# Learning Diffusion Policies from Demonstrations For Compliant Contact-rich Manipulation

Malek Aburub[1,2*], Cristian C. Beltran-Hernandez[1*], Tatsuya Kamijo[1], Masashi Hamaya[1]

*Abstract*— **Robots excel at repetitive or hazardous tasks, but achieving human-like dexterity in contact-rich environments remains challenging. Rigid robots struggle with stable contact and force application. Learning from Demonstration helps, but it faces difficulties in intricate tasks like powder grinding. This paper introduces Diffusion Policies For Compliant Manipulation (DIPCOM), a diffusion-based framework for compliant control. By leveraging generative diffusion models, our policy predicts Cartesian end-effector poses and adjusts arm stiffness for precise force control. We improve multimodal force modeling, enhance diffusion policy integration in compliance control, and demonstrate effectiveness in real-world tasks by comparing DIPCOM with existing methods. See our project page for the supplemental video `omron-sinicx.github.io/DIPCOM`**

## I. Introduction

Robots have great potential to improve daily life by handling repetitive or hazardous tasks. However, achieving human-like dexterity remains challenging, especially for precise, contact-rich manipulation in dynamic environments. Rigid robots struggle to maintain stable surface contact and apply consistent force, making force-intensive tasks difficult.

Learning from Demonstration (LfD) techniques [1] enable robots to learn complex tasks by observing human experts. However, tasks requiring high force, such as powder grinding, present unique challenges. Rigid robots often need mechanical compliance mechanisms [2] to ensure safe interactions, but this makes precise tool positioning harder.

To address these challenges, we adopt compliance control schemes that regulate forces via external sensors [3]. Our prior work [4] introduced Comp-ACT, which combines force information and compliance control with VAE-ACT policies. While effective, it struggles with long-horizon tasks requiring repetitive behaviors. To improve performance, we incorporate diffusion models, allowing adaptive force regulation while maintaining precision and stability.

We propose Diffusion Policies For Compliant Manipulation (DIPCOM), a diffusion-based framework for compliant control. Diffusion models capture multimodal action distributions and generate diverse behaviors [5]. By leveraging these properties, our policy predicts both Cartesian EE pose and stiffness adjustments, enabling robust force application.

Our contributions: First, a diffusion-based framework for rigid robots to learn contact-rich manipulation via compliance control, enhancing force regulation. Second, A comparison of DIPCOM and Comp-ACT [4], demonstrating the strengths of diffusion-based policies in force-intensive tasks.

* Equal contribution.
[1]OMRON SINIC X Corporation, Tokyo, Japan
[2]Department of Engineering Science, Osaka University, Japan
Corresponding authors:
`cristian.beltran [at] sinicx.com`
This work was partly supported by the JST-Mirai Program (Grant Number JPMJMI21G2).

## II. Related Works

### A. Learning from Demonstrations for Contact-rich Manipulation

Learning from Demonstration (LfD) enables robots to acquire complex contact-rich skills [6], leveraging force/torque sensing for tasks like grasping [7], ironing [8], pouring [9], and insertion [10]. Recent methods improve sample efficiency using transformer-based models and teleoperation systems like ALOHA [11] and UMI [12].

Most LfD approaches focus on position control or mechanical compliance, whereas we emphasize active compliance control. Building on Comp-ACT [4], we introduce DIPCOM, leveraging diffusion models for enhanced force-aware manipulation.

Inspired by prior work [13], [14] comparing imitation learning methods, we evaluate DIPCOM against Comp-ACT on real-world contact-rich tasks with rigid robots.

### B. Diffusion Policies

Diffusion models, originally proposed by Ho et al. [15], generate samples by refining noise through a stochastic process. For a detailed survey, see [16].

In robotics, diffusion models capture multi-modal actions and have been applied to motion planning [17], navigation [18], human-robot interaction [19], and grasping [20]. Chi et al. [21] demonstrated their efficacy in visuomotor policy learning. Further advancements include skill acquisition from language-annotated play data [22] and goal-conditioned diffusion policies [23].

