Does Distance Heal Time?

Jacob Engelberg[1] **Preston Futaba**[1] **Eshanika Ray**[1] **Marvin Wong**[1]
[1]UCLA, United States
jengelberg@g.ucla.edu, gambitnf6@g.ucla.edu,
eshanika@g.ucla.edu, marvinwong25@g.ucla.edu

Abstract: Retrieval-augmented imitation learning improves data efficiency but is highly sensitive to temporal misalignment in human demonstrations. We replace Subsequence Dynamic Time Warping (SDTW) in the STRAP retrieval pipeline with a trajectory-level Optimal Transport (OT) similarity metric that ranks full offline trajectories using a hybrid feature and structural distance. Our OT-based retrieval reduces mean retrieval cost by over 93% relative to SDTW and retrieves longer, more geometrically coherent skill executions. On the long-horizon LIBERO stove-pot task, policies trained on OT-retrieved data achieve a 90% success rate compared to 70% for SDTW and exhibit significantly improved robustness under observation noise. These results demonstrate that geometric distributional matching provides a more reliable retrieval signal than temporal alignment for real-world teleoperation.

Keywords: Imitation Learning, Retrieval, Optimal Transport

1 Introduction

Retrieval-augmented imitation learning improves data efficiency by identifying relevant examples from large offline datasets. STRAP [1] performs such retrieval using Subsequence Dynamic Time Warping (SDTW), but SDTW's reliance on temporal alignment makes it sensitive to timing variability in human tele-operation, where pauses and inconsistent execution speeds often lead to unreliable matches.

We replace SDTW's subsequence search with a trajectory-level similarity metric based on optimal transport (OT) that compares segmented query behaviors against complete offline trajectories. Our hybrid metric combines feature-level OT with a structural consistency term that captures intra-trajectory geometric relationships, yielding a distance measure that reflects spatial coherence rather than temporal alignment.

Integrated into STRAP, this OT-based retrieval produces more stable similarity scores, retrieves trajectories with higher geometric coherence, and yields downstream behavior cloning policies with improved task performance. These results show that trajectory-level OT distance offers a more reliable retrieval signal than subsequence-based temporal alignment for real-world teleoperation.

2 Related Work

Dynamic Time Warping: Dynamic Time Warping (DTW) [2] is a classical method to align sequences by computing a monotone warping path. Although effective under consistent timing, DTW is highly sensitive to pauses, hesitations, and other irregularities common in human demonstrations.

Subsequence DTW for Retrieval: STRAP [1] extends DTW through Subsequence Dynamic Time Warping (SDTW), allowing retrieval of skill-level segments even when demonstrations differ in duration. However, SDTW still enforces temporal alignment and thus inherits DTW's sensitivity to timing variations.

Optimal Transport for Trajectory Similarity. Optimal Transport (OT) [3] measures distributional similarity rather than pointwise temporal correspondence, making it naturally robust to timing distortions. However, OT has seen limited use in demonstration retrieval pipelines.

OT-Based Retrieval: We replace SDTW's subsequence search with a trajectory-level OT similarity metric. Our hybrid distance combines feature-level OT with a structural consistency term that preserves pairwise geometric relationships. This avoids temporal alignment and computationally costly windowed search, enabling more stable retrieval under natural teleoperation variability.

3 Problem Formulation

We consider retrieval-augmented imitation learning in which a small set of task demonstrations $\mathcal{D}_{\text{task}} = \{\tau^{(i)}\}_{i=1}^{N}$ is used to extract relevant behaviors from a large offline dataset $\mathcal{D}_{\text{offline}} = \{\tau^{(j)}\}_{j=1}^{M}$. Each trajectory $\tau = (o_{1:T}, a_{1:T})$ is mapped into embeddings $E(\tau) = (e_1, \dots, e_T)$ with $e_t = \phi(o_t) \in \mathbb{R}^d$ for a pretrained encoder ϕ.

Each demonstration is segmented into contiguous sub-trajectories $\mathcal{S}(\tau) = \{Q_1, \dots, Q_K\}$, where a segment $Q = (q_1, \dots, q_m) \in \mathbb{R}^{m \times d}$ represents a coherent skill.

