Adaptive contact-rich manipulation through few-shot imitation learning with Force-Torque feedback and pre-trained object representations

Chikaha Tsuji[1*], Enrique Coronado[2], Pablo Osorio[3] and Gentiane Venture[4,2*]

***Abstract*— Imitation learning offers a pathway for robots to perform repetitive tasks, allowing humans to focus on more engaging and meaningful activities. However, challenges arise from the need for extensive demonstrations and the disparity between training and real-world environments. This paper focuses on contact-rich tasks involving deformable objects, such as wiping with a sponge, requiring adaptive force control to handle variations in wiping surface heights and the sponge's physical properties. To address these challenges, we propose a novel method that integrates real-time time-series force-torque (FT) feedback with pre-trained object representations, allowing robots to dynamically adjust to previously unseen changes in wiping surface heights and sponge's physical properties.**

In real-world experiments, our method achieved 96% accuracy in applying reference forces, significantly outperforming the previous method that lacked an FT feedback loop, which only achieved 4% accuracy. To evaluate the adaptability of our approach, we conducted experiments under different conditions from the training setup, involving 40 scenarios using 10 sponges with varying physical properties and 4 types of wiping surface heights, demonstrating significant improvements in the robot's adaptability by analyzing force trajectories. The video of our work is available at: https://sites.google.com/view/adaptive-wiping.

I. INTRODUCTION

Robots are crucial for handling mundane tasks, but pre-programming each task is impractical, leading to increased interest in imitation learning [1]. Despite its benefits, challenges like the need for extensive demonstrations and discrepancies between training and real-world environments persist [2]. Thus, robots must not merely mimic but adapt to new environments, even with limited demonstration data.

A challenging aspect of robotic manipulation is executing contact-rich tasks, which involve extensive physical interactions. Interestingly, those involving deformable objects pose particular challenges due to the need for precise force control and adaptation to changes [3]. Wiping tasks, for example, demand careful force adjustments based on wiping surface height and sponge's physical properties.

Therefore, in this paper, we address the challenge – *Could robots learn a versatile manipulation policy via few-shot imitation learning capable of adapting to environmental changes: the height of manipulating surface and the physical properties of manipulated objects?*

[1] Department of Mechano-Informatics, The University of Tokyo, Japan. tsujchik@g.ecc.u-tokyo.ac.jp

[2] National Institute of Advanced Industrial Science and Technology, Japan

[3] Department of Mechanical Systems Engineering, Tokyo University of Agriculture and Technology, Japan.

[4] Department of Mechanical Engineering, The University of Tokyo, Japan. venture@g.ecc.u-tokyo.ac.jp *Corresponding authors

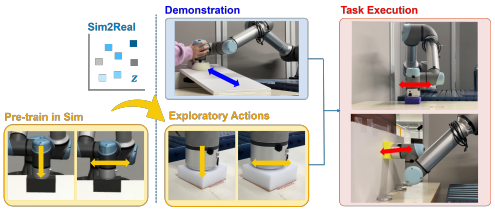


Fig. 1: Wiping tasks: Pretraining in simulation (left), real-world data collection (middle), and task execution (right).

II. RELATED WORKS AND CONTRIBUTION

Learning-based methods have addressed contact-rich tasks. Reinforcement learning enables such tasks by defining reward functions, as in [4], [5], but suffers from sample inefficiency and sensitivity to reward design.

Imitation learning, in contrast, learns from demonstrations and is more sample-efficient. Rozo et al. [6] used Gaussian mixture models for cooperative tasks, while Yamane et al. [7] employed bilateral control to decouple applied and environmental forces for grasping.

Manipulating deformable objects is especially challenging due to the need for precise force control and adaptability. To address this, prior works have used self-supervised representation learning with visual or tactile observations to encode object properties [8]–[10], while other studies highlight the effectiveness of haptic time-series data in capturing physical properties [11].