Most prior approaches rely on position or velocity controllers, limiting effectiveness in contact-rich tasks. To overcome this, we propose a force-conditioned diffusion policy integrated with a compliance controller, enabling more precise force regulation. Our framework requires fewer demonstrations than previous methods lacking compliance control [14], significantly improving task performance.

## III. Methodology

We introduce Diffusion Policies For Compliant Manipulation (DIPCOM), a novel method for learning variable compliance control from demonstrations using diffusion models. Our approach predicts target EE poses and robot stiffness parameters conditioned on current observations, including the contact force data. A compliance controller uses these

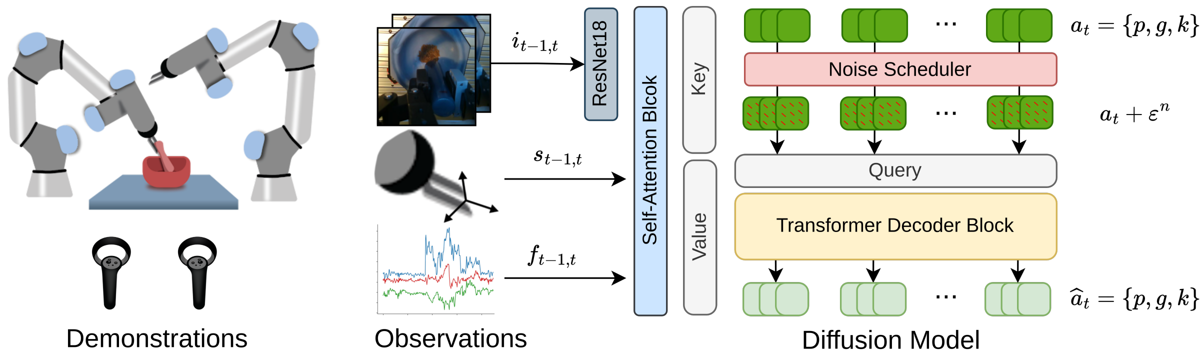


Fig. 1: **Policy Framework:** Left: Dataset collection framework. Middle: Observations $O$ include images $i_{t-1,t}$, robot Cartesian pose $s_{t-1,t}$, and measured force/torque $f_{t-1,t}$, all encoded using a self-attention transformer. Right: During training, actions $a_t$—comprising the end-effector pose $p$, gripper pose $g$, and stiffness $K$—are processed through a noise scheduler that adds Gaussian noise $\epsilon$ over time steps $n$. These noisy actions are then input into the transformer decoder block. During inference, Gaussian noise replaces the training noise, and the transformer decoder block predicts the actions $\hat{a}_t$

predictions to compute the final joint position commands that allow robots to move compliantly at the predicted stiffness. Fig. 1 illustrates the architecture of DIPCOM.

### A. Problem Formulation

Learning from demonstration (LfD) aims to enable robots to acquire new skills by autonomously observing and imitating human-provided demonstrations. In our approach, the policy learns to predict a sequence of absolute Cartesian EE pose and stiffness parameters given current observation $\mathcal{O}$, including RGB images $\mathcal{I} \in \mathbb{R}^{H\times W\times 3}$, the latest F/T sensor reading $\mathcal{F} \in \mathbb{R}^6$, and proprioception data $\mathcal{S} \in \mathbb{R}^9$. Demonstrations are collected using the VR teleoperation interface introduced in [4], which streams reference actions $\mathcal{A}$ and robot observations $\mathcal{O}$ during execution.

The predicted action $\mathcal{A} = \{\mathbf{p}, \mathbf{g}, \mathbf{k}\}$ comprises the absolute EE pose, the gripper action, and the stiffness parameter. The absolute EE pose, denoted as $\mathbf{p} = \{\mathbf{r}, \mathbf{o}\} \in \mathbb{R}^9$, includes a position vector $\mathbf{r} \in \mathbb{R}^3$ and an orientation vector $\mathbf{o} \in \mathbb{R}^6$. To handle discontinuities in the axis-angle representation, we use the 6D continuous orientation representation from Zhou et al. [24], which offers a smooth and unique encoding suitable for learning. The gripper action $\mathbf{g} \in \mathbb{R}^1$ represents the desired gripper width, and the stiffness parameter $\mathbf{k} \in \mathbb{R}^6$ corresponds to the diagonal of the stiffness matrix. Each action $\mathcal{A}$ is executed through a compliance controller, and during teleoperation, the operator can switch between two predefined stiffness modes using the VR grip button.