For retrieval, each query segment Q is compared against each full offline trajectory $Y^{(j)} = (y_1^{(j)}, \dots, y_{T_j}^{(j)}) \in \mathbb{R}^{T_j \times d}$ using a hybrid distance:

$$C(Q, Y^{(j)}) = \alpha\, C_{\text{feat}}(Q, Y^{(j)}) + (1-\alpha)\, C_{\text{struct}}(Q, Y^{(j)}), \tag{3.1}$$

where C_{feat} is an optimal-transport–based divergence between the empirical distributions of $\{q_i\}$ and $\{y_t^{(j)}\}$, and C_{struct} measures intra-trajectory geometric consistency using the pairwise distance matrices $C_Q = [\|q_i - q_{i'}\|_2]$ and $C_{Y^{(j)}} = [\|y_t^{(j)} - y_{t'}^{(j)}\|_2]$.

For each segment $Q \in \mathcal{S}(\tau)$ we rank all offline trajectories by $C(Q, Y^{(j)})$. A total retrieval budget k is allocated uniformly across segments:

$$\mathcal{R}(Q) = \text{Top-}\left\lceil \frac{k}{|\mathcal{S}(\tau)|} \right\rceil \{(j,\, C(Q, Y^{(j)})) \mid \tau^{(j)} \in \mathcal{D}_{\text{offline}}\}. \tag{3.2}$$

The final retrieved dataset is

$$\mathcal{D}_{\text{retrieved}} = \bigcup_{\tau \in \mathcal{D}_{\text{task}}} \bigcup_{Q \in \mathcal{S}(\tau)} \{\tau^{(j)} \mid (j, \cdot) \in \mathcal{R}(Q)\}. \tag{3.3}$$

Problem statement: Given segmented task demonstrations and a large offline dataset, the goal is to define a trajectory-level similarity metric that retrieves complete trajectories whose geometric structure is most semantically aligned with each query segment. The formulation must be robust to timing variability, avoiding explicit temporal alignment and enabling retrieval based on spatial and structural consistency.

4 Method

We integrate an optimal-transport–based trajectory similarity metric into the STRAP retrieval framework to address the limitations of temporal alignment methods under timing variability in human teleoperation. Our method retains STRAP's embedding and segmentation pipeline but replaces its subsequence matching module with a trajectory-level similarity computation based on distributional and structural consistency.

Retrieval Pipeline: All task and offline demonstrations are embedded using a pretrained vision encoder (DINOv2-base in our experiments). Following STRAP, each task demonstration is segmented into skill-level sub-trajectories using velocity-based change points with a minimum length of 20 frames. For each query segment, we compute its similarity to every *full* offline trajectory. Unlike

SDTW, which searches for aligned subsequences *within* trajectories, our approach compares query segments against *complete* trajectories, avoiding the computational overhead of windowed subsequence search and enabling distributional comparison of segment-level behavior. For each segment $Q \in \mathcal{S}(\tau)$, we rank all offline trajectories by distance and retrieve the top $\lceil k/|\mathcal{S}(\tau)|\rceil$ matches, where k is the total retrieval budget (we use $k = 100$ in our experiments). This allocates retrieval uniformly across all skill segments.

Hybrid Optimal Transport Distance: For a query segment $Q \in \mathbb{R}^{m\times d}$ and offline trajectory $Y \in \mathbb{R}^{T\times d}$, we define a hybrid distance

$$C(Q, Y) = \alpha\, C_{\text{feat}}(Q, Y) + (1-\alpha)\, C_{\text{struct}}(Q, Y), \tag{4.1}$$

with $\alpha = 0.5$. The feature-level distance C_{feat} is the Sinkhorn divergence, an entropic-regularized optimal transport distance computed via GPU-accelerated Sinkhorn iterations [4] with Euclidean ground cost ($p = 2$) and blur parameter $\varepsilon = 0.05$. This quantifies similarity between embedding distributions without requiring temporal alignment, naturally handling pauses and execution-speed variation. The structural term preserves geometric coherence by comparing mean intra-trajectory pairwise distances:

$$C_{\text{struct}}(Q, Y) = \left| \frac{1}{m^2} \sum_{i,i'} \|q_i - q_{i'}\|_2 - \frac{1}{T^2} \sum_{j,j'} \|y_j - y_{j'}\|_2 \right|. \tag{4.2}$$

This efficiently captures pairwise geometric structure without the cost of full Gromov–Wasserstein optimization.

