discrete actions. The optimization took the $\ell 2$ norm of the distance from the finger center to the mean position prediction as the loss and backpropagated the gradient to the input action as the direction to improve the original action. Adam optimizer (Kingma & Ba, 2014) was adopted and the optimization was conducted for 100 iterations with a learning rate of 0.0005. Both experiments were conducted 30 times to obtain the expected performance. The predicted and actual distances from the contact center to the finger center in each step are shown in Fig. 5. The results demonstrate that both methods can regulate the contact to approach the finger center. Still, the additional gradient-based optimization can quickly reduce the distance within the millimeter level. The improvement with the optimization also suggests that the learned dynamics model not only is capable of predicting the outcome of the original discrete actions but also exhibits good performance on the continuous action space, making it possible to generalize to unseen actions owing to the neural network's interpolation ability.

## 5 Limitation

The prediction horizon is limited to 1 in this work, which is sufficient for simple pose regulations but may not yield optimal behavior and may get stuck in complex pose objectives. We plan to extend the prediction horizon to enable multi-step predictions. This will allow for more flexible planning and control, which may lead to more advanced manipulation tasks. In this work, only two fingers were considered to grasp the stick, which limited the controllability over the object's orientation. We will explore using additional fingers to control the full pose of the object for increased dexterity. Although the modules for this framework are tailored for the stick manipulation task, the same philosophy of feasible robot configuration classification and discretization, model-based tactile perception and uncertainty-aware dynamics model can be applied to many other tasks for improving the learning efficiency. Overall, the key contribution of this work is to provide a promising modular framework for future research on real hardware dexterous manipulation with limited data.

## 6 Conclusion

This work presented a data-efficient framework on real hardware for learning to manipulate a small stick with high accuracy. The approach involves reducing the feasible action space, leveraging model-based tactile signal extraction, and utilizing an uncertainty-aware contact dynamics model. The experimental results demonstrate that the framework achieves sub-millimeter precision with only 2.3% of the data required by previous work. Using the learned contact prediction model, our robotic hand can manipulate the rolling object with precision and safety.

## Declarations

### 6.1 Funding

This work was supported by the Project of Hetao Shenzhen-Hong Kong Science and Technology Innovation Cooperation Zone, under Grant HZQB-KCZYB-2020083. Partial financial support was received from Google.

### 6.2 Competing interests

The authors have no competing interests to declare that are relevant to the content of this article.

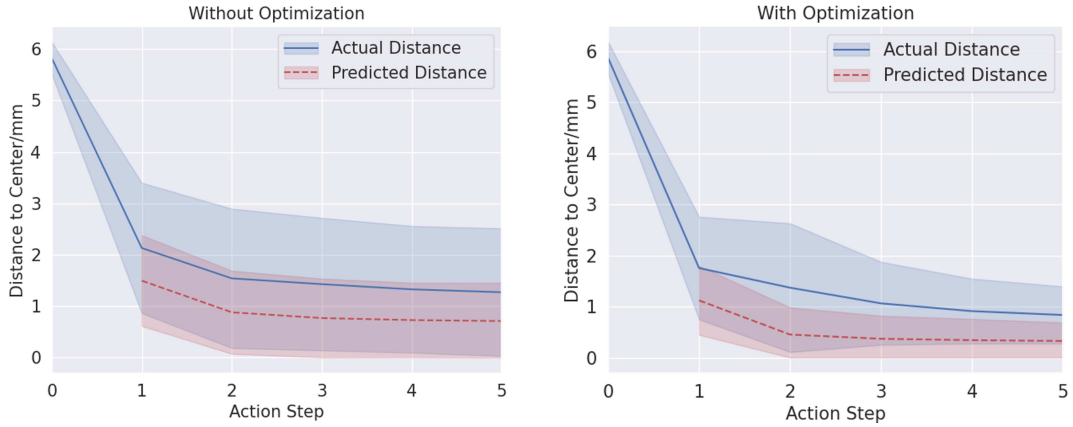


**Fig. 5** Predicted and actual contact distances to the finger center by applying the best actions iteratively. On the right, a gradient-based optimization was added to further improve the actions

## 6.3 Ethics approval

Not applicable.

## 6.4 Consent to participate

Not applicable.

## 6.5 Consent for publication

Not applicable.

## 6.6 Availability of data

The collected data for this project is available here.

## 6.7 Code availability

This project's source code is available here.

## 6.8 Authors' contributions

The principal author is Yipai Du while other authors are involved by supervision, guidance and discussion.

# References

Andrychowicz, O.M., Baker, B., Chociej, M., Józefowicz, R., McGrew, B., Pachocki, J., ... Zaremba, W. (2020). Learning dexterous in-hand manipulation. *The International Journal of Robotics Research*, *39*(1), 3-20, https://doi.org/10.1177/0278364919887447 Retrieved from https://doi.org/10.1177/0278364919887447 https://doi.org/10.1177/0278364919887447

Arunachalam, S.P., Silwal, S., Evans, B., Pinto, L. (2023). Dexterous imitation made easy: A learning-based framework for efficient dexterous manipulation. *2023 ieee international conference on robotics and automation (icra)* (pp. 5954–5961).

Chen, Q., Van Wyk, K., Chao, Y.-W., Yang, W., Mousavian, A., Gupta, A., Fox, D. (2023). Learning robust real-world dexterous grasping policies via implicit shape augmentation. *Conference on robot learning* (pp. 1222–1232).

