

SPREAD: Subspace Representation Distillation for Lifelong Imitation Learning

Kaushik Roy¹, Giovanni D’urso¹, Nicholas Lawrance¹, Brendan Tidd¹, Peyman Moghadam^{1,2}

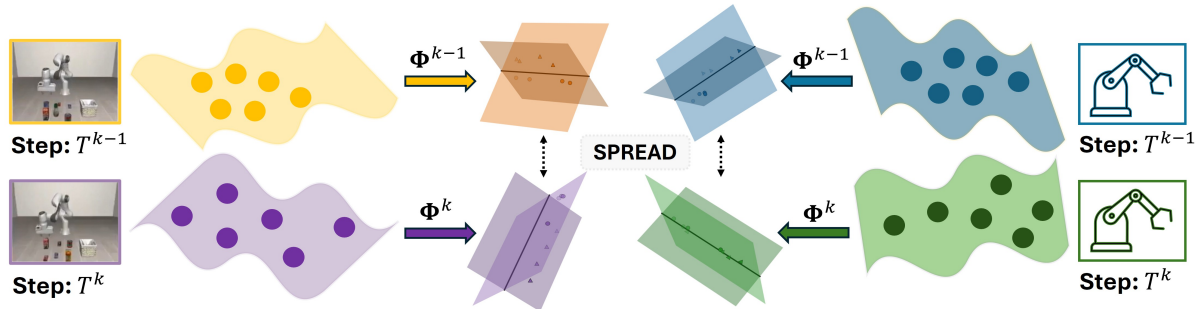


Fig. 1: Geometrical interpretation of the Subspace Representation Distillation loss. *SPREAD* maximizes similarity between consecutive LIL policies by minimizing discrepancies in projected feature representations across different input modalities within each policy’s subspace.

Abstract—A key challenge in lifelong imitation learning (LIL) is enabling agents to acquire new skills from expert demonstrations while retaining prior knowledge. This requires preserving the low-dimensional manifolds and geometric structures that underlie task representations across sequential learning. Existing distillation methods, which rely on L_2 -norm feature matching in raw feature space, are sensitive to noise and high-dimensional variability, often failing to preserve intrinsic task manifolds. To address this, we introduce *SPREAD*, a geometry-preserving framework that employs singular value decomposition (SVD) to align policy representations across tasks within low-rank subspaces. This alignment maintains the underlying geometry of multimodal features, facilitating stable transfer, robustness, and generalization. Additionally, we propose a confidence-guided distillation strategy that applies a Kullback–Leibler divergence loss restricted to the top- M most confident action samples, emphasizing reliable modes and improving optimization stability. Experiments on the LIBERO, lifelong imitation learning benchmark, show that *SPREAD* substantially improves knowledge transfer, mitigates catastrophic forgetting, and achieves state-of-the-art performance.

I. INTRODUCTION

Robotic agents are increasingly required to operate in open-world environments where new tasks arise sequentially and previously acquired skills must be retained over long horizons. While imitation learning [1]–[5] has demonstrated strong performance in teaching robots complex behaviors from demonstrations, its standard formulations assume a fixed task distribution and thus fail under continual task acquisition. When exposed to sequential learning, policy networks suffer from catastrophic forgetting [6], [7], where adapting to new demonstrations degrades representations

required for prior skills. Addressing this challenge is central to enabling scalable and reliable lifelong imitation learning.

A variety of strategies have been proposed to alleviate catastrophic forgetting in lifelong imitation learning. Experience Replay (ER) [8] addresses forgetting by replaying samples from previous tasks, but imbalanced data distributions often shift the representation space toward the current task, degrading performance on earlier tasks. Hierarchical skill learning methods such as LOTUS [9] enhance scalability through a dynamic skill library, but remain vulnerable to degradation of past knowledge. Distillation-based approaches, such as M2Distill [10], transfer multimodal knowledge between policies; however, their reliance on L_2 feature alignment in raw feature space neglects the underlying geometric structure of task representations. This limitation can lead to overly rigid alignment that impedes adaptation. These shortcomings highlight the need for subspace-preserving, geometry-aware distillation strategies to enable robust and scalable lifelong imitation learning.

To this end, we propose *SPREAD*, a lifelong imitation learning framework that explicitly aligns the low-rank subspace representations of consecutive policies. By enforcing consistency in how features are projected onto dominant subspaces, our approach preserves the intrinsic directions that define task manifolds while leaving orthogonal directions available for novel skill acquisition (see Fig. 1). The motivation for subspace alignment arises from the observation that neural representations are often concentrated in low-dimensional subspaces within high-dimensional feature spaces. Prior work in representation learning has demonstrated that these subspaces capture the most discriminative and transferable structure of task-relevant features [11], while also exhibiting robustness to noise [12]. Leveraging this property allows *SPREAD* to distill knowledge more

¹ CSIRO Robotics, DATA61, CSIRO, Australia. E-mails: firstname.lastname@csiro.au

² Queensland University of Technology (QUT), Brisbane, Australia. E-mails: firstname.lastname@qut.edu.au

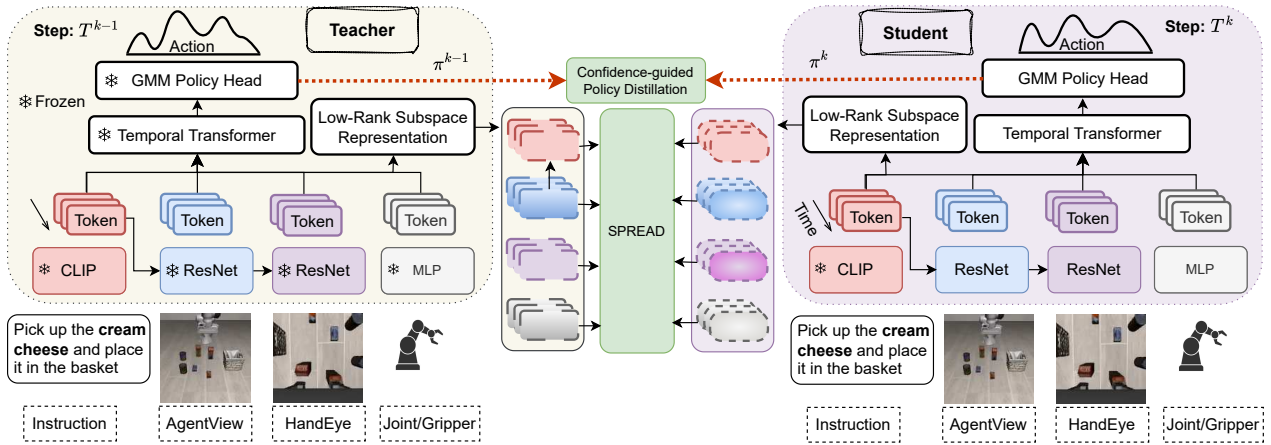


Fig. 2: Overview of our proposed *SPREAD* method. Subspace Representation Distillation aligns the latent representations from different input modality encoders (e.g., Task, AgentView, HandEye, Joint, and Gripper information), while confidence-guided policy distillation maps the action distribution of the GMM policy between incremental steps T^{k-1} and T^k .

effectively, avoiding over-constraining raw activations while maintaining the geometry critical for skill retention. This geometry-aware distillation provides a principled balance between stability and plasticity: the representational essence of past tasks is retained, yet sufficient capacity remains to encode new skills without interference. Complementing this, our confidence-guided policy distillation applies Kullback-Leibler (KL) divergence [13] to top- K action samples, ensuring robust behavioral transfer.

