Sparse Diffusion Policy: A Sparse, Reusable, and Flexible Policy for Robot Learning

Anonymous Author(s)
Affiliation
Address
email

1 Introduction

The supplement includes five videos and one description document. This document provides a detailed description of the supplementary materials for the manuscript entitled "Sparse Diffusion Policy: A Sparse, Reusable, and Flexible Policy for Robot Learning." The primary materials contained are as follows:

- **Sparsity of SDP.mp4** This video demonstrates the activation of experts across different layers within our Sparse Diffusion Policy (SDP) during task execution. It is evident that only a select few experts are activated during inference, showcasing the **sparsity** of SDP. This feature highlights SDP's efficiency in utilizing a few number of components to effectively perform tasks.
- **Multitask - Simulation.mp4** This video presents examples of our model policy across eight tasks in MimicGen [1], demonstrating the superior performance of the Sparse Diffusion Policy (SDP) in multitask learning.
- **Multitask - FANUC Robotic Arm.mp4** This video showcases our model policy implemented in four tasks during real robot experiments, specifically using the FANUC LRMate 200iD/7L robotic arm. It demonstrates the superior performance of our SDP in real-world multitask learning.
- **Multitask - Comparison with TCD.mp4** This video presents a comparison between our framework, SDP, and the Task-Conditioned Diffusion (TCD) model in a real-world *Pick and Place* task experiment. It highlights the superior capability of SDP to preserve a multimodal policy capable of handling various types of tasks. Conversely, the baseline model struggles with completing all the tasks within a single model.
- **Task Transfer.mp4** This video corresponds to the experiment detailed in Section 4.3 of our study, illustrating the changes in selection probability among different experts when the Sparse Diffusion Policy (SDP) executes the *Coffee Preparation* task [1]. We observe that experts trained specifically on the *Coffee* task [1] are more likely to be selected when actions require information about the coffee machine. This finding provides valuable insights, suggesting that the combination of experts embodies distinct skills, and the router acts as a skill planner.

In the following sections, we provide detailed explanations for each video.

2 Sparsity of SDP

2.1 Sparsity of SDP.mp4

In this video, we display the activated experts in each layer of the Sparse Diffusion Policy (SDP) during task execution. We use an orange block to represent the top expert and a pink block for the second highest-ranked expert. For clear illustration, we connect these blocks with arrows. It is important to note that in actual implementation, the contributions of the selected experts in each layer are weighted and summed according to the router's weights (details are in Section 3.2 and Equation 2). Considering that semantic features typically emerge in the final few diffusion timesteps within the computer vision domain [2, 3], we have selected the last fifth timestep of the diffusion model to visualize the results. From this video, we can see during the inference, activated experts are sparse which demonstrates the **sparsity** of our model.

3 Multitask Learning

3.1 Multitask - Simulation.mp4

In this video, we visualize the performance of our Sparse Diffusion Policy (SDP) across eight tasks in MimicGen [1]: *Square*, *Stack*, *Coffee*, *Hammer*, *Mug*, *Nut*, *Stack Three*, and *Thread*. The video demonstrates that SDP is capable of performing effectively in diverse different tasks.

3.2 Multitask - FANUC Robotic Arm.mp4

In this video, we visualize the performance of our Sparse Diffusion Policy (SDP) across three tasks in real robot experiments: *Push*, *Pick and Place*, and *Hang*. The *Hang* task involves two distinct goals: hanging a cup on a lower stick and hanging it on a higher stick. We collected 20 demonstrations for both the *Push* and *Pick and Place* tasks, and 40 demonstrations for the *Hang* task (20 for each goal). Our results demonstrate that SDP achieves superior performance in these tasks.

3.3 Multitask - Comparison with TCD.mp4

In this video, we compare the performance of our Sparse Diffusion Policy (SDP) with Task-Conditioned Diffusion (TCD) [4, 5]. Our SDP exhibits stable and precise performance across diverse tasks. In contrast, TCD struggles to capture the multimodality of the task-specific policies and distinguish between different task behaviors, leading to failures. In detail, TCD always rotates the end effectors in the *Pick and Place* task, a behavior not observed in the demonstrations. Instead, rotations are necessary only for the *Hang* task. This observation indicates that TCD tends to conflate policies from different tasks and struggles to capture multimodal action distributions across diverse tasks.