Implementation: All computations are performed on GPU using PyTorch tensors. Sinkhorn divergence is evaluated in batches using the GeomLoss library, enabling efficient computation of $C(Q, Y)$ across the entire offline dataset. Retrieval is obtained by sorting the resulting distance values for each query segment.

Training Data Construction: Retrieved trajectories are stored in an HDF5 dataset compatible with `robomimic`. For each retrieved element, we save the full trajectory and associated metadata (retrieval cost, trajectory identifier, and source file path). The final dataset is the union of all retrieved trajectories across all segments and all task demonstrations.

Policy Training: We train a Behavior Cloning Transformer (BCT) policy on the retrieved dataset using `robomimic` with frame stacking of 5 observations and action chunking of 5 steps, matching STRAP's training configuration. An otherwise identical BCT policy is trained on a dataset retrieved using STRAP's SDTW module with the same query demonstrations, retrieval budget, architecture, hyperparameters, optimizer settings, and random seed. Because the policies differ only in the retrieval method, any differences in downstream performance directly reflect differences in retrieval quality. Appendix F provides retrieval distribution plots which further illustrate how OT consistently retrieves longer and more centralized skill segments compared to SDTW.

5 Experiments

We evaluate our trajectory-level optimal transport (OT) retrieval method through (1) offline retrieval analysis and (2) downstream policy learning on a long-horizon manipulation task. All experiments follow a paired evaluation protocol to isolate the effect of the retrieval strategy.

Dataset and Task: We use the LIBERO benchmark [5]. Our primary evaluation task is the LIBERO-10 long-horizon task *"turn on the stove and put the moka pot on it"* ("stove-pot"), also used in the original STRAP evaluation. We use 5 demonstrations from LIBERO-10 as task queries and retrieve from the LIBERO-90 offline dataset. All demonstrations are embedded with DINOv2-base and segmented as described in Section 4. The total retrieval budget is fixed to $k = 100$ trajectories, distributed uniformly across query segments.

Baselines: We compare against STRAP's original Subsequence Dynamic Time Warping (SDTW) retrieval method. Both OT and SDTW use identical query demonstrations, segmentation procedures,

Table 1: Retrieval quality metrics comparing SDTW and OT on identical query segments. OT achieves substantially lower and more stable retrieval costs, indicating better geometric consistency, but retrieves from fewer unique source trajectories.

Method	Mean Cost	Std Cost	Unique Sources	Entropy
SDTW	189.46	137.84	49 (50.0%)	3.63
OT	11.43	3.37	33 (33.7%)	3.21

embedding models, and retrieval budgets. The only difference is the similarity metric used for ranking offline trajectories.

Policy Training: For each retrieval method, we train a Behavior Cloning Transformer (BCT) policy using the `robomimic` framework. Both policies share identical network architectures, optimizer settings, training epochs, and hyperparameters, with frame stacking of 5 observations and action chunking of 5 steps. All runs use fixed random seeds to eliminate stochastic variation.

Evaluation Protocol: Each trained policy is evaluated over 50 rollout episodes with fixed random seeds for environment initialization, horizon limits, NumPy, PyTorch, and simulator stochasticity. Offscreen rendering is disabled for computational efficiency. This paired setup ensures both OT- and SDTW-trained policies are tested under identical conditions.

Robustness Tests: We evaluate under two perturbation settings: (i) *Observation noise*, where zero-mean Gaussian noise with $\sigma = 0.02$ is injected into proprioceptive observations at each timestep (vision inputs remain unperturbed); and (ii) *Layout perturbations*, where object placement bounds are expanded by $\delta = 0.1$m for all objects. Identical paired random seeds are used across methods for both tests.

Evaluation Metrics: For retrieval analysis, we report mean and variance of retrieval costs, retrieval diversity via unique source trajectories and entropy, consistency between OT and SDTW retrievals, and temporal extent differences (start index, end index, and segment length). For downstream control, we report success rate (primary metric), average episode length, and cumulative return using `robomimic`'s standardized evaluation utilities. Robustness is assessed by success-rate degradation under perturbations.

6 Results and Discussion

6.1 Retrieval Quality Analysis

We first analyze offline retrieval behavior to understand how OT and SDTW differ in selecting demonstrations from the offline dataset.