### B. Diffusion Policies For Compliant Manipulation

Our dataset is inherently multi-modal, containing diverse observations and corresponding actions. We aim to learn a policy distribution $\pi(\mathcal{A}|I, F, S)$ from task-specific demonstrations using Diffusion Policies For Compliant Manipulation, a classifier-free conditional diffusion model that generates actions $\mathcal{A}$ from observations $\mathcal{O}$.

The diffusion process consists of a forward stage that progressively adds noise and an inverse stage that denoises the data conditioned on observations. We adopt the Denoising Diffusion Implicit Model (DDIM) formulation [25] for deterministic denoising, enabling efficient inference with adjustable steps.

Diffusion Policies For Compliant Manipulationfollows the architecture of [21], employing an encoder-decoder transformer [26] with a ResNet18 vision backbone (without pretraining). Processed images, concatenated with force and robot state data, are input into the transformer encoder, while cross-attention is applied to noisy actions in the decoder.

The model is trained with mean squared error loss:

$$L_{\text{sample}} = \|a_t^0 - \hat{a}_t^0\|^2$$

During inference, the model estimates the original action iteratively:

$$a_t^{n-1} = \sqrt{\beta_{n-1}}\hat{a}_t^0 + \sqrt{1-\beta_{n-1}} \cdot \frac{a_t^n - \sqrt{\beta}\hat{a}_t^n}{\sqrt{1-\beta}}$$

where $a_t^n$ is the noisy data at step $n$, $\beta$ is the cumulative noise scale, and $\hat{a}_t^0$ is the estimated original action.

### C. Action Sequence Generation

Applying diffusion models to contact-rich manipulation is challenging due to prolonged surface interactions, where inconsistent action predictions can degrade performance. Prior works [21], [12] predict a fixed number of actions, discarding part of the horizon, which can cause jerky transitions—acceptable for simple tasks but problematic in contact-rich scenarios.

To improve stability, we predict at a higher frequency than previous methods, which typically generate 16 actions per horizon at 20 Hz. Our approach averages 48 actions per horizon, with some tasks requiring longer sequences, enabling finer control.

Diffusion over longer horizons can accumulate errors when only a portion of predictions is applied. To mitigate this, we extend the Temporal Ensemble method from [11], applying remaining actions to subsequent steps, smoothing