We evaluate *SPREAD* on diverse robotic manipulation benchmarks under sequential learning settings. Results show that *SPREAD* significantly outperforms conventional feature-level distillation and state-of-the-art lifelong imitation learning baselines. Beyond quantitative gains, our analysis reveals that *SPREAD* maintains the intrinsic manifold structure of task features, thereby yielding more robust transfer and retention across tasks.

In summary, this work makes the following contributions:

- We introduce *SPREAD*, a novel framework that preserves the low-dimensional subspace geometry of task features in lifelong imitation learning. We provide theoretical justification for why subspace-level alignment better preserves intrinsic task manifolds compared to feature-level distillation.
- We present a confidence-guided policy distillation strategy, using KL divergence on top- K actions to enhance behavioral robustness.
- We demonstrate through extensive experiments that *SPREAD* achieves state-of-the-art performance in mitigating catastrophic forgetting, while enabling efficient adaptation to new robotic skills.

II. RELATED WORK

Imitation learning (IL) enables robots to replicate expert behaviors by mapping observations to actions [1]–[3], [14]. Extending IL to sequential settings, Lifelong Imitation Learning (LIL) aims to continuously acquire new skills while retaining past ones, thereby addressing catastrophic

forgetting [6], [7]. Existing LIL approaches draw inspiration from lifelong learning in vision [15], [16] and language [17], adapting regularization [18], [19], memory-replay [6], [20]–[22], and dynamic architecture [23], [24] based strategies to mitigate catastrophic forgetting. Experience Replay (ER) [8] retains a limited subset of trajectories from past tasks and replays them alongside new demonstrations, but imbalanced replay distributions can bias the learned policy toward more recent skills. Similarly, CRIL [22] leverages deep generative replay (DGR) [25] to generate pseudo-demonstrations through action-conditioned video prediction, however, the instability and inaccuracy of generated videos degrade performance on complex LIL tasks. EWC [6] uses a Fisher information-based measure to constrain updates on parameters essential for earlier tasks to balance plasticity and stability, but struggles to maintain consistent performance in LIL. By decomposing complex behaviors into modular skills, BUDS [26] enables effective skill reuse through autonomous segmentation of demonstration trajectories. LOTUS [9] advances this paradigm by coupling unsupervised skill discovery with open-vocabulary vision models to build a continually expanding skill library. Despite their scalability benefits, both BUDS and LOTUS face challenges in maintaining efficiency and representational consistency as the skill library scales. Tokenization-based frameworks such as T2S [27] introduce parameter-efficient skill tokens that extend transformer architectures without uncontrolled parameter growth, while SPECI [28] adopts a prompt-driven skill codebook for efficient skill reuse and adaptation. These approaches improve scalability but often depend on strong architectural assumptions. Knowledge distillation strategies like PolyTask [29] consolidates task-specific policies into a single generalist model. M2Distill [10] employs multi-modal distillation to align visual, textual, and motor features between old and new policies, thereby reducing modality drift. However, most distillation-based approaches rely on L_2 matching in raw feature space, neglecting the intrinsic low-dimensional manifold structure of task features. Our work

addresses this gap by introducing *SPREAD*, a subspace representation distillation framework that explicitly respects the geometry of task representations.

III. PROPOSED METHODOLOGY

A. Problem Formulation

In lifelong imitation learning (LIL), a robot sequentially learns a policy π over a sequence of tasks $\{T^1, \dots, T^K\}$ to mimic expert demonstrations while retaining and transferring knowledge across tasks. Each task $T^k = (\mu_0^k, g^k, l^k)$ is defined by an initial state distribution μ_0^k , a goal predicate $g^k : \mathcal{S} \rightarrow \{0, 1\}$, and a language description l^k , within a shared Markov Decision Process $\mathcal{M} = (\mathcal{S}, \mathcal{A}, \mathcal{T}, H)$, where \mathcal{S} is the state space, \mathcal{A} is the action space, $\mathcal{T} : \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{S}$ is the transition function, and H is the episode horizon. For each task T^k , the robot receives N expert demonstration trajectories $D^k = \{\tau_i^k\}_{i=1}^N$, along with an associated language description l^k . Each demonstration trajectory τ_i^k consists of a sequence of state-action pairs $\{(s_t, a_t)\}_{t=1}^{L^k}$, where $L^k < H$ and H denotes the episode horizon. These demonstrations are typically collected via teleoperation and serve as supervision for training a policy π . The policy is optimized through behavioral cloning [30], which minimizes the negative log-likelihood of expert actions given the corresponding states:

$$\min_{\pi} J(\pi) = \frac{1}{K} \sum_{k=1}^K \mathbb{E}_{s_t, a_t \sim D^k} \left[\sum_{t=1}^{L^k} -\log \pi(a_t | s_{\leq t}; T^k) \right]. \quad (1)$$

A central challenge in this setting is that the robot, while progressing through tasks sequentially, no longer has direct access to demonstrations from earlier tasks. Consequently, the policy must not only sequentially adapt to new tasks but also preserve knowledge from prior tasks, thereby mitigating catastrophic forgetting and enabling effective transfer of skills across the task sequence.

In lifelong imitation learning (LIL), preserving the intrinsic low-dimensional manifold and geometric structure of task features across sequential tasks is critical to mitigate catastrophic forgetting while enabling adaptation to new tasks. Unlike conventional L2-norm-based distillation, which is sensitive to noise and high-dimensional variations, we propose a *subspace representation distillation* loss that aligns low-rank subspace projections of teacher and student features. This approach leverages singular value decomposition (SVD) [31] to capture the principal geometric structure of multimodal features, ensuring robust knowledge transfer. Additionally, for policy distillation, we introduce a confidence-guided Kullback-Leibler (KL) divergence loss that prioritizes the top K samples with high log probability scores, enhancing stability by focusing on reliable action distributions. Fig. 2 shows the complete architecture of the proposed *SPREAD* framework.