Chen, T., Tippur, M., Wu, S., Kumar, V., Adelson, E., Agrawal, P. (2023). Visual dexterity: In-hand dexterous manipulation from depth. *Icml workshop on new frontiers in learning, control, and dynamical systems.*

Chen Chen, F., Appendino, S., Battezzato, A., Favetto, A., Mousavi, M., Pescarmona, F. (2013). Constraint study for a hand exoskeleton: human hand kinematics and dynamics. *Journal of Robotics*, *2013*, ,

Dafle, N.C., Rodriguez, A., Paolini, R., Tang, B., Srinivasa, S.S., Erdmann, M., ... Fuhlbrigge, T. (2014, May). Regrasping objects using extrinsic dexterity. *Proceedings of (icra) international conference on robotics*

*and automation* (p. 2560 - 2566).

Dong, S., et al. (2021). *High-resolution tactile sensing for reactive robotic manipulation* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Du, Y., Zhang, G., Wang, M.Y. (2022). 3d contact point cloud reconstruction from vision-based tactile flow. *IEEE Robotics and Automation Letters*, *7*(4), 12177–12184,

Handa, A., Allshire, A., Makoviychuk, V., Petrenko, A., Singh, R., Liu, J., ... others (2023). Dextreme: Transfer of agile in-hand manipulation from simulation to reality. *2023 ieee international conference on robotics and automation (icra)* (pp. 5977–5984).

Ishihara, T., Namiki, A., Ishikawa, M., Shimojo, M. (2006). Dynamic pen spinning using a high-speed multifingered hand with high-speed tactile sensor. *2006 6th ieee-ras international conference on humanoid robots* (pp. 258–263).

Khandate, G., Mehlman, C., Wei, X., Ciocarlie, M. (2023). Value guided exploration with sub-optimal controllers for learning dexterous manipulation. *arXiv preprint arXiv:2303.03533*, ,

Kingma, D.P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, ,

Kober, J., Bagnell, J.A., Peters, J. (2013). Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research*, *32*(11), 1238–1274,

Kumar, V., Xu, Z., Todorov, E. (2013). Fast, strong and compliant pneumatic actuation for dexterous tendon-driven hands. *2013 ieee international conference on robotics and automation* (pp. 1512–1519).

Lakshminarayanan, B., Pritzel, A., Blundell, C. (2017). Simple and scalable predictive uncertainty estimation using deep ensembles. *Advances in neural information processing systems*, *30*, ,

Lambeta, M., Chou, P.-W., Tian, S., Yang, B., Maloon, B., Most, V.R., ... others (2020). Digit: A novel design for a low-cost compact high-resolution tactile sensor with application to in-hand manipulation. *IEEE Robotics and Automation Letters*, *5*(3), 3838–3845,

Nagabandi, A., Konolige, K., Levine, S., Kumar, V. (2020). Deep dynamics models for learning dexterous manipulation. *Conference on robot learning* (pp. 1101–1112).

Qi, H., Kumar, A., Calandra, R., Ma, Y., Malik, J. (2023). In-hand object rotation via rapid motor adaptation. *Conference on robot learning* (pp. 1722–1732).

Sferrazza, C., & D'Andrea, R. (2019). Design, motivation and evaluation of a full-resolution optical tactile sensor. *Sensors*, *19*(4), 928,

Shaw, K., Bahl, S., Pathak, D. (2023). Videodex: Learning dexterity from internet videos. *Conference on robot learning* (pp. 654–665).

She, Y., Wang, S., Dong, S., Sunil, N., Rodriguez, A., Adelson, E. (2021). Cable manipulation with a tactile-reactive gripper. *The International Journal of Robotics Research*, *40*(12-14), 1385-1401, https://doi.org/10.1177/02783649211027233 Retrieved from https://doi.org/10.1177/02783649211027233 https://doi.org/10.1177/02783649211027233

Shimonomura, K. (2019). Tactile image sensors employing camera: A review. *Sensors*, *19*(18), 3933,

Solak, G., & Jamone, L. (2019). Learning by demonstration and robust control of dexterous in-hand robotic manipulation skills. *2019 ieee/rsj international conference on intelligent robots and systems (iros)* (pp. 8246–8251).

Taylor, I.H., Dong, S., Rodriguez, A. (2022). Gelslim 3.0: High-resolution measurement of shape, force and slip in a compact tactile-sensing finger. *2022 international conference on robotics and automation (icra)* (pp. 10781–10787).

Van Hoof, H., Hermans, T., Neumann, G., Peters, J. (2015). Learning robot in-hand manipulation with tactile features. *2015 ieee-ras 15th international conference on humanoid robots (humanoids)* (pp. 121–127).

Ward-Cherrier, B., Pestell, N., Cramphorn, L., Winstone, B., Giannaccini, M.E., Rossiter, J., Lepora, N.F. (2018). The tactip family: Soft optical tactile sensors with 3d-printed biomimetic morphologies. *Soft robotics*, *5*(2), 216–227,

Yin, Z.-H., Huang, B., Qin, Y., Chen, Q., Wang, X. (2023). Rotating without seeing: Towards in-hand dexterity through touch. *arXiv preprint arXiv:2303.10880*, ,

Yuan, W., Dong, S., Adelson, E.H. (2017). Gelsight: High-resolution robot tactile sensors for estimating geometry and force. *Sensors*, *17*(12), 2762,

Zhou, Y., Barnes, C., Lu, J., Yang, J., Li, H. (2019). On the continuity of rotation representations in neural networks. *Proceedings of the ieee/cvf conference on computer vision and pattern recognition* (pp. 5745–5753).