Retrieval Cost: OT achieves dramatically lower and more stable retrieval costs than SDTW (Table 1). The mean OT cost is 11.43 (std=3.37) compared to SDTW's 189.46 (std=137.84), a 93.97% reduction. OT yields lower cost than SDTW on 94.90% of queries, with a mean paired difference of -178.02. This confirms that OT consistently identifies trajectories with stronger geometric correspondence to the query segments.

Retrieval Diversity: SDTW retrieves from 49 unique source demonstrations (50.0%), while OT retrieves from 33 (33.7%), with entropies of 3.63 and 3.21 respectively. This indicates that OT concentrates retrieval on a smaller subset of higher-quality trajectories. Only 5.10% of retrievals match exactly between methods, showing that OT and SDTW select fundamentally different behavioral examples.

Temporal Extent: OT retrieves significantly longer segments than SDTW, with a mean length increase of 90.5 frames (std=60.6). OT segments begin earlier (mean shift: -64.7 frames) and terminate later (mean shift: $+25.8$ frames), indicating that OT recovers more complete skill executions while SDTW fragments behavior into shorter subsequences. This structural difference directly explains the improved coherence observed in OT-trained policies.

6.2 Policy Performance

Table 2: Quantitative comparison of SDTW and OT policies. OT achieves higher success and lower horizons in baseline and noisy settings, and remains more reliable under layout perturbations.

Method	σ	δ (m)	Success	Horizon	Return
SDTW	0.00	0.0	0.70	395.5	0.70
OT	0.00	0.0	0.90	306.1	0.90
SDTW	0.02	0.0	0.50	481.9	0.50
OT	0.02	0.0	0.85	319.7	0.85
SDTW	0.00	0.1	0.76	265.9	0.76
OT	0.00	0.1	0.82	266.2	0.82

Baseline. Under noise-free conditions, OT outperforms SDTW across all metrics, achieving a 90% success rate vs. 70% and reducing average horizon from 395.5 to 306.1 steps. This indicates faster, more decisive task execution from OT-derived demonstrations.

Observation Noise. With Gaussian noise ($\sigma = 0.02$), SDTW degrades sharply (success $0.70 \rightarrow 0.50$, horizon $395.5 \rightarrow 481.9$), while OT remains stable ($0.90 \rightarrow 0.85$, $306.1 \rightarrow 319.7$). This confirms that OT retrieval produces smoother and more geometrically consistent training signals that generalize under sensor uncertainty.

Layout Perturbation. Under spatial shifts ($\delta = 0.1$m), OT drops moderately ($0.90 \rightarrow 0.82$), as expected for geometric matching. SDTW improves anomalously ($0.70 \rightarrow 0.76$), with both methods converging to $\sim$266 steps. This suggests SDTW's temporal elasticity may intermittently compensate for spatial shifts, though without overall stability.

Qualitative Behavior. OT rollouts remain smooth and confident under noise, while SDTW exhibits twitchy, unstable corrections that amplify small misalignments learned during subsequence matching. The unexpected SDTW improvement likely stems from seed-specific object proximity, temporal elasticity, and stochastic suppression of inefficient dithering.

7 Conclusion

We replaced Subsequence Dynamic Time Warping (SDTW) with Optimal Transport (OT) within the STRAP retrieval pipeline, shifting retrieval from temporal alignment to geometric structural matching. OT-based retrieval significantly improves demonstration matching quality, reducing mean retrieval cost from 189 to 11 and recovering fuller, more geometrically consistent executions.

On the LIBERO benchmark, the OT-trained policy outperforms SDTW under standard conditions (90% vs. 70% success) and exhibits strong robustness to observation noise, maintaining smooth and stable execution where SDTW becomes erratic. This confirms that geometrically aligned demonstrations provide a more consistent and reliable training signal under timing variability.

We observe a clear trade-off under spatial perturbations: OT degrades modestly (90% $\rightarrow$ 82%), reflecting sensitivity to geometric shifts, while SDTW improves unexpectedly due to its temporal elasticity. This suggests that temporal flexibility may offer resilience to spatial variation, albeit with reduced overall stability.

Future work will evaluate both methods across a broader set of long-horizon LIBERO tasks and explore hybrid retrieval schemes that combine OT's geometric stability with SDTW's temporal elasticity. We further plan to study OT's limits under large out-of-distribution spatial shifts.

References

[1] M. Memmel, J. Berg, B. Chen, A. Gupta, and J. Francis. Strap: Robot sub-trajectory retrieval for augmented policy learning, 2025. URL https://arxiv.org/abs/2412.15182.