B. Subspace Representation Distillation

Knowledge distillation [32] aims to transfer learned representations from a teacher model to a student model, preserv-

ing performance while supporting efficient adaptation. We introduce *Subspace Representation Distillation (SPREAD)*, a geometry-aware distillation method that aligns the principal subspaces of feature representations between teacher and student. Given a feature matrix $f \in \mathbb{R}^{m \times n}$, we compute its reduced SVD [33]: $f = U \Sigma V^T$, where $U \in \mathbb{R}^{m \times r}$ contains the top- r left singular vectors spanning the dominant feature subspace. The projection of f onto this r -dimensional subspace is given by $U U^T f$, which captures the most informative directions in the representation.

For teacher and student feature matrices f_t and f_s , *SPREAD* minimizes the discrepancy between their subspace-projected features using the Frobenius norm:

$$\mathcal{L}_{\text{SPREAD}} = \|U_t U_t^T f_t - U_s U_s^T f_s\|_F^2 + \|U_t U_t^T f_s - U_s U_s^T f_t\|_F^2, \quad (2)$$

where U_t and U_s are the top- r left singular vectors of f_t and f_s , respectively. This symmetric formulation encourages both alignment of the subspace bases (*i.e.*, $U_s \approx U_t$) and consistency of feature content within those subspaces. Although subspace representation distillation introduces an $\mathcal{O}(mnr)$ computational overhead, the overhead is limited in practice, and the resulting performance improvements justify the computation.

Theoretical Significance. *SPREAD* leverages the geometric structure of feature representations by focusing on the principal subspaces, which encapsulate the dominant patterns in the data. The projection $U U^T f$ retains the most informative components of f , as determined by the largest singular values in Σ . By minimizing $\mathcal{L}_{\text{SPREAD}}$, the student model (at step k) aligns its feature subspace with that of the teacher (at step $k-1$), ensuring the preservation of critical task-relevant structures. This approach is robust to noise and model-specific artifacts, unlike direct feature matching (*e.g.*, $\|f_t - f_s\|_F^2$), as it prioritizes the principal directions of variation. The *SPREAD* loss further optimizes both the alignment of subspace bases ($U_s \approx U_t$) and the content within those subspaces ($U_s U_s^T f_s \approx U_t U_t^T f_t$). This can be interpreted as minimizing the principal angles between the column spaces of f_t and f_s , providing a geometrically sound framework for knowledge transfer. Moreover, *SPREAD* is invariant to differences in feature magnitude or dimensionality, making it particularly effective for distillation across incremental steps in lifelong imitation learning.

In incremental learning, at step k , the frozen model from step $k-1$ acts as the teacher, and the model at step k learning task T^k is the student. For a batch of size N and sequence length L , each modality’s latent tensor is reshaped into a $D \times (NL)$ matrix by flattening the batch and temporal dimensions, yielding compact feature matrices for distillation.

Visual Modalities (wrist-mounted (HandEye) and overhead (AgentView) camera). Let $F^{k-1, \epsilon}, F^{k, \epsilon} \in \mathbb{R}^{D \times (NL)}$ denote the reshaped feature matrices extracted using Resnet18 [34] for modality $\epsilon \in \{\text{HandEye}, \text{AgentView}\}$ at steps $k-1$ and k , respectively. The image modality

distillation loss is:

$$\mathcal{L}_{\text{image}} = \sum_{\epsilon \in \{\text{HandEye}, \text{AgentView}\}} \mathcal{L}_{\text{SPREAD}}(F^{k-1, \epsilon}, F^{k, \epsilon}). \quad (3)$$

Language (Text). For text features $G^{k-1}, G^k \in \mathbb{R}^{D \times (NL)}$ extracted using CLIP [35] language encoder and subsequently passed through an MLP at steps $k-1$ and k , we apply subspace representation distillation analogously:

$$\mathcal{L}_{\text{text}} = \mathcal{L}_{\text{SPREAD}}(G^{k-1}, G^k). \quad (4)$$

Extra Modalities (e.g., Joints, Gripper). For proprioceptive signals such as joint angles and gripper state, let $H^{k-1, \epsilon}, H^{k, \epsilon} \in \mathbb{R}^{D \times (NL)}$ denote the reshaped features extracted via an MLP for $\epsilon \in \{\text{joint}, \text{gripper}\}$ at steps $k-1$ and k . The distillation loss for these modalities is:

$$\mathcal{L}_{\text{extra}} = \sum_{\epsilon \in \{\text{joint}, \text{gripper}\}} \mathcal{L}_{\text{SPREAD}}(H^{k-1, \epsilon}, H^{k, \epsilon}). \quad (5)$$

C. Confidence-guided Policy Distillation

To ensure consistent retention of prior knowledge in lifelong imitation learning, we align the action distributions of previously learned tasks, parameterized as a Gaussian mixture policy. While the Kullback–Leibler (KL) divergence is the standard measure for distributional alignment, a closed-form solution is intractable for mixture models [36]; hence, we employ a Monte Carlo approximation. However, uniform sampling introduces variance, as low-probability regions of the prior policy can disproportionately affect estimation and destabilize optimization. To address this, we use high-confidence samples from the prior policy, focusing knowledge transfer on statistically reliable and behaviorally meaningful regions. This selective distillation reduces spurious influence, preserves prior policy structure, and improves the stability of lifelong adaptation.

Let π^k and π^{k-1} be the policies at steps k and $k-1$. For policy distillation, we align action distributions by minimizing a confidence-weighted KL divergence between the current GMM policy π^k and the previous policy π^{k-1} . We uniformly sample $\{\mathbf{a}^s\}_{s=1}^B \sim \pi^{k-1}$ and select the top $M = \lfloor 0.9B \rfloor$ samples with the highest log probabilities $\log \pi^{k-1}(\mathbf{a}^s)$, indexed by \mathcal{S}_M . The distillation loss is:

$$\mathcal{L}_{\text{policy}} = \frac{1}{M} \sum_{s \in \mathcal{S}_M} (\log \pi^k(\mathbf{a}^s) - \log \pi^{k-1}(\mathbf{a}^s)), \quad (6)$$

Eq. 6 focuses distillation on confident regions of π^{k-1} (high density under the prior policy), reducing variance and avoiding misleading gradients from low-likelihood samples.