Most closely related to our work, Aoyama et al. [12] pre-trained haptic representations from force-torque data in simulation and transferred them to reality to wipe a table with deformable sponges. However, their approach used open-loop control and could not adapt to changes in table height. In contrast to learning-based methods, non-learning-based methods such as impedance [13] and admittance control [14] offer closed-loop control but require predefined target forces or positions, which are unknown in our setting due to varying sponge properties and wiping surface heights. Thus, we addressed these challenges with three contributions:

- We propose a framework that combines simulated pre-training of object properties with real-time time-series force-torque (FT) feedback, enabling a robot to adapt to changes in wiping surface heights and sponge's physical properties from only a few demonstrations (Fig. 1).
- Unlike [12], we incorporate closed-loop control without requiring prior target force or position information.
- We validate our method on real hardware by testing it on various wiping surface heights and sponges, showing effective adaptation through force measurements.

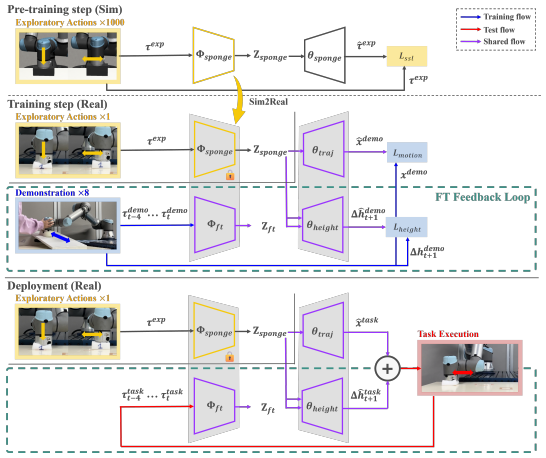


Fig. 2: Overview of our proposed framework.

III. Methods

The proposed method consists of two steps: a pre-training step using a simulator and a training step using a real robot, before being deployed (Fig. 2), each step is detailed below.

A. Pre-training step

We pre-train the sponge properties encoder ϕ_{sponge} on simulated unlabeled data $D_{\text{sim}} = \{(\tau^{\text{exp}})_1, \ldots, (\tau^{\text{exp}})_M\}$, collected by performing pre-defined exploratory actions (detailed in IV-B), to capture the sponges' physical properties as the latent space Z_{sponge} covering a wide range of the underlying distribution. We use a self-supervised learning framework inspired by [12] but with a modified architecture.

Using a Variational Autoencoder (VAE) [15] approach, the VAE encoder-decoder model $\phi_{\text{sponge}} - \theta_{\text{sponge}}$ takes FT trajectory τ^{exp} from D_{sim} as inputs and outputs reconstructed FT trajectory $\hat{\tau}^{\text{exp}}$, treating the latent space Z_{sponge} as a Gaussian distribution with five dimensions.

The VAE encoder-decoder model $\phi_{\text{sponge}} - \theta_{\text{sponge}}$ consists of 2 fully connected encoder layers, 1 sampling step, and 2 fully connected decoder layers. To flatten the six sensors' time-series data $\tau^{\text{exp}} \in \mathbb{R}^{400\times6}$, we employ 2 fully connected layers each for the encoder ϕ_{sponge} and decoder θ_{sponge}. The encoder ϕ_{sponge} comprises 1 fully connected layer of 5 hidden dimensions followed by the flattening step and 1 fully connected layer. Whereas the decoder θ_{sponge} comprises 1 fully connected layer with Rectified Linear Unit (ReLU) as an activation function and a dropout rate of 0.1 followed by a reshaping step and 1 fully connected layer of 5 hidden dimensions. The latent space dimension $Z_{\text{sponge}} \in \mathbb{R}^5$ is designed to capture sponges' stiffness, friction, and other non-intuitive physical properties. We adopt a loss function L_{ssl} shown in Eq. (1), with $\beta = 0.06$.

