Novelty Assessment Report

Paper: Robust Fine-tuning of Vision-Language-Action Robot Policies via Parameter Merging
PDF URL: https://openreview.net/pdf?id=uWJwQ5SZoM
Venue: ICLR 2026 Conference Submission
Year: 2026
Report Generated: 2026-01-01

Abstract

Generalist robot policies, trained on large and diverse datasets, have demonstrated the ability to generalize across a wide spectrum of behaviors, enabling a single policy to act in varied real-world environments. However, they still fall short on new tasks not covered in the training data. When finetuned on limited demonstrations of a new task, these policies often overfit to the specific demonstrations---not only losing their prior abilities to solve a wide variety of generalist tasks but also failing to generalize within the new task itself. In this work, we aim to develop a method that preserves the generalization capabilities of the generalist policy during finetuning, allowing a single policy to robustly incorporate a new skill into its repertoire. Our goal is a single policy that both learns to generalize to variations of the new task and retains the broad competencies gained from pretraining. We show that this can be achieved through a simple yet effective strategy: interpolating the weights of a finetuned model with that of the pretrained model. We show, across extensive simulated and real-world experiments, that such model merging produces a single model that inherits the generalist abilities of the base model and learns to solve the new task robustly, outperforming both the pretrained and finetuned model on out-of-distribution variations of the new task. Moreover, we show that model merging enables continual acquisition of new skills in a lifelong learning setting, without sacrificing previously learned generalist abilities.

Disclaimer

This report is **AI-GENERATED** using Large Language Models and WisPaper (a scholar search engine). It analyzes academic papers' tasks and contributions against retrieved prior work. While this system identifies **POTENTIAL** overlaps and novel directions, **ITS COVERAGE IS NOT EXHAUSTIVE AND JUDGMENTS ARE APPROXIMATE**. These results are intended to assist human reviewers and **SHOULD NOT** be relied upon as a definitive verdict on novelty.

Note that some papers exist in multiple, slightly different versions (e.g., with different titles or URLs). The system may retrieve several versions of the same underlying work. The current automated pipeline does not reliably align or distinguish these cases, so human reviewers will need to disambiguate them manually.

If you have any questions, please contact: mingzhang23@m.fudan.edu.cn

Core Task Landscape

This paper addresses: **Continual Learning and Robust Generalization in Vision-Language-Action Robot Policy Finetuning**

A total of **7 papers** were analyzed and organized into a taxonomy with **5 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Parameter-Efficient Adaptation and Catastrophic Forgetting Mitigation**
- **Reinforcement Learning and Self-Improvement for VLA Adaptation**
- **Explicit Reasoning and Visual Latent Planning**

Complete Taxonomy Tree

- Continual Learning and Robust Generalization in Vision-Language-Action Robot Policy Finetuning Survey Taxonomy
- Parameter-Efficient Adaptation and Catastrophic Forgetting Mitigation
  - Weight Interpolation and Model Merging ★ (2 papers)
  - [0] Robust Fine-tuning of Vision-Language-Action Robot Policies via Parameter Merging (Anon et al., 2026) View paper
  - [3] Actions as language: Fine-tuning vlms into vlas without catastrophic forgetting (Wu, 2025) View paper
  - Progressive Expert Libraries and Knowledge-Driven Continual Learning (2 papers)
  - [5] Continually Evolving Skill Knowledge in Vision Language Action Model (Yuxuan Wu, 2025) View paper
  - [7] Dynamic Mixture of Progressive Parameter-Efficient Expert Library for Lifelong Robot Learning (Lei, 2025) View paper
- Reinforcement Learning and Self-Improvement for VLA Adaptation
  - World Model-Based Policy Optimization (1 papers)
  - [1] WMPO: World Model-based Policy Optimization for Vision-Language-Action Models (Fangqi Zhu, 2025) View paper
  - Self-Reflective and Dual-System Adaptation (2 papers)
  - [4] Reflection-based task adaptation for self-improving vla (Li Baicheng, 2025) View paper
  - [6] A Self-Correcting Vision-Language-Action Model for Fast and Slow System Manipulation (Li Chen-Xuan, 2024) View paper
- Explicit Reasoning and Visual Latent Planning (1 papers)
  - [2] Thinkact: Vision-language-action reasoning via reinforced visual latent planning (Wu, 2025) View paper