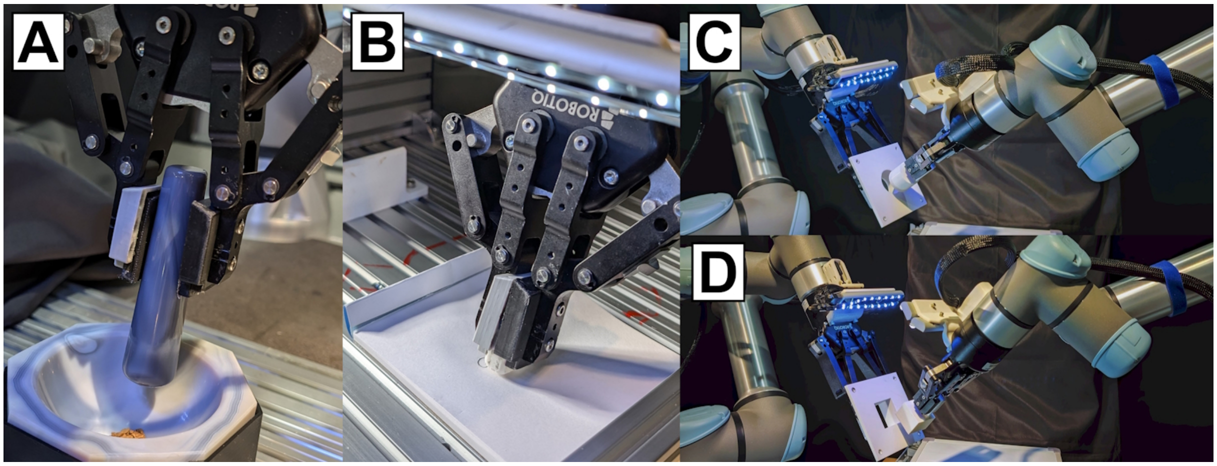


Fig. 2: Contact-rich manipulation tasks used for evaluation. A - Powder grinding. B - Pencil eraser. C - Bimanual round peg insertion. D - Bimanual cuboid peg insertion.

TABLE I: Task conditions

| Task | Task Conditions | | | | | | |
|---|---|---|---|---|---|---|---|
| | # of demos. | # of views | # of arms | Stiffness Modes | | | |
| | | | | Position | | Rotation | |
| | | | | low | high | low | high |
| **A** | 40 | 1 | 1 | 300 | 800 | 100 | 150 |
| **B** | 60 | 2 | 1 | 800 | 1200 | 150 | 300 |
| **C / D** | 60 | 3 | 2 | R1: 800<br>R2: 200 | 1200<br>800 | 150<br>100 | 300<br>150 |

transitions and ensuring stable performance in contact-rich tasks.

## IV. Experiments

We evaluate DIPCOM on four contact-rich manipulation tasks requiring precise force application, shown in Fig. 2. Our method is compared against Comp-ACT [4], a CVAE-based policy that uses a transformer encoder to process observations and a decoder to predict Cartesian EE poses and stiffness. Both policies use RGB images, F/T readings, and Cartesian EE poses for action prediction.

Experiments were performed on a dual-arm setup using UR5e robots with wrist-mounted F/T sensors. Each robot was equipped with an Intel RealSense SR305 camera, with a third static camera providing an external view. Demonstrations were collected from three co-authors using the teleoperation setup in [4] to ensure diverse motion strategies.

Each policy was trained per task, and task conditions are summarized in Table I. Results consistently show that DIPCOM outperforms Comp-ACT in fine-grained contact manipulation.

We now describe each task in detail and discuss the comparative performance of both methods.

- **A - Powder grinding:** The robot grinds powder using a ceramic pestle and pauses periodically to check the powder's state. Wrist camera images guide the grinding process. Success is measured by the percentage of fine powder produced (Table II). DIPCOM achieved 56% compared to Comp-ACT's 10%, despite similar applied force (Fig. 4). This improvement stems from DIPCOM's ability to reproduce circular grinding motions and evaluate intermediate outcomes (Fig. 3).
- **B - Pencil eraser:** The robot uses a rubber eraser to remove pencil marks from a notepad. Success is measured by full mark removal and percentage erased (Table III). DIPCOM achieved a 52.3% success rate, while Comp-ACT failed all 20 rollouts. As shown in Fig. 5, Comp-ACT applied insufficient force and suffered from poor alignment, while DIPCOM adapted its force and trajectory to better match the target.
- **C / D - Bimanual insertion tasks:** One robot arm holds a peg while the other positions the mating part. The task requires precise coordination for successful insertion and release (Table IV). While both methods had similar success rates, DIPCOM displayed more adaptive behavior, dynamically adjusting the pose of the supporting arm during insertion. This led to more robust executions and novel, effective strategies not seen in the original demonstrations.

TABLE II: A - Powder grinding results

| Percentage of fine powder produced | | |
|---|---|---|
| | Average | Standard Deviation |
| Human demonstrations | 76.67% | 8.4% |
| DIPCOM | **55.88%** | 13.54% |
| CompACT | 9.96% | 1.39% |

TABLE III: B - Pencil eraser task results

| Method | Percentage Erased Average (SD) | Success Rate |
|---|---|---|
| Comp-ACT | 26.0% (16.6%) | 0.0% |
| DIPCOM | **77.32% (19.48%)** | **52.3%** |