[2] M. Müller. Dynamic time warping. *Information Retrieval for Music and Motion*, 2:69–84, 2007. doi:10.1007/978-3-540-74048-3_4.

[3] C. Villani. *Optimal Transport: Old and New*, volume 338 of *Grundlehren der mathematischen Wissenschaften*. Springer, 2009. doi:10.1007/978-3-540-71050-9.

[4] J. Feydy, T. Séjourné, F.-X. Vialard, S.-i. Amari, A. Trouvé, and G. Peyré. Interpolating between optimal transport and MMD using sinkhorn divergences. In *Proceedings of AISTATS*, PMLR, pages 2681–2690, 2019. URL https://proceedings.mlr.press/v89/feydy19a.html.

[5] B. Liu, Y. Zhu, C. Gao, Y. Feng, Q. Liu, Y. Zhu, and P. Stone. Libero: Benchmarking knowledge transfer for lifelong robot learning, 2023. URL https://arxiv.org/abs/2306.03310.

Appendix A Retrieval Cost Distribution Comparison

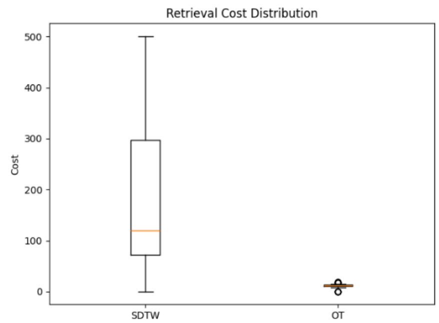

Figure 1: Box plot comparing retrieval costs. OT (right) shows dramatically lower median and tighter variance compared to SDTW (left), confirming superior geometric consistency.

Appendix B Cost Difference Distribution

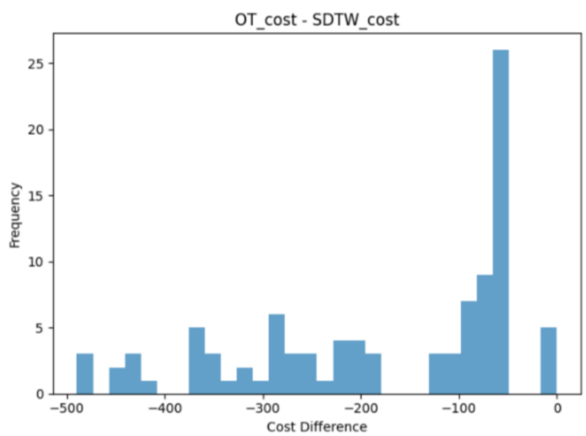

Figure 2: Histogram of pairwise cost differences (OT - SDTW). The heavily negative-skewed distribution confirms OT produces lower costs on 94.90% of queries.

Appendix C Top Retrieval Sources - OT

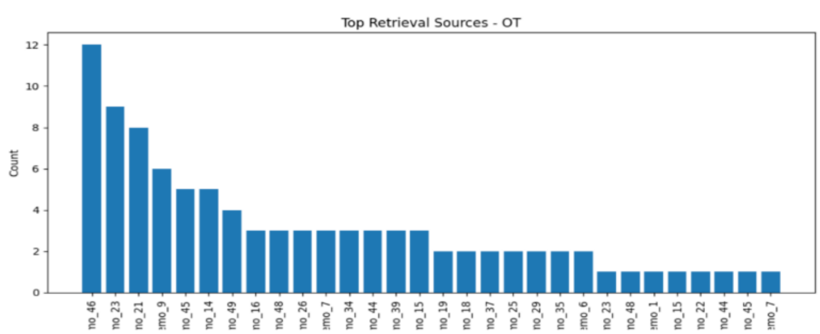

Figure 3: Retrieval frequency across source trajectories for OT. The method concentrates on fewer, higher-quality sources (33 unique, entropy 3.21).