$$\begin{aligned} L_{ssl} &= E_{\text{MSE}}(\hat{\tau}^{\text{exp}}, \tau^{\text{exp}}) \\ &+ \beta D_{\text{KL}}(q_{\phi_{\text{sponge}}}(z \mid \tau^{\text{exp}}) \,||\, p_{\phi_{\text{sponge}}}(z)) \end{aligned} \tag{1}$$

B. Training step

We train the motion trajectory decoder θ_{traj} and the FT feedback loop $\phi_{\text{ft}} - \theta_{\text{height}}$ on real-world unlabeled data $D_{\text{real}} = \{\tau^{\text{exp}}\}$, collected by the same pre-defined exploratory actions with III-A, and few-shot human demonstration data $D_{\text{demo}} = \{(x^{\text{demo}}, \Delta h^{\text{demo}}, \tau^{\text{demo}})_1, \ldots, (x^{\text{demo}}, \Delta h^{\text{demo}}, \tau^{\text{demo}})_N\}$.

1) Motion trajectory decoder θ_{traj}

We train the wiping motion trajectory decoder θ_{traj} using Learning from Demonstration (LfD) [12] to generate the wiping motion $\hat{x}^{\text{task}}$ according to sponge properties.

The encoder-decoder model $\phi_{\text{sponge}} - \theta_{traj}$ takes FT trajectory τ^{exp} from D_{real} as inputs and outputs the corresponding motion trajectory $\hat{x}^{\text{demo}}$. Here, the encoder ϕ_{sponge} is pre-trained on simulated data D_{sim}, with its weights frozen during training on real data, and then deployed in the real world (Sim2Real).

The motion trajectory decoder θ_{traj} consists of 1 fully connected layer with a dropout rate of 0.1. We adopt the Mean Squared Error L_{traj} between the generated motion trajectory $\hat{x}^{\text{demo}}$ and the demonstrated one x^{demo} represented in the absolute coordinate from the base link (Eq. (2)).

$$L_{motion} = E_{\text{MSE}}(\hat{x}^{\text{demo}}, x^{\text{demo}}) \tag{2}$$

2) FT feedback loop $\phi_{ft} - \theta_{height}$

We train an FT feedback loop $\phi_{\text{ft}} - \theta_{\text{height}}$ composed of the FT encoder ϕ_{ft} and the end-effector's vertical position decoder θ_{height} to obtain a control input of the next time step's vertical position according to the contact state and the manipulated sponge.

The FT encoder ϕ_{ft} processes the FT history from the demonstrations $D_{\text{demo\_ft}} = \{\tau^{\text{demo}}_{\text{t-4}}, \ldots, \tau^{\text{demo}}_{\text{t}}\}$, encoding it into the latent space $Z_{\text{ft}} \in \mathbb{R}^6$, which is designed to represent the forces and torques along the x, y, and z axes. The end-effector's vertical position decoder θ_{height} takes the concatenated latent spaces Z_{sponge} from the sponge properties encoder ϕ_{sponge} and Z_{ft} from the FT encoder ϕ_{ft} as inputs, and outputs the next time step's vertical displacement $\Delta\hat{h}^{\text{demo}}_{\text{t+1}}$.

The FT encoder ϕ_{ft} consists of 2 layers of temporal convolutional network (TCN) [16] with 25 hidden channels each and a dropout rate of 0.1. Inspired by [17], which suggests that TCN has advantages in training efficiency and training time over gated recurrent units (GRU) [18], we adopt TCN as our sequence model. The end-effector's vertical position decoder θ_{height} consists of 2 fully connected layers: the first fully connected layer of 128 hidden dimensions with ReLU as an activation function and a dropout rate of 0.1 followed by the final layer (the second fully connected layer). We adopt the Mean Squared Error L_{height} between the predicted vertical displacement in the next time step $\Delta\hat{h}^{\text{demo}}_{\text{t+1}}$ and that of the ground truth $\Delta h^{\text{demo}}_{\text{t+1}}$ (Eq. (3)).