Final Optimization Objective. Integrating the task loss with the modality-specific distillation objectives, the final optimization objective updates the policy π by minimizing:

$$J(\pi) = \frac{1}{K} \sum_{k=1}^K \mathbb{E}_{s_t, a_t \sim \mathcal{D}^k \cup \hat{\mathcal{D}}^k} \left[\sum_{t=0}^{L^k} -\log \pi(a_t \mid (s_t)_{\leq t}; T^k) + \lambda_i \mathcal{L}_{\text{image}} + \lambda_t \mathcal{L}_{\text{text}} + \lambda_e \mathcal{L}_{\text{extra}} + \lambda_p \mathcal{L}_{\text{policy}} \right]. \quad (7)$$

where $\hat{\mathcal{D}}^k$ refers to memory exemplars uniformly sampled from previously learned tasks. $\lambda_i, \lambda_t, \lambda_e$, and λ_p are weighting coefficients that balance the trade-off between stability (preserving past knowledge) and plasticity (adapting to new tasks) during the learning process. This joint optimization takes inspiration from experience replay, to ensure that the policy not only adapts to the current task but also maintains consistency with previously acquired knowledge.

IV. EXPERIMENTAL SETTINGS

A. Training and Implementation Details

We train our subspace representation distillation framework on an NVIDIA H100 GPU, employing the same data augmentation strategy and ResNet–Transformer backbone as the ResNet-T baseline [37]. Each incremental task is trained for 50 epochs. To account for varying levels of task complexity, we adopt different regularization weights across the LIBERO suites. For **LIBERO-OBJECT**, we set $\lambda_p = 0.003$ and $\lambda_t = \lambda_i = \lambda_e = 0.005$. For **LIBERO-SPATIAL**, the weights are increased to $\lambda_p = 0.005$ and $\lambda_t = \lambda_i = \lambda_e = 0.03$. Finally, for **LIBERO-GOAL**, due to its higher difficulty, we employ $\lambda_p = 0.01$ and $\lambda_t = \lambda_i = \lambda_e = 0.3$. We fix the subspace rank to $r = 48$ for all experiments. We evaluate our method against the following baselines: (1) **SEQUENTIAL** [37], which naively fine-tunes tasks sequentially using the ResNet-Transformer; (2) **EWC** [6], a regularization method that selectively updates less critical weights to preserve prior task knowledge; (3) **ER** [8], a rehearsal-based approach with a 1000-trajectory memory buffer to retain prior task samples; (4) **BUDS** [26], a hierarchical policy method using multitask skill discovery; (5) **LOTUS** [9], a hierarchical imitation learning framework with open-vocabulary vision and experience replay for unsupervised skill discovery; and (6) **M2Distill** [10], a multi-modal distillation method maintaining consistent latent spaces across vision, language, and action modalities. Baseline results are sourced from M2Distill [10].

B. Datasets

For our evaluations, we utilize the LIBERO benchmark [37], designed for lifelong imitation learning in robotic manipulation. LIBERO encompasses diverse, language-conditioned tasks with varied objects, sparse rewards, and long-horizon objectives, making it ideal for assessing continual learning. We focus on three task suites, each comprising 10 sequential tasks: LIBERO-OBJECT, LIBERO-GOAL, and LIBERO-SPATIAL. These suites evaluate knowledge transfer in distinct dimensions: object-specific declarative knowledge, procedural goal-oriented behaviors, and spatial relational understanding, respectively. LIBERO-OBJECT tasks require continual learning to manipulate distinct objects. LIBERO-GOAL tasks involve identical objects with different spatial arrangements and task objectives, necessitating adaptation of motions and behaviors. LIBERO-SPATIAL tasks challenge the robot to distinguish identical objects based solely on spatial context, demanding persistent learning of spatial relationships. This setup rigorously tests our subspace

TABLE I: Experimental results across three different LIBERO task suites. The reported values are averages from three seeds, including the mean and standard error. The best values are highlighted in bold, and the second-best values are underlined. The dash (-) indicates a failure to reproduce results. All metrics are measured based on success rates (%).

Method	LIBERO-OBJECT			LIBERO-GOAL			LIBERO-SPATIAL		
	FWT (\uparrow)	NBT (\downarrow)	AUC (\uparrow)	FWT (\uparrow)	NBT (\downarrow)	AUC (\uparrow)	FWT (\uparrow)	NBT (\downarrow)	AUC (\uparrow)
Sequential	62.0 (\pm 1.0)	63.0 (\pm 2.0)	30.0 (\pm 1.0)	55.0 (\pm 1.0)	70.0 (\pm 1.0)	23.0 (\pm 1.0)	<u>72.0</u> (\pm 1.0)	81.0 (\pm 1.0)	20.0 (\pm 1.0)
EWC [6]	56.0 (\pm 3.0)	69.0 (\pm 2.0)	16.0 (\pm 2.0)	32.0 (\pm 2.0)	48.0 (\pm 3.0)	6.0 (\pm 1.0)	23.0 (\pm 1.0)	33.0 (\pm 1.0)	6.0 (\pm 1.0)
ER [8]	56.0 (\pm 1.0)	24.0 (\pm 1.0)	49.0 (\pm 1.0)	53.0 (\pm 1.0)	36.0 (\pm 1.0)	47.0 (\pm 2.0)	65.0 (\pm 3.0)	27.0 (\pm 3.0)	56.0 (\pm 1.0)
BUDS [26]	52.0 (\pm 2.0)	21.0 (\pm 1.0)	47.0 (\pm 1.0)	50.0 (\pm 1.0)	39.0 (\pm 1.0)	42.0 (\pm 1.0)	-	-	-
LOTUS [9]	74.0 (\pm 3.0)	11.0 (\pm 1.0)	65.0 (\pm 3.0)	61.0 (\pm 3.0)	30.0 (\pm 1.0)	56.0 (\pm 1.0)	-	-	-
M2Distill [10]	<u>75.0</u> (\pm 3.0)	<u>8.0</u> (\pm 5.0)	<u>69.0</u> (\pm 4.0)	<u>71.0</u> (\pm 1.0)	<u>20.0</u> (\pm 3.0)	<u>57.0</u> (\pm 2.0)	74.0 (\pm 1.0)	11.0 (\pm 1.0)	61.0 (\pm 2.0)
<i>SPREAD</i> (Ours)	81.0 (\pm 2.0)	8.0 (\pm 1.0)	73.0 (\pm 3.0)	78.0 (\pm 1.0)	9.0 (\pm 5.0)	72.0 (\pm 1.0)	71.0 (\pm 1.0)	8.0 (\pm 5.0)	66.0 (\pm 2.0)

representation distillation framework’s ability to preserve geometric structures and action distributions across tasks.