## V. Discussion

Both **DIPCOM** and **Comp-ACT** achieved similar success in bimanual insertion tasks but exhibited distinct behaviors, especially in long-horizon tasks like powder grinding and pencil erasure. These tasks require repetitive, adaptive

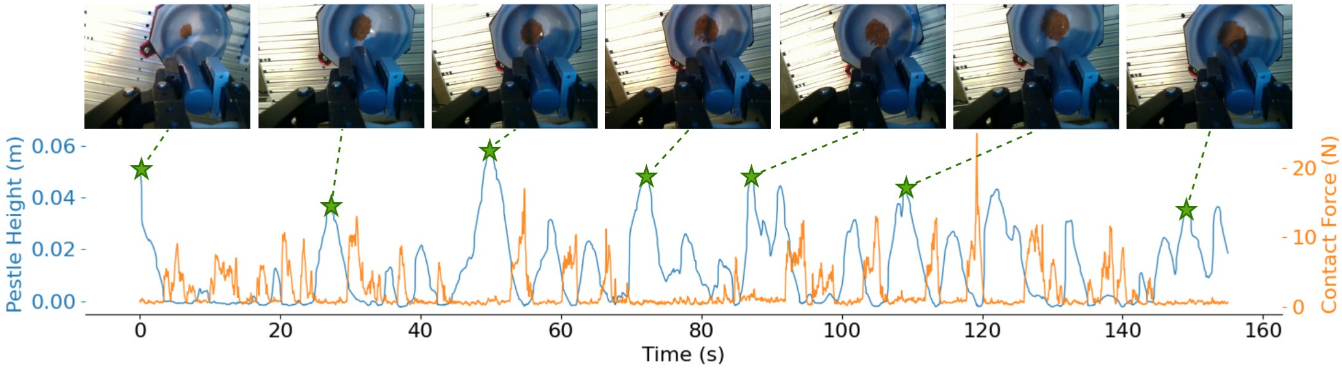


Fig. 3: Powder Grinding performance by the DIPCOM policy. The policy imitates the demonstrated behavior of pausing every few seconds to look at the powder's state before continuing the grinding process. Position Z indicates the height of the tip of the pestle relative to the mortar.

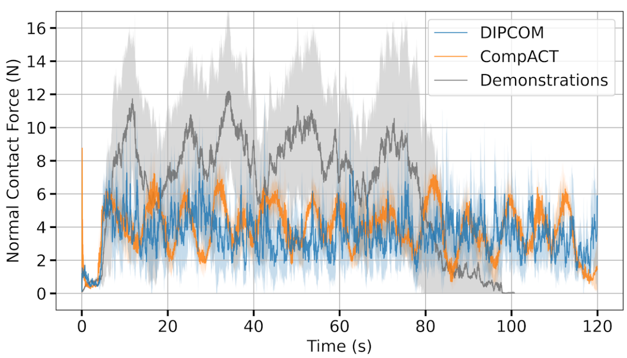


Fig. 4: A - Powder Grinding: Force profile comparison between the demonstrations, DIPCOM, and Comp-ACT. Bold lines and shaded areas represent the average and standard deviation of normal contact force.

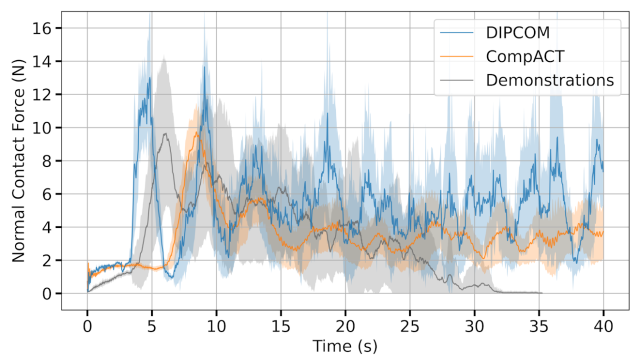


Fig. 5: B - Pencil Eraser: Force profile comparison between demonstrations, DIPCOM, and Comp-ACT.