$$L_{height} = E_{\text{MSE}}(\Delta\hat{h}^{\text{demo}}_{\text{t+1}}, \Delta h^{\text{demo}}_{\text{t+1}}) \tag{3}$$

C. Deployment

In the task execution, the robot performs a wiping motion by combining offline horizontal (x, y) motion of $\hat{x}^{\text{task}}$ and online vertical (z) motion of $\Delta\hat{h}^{\text{task}}_{\text{t+1}}$. First, the robot collects unlabeled data $D_{\text{task}} = \{\tau^{\text{exp}}\}$ of the sponge being used in the task through pre-defined exploratory actions. Then it

generates (x,y) planar motion $\hat{x}^{\text{task}}$ from D_{task} and replays the motion offline. The FT feedback loop actively infers the next vertical position $\Delta\hat{h}^{\text{task}}_{\text{t+1}}$ from the previous $\sim$ current force and torque history $D_{\text{task\_ft}} = \{\tau^{\text{task}}_{\text{t-4}}, \ldots, \tau^{\text{task}}_{\text{t}}\}$, and adapts online.

IV. Experiment Setup

A. Wiping Task and Setup

We evaluated our method on a contact-rich wiping task requiring adaptation to wiping surface height and sponge properties. A 6-DoF UR5e robot arm equipped with a 6-axis FT sensor and a sponge at its end-effector was controlled via position commands. Demonstrations were collected kinesthetically in free-drive mode.

B. Dataset

Unlabeled data: We performed two exploratory actions [12] to capture sponge stiffness and friction: (i) pressing at $0.01\ \mathrm{m/s}$ for 2s, (ii) lateral motion at $0.05\ \mathrm{m/s}$ for 1s in each direction. During these, 3-axis FT data was recorded for 4s to obtain the FT trajectory $\tau^{\text{exp}} \in \mathbb{R}^{400\times 6}$. We collected 1000 unlabeled data in simulation using dynamics domain randomization with friction $\mu \in [0.0, 3.5]$, stiffness $k \in [0.5, 1000]$ N/m, and damping width $d \in [0.02, 0.3]$ m. One additional real-world unlabeled data τ^{exp} was collected using the same sponge as in the demonstrations.

Demonstration data: A human demonstrator wiped an inclined table (different from experiment settings) with maximal force. 8 demonstrations using a sponge were collected. We recorded the end-effector position and FT data for 10s to obtain the motion trajectory $x^{\text{demo}} \in \mathbb{R}^{25\times 2}$, vertical displacement trajectory $\Delta h^{\text{demo}} \in \mathbb{R}^{25}$, and FT trajectory $\tau^{\text{demo}} \in \mathbb{R}^{6\times 25}$.

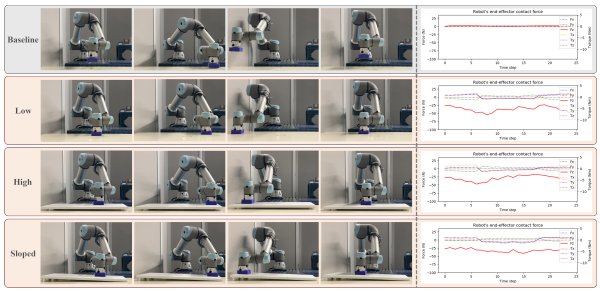
C. Model Training

We applied Butterworth low-pass filtering to the data, followed by normalization to the range $[0.0, 0.9]$. In the pre-training step, the sponge properties encoder ϕ_{sponge} was trained on 1000 unlabeled simulation data for 200 epochs using the Adam optimizer with a learning rate of 0.0001. In the training step, the motion trajectory decoder θ_{traj} was trained on 8 demonstrated motion trajectories x^{demo} for 10000 epochs, while the FT feedback loop $\phi_{\text{ft}} - \theta_{\text{height}}$ was trained on 8 demonstrated vertical displacement trajectories Δh^{demo} and FT trajectories τ^{demo} as time-series data with a window size of 5 for 2000 epochs, where both used the Adam optimizer with a learning rate of 0.001.