Narrative

Core task: continual learning and robust generalization in vision-language-action robot policy finetuning. The field addresses how to adapt large-scale vision-language-action (VLA) models to new robotic tasks without losing previously acquired capabilities or overfitting to narrow domains. The taxonomy reveals three main branches: Parameter-Efficient Adaptation and Catastrophic Forgetting Mitigation focuses on lightweight tuning strategies and techniques to preserve base model knowledge during specialization; Reinforcement Learning and Self-Improvement for VLA Adaptation explores online learning and iterative refinement methods that leverage environmental feedback; and Explicit Reasoning and Visual Latent Planning emphasizes structured decision-making and internal world models. These branches reflect complementary perspectives on the same challenge—balancing plasticity for new skills with stability of existing competencies—while differing in whether they prioritize efficient parameter updates, interactive learning loops, or interpretable reasoning pathways.

Several active lines of work illustrate key trade-offs in this landscape. Some approaches emphasize self-correction and reflection mechanisms (Self-Correcting VLA[6], Reflection Task Adaptation[4]) to improve generalization through iterative reasoning, while others focus on modular skill libraries or expert ensembles (Dynamic Expert Library[7], Continual Skill Knowledge[5]) to compartmentalize knowledge and reduce interference. Within the Parameter-Efficient Adaptation branch, Robust VLA Finetuning[0] sits alongside Actions as Language[3] in exploring weight interpolation and model merging strategies that blend task-specific adaptations with pretrained

representations. Compared to WMPO[1], which integrates preference optimization into the adaptation process, Robust VLA Finetuning[0] appears more concerned with maintaining robustness across diverse deployment conditions through careful weight combination. This positioning highlights an ongoing question: whether continual learning is best achieved by merging separately trained experts or by designing unified training objectives that inherently balance old and new knowledge.

Related Works in Same Category

The following **1 sibling papers** share the same taxonomy leaf node with the original paper:

1. Actions as language: Fine-tuning vlms into vlas without catastrophic forgetting

Authors: Wu, Xindi, Asher Hancock, Zha, Lihan, et al. (13 authors total) | **Year/Venue**: 2025 | **URL**: View paper

Abstract

Fine-tuning vision-language models (VLMs) on robot teleoperation data to create vision-language-action (VLA) models is a promising paradigm for training generalist policies, but it suffers from a fundamental tradeoff: learning to produce actions often diminishes the VLM's foundational reasoning and multimodal understanding, hindering generalization to novel scenarios, instruction following, and semantic understanding. We argue that this catastrophic forgetting is due to a distribution mismatch b...

Relationship Analysis

Both papers belong to the Weight Interpolation and Model Merging category, addressing catastrophic forgetting in VLA policy finetuning. While the original paper (RETAIN) focuses on interpolating pretrained and finetuned model weights in parameter space to balance generalist abilities with new task learning, the candidate paper (VLM2VLA) takes a different approach by representing actions as natural language to align robot data with VLM pretraining distributions, then using LoRA for minimal backbone modification. The key difference is that RETAIN merges weights post-finetuning to preserve generalization, whereas VLM2VLA prevents forgetting through data-level alignment and parameter-efficient training from the start.

Contributions Analysis

Overall novelty summary. The paper proposes weight interpolation between pretrained and finetuned vision-language-action models to preserve generalist capabilities while learning new tasks. It resides in the 'Weight Interpolation and Model Merging' leaf, which contains only two papers total, indicating a relatively sparse research direction within the broader taxonomy. This leaf sits under 'Parameter-Efficient Adaptation and Catastrophic Forgetting Mitigation', one of three major branches addressing continual learning in VLA policies. The small sibling count suggests this specific approach—direct parameter blending—is less explored compared to modular expert libraries or reinforcement learning-based adaptation methods.