Appendix D   Top Retrieval Sources - SDTW

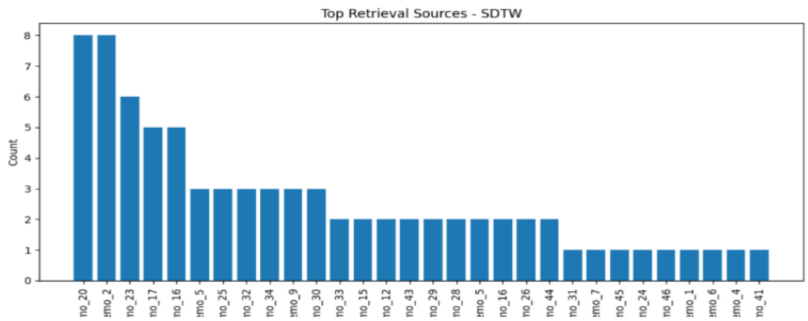

Figure 4: Retrieval frequency for SDTW. More uniform distribution across sources (49 unique, entropy 3.63) but with higher overall costs.

Appendix E   Entropy Comparison

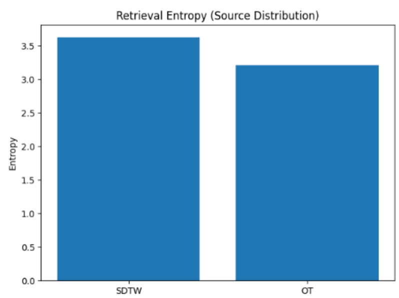

Figure 5: Direct entropy comparison. SDTW (left, 3.63) samples more broadly; OT (right, 3.21) concentrates on higher-quality trajectories.

Appendix F   Temporal Extent Analysis

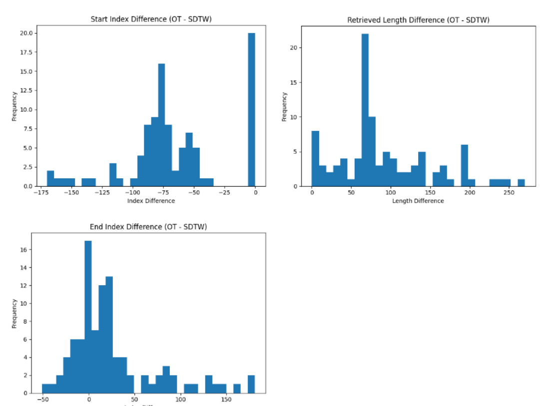

Figure 6: Temporal extent differences (OT - SDTW). Left: Start index (OT begins 64.7 frames earlier). Top right: Retrieved length (OT segments 90.5 frames longer). Bottom: End index (OT terminates 25.8 frames later). These distributions confirm OT recovers more complete skill executions.

Appendix G Team Contributions

Jacob Engelberg: Jacob researched retrieval metrics and comparison techniques used in the original STRAP paper to establish the criteria we should use for sub-trajectory quality. He worked with Eshanika to design, calculate, and analyze metrics for the SDTW and OT retrieval comparison. Jacob also contributed significantly to the presentation slide deck by organizing the content and making sure the results were clearly explained.

Preston Futaba: Preston handled the project's setup and data preparation workflow after the demonstration HDF5s were produced. He configured Robosuite and LIBERO Mini, ensured that the datasets loaded correctly, and validated sub-trajectory segmentation through visualization and consistency checks. In addition to environment setup, he implemented the Gaussian noise augmentation used for the robustness experiments, enabling controlled temporal and observational perturbations during evaluation. Preston also identified and resolved several issues within the STRAP directory, including fixing pathing and data-handling bugs that interfered with retrieval experiments.

Eshanika Ray: Eshanika developed the formulation of the optimal transport (OT) based retrieval method and implemented the full OT retrieval pipeline, including feature-level and structural cost computation, Sinkhorn divergence integration, and retrieval ranking. She designed and computed all retrieval quality and policy evaluation metrics, including cost statistics, entropy, diversity, and temporal extent measures. Eshanika generated the full HDF5 retrieval datasets used for both SDTW and OT policy training, and collaborated closely with Jacob on the statistical analysis and interpretation of the retrieval results. She also contributed to shaping the experimental narrative, results discussion, and overall presentation of the project.

Marvin Wong: Marvin implemented the STRAP baseline by adapting and executing the original STRAP GitHub codebase. He generated the baseline HDF5 datasets used for SDTW-based retrieval and downstream policy training. Marvin collaborated closely with Preston on policy training and evaluation within the LIBERO environment, and was primarily responsible for implementing and analyzing the layout perturbation experiments used to assess spatial robustness. He also contributed to the project presentation and report writing.