V. Results and Discussion

To evaluate our method, we conducted experiments across 40 scenarios with varying table heights (V-A) and sponge types (V-B) as illustrated in Fig. 3. We compared our method with two baselines: (1) an imitation learning-based method without FT feedback [12] (baseline), and (2) an admittance control (AC) method. Since AC requires a predefined target force, which is unavailable in our setting, we defined it as the force applied when the sponge is compressed by 1cm. AC computes vertical displacement Δh using Eq. (4) [19]:

(a)



(b)

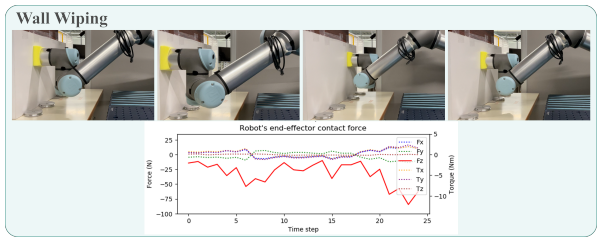


Fig. 3: Wiping under various settings: (a) low, high, and sloped (table wiping); (b) vertical wall (wall wiping).

$$\Delta h = \frac{FT^2 + BT\,\Delta h_{t-1} + M\,(2\,\Delta h_{t-1} - \Delta h_{t-2})}{M + BT + KT^2} \tag{4}$$

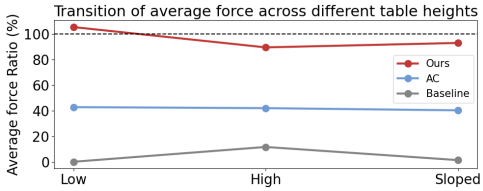
where $M = 0.5\,\mathrm{kg}$, $B = 5\,\mathrm{N/(m/s)}$, $K = 15\,\mathrm{N/m}$, and $T = 0.4\,\mathrm{s}$. We also tested our model on a vertical wall using the same model trained on table wiping. For each condition, we evaluated the contact ratio, the average vertical force, and its ratio to the reference force. The reference force for each sponge was obtained by performing demonstrations using the same procedure as in IV-B, and using the average vertical force applied during those demonstrations as the reference.

A. Verification of the ability to adapt to changes in height

We varied the wiping table heights (low, high, sloped) from the height used in the demonstrations (inclined table). The results are shown in Fig. 4 (a).

To adapt wiping motions to changes in the wiping surface height, the robot should apply a consistent force to the sponge regardless of the height. With the same sponge, the robot should wipe with as much force as possible to ensure effective wiping. With the baseline method, the sponge was in contact with the table only 0-44% of the time, and the average force reached merely 4% of the desired reference force. Specifically, in some cases with the low and sloped tables, the average force turned positive because the sponge did not contact the table, and the influence of gravitational force from the sponge's own weight became dominant. This indicates that a robot did not effectively *'wipe'* and was unable to adapt to changes in the wiping surface height. In contrast, both AC and our proposed method maintained constant contact in all 30 cases. However, AC applied only an average of 42% of the reference force, whereas our proposed method successfully maintained an appropriate average force on the sponge across all heights, averaging 96% of the reference force. Furthermore, the applied force did not significantly vary with changes in table height as shown in Fig. 4 (a), with the standard deviation being only about 5% larger than that of human demonstrations. This

(a)



(b)

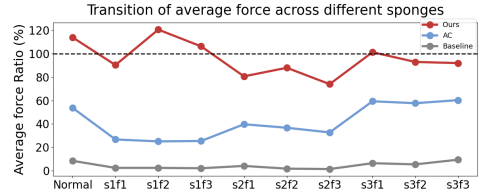


Fig. 4: Transition in the ratio of average force to human-demonstrated reference force (100% dotted line): (a) across table heights, (b) across sponges.

indicates the robot's ability to successfully adapt to height variations.