The taxonomy reveals neighboring directions that tackle similar challenges through different mechanisms. The sibling branch 'Progressive Expert Libraries and Knowledge-Driven Continual Learning' uses modular components rather than weight merging, while the parallel 'Reinforcement Learning and Self-Improvement' branch emphasizes online learning and world models. The 'Explicit Reasoning and Visual Latent Planning' branch decouples high-level reasoning from action execution entirely. The paper's approach contrasts with these alternatives by operating directly on parameter space without requiring modular architectures, explicit reasoning systems, or environmental interaction during adaptation.

Among thirty candidates examined across three contributions, none were identified as clearly refuting the proposed methods. The RETAIN method for parameter merging examined ten candidates with zero refutable matches, as did the modality-specific merging strategy and the continual skill acquisition framework. This absence of overlapping prior work within the limited search scope suggests the specific combination of weight interpolation for VLA policies, modality-aware merging, and sequential skill accumulation may represent a relatively unexplored configuration. However, the search examined only top-K semantic matches and citations, not the entire literature.

The analysis indicates the paper occupies a sparsely populated research direction within a broader field that favors modular or reinforcement learning-based solutions. The limited search scope—thirty candidates across three contributions—provides evidence of novelty within examined papers but cannot confirm exhaustive originality. The taxonomy structure suggests the field is actively exploring diverse continual learning strategies, with weight interpolation representing one of several competing paradigms for balancing plasticity and stability in robot policy adaptation.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: RETAIN method for robust robot policy finetuning via parameter merging

Description: The authors propose a method that interpolates the weights of pretrained and finetuned generalist robot policies in weight space. This simple weight merging strategy enables finetuned policies to generalize better to out-of-distribution variations of new tasks while retaining the broad competencies from pretraining.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Robust Policy Generalization and Risk Quantification in Autonomous Robotics

URL: View paper

Brief Assessment

Robust Policy Generalization[16] focuses on risk quantification and policy generalization in autonomous robotics, not on weight interpolation or parameter merging for robot policy finetuning. The candidate's full text context contains only fragmentary phrases about interpolation and policy parameters without substantive technical content about the RETAIN method's core contribution.

---

2. Robust imitation of diverse behaviors

URL: View paper

Brief Assessment

Robust Imitation[13] focuses on learning diverse imitation behaviors through VAE-based trajectory embeddings and adversarial training for locomotion/manipulation tasks, not on weight interpolation between pretrained and finetuned robot policies for robust generalization.

---

3. Poco: Policy composition from and for heterogeneous robot learning

URL: View paper

Brief Assessment

Poco[10] focuses on policy composition for heterogeneous robot learning by interpolating task-conditioned policies, not on finetuning pretrained generalist policies to preserve generalization capabilities on new tasks.

---

4. Rewarded soups: towards pareto-optimal alignment by interpolating weights fine-tuned on diverse rewards

URL: View paper

Brief Assessment

Rewarded Soups[8] focuses on multi-objective alignment across diverse rewards in foundation models (text, vision, control), while the original paper addresses robot policy finetuning to preserve generalization on new tasks. The technical contexts differ fundamentally: Rewarded Soups[8] targets Pareto-optimal trade-offs between conflicting reward objectives, whereas the original work aims to prevent overfitting during robot policy adaptation.

---

5. Policy-space interpolation for physics-based characters

URL: View paper

Brief Assessment

Policy Space Interpolation[9] focuses on physics-based character animation controllers for humanoid motion synthesis, not robot manipulation policies. The candidate addresses combining motion controllers for animation tasks, while the original work targets vision-language-action robot policies for manipulation tasks with different objectives and domains.

---

6. A survey on diffusion policy for robotic manipulation: Taxonomy, analysis, and future directions

URL: View paper

Brief Assessment

Diffusion Policy Survey[12] focuses on diffusion-based architectures for robotic manipulation, not on weight interpolation or parameter merging strategies for finetuning generalist robot policies. The survey discusses diffusion policy architectures rather than the weight-space merging approach that RETAIN proposes.

---

7. RoFt-Mol: Benchmarking Robust Fine-Tuning with Molecular Graph Foundation Models

URL: View paper

Brief Assessment

RoFt-Mol[15] focuses on molecular graph foundation models for chemistry tasks, not robot policies. The domain (molecular graphs vs. robotic manipulation), model architectures (molecular GNNs vs. vision-language-action models), and application contexts are fundamentally different.