C. Evaluation Metrics

To assess policy performance in lifelong imitation learning for robotic manipulation tasks, we adopt three key metrics: Forward Transfer (FWT), Negative Backward Transfer (NBT), and Area Under the Success Rate Curve (AUC), following established protocols [9], [37]. These metrics, based on task success rates, offer a robust evaluation compared to training loss, which is less reliable for manipulation tasks. FWT, defined as $FWT = \frac{1}{K} \sum_{k=1}^K FWT_k$ where $FWT_k = \frac{1}{51} \sum_{e=0}^{50} c_{k,k,e}$, measures adaptation to new tasks using prior knowledge, with higher values indicating better transfer. NBT, $NBT = \frac{1}{K} \sum_{k=1}^K \frac{1}{K-k} \sum_{\tau=k+1}^K (c_{k,k} - c_{\tau,k})$, quantifies forgetting of prior tasks, with lower values reflecting stronger retention. AUC, $AUC = \frac{1}{K} \sum_{k=1}^K \frac{1}{K-k+1} (FWT_k + \sum_{\tau=k+1}^K c_{\tau,k})$, assesses overall task performance, with higher scores denoting superior success. Here, K is the number of tasks, $c_{k,k,e}$ is the success rate of task k at epoch e , and $c_{\tau,k}$ is the success rate of task k after training on task τ [37].

V. EXPERIMENTAL EVALUATION

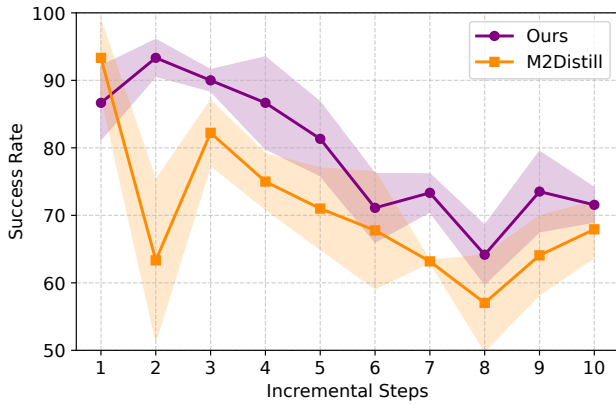
Comparison to SOTA. We evaluate our Subspace Representation Distillation framework, which aligns low-rank feature projections and employs confidence-guided policy distillation with top- K samples, by comparing it against state-of-the-art (SOTA) methods on the LIBERO benchmark [37], comprising three task suites (LIBERO-OBJECT, LIBERO-GOAL, LIBERO-SPATIAL), each with 10 sequential robotic manipulation tasks. Following prior work, we report FWT, NBT, and AUC, with results averaged over three random seeds (see Table I). Across all benchmarks, *SPREAD* achieves the best or competitive performance compared to strong baselines such as LOTUS and M2Distill. On **LIBERO-OBJECT**, *SPREAD* obtains the highest FWT (81.0% \pm 2.0), AUC (73.0% \pm 3.0), while also maintaining the lowest NBT (8.0% \pm 1.0), outperforming M2Distill by +6% in FWT and +4% in AUC. On **LIBERO-GOAL**, which presents more challenging goal-conditioned tasks, *SPREAD* surpasses prior methods by a large margin, achieving an FWT of 78.0% \pm 1.0 and an NBT of only 9.0% \pm 5.0; in contrast, LOTUS and M2Distill suffer significantly higher forgetting (30.0% \pm 1.0 and 20.0% \pm 3.0, respectively). Finally, on **LIBERO-SPATIAL**, *SPREAD* reaches the best AUC (66.0% \pm 2.0) while keeping NBT low (8.0%

\pm 5.0), outperforming M2Distill by +5% in AUC with comparable FWT. These results demonstrate *SPREAD*’s superior adaptation (FWT) and retention (NBT), driven by its low-rank subspace alignment and confidence-guided policy distillation, which preserve intrinsic task manifolds and robust action distributions across diverse LIL scenarios.

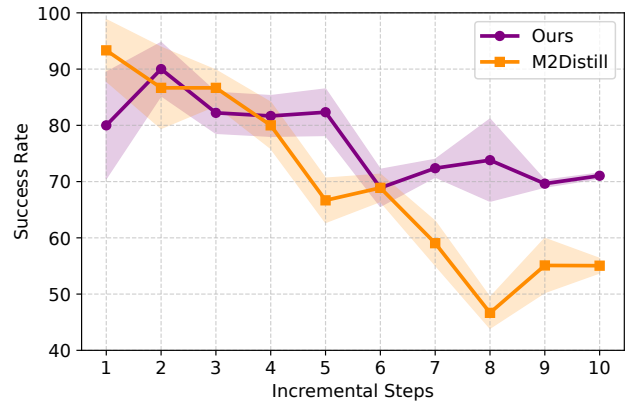
Success Rate Analysis. We evaluate our subspace representation distillation and M2Distill baseline across ten incremental steps on the LIBERO-OBJECT and LIBERO-GOAL datasets to assess robustness in continual learning and present the result in Fig. 3. On LIBERO-OBJECT, our method achieved success rates from 64.0% \pm 4.0% to a peak of 93.0% \pm 3.0% at step 2, staying above 70.0%, while M2Distill started at 93.0% \pm 5.0% but dropped to 68.0% \pm 4.0% by step 10, with variability like 63.0% \pm 12.0% at step 2. On LIBERO-GOAL, our approach maintained rates from 69.0% \pm 3.0% to 90.0% \pm 5.0%, peaking at step 2 from an 80.0% \pm 9.0% start, whereas M2Distill fell from 93.0% \pm 5.0% to 55.0% \pm 1.0% by step 10, hitting 47.0% \pm 3.0% at step 8. These results highlight our method’s superior consistency and knowledge retention over M2Distill in continual learning.

FWT Analysis. We assess the Forward Transfer (FWT) of our subspace representation distillation strategy and the M2Distill baseline across 10 incremental steps on the LIBERO-OBJECT and LIBERO-GOAL datasets to evaluate their ability to utilize prior knowledge for new tasks in continual learning. Results are illustrated in Fig. 4. FWT quantifies performance improvement on new tasks from prior learning, with our method achieving robust FWT values of 63.0% \pm 5.0% to 83.0% \pm 2.0% on LIBERO-OBJECT, peaking at step 6 with standard errors of 1.0% to 5.0%, compared to M2Distill’s 60.0% \pm 9.0% to 75.0% \pm 3.0%, peaking at step 10 with higher variability up to 9.0%. On LIBERO-GOAL, our approach maintained stable FWT values of 73.0% \pm 1.0% to 83.0% \pm 1.0%, peaking at step 1 with errors of 1.0% to 3.0%, while M2Distill ranged from 72.0% \pm 1.0% to 80.0% \pm 1.0%, peaking at step 3 with errors up to 4.0%. These findings highlight our strategy’s consistent superiority over M2Distill in knowledge transfer, enhancing continual learning performance in LIL.