B. Verification of the ability to adapt to changes in sponge

We varied the sponge properties from the sponge used in the demonstrations (*normal*). We denote a sponge with stiffness level m and friction level n as $\texttt{s}m\texttt{f}n$, where $m, n \in 1, 2, 3$. The results are shown in Fig. 4 (b). Adapting the wiping motions to changes in the sponges' physical properties requires adjusting the force applied to the sponge accordingly. With the baseline method, the robot failed to maintain contact with the table when using sponges with unseen properties. Specifically, with the low table, the contact ratio was 0% for all 9 unseen sponges. Moreover, the average force applied was less than 25% of the expected force (Fig. 4 (b)), averaging only 4% of the reference force. Therefore, the baseline is unable to adapt to unseen sponges.

In contrast, both AC and our proposed method successfully maintained contact at all time steps in all 30 cases. However, AC merely maintained the predefined target force without considering the sponge's physical properties, resulting in only 23-63% of the expected force and an average of 42% of the reference force being exerted. Our proposed method, on the other hand, applied an average force comparable to the expected force, achieving over 63% and an average of 96% of the reference force, according to the type of sponge (Fig. 4 (b)). This demonstrates that our method successfully enables the robot to adapt to unseen sponge properties.

C. Wall Wiping

In real-world scenarios, cleaning involves more than just wiping horizontal surfaces like tables; it may include tasks such as wiping walls and other vertical surfaces. A key challenge for robots in these tasks is the ability to adapt to the physical properties of sponges and adjust the applied force in real time as surface conditions change. Our method achieves this adaptiveness independently of gravitational effects. In previous tasks (V-A and V-B), the direction of the forces applied to the sponge was aligned with gravitational acceleration. To further demonstrate that our method is effective regardless of gravity's influence, we tested our method in a gravity-neutral setting—wall wiping (Fig. 3 (b))—where gravitational forces do not affect the applied forces during the task.

TABLE I: Wall wiping results

Sponge	Contact Ratio	Avg. Force [N] (Ratio)
Normal	100%	-14.5 (115%)
s1f1	100%	-23.7 (104%)
s1f2	100%	-29.5 (138%)
s1f3	100%	-25.4 (119%)
s2f1	100%	-29.0 (120%)
s2f2	100%	-32.2 (107%)
s2f3	100%	-28.9 (84%)
s3f1	100%	-27.0 (87%)
s3f2	100%	-33.5 (95%)
s3f3	100%	-27.3 (74%)

TABLE II: The average ratio of the applied force in the z-direction to the reference force exerted by the demonstrator.

	Layer		Window Size		Demo		proposed
	Fewer	More	Smaller	Larger	Fewer	More	
Average (%)	159	152	182	170	190	114	**97**

We evaluated the same model as V-A and V-B, trained using the same demonstration data of table wiping. Due to the setting changes, the end-effector's frame rotated 90 degrees and the base-link's x-axis came vertically to the end-effector. We swapped the position outputs of the x-axis and z-axis based on the base-link, and introduced an offset to the z-axis positions. Although this might appear as a mere transformation of output trajectories, the core challenge lies in the method's ability to adjust applied forces in a gravity-independent manner. The results are shown in Table I.

Our method maintained contact with a wall in all 10 cases and the applied forces were comparable to that expected, averaging 104% of the reference force. This indicates that a robot can adapt to wall wiping even with unseen sponges.