---

8. Waypoint-based imitation learning for robotic manipulation

URL: View paper

Brief Assessment

Waypoint Imitation Learning[11] focuses on automatic waypoint extraction from demonstrations to reduce compounding errors in behavioral cloning, not on weight interpolation or parameter merging strategies for finetuning pretrained robot policies.

---

9. Robust Fine-tuning for Pre-trained 3D Point Cloud Models

URL: View paper

Brief Assessment

Robust Point Cloud[17] focuses on 3D point cloud models for computer vision tasks, not robot policies. The domain (point cloud processing vs. vision-language-action robot control) and application context are fundamentally different.

---

10. DaWin: Training-free Dynamic Weight Interpolation for Robust Adaptation

URL: View paper

Brief Assessment

DaWin[14] focuses on visual recognition tasks and multi-task learning in computer vision, not robot policy finetuning. The candidate addresses weight interpolation for foundation models on image classification benchmarks, while the original work specifically targets vision-language-action robot policies and robotic manipulation tasks.

Contribution 2: Modality-specific parameter merging for vision-language-action policies

Description: The authors extend their merging approach to use different interpolation coefficients for the vision encoder, language model backbone, and action expert components of VLA policies. They demonstrate that merging only the language model parameters can be sufficient for effective policy adaptation.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Visual grounding with multi-modal conditional adaptation

URL: View paper

Brief Assessment

Visual Grounding Adaptation[21] focuses on visual grounding tasks using multi-modal conditional adaptation for weight updates in visual encoders, not on vision-language-action robot policies or parameter merging strategies for different modality components in VLA models.

---

2. A dynamic weighted fusion model for multimodal sentiment analysis

URL: View paper

Brief Assessment

Dynamic Weighted Fusion[19] focuses on multimodal sentiment analysis with dynamic weighting of text, visual, and audio features. This is fundamentally different from the original paper's contribution of parameter merging in vision-language-action robot policies, which involves interpolating model weights across different components (vision encoder, language model, action expert) during policy adaptation.

---

3. Remedy: Recipe merging dynamics in large vision-language models

URL: View paper

Brief Assessment

Remedy[18] focuses on merging vision-language models for multi-modal understanding tasks (VQA, captioning), not vision-language-action policies for robotic control. The candidate addresses static vision-language comprehension, while the original work targets dynamic robotic manipulation with action prediction.

---

4. Deciphering Cross-Modal Alignment in Large Vision-Language Models via Modality Integration Rate

URL: View paper

Brief Assessment

Modality Integration Rate[22] focuses on measuring cross-modal alignment quality in vision-language models during pre-training, not on parameter merging strategies for vision-language-action policies. The candidate addresses distribution distance measurement between modalities, while the original contribution concerns weight interpolation coefficients for different components of VLA policies.

---

5. Multi-modality expansion and retention for llms through parameter merging and decoupling

URL: View paper

Brief Assessment

Multi-modality Expansion[24] focuses on merging multimodal LLMs for vision-language tasks, not vision-language-action robot policies. The candidate does not address action prediction or robotic manipulation.

---

6. Parameter efficient merging for multimodal large language models with complementary parameter adaptation

URL: View paper

Brief Assessment

Parameter Efficient Merging[20] focuses on multimodal large language models with complementary parameter adaptation, not specifically on vision-language-action robot policies or the robotics domain that the original paper addresses.

---

7. X-Fusion: Introducing New Modality to Frozen Large Language Models

URL: View paper

Brief Assessment

X-Fusion[23] addresses modality-specific parameter handling in vision-language models for multimodal understanding and generation tasks, not vision-language-action robot policies. The candidate focuses on frozen LLM parameters with modality-specific weights for image-text tasks, while the original work merges parameters across vision encoder, language model, and action expert components specifically for robotic manipulation policies.

---

8. An Empirical Study of Multimodal Model Merging

URL: View paper

Brief Assessment

Multimodal Model Merging[25] focuses on merging vision and language transformers for multimodal understanding tasks (VQA, retrieval, segmentation), not vision-language-action policies for robot control. The domains and applications are fundamentally different.