NBT Analysis. We assess the Negative Backward Transfer (NBT) of our subspace representation distillation strategy and the M2Distill baseline across ten incremental steps on the LIBERO-GOAL dataset to evaluate forgetting in continual learning. Results are depicted in Fig. 5. NBT measures performance decline on prior tasks with new task introduction.

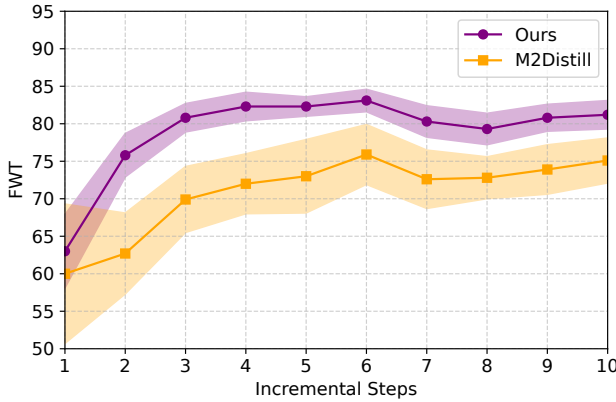


(a) LIBERO-OBJECT

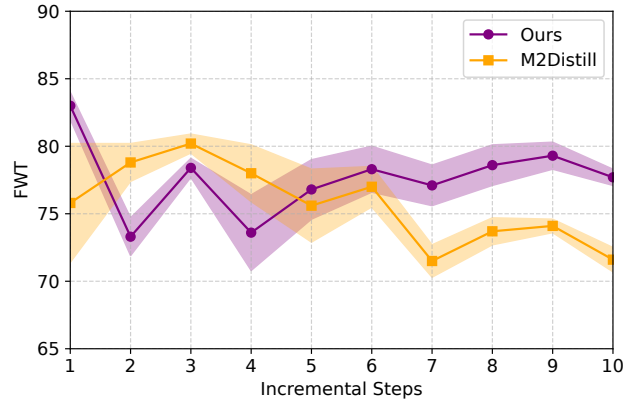


(b) LIBERO-GOAL

Fig. 3: Average success rate across incremental tasks on (a) LIBERO-OBJECT and (b) LIBERO-GOAL (Higher numbers are better).



(a) LIBERO-OBJECT



(b) LIBERO-GOAL

Fig. 4: Forward Transfer (FWT) across incremental tasks on (a) LIBERO-OBJECT and (b) LIBERO-GOAL (Higher numbers are better).

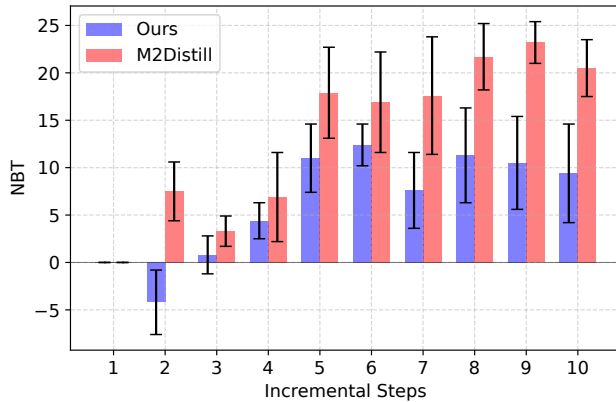


Fig. 5: Negative Backward Transfer (NBT) across incremental tasks on LIBERO-GOAL (lower is better).

Our method shows NBT values from -4.0 ± 3.0 at step 2 to 12.0 ± 2.0 at step 6, featuring modest positive values (1.0 to 12.0) and errors of 2.0% to 5.0%, indicating minimal forgetting and a gain at step 2. Conversely, M2Distill displayed higher NBT values, ranging from 3.0 ± 2.0 at step 3 to 23.0 ± 2.0 at step 9, with errors up to 6.0%, reflecting greater forgetting and a peak at step 9. These findings highlight our strategy’s superior stability in reducing forgetting compared to M2Distill on the LIBERO-GOAL dataset.

Drift Analysis. We evaluate representation drift, defined as the change in learned feature embeddings between two consecutive tasks, measured as the mean distance between their representations. The evaluation is performed across language, wrist-mounted camera (HandEye), and overhead camera (AgentView) embedding spaces over successive incremental learning steps on the LIBERO-OBJECT task suite with a fixed random seed of 100. Overall, we observe that language embeddings remain the most stable across incremental steps, while hand-eye and agent-view embeddings exhibit higher drift. As shown in Fig. 6, our proposed subspace representation distillation method significantly outperforms M2Distill in preserving feature stability across all modalities by maintaining similar subspaces. In the language embedding space, *SPREAD* reduces drift by more than 75%, demonstrating strong preservation of semantic representations. In the visual modalities, it effectively suppresses the large drift spikes observed in M2Distill, keeping drift in the HandEye embedding space below 0.5 compared to the baseline peak of over 2.7. For AgentView, *SPREAD* again outperforms with a mean drift of 0.095 against 0.356 for M2Distill. Overall, these results demonstrate that subspace representation distillation effectively reduces drift across modalities, mitigating catastrophic forgetting in lifelong imitation learning.

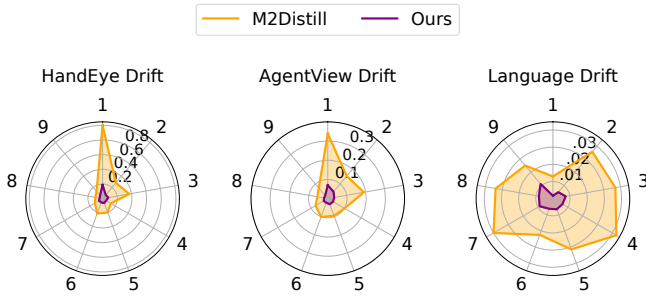


Fig. 6: Drift analysis across incremental tasks for different modalities on LIBERO-OBJECT. The radial axes (1-9) indicate the incremental steps in LIL. Lower representation drift indicates better preservation of previously learned knowledge across tasks in lifelong imitation learning.