D. Ablation Study

We conducted ablation studies on (1) the number of layers in the FT feedback loop, (2) the TCN window size, and (3) the number of demonstrations. The results are summarized in Table II. As shown, performance degraded when the model had either fewer or more layers, or when the window size was smaller or larger than the proposed setting. The best performance was achieved using a two-layer feedback loop with a window size of 5. Furthermore, 8 demonstrations were sufficient for the model to learn adaptive force control comparable to using more data.

VI. Conclusion

We proposed a method for contact-rich manipulation of deformable objects that enables robots to adapt to environmental and object property changes using only a few demonstrations, by combining real-time time-series FT feedback with pre-trained object representations. Experimental results on wiping tasks with varying wiping surface heights and sponge properties demonstrated superior adaptability compared to the baseline and AC methods. In future work, we aim to extend the method to a wider range of objects by leveraging large-scale multimodal pre-trained models.

REFERENCES

[1] A. Hussein and et al., "Imitation learning: A survey of learning methods," *ACM Computing Surveys (CSUR)*, vol. 50, no. 2, pp. 1–35, 2017.

[2] Y. Duan and et al., "One-shot imitation learning," in *Advances in Neural Information Processing Systems*, I. Guyon, U. von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds. Curran Associates, Inc.

[3] A. Billard and D. Kragic, "Trends and challenges in robot manipulation," *Science*, vol. 364, no. 6446, p. eaat8414, 2019.

[4] R. Martín-Martín and et al., "Variable impedance control in end-effector space: An action space for reinforcement learning in contact-rich tasks," in *IEEE/RSJ Int. Conf. on intelligent robots and systems*, 2019, pp. 1010–1017.

[5] O. Spector and M. Zacksenhouse, "Learning contact-rich assembly skills using residual admittance policy," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2021, pp. 6023–6030.

[6] L. Rozo, D. Bruno, S. Calinon, and D. G. Caldwell, "Learning optimal controllers in human-robot cooperative transportation tasks with position and force constraints," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2015, pp. 1024–1030.

[7] K. Yamane, Y. Saigusa, S. Sakaino, and T. Tsuji, "Soft and rigid object grasping with cross-structure hand using bilateral control-based imitation learning," *IEEE Robotics and Automation Letters*, 2023.

[8] P. Florence and et al., "Self-supervised correspondence in visuomotor policy learning," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 492–499, 2019.

[9] J. Pari and et al., "The surprising effectiveness of representation learning for visual imitation," *CoRR*, vol. abs/2112.01511, 2021.

[10] I. Guzey and et al., "Dexterity from touch: Self-supervised pre-training of tactile representations with robotic play," *arXiv preprint arXiv:2303.12076*, 2023.

[11] O. Kroemer and et al., "Learning dynamic tactile sensing with robust vision-based training," *IEEE Trans. on robotics*, vol. 27, no. 3, pp. 545–557, 2011.

[12] M. Y. Aoyama and et al., "Few-shot learning of force-based motions from demonstration through pre-training of haptic representation," in *Proc. of the IEEE Int. Conf. on Robotics and Automation*, 2023.

[13] N. Hogan, "Impedance control: An approach to manipulation: Part ii—implementation," 1985.

[14] H. Seraji, "Adaptive admittance control: An approach to explicit force control in compliant motion," in *Proceedings of the 1994 IEEE Int. Conf. on Robotics and Automation*, 1994, pp. 2705–2712.

[15] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," *arXiv preprint arXiv:1312.6114*, 2013.

[16] S. Bai and et al., "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," *arXiv:1803.01271*, 2018.

[17] J. Lee and et al., "Learning quadrupedal locomotion over challenging terrain," *Science Robotics*, vol. 5, no. 47, p. eabc5986, 2020.

[18] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," in *Proc. NeurIPS Workshop Deep Learn.*, 2014.

[19] P. Song, Y. Yu, and X. Zhang, "A tutorial survey and comparison of impedance control on robotic manipulation," *Robotica*, vol. 37, no. 5, p. 801–836, 2019.