---

9. MAKE: Vision-Language Pre-training based Product Retrieval in Taobao Search

URL: View paper

Brief Assessment

MAKE[27] focuses on vision-language pre-training for product retrieval in e-commerce search, not on parameter merging for vision-language-action robot policies. The paper addresses cross-modal fusion for matching user queries with product information, which is a fundamentally different domain and technical problem from robotic policy adaptation.

---

10. Token Expand-Merge: Training-Free Token Compression for Vision-Language-Action Models

URL: View paper

Brief Assessment

Token Expand-Merge[26] focuses on token-level compression during inference through dynamic expansion and merging mechanisms, not on parameter merging strategies across different modalities (vision encoder, language model, action expert) during model training or fine-tuning.

Contribution 3: Continual skill acquisition framework through sequential parameter merging

Description: The authors show that RETAIN can be applied iteratively to sequentially add new tasks into a pretrained checkpoint by merging finetuned weights into the base model. This enables lifelong learning where a single policy accumulates new capabilities without sacrificing previously learned generalist abilities.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. A Novel Robotic Skill Learning Approach for Assembly Task With Dynamical System and Broad Learning

URL: View paper

Brief Assessment

Broad Learning Assembly[29] focuses on peg-hole assembly tasks using incremental learning for position/attitude adjustments during insertion, not on sequential parameter merging for continual skill acquisition in generalist robot policies.

---

2. A trajectory and force dual-incremental robot skill learning and generalization framework using improved dynamical movement primitives and adaptive neural â¦

URL: View paper

Brief Assessment

Trajectory Force Dual[30] focuses on trajectory and force learning using dynamical movement primitives and adaptive neural networks for incremental skill updates in robot manipulation. This is fundamentally different from RETAIN's approach of merging pretrained and finetuned policy weights in vision-language-action models to enable continual task acquisition while preserving generalist capabilities.

---

3. Preserving and combining knowledge in robotic lifelong reinforcement learning

URL: View paper

Brief Assessment

Preserving Robotic Knowledge[28] focuses on robotic lifelong reinforcement learning using a Bayesian non-parametric knowledge space (DPMM) for task inference and knowledge preservation, not on parameter merging of pretrained and finetuned models. The technical approaches are fundamentally different.

---

4. Active incremental learning of robot movement primitives

URL: View paper

Brief Assessment

Active Incremental Learning[33] focuses on incrementally learning movement primitives for robot reaching tasks using Gaussian processes and DMPs, with active learning to request demonstrations. This differs fundamentally from RETAIN's approach of merging pre-trained and fine-tuned policy parameters for vision-language-action models in a continual learning setting.

---

5. Analysis of methods for incremental policy refinement by kinesthetic guidance

URL: View paper

Brief Assessment

Kinesthetic Guidance Analysis[36] focuses on incremental refinement of robot policies through kinesthetic guidance and recursive/batch regression techniques for trajectory modification, not on sequential parameter merging for continual skill acquisition in pretrained vision-language-action models.

---

6. iManip: Skill-Incremental Learning for Robotic Manipulation

URL: View paper

Brief Assessment

iManip[35] addresses skill-incremental learning in robotic manipulation using temporal replay and extendable action prompts, not sequential parameter merging. The candidate focuses on mitigating catastrophic forgetting through replay strategies rather than weight interpolation methods.

---

7. Incremental learning of subtasks from unsegmented demonstration

URL: View paper

Brief Assessment

Incremental Subtask Learning[34] focuses on segmenting demonstrations into subtasks and learning subtask policies simultaneously using an infinite mixture of experts approach. This is fundamentally different from RETAIN's sequential parameter merging framework for continual skill acquisition in pretrained robot policies.

---

8. SPECI: Skill Prompts based Hierarchical Continual Imitation Learning for Robot Manipulation

URL: View paper

Brief Assessment

SPECI[31] focuses on hierarchical continual imitation learning with skill prompts and implicit skill codebooks for robot manipulation, not on sequential parameter merging of pretrained checkpoints. The technical approaches are fundamentally different.