4 Task Transfer

4.1 Task Transfer.mp4

This video illustrates the performance of our Sparse Diffusion Policy (SDP) in the *Coffee Preparation* task [1] and the changes in expert selection probability during task execution. We initially trained our SDP on the *Coffee* and *Mug Cleanup* tasks [1]. Subsequently, we froze all the experts and trained a new, lightweight router (comprising less than 0.4% of the total parameters) specifically for the *Coffee Preparation* task.

In the *Coffee* task, the robot is required to place the pod into the holder and close it. For the *Mug Cleanup*, the robot must open a drawer, place the mug inside, and close the drawer. The *Coffee Preparation* task involves placing the mug into the drip tray, opening the holder, and the drawer, placing the pod into the holder, and then closing it. This task can be seen as a composite of the *Coffee* and *Mug Cleanup* tasks, but it includes unique actions, such as moving the mug to the drip tray, which are not present in the initial training tasks.

SDP thus needs to utilize the **frozen** experts developed for previous two tasks, and learn task-specific routers combining these to acquire new skills (e.g., moving the mug to the drip tray). To better illustrate the router's functionality, we set eight experts per layer and selected only the top two experts for activation, enabling the use of entirely distinct experts for *Coffee* and *Mug Cleanup*, with Mutual Information loss (described in Section 3.3). From Figure 6 in the manuscript, we observe that the *Coffee* task primarily activates experts 0127, while the *Mug Cleanup* task activates experts 3456. It indicates that there are no shared experts in these two tasks. In the video, a light green area indicates the experts used for the *Coffee* task, and a light red area for the *Mug Cleanup*. According to the findings in [2, 3], We choose the last fifth diffusion timestep and the last layer of the transformer for visualization.

We observe that experts trained on the Coffee task are more likely to be selected when the task requires information about the coffee machine (not appear in the *Mug Cleanup* task). For instance, experts related to Coffee are activated (highlighted in red) when the action involves placing the pod. This observation underscores that the router acts as a skill planner, and experts function as the skills. The router can effectively composite the skills across the tasks and complete the complex and unseen tasks.

Interestingly, we find that **new skills** can be learned through the combination of experts from different tasks. For example, when the robot moves the mug to the coffee machine's drip tray—a new skill—it must ascertain the location of the coffee machine, prompting the activation of *Coffee*-related experts (highlighted in red). Another observation is that in most cases, experts from the *Mug Cleanup* are activated, showing that their combination is enough to address the majority of the *Coffee Preparation* task, except for actions that uniquely require information specific to the *Coffee* task. This demonstrates that such expert composition spans a broad range of skills, highlighting the powerful **expressive capability** of our SDP.

References

[1] A. Mandlekar, S. Nasiriany, B. Wen, I. Akinola, Y. Narang, L. Fan, Y. Zhu, and D. Fox. Mimicgen: A data generation system for scalable robot learning using human demonstrations. In *7th Annual Conference on Robot Learning*, 2023.

[2] L. Tang, M. Jia, Q. Wang, C. P. Phoo, and B. Hariharan. Emergent correspondence from image diffusion. *Advances in Neural Information Processing Systems*, 36:1363–1389, 2023.

[3] G. Luo, L. Dunlap, D. H. Park, A. Holynski, and T. Darrell. Diffusion hyperfeatures: Searching through time and space for semantic correspondence. *Advances in Neural Information Processing Systems*, 36, 2024.

[4] A. Ajay, Y. Du, A. Gupta, J. Tenenbaum, T. Jaakkola, and P. Agrawal. Is conditional generative modeling all you need for decision-making? *arXiv preprint arXiv:2211.15657*, 2022.

[5] Z. Liang, Y. Mu, H. Ma, M. Tomizuka, M. Ding, and P. Luo. Skilldiffuser: Interpretable hierarchical planning via skill abstractions in diffusion-based task execution. *arXiv preprint arXiv:2312.11598*, 2023.