TABLE IV: Success rate for bimanual insertion tasks

| Task Name | Comp-ACT | DIPCOM |
|---|---|---|
| C - Bimanual Round Insertion | 100% | 100% |
| D - Bimanual Cuboid Insertion | 95% | 95% |

actions rather than linear sequences. **Comp-ACT** initially performed well but struggled with maintaining fluid motion, often freezing mid-task during up-and-down erasing or circular grinding. In contrast, **DIPCOM** adapted more flexibly, sustaining smooth, continuous actions despite greater execution variance. These findings align with Jia et al. [5], reinforcing the behavioral distinctions between VAE-based and diffusion-based methods.

In force application, both policies applied force more conservatively than human demonstrators. While **Comp-ACT** was more consistent, it often fell into cyclic force patterns. **DIPCOM**, despite greater variance, better matched human force patterns due to its diffusion-based modeling.

*Limitations*

**DIPCOM** is more sensitive to hyperparameters than **Comp-ACT** and demands higher computational resources, making it susceptible to control frequency fluctuations. Future work will explore hyperparameter tuning and alternative action spaces, such as relative trajectories [12].

This study also used a relatively small number of demonstrations compared to similar works [21], [14]. Scaling up datasets could improve generalization, enabling a single policy to handle variations within a task—for example, learning a general powder grinding policy rather than training separately for each instance.

## VI. Conclusions

This work introduced **Diffusion Policies For Compliant Manipulation** (DIPCOM), a diffusion-based framework for compliant control in rigid robots. By leveraging diffusion models, our approach captures multimodal action distributions, predicting Cartesian end-effector poses while adjusting stiffness to regulate contact forces. We provided guidelines for diffusion-based compliant control and demonstrated DIPCOM's advantages over prior methods in contact-rich tasks.

Future directions include expanding datasets for improved generalization, refining force-processing mechanisms, and enhancing policy architectures for more precise force-aware inference.

## References


[1] H. Ravichandar, A. S. Polydoros, S. Chernova, and A. Billard, "Recent advances in robot learning from demonstration," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 3, pp. 297–330, 2020. [Online]. Available: https://doi.org/10.1146/annurev-control-100819-063206

[2] Y. Nakajima, M. Hamaya, Y. Suzuki, T. Hawai, F. von Drigalski, K. Tanaka, Y. Ushiku, and K. Ono, "Robotic powder grinding with a soft jig for laboratory automation in material science," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2022, pp. 2320–2326.

[3] A. Calanca, R. Muradore, and P. Fiorini, "A review of algorithms for compliant control of stiff and fixed-compliance robots," *IEEE/ASME Transactions on Mechatronics*, vol. 21, no. 2, pp. 613–624, 2015.

[4] T. Kamijo, C. C. Beltran-Hernandez, and M. Hamaya, "Learning variable compliance control from a few demonstrations for bimanual robot with haptic feedback teleoperation system," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2024.

[5] X. Jia, D. Blessing, X. Jiang, M. Reuss, A. Donat, R. Lioutikov, and G. Neumann, "Towards diverse behaviors: A benchmark for imitation learning with human demonstrations," in *International Conference on Learning Representations (ICLR)*, 2024. [Online]. Available: https://openreview.net/forum?id=6pPYRXKPpw

[6] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, "Robot programming by demonstration," in *Springer Handbook of Robotics*, 2008, pp. 1371–1394. [Online]. Available: https://doi.org/10.1007/978-3-540-30301-5_60

[7] A. M. Schmidts, D. Lee, and A. Peer, "Imitation learning of human grasping skills from motion and force data," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2011, pp. 1002–1007.

[8] P. Kormushev, S. Calinon, and D. G. Caldwell, "Imitation learning of positional and force skills demonstrated via kinesthetic teaching and haptic input," *Advanced Robotics*, vol. 25, no. 5, pp. 581–603, 2011.

[9] L. Rozo, P. Jiménez, and C. Torras, "A robot learning from demonstration framework to perform force-based manipulation tasks," *Intelligent Service Robotics*, vol. 6, no. 1, pp. 33–51, 2013.