VI. ABLATION STUDY

Influence of each loss component. We further investigate the contribution of individual loss components through an ablation study on the LIBERO-GOAL task suite, with results summarized in Table II. The removal of the text loss $\mathcal{L}_{\text{text}}$ yields a slight reduction in AUC (0.72 \rightarrow 0.69), indicating that language alignment provides auxiliary benefits but is not the dominant factor. In contrast, excluding the image loss $\mathcal{L}_{\text{image}}$ leads to a substantial degradation in performance, with AUC dropping from 0.72 to 0.57 and NBT increasing from 0.09 to 0.20, thereby underscoring the critical role of visual representation preservation in mitigating forgetting. Eliminating the action loss $\mathcal{L}_{\text{policy}}$ results in decreased forward transfer (0.78 \rightarrow 0.70), despite a marginal improvement in NBT, suggesting that action alignment is essential for retaining transferable policies across tasks. Finally, the exclusion of the extra-modality loss $\mathcal{L}_{\text{extra}}$ produces a minor decline in both FWT and AUC, confirming its complementary role in enhancing representational consistency. Collectively, these findings demonstrate that while all components contribute to the robustness of the proposed framework, image feature alignment emerges as the most influential factor for effective lifelong imitation learning.

Dimensionality of Subspace Representation Distillation.

We investigate the impact of subspace dimensionality in *SPREAD* by varying the projection rank r on the LIBERO-GOAL task suite (Table III). The results demonstrate that the choice of rank has a significant effect on both knowledge transfer and forgetting. Although the full-rank ($r = 64$) baseline attains competitive FWT, it suffers from higher

TABLE II: Ablation studies on the contribution of each component in our method. Experiments were performed on the LIBERO-GOAL task suite. Reported results are the mean across three different seed values.

$\mathcal{L}_{\text{text}}$	$\mathcal{L}_{\text{image}}$	$\mathcal{L}_{\text{action}}$	$\mathcal{L}_{\text{extra}}$	LIBERO-GOAL		
				FWT \uparrow	NBT \downarrow	AUC \uparrow
✓	✓	✓	✓	0.78	0.09	0.72
✗	✓	✓	✓	0.78	0.11	0.69
✓	✗	✓	✓	0.67	0.20	0.57
✓	✓	✗	✓	0.70	0.02	0.70
✓	✓	✓	✗	0.74	0.03	0.71

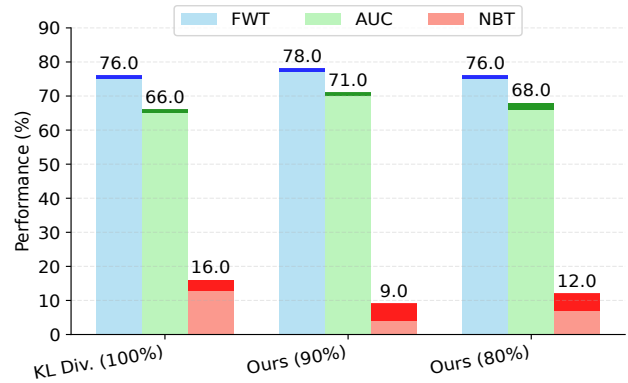


Fig. 7: Ablation study on policy distillation strategies using LIBERO-GOAL task suite. Bar represents mean across three random seeds while standard error is represented with darker color. Our top- M confidence-weighted approach outperforms baselines in FWT, NBT and AUC.

TABLE III: Ablation study on the dimensionality/rank of subspace representation distillation strategies using LIBERO-GOAL task suite. Our proposed *SPREAD* loss with 75%-rank performs best in FWT, NBT and AUC.

Rank of Subspace	FWT (\uparrow)	NBT (\downarrow)	AUC (\uparrow)
<i>SPREAD</i> (Full-Rank)	76.0 (\pm 1.0)	16.0 (\pm 8.0)	66.0 (\pm 5.0)
<i>SPREAD</i> (Rank = 48)	78.0 (\pm 1.0)	9.0 (\pm 5.0)	72.0 (\pm 1.0)
<i>SPREAD</i> (Rank = 32)	74.0 (\pm 1.0)	6.0 (\pm 4.0)	71.0 (\pm 2.0)
<i>SPREAD</i> (Rank = 16)	76.0 (\pm 3.0)	9.0 (\pm 4.0)	69.0 (\pm 2.0)

forgetting and a lower AUC. In contrast, the 75%-rank configuration ($r = 48$) achieves the best overall performance, with the highest AUC (71.0% \pm 1.0) and FWT (78.0% \pm 1.0), outperforming the full-rank *SPREAD* by 5% in AUC and 7% in NBT. This suggests that a moderately compressed subspace effectively retains transferable information while mitigating redundancy. However, further reducing the rank ($r = 32, 16$) causes noticeable information loss, leading to degraded FWT and AUC. Following these results, we set the 75%-rank as the default in all subsequent experiments.

Policy Distillation with Top- M Confident Samples.

We evaluate the effectiveness of our confidence-based policy distillation strategy, which selects the top M samples with the highest log-probability scores under the previous policy π^{k-1} . Experiments are conducted on the LIBERO-GOAL suite, comprising 10 consecutive robotic manipulation tasks. The policy distillation loss, $\mathcal{L}_{\text{policy}}$, complements subspace representation distillation by enforcing alignment of action distributions. As shown in Fig. 7, our method—combining low-rank subspace projection with top- $M = \lfloor 0.9B \rfloor$ confident samples—achieves the best results, with an AUC of 71.0% \pm 1.0, FWT of 78.0% \pm 1.0, and NBT of 9.0% \pm 5.0. This outperforms standard KL divergence using all samples by 5% in AUC and 7% in NBT. In contrast, a variant with $M = \lfloor 0.8B \rfloor$ samples shows reduced performance (AUC: 68.0% \pm 2.0, NBT: 12.0% \pm 5.0), suggesting that $M = \lfloor 0.9B \rfloor$ provides a good balance between sample diversity and confidence. Overall, this confidence-based top- M selection enhances robustness by prioritizing reliable

action distributions, complementing geometric preservation in feature distillation, and improving both stability and generalization across sequential tasks.

VII. CONCLUSION

This work presented *SPREAD*, a modality-aware framework for lifelong imitation learning that explicitly leverages subspace geometry to mitigate catastrophic forgetting. By aligning features of two consecutive policies through low-rank projections derived via singular value decomposition, *SPREAD* preserves the intrinsic low-dimensional structure of multimodal task embeddings more effectively than conventional feature-space distillation. Complementary to this, the proposed confidence-guided policy distillation selectively emphasizes high-confidence action samples, yielding a more stable and reliable transfer of behavioral priors. On the LIBERO benchmark, *SPREAD* achieves AUC scores of 73.0%, 72.0%, and 66.0% on LIBERO-OBJECT, LIBERO-GOAL, and LIBERO-SPATIAL, outperforming state-of-the-art M2Distill by 4–15% in AUC and reducing negative backward transfer by 12%. These findings highlight both the empirical benefits and the theoretical advantages of subspace-level alignment for continual representation learning. Future research will explore *SPREAD*'s application to long-horizon incremental task learning and addressing extended task sequences for real-world adaptability.