---

9. Toward Scalable Multirobot Control: Fast Policy Learning in Distributed MPC

URL: View paper

Brief Assessment

Scalable Multirobot Control[32] focuses on distributed model predictive control for multirobot systems through policy learning, not on continual skill acquisition through sequential parameter merging in robot policies. The candidate addresses online policy learning for cooperative control tasks, while the original contribution concerns merging finetuned weights to accumulate new capabilities in pretrained models.

---

10. Dynamic Mixture of Progressive Parameter-Efficient Expert Library for Lifelong Robot Learning

URL: View paper

Brief Assessment

Dynamic Expert Library[7] focuses on building a progressive library of low-rank experts with dynamic routing for lifelong learning, rather than sequential weight interpolation between pretrained and finetuned models as in RETAIN.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] Robust Fine-tuning of Vision-Language-Action Robot Policies via Parameter Merging View paper
- [1] WMPO: World Model-based Policy Optimization for Vision-Language-Action Models View paper
- [2] Thinkact: Vision-language-action reasoning via reinforced visual latent planning View paper
- [3] Actions as language: Fine-tuning vlms into vlas without catastrophic forgetting View paper
- [4] Reflection-based task adaptation for self-improving vla View paper
- [5] Continually Evolving Skill Knowledge in Vision Language Action Model View paper
- [6] A Self-Correcting Vision-Language-Action Model for Fast and Slow System Manipulation View paper
- [7] Dynamic Mixture of Progressive Parameter-Efficient Expert Library for Lifelong Robot Learning View paper
- [8] Rewarded soups: towards pareto-optimal alignment by interpolating weights fine-tuned on diverse rewards View paper
- [9] Policy-space interpolation for physics-based characters View paper
- [10] Poco: Policy composition from and for heterogeneous robot learning View paper
- [11] Waypoint-based imitation learning for robotic manipulation View paper
- [12] A survey on diffusion policy for robotic manipulation: Taxonomy, analysis, and future directions View paper
- [13] Robust imitation of diverse behaviors View paper
- [14] DaWin: Training-free Dynamic Weight Interpolation for Robust Adaptation View paper
- [15] RoFt-Mol: Benchmarking Robust Fine-Tuning with Molecular Graph Foundation Models View paper
- [16] Robust Policy Generalization and Risk Quantification in Autonomous Robotics View paper
- [17] Robust Fine-tuning for Pre-trained 3D Point Cloud Models View paper
- [18] Remedy: Recipe merging dynamics in large vision-language models View paper
- [19] A dynamic weighted fusion model for multimodal sentiment analysis View paper
- [20] Parameter efficient merging for multimodal large language models with complementary parameter adaptation View paper
- [21] Visual grounding with multi-modal conditional adaptation View paper
- [22] Deciphering Cross-Modal Alignment in Large Vision-Language Models via Modality Integration Rate View paper
- [23] X-Fusion: Introducing New Modality to Frozen Large Language Models View paper
- [24] Multi-modality expansion and retention for llms through parameter merging and decoupling View paper
- [25] An Empirical Study of Multimodal Model Merging View paper

- [26] Token Expand-Merge: Training-Free Token Compression for Vision-Language-Action Models View paper
- [27] MAKE: Vision-Language Pre-training based Product Retrieval in Taobao Search View paper
- [28] Preserving and combining knowledge in robotic lifelong reinforcement learning View paper
- [29] A Novel Robotic Skill Learning Approach for Assembly Task With Dynamical System and Broad Learning View paper
- [30] A trajectory and force dual-incremental robot skill learning and generalization framework using improved dynamical movement primitives and adaptive neural â¦ View paper
- [31] SPECI: Skill Prompts based Hierarchical Continual Imitation Learning for Robot Manipulation View paper
- [32] Toward Scalable Multirobot Control: Fast Policy Learning in Distributed MPC View paper
- [33] Active incremental learning of robot movement primitives View paper
- [34] Incremental learning of subtasks from unsegmented demonstration View paper
- [35] iManip: Skill-Incremental Learning for Robotic Manipulation View paper
- [36] Analysis of methods for incremental policy refinement by kinesthetic guidance View paper