[10] Y. Wang, C. C. Beltran-Hernandez, W. Wan, and K. Harada, "Robotic imitation of human assembly skills using hybrid trajectory and force learning," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 11 278–11 284.

[11] T. Z. Zhao, V. Kumar, S. Levine, and C. Finn, "Learning fine-grained bimanual manipulation with low-cost hardware," in *Robotics: Science and Systems (RSS)*, 2023. [Online]. Available: https://doi.org/10.15607/RSS.2023.XIX.016

[12] C. Chi, Z. Xu, C. Pan, E. Cousineau, B. Burchfiel, S. Feng, R. Tedrake, and S. Song, "Universal manipulation interface: In-the-wild robot teaching without in-the-wild robots," *arXiv preprint arXiv:2402.10329*, 2024.

[13] M. Drolet, S. Stepputtis, S. Kailas, A. Jain, J. Peters, S. Schaal, and H. Ben Amor, "A comparison of imitation learning algorithms for bimanual manipulation," *IEEE Robotics and Automation Letters (RA-L)*, vol. 9, no. 10, pp. 8579–8586, 2024.

[14] T. Z. Zhao, J. Tompson, D. Driess, P. Florence, S. K. S. Ghasemipour, C. Finn, and A. Wahid, "ALOHA Unleashed: A simple recipe for robot dexterity," in *Annual Conference on Robot Learning (CoRL)*, 2024.

[15] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 33, pp. 6840–6851, 2020.

[16] L. Yang, Z. Zhang, Y. Song, S. Hong, R. Xu, Y. Zhao, W. Zhang, B. Cui, and M.-H. Yang, "Diffusion models: A comprehensive survey of methods and applications," *ACM Computing Surveys*, vol. 56, no. 4, pp. 1–39, 2023.

[17] J. Carvalho, A. T. Le, M. Baierl, D. Koert, and J. Peters, "Motion planning diffusion: Learning and planning of robot motions with diffusion models," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2023, pp. 1916–1923.

[18] H. Ryu, J. Kim, H. An, J. Chang, J. Seo, T. Kim, Y. Kim, C. Hwang, J. Choi, and R. Horowitz, "Diffusion-edfs: Bi-equivariant denoising generative modeling on se (3) for visual robotic manipulation," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024, pp. 18 007–18 018.

[19] E. Ng, Z. Liu, and M. Kennedy, "Diffusion co-policy for synergistic human-robot collaborative tasks," *IEEE Robotics and Automation Letters (RA-L)*, vol. 9, no. 1, pp. 215–222, 2024.

[20] J. Urain, N. Funk, J. Peters, and G. Chalvatzaki, "Se (3)-diffusionfields: Learning smooth cost functions for joint grasp and motion optimization through diffusion," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 5923–5930.

[21] C. Chi, S. Feng, Y. Du, Z. Xu, E. Cousineau, B. Burchfiel, and S. Song, "Diffusion policy: Visuomotor policy learning via action diffusion," in *Robotics: Science and Systems (RSS)*, 2023. [Online]. Available: https://doi.org/10.15607/RSS.2023.XIX.026

[22] W. Liu, Y. Du, T. Hermans, S. Chernova, and C. Paxton, "Structdiffusion: Language-guided creation of physically-valid structures using unseen objects," in *Robotics: Science and Systems (RSS)*, 2023. [Online]. Available: https://doi.org/10.15607/RSS.2023.XIX.031

[23] M. Reuss, M. Li, X. Jia, and R. Lioutikov, "Goal-conditioned imitation learning using score-based diffusion policies," in *Robotics: Science and Systems (RSS)*, 2023. [Online]. Available: https://doi.org/10.15607/RSS.2023.XIX.028

[24] Y. Zhou, C. Barnes, J. Lu, J. Yang, and H. Li, "On the continuity of rotation representations in neural networks," in *IEEE/CVF conference on computer vision and pattern recognition (CVPR)*, 2019, pp. 5745–5753.

[25] J. Song, C. Meng, and S. Ermon, "Denoising diffusion implicit models," in *International Conference on Learning Representations (ICLR)*, 2021.

[26] A. Vaswani, "Attention is all you need," *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.