REFERENCES

- [1] A. Hussein, M. M. Gaber *et al.*, "Imitation learning: A survey of learning methods," *ACM Computing Surveys (CSUR)*, vol. 50, no. 2, pp. 1–35, 2017.
- [2] S. Schaal, "Learning from demonstration," *Advances in neural information processing systems*, vol. 9, 1996.
- [3] A. G. Billard, S. Calinon, and R. Dillmann, "Learning from humans," *Springer handbook of robotics*, pp. 1995–2014, 2016.
- [4] S. Stepputtis, J. Campbell *et al.*, "Language-conditioned imitation learning for robot manipulation tasks," *Advances in Neural Information Processing Systems*, vol. 33, pp. 13 139–13 150, 2020.
- [5] A. Xie, L. Lee *et al.*, "Decomposing the generalization gap in imitation learning for visual robotic manipulation," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 3153–3160.
- [6] J. Kirkpatrick, R. Pascanu *et al.*, "Overcoming catastrophic forgetting in neural networks," *Proceedings of the National Academy of Sciences*, vol. 114, no. 13, pp. 3521–3526, 2017.
- [7] R. M. French, "Catastrophic forgetting in connectionist networks," *Trends in cognitive sciences*, vol. 3, no. 4, pp. 128–135, 1999.
- [8] A. Chaudhry, M. Rohrbach *et al.*, "On tiny episodic memories in continual learning," *arXiv preprint arXiv:1902.10486*, 2019.
- [9] W. Wan, Y. Zhu *et al.*, "Lotus: Continual imitation learning for robot manipulation through unsupervised skill discovery," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 537–544.
- [10] K. Roy, A. Dissanayake *et al.*, "M2Distill: Multi-modal distillation for lifelong imitation learning," in *2025 IEEE International Conference on Robotics and Automation (ICRA)*, 2025, pp. 1429–1435.
- [11] K. Roy, C. Simon *et al.*, "Subspace distillation for continual learning," *Neural Networks*, vol. 167, pp. 65–79, 2023.
- [12] S. Cheng, Y. Wang *et al.*, "Nbnnet: Noise basis learning for image denoising with subspace projection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 4896–4906.
- [13] S. Kullback and R. A. Leibler, "On information and sufficiency," *The annals of mathematical statistics*, vol. 22, no. 1, pp. 79–86, 1951.
- [14] E. Jang, A. Irpan *et al.*, "Bc-z: Zero-shot task generalization with robotic imitation learning," in *Conference on Robot Learning*. PMLR, 2022, pp. 991–1002.
- [15] K. Shmelkov, C. Schmid, and K. Alahari, "Incremental learning of object detectors without catastrophic forgetting," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 3400–3409.
- [16] L. Mahmoodi, P. Moghadam *et al.*, "Flashbacks to harmonize stability and plasticity in continual learning," *Neural Networks*, vol. 190, p. 107616, 2025.
- [17] K. Huang, P. Li *et al.*, "Knowledge transfer in incremental learning for multilingual neural machine translation," in *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, 2023, pp. 15 286–15 304.
- [18] H. Ahn, S. Cha *et al.*, "Uncertainty-based continual learning with adaptive regularization," *Advances in neural information processing systems*, vol. 32, 2019.
- [19] K. Roy, P. Moghadam, and M. Harandi, "L3DMC: Lifelong learning using distillation via mixed-curvature space," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2023, pp. 123–133.
- [20] S.-A. Rebuffi, A. Kolesnikov *et al.*, "icarl: Incremental classifier and representation learning," in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, 2017, pp. 2001–2010.
- [21] D. Rolnick, A. Ahuja *et al.*, "Experience replay for continual learning," *Advances in neural information processing systems*, vol. 32, 2019.
- [22] C. Gao, H. Gao *et al.*, "CRIL: Continual robot imitation learning via generative and prediction model," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2021, pp. 6747–5754.
- [23] J. Yoon, E. Yang *et al.*, "Lifelong learning with dynamically expandable networks," in *International Conference on Learning Representations (ICLR)*, 2018.
- [24] A. Douillard, A. Ramé *et al.*, "Dytox: Transformers for continual learning with dynamic token expansion," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 9285–9295.
- [25] H. Shin, J. K. Lee *et al.*, "Continual learning with deep generative replay," *Advances in neural information processing systems*, vol. 30, 2017.
- [26] Y. Zhu, P. Stone, and Y. Zhu, "Bottom-up skill discovery from unsegmented demonstrations for long-horizon robot manipulation," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 4126–4133, 2022.
- [27] H. Zhang, J. Gong *et al.*, "T2S: Tokenized skill scaling for lifelong imitation learning," *arXiv preprint arXiv:2508.01167*, 2025.
- [28] J. Xu and X. Nie, "SPECI: Skill prompts based hierarchical continual imitation learning for robot manipulation," *arXiv preprint arXiv:2504.15561*, 2025.
- [29] S. Haldar and L. Pinto, "Polytask: Learning unified policies through behavior distillation," *arXiv preprint arXiv:2310.08573*, 2023.
- [30] M. Bain and C. Sammut, "A framework for behavioural cloning," in *Machine Intelligence 15*, 1995, pp. 103–129.
- [31] M. E. Wall, A. Rechtsteiner, and L. M. Rocha, "Singular value decomposition and principal component analysis," *A practical approach to microarray data analysis*, pp. 91–109, 2003.
- [32] G. Hinton, O. Vinyals *et al.*, "Distilling the knowledge in a neural network," *arXiv preprint arXiv:1503.02531*, vol. 2, no. 7, 2015.
- [33] C. Eckart and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrika*, vol. 1, no. 3, pp. 211–218, 1936.
- [34] K. He, X. Zhang *et al.*, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [35] A. Radford, J. W. Kim *et al.*, "Learning transferable visual models from natural language supervision," in *International conference on machine learning*. PMLR, 2021, pp. 8748–8763.
- [36] J. R. Hershey and P. A. Olsen, "Approximating the Kullback Leibler divergence between Gaussian mixture models," in *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07*, vol. 4. IEEE, 2007, pp. IV–317.
- [37] B. Liu, Y. Zhu *et al.*, "Libero: Benchmarking knowledge transfer for lifelong robot learning," *Advances in Neural Information Processing Systems*, vol. 36, 2024.