Novelty Assessment Report

Paper: Command-V: Training-Free Representation Finetuning Transfer
PDF URL: https://openreview.net/pdf?id=oRYzpI3cmJ
Venue: ICLR 2026 Conference Submission
Year: 2026
Report Generated: 2025-12-29

Abstract

Retrofitting large language models (LLMs) with new behaviors typically requires full finetuning or distillation—costly steps that must be repeated for every architecture. In this work, we introduce ⌘V (Command-V), a backpropagation-free behavior transfer method that copies an existing residual representation adapter from a donor model and pastes its effect into an architecturally different recipient model. ⌘V profiles layer activations on a small prompt set, derives linear converters between corresponding layers, and applies the donor intervention in the recipient's activation space. This process does not require access to the original training data and needs minimal compute. In three case studies—safety-refusal enhancement, jailbreak facilitation, and automatic chain-of-thought reasoning—⌘V matches the performance of direct finetuning while using orders of magnitude less resources.

Core Task Landscape

This paper addresses: **Training-Free Behavior Transfer Across Different Language Model Architectures**

A total of **46 papers** were analyzed and organized into a taxonomy with **30 categories**.

Taxonomy Overview

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

- **Cross-Architecture Adapter and Module Transfer**
- **Cross-Lingual and Multilingual Transfer**
- **Prompt-Based and Memory-Augmented Transfer**
- **Domain and Task Adaptation Without Training**
- **Behavioral and Reinforcement Learning Transfer**
- **Specialized Transfer and Distillation Methods**
- **Foundation Model Composition and Unified Frameworks**
- **Emerging and Experimental Transfer Paradigms**

Complete Taxonomy Tree

- Training-Free Behavior Transfer Across Different Language Model Architectures Survey Taxonomy
- Cross-Architecture Adapter and Module Transfer
  - Low-Rank Adapter Transfer (2 papers)
  - [1] Lora: Low-rank adaptation of large language models. (Hu, 2022) View paper
  - [13] Cross-LoRA: A Data-Free LoRA Transfer Framework across Heterogeneous LLMs (Liao Mingyang, 2025) View paper
  - Activation-Space Representation Transfer ★ (2 papers)
  - [0] Command-V: Training-Free Representation Finetuning Transfer (Anon et al., 2026) View paper
  - [23] Activation Manifold Projection: Liberating Task-Specific Behaviors from LLM Architectures (Kari, 2025) View paper
  - Linear-Cost Architecture Transfer (1 papers)
  - [11] Cross-Architecture Transfer Learning for Linear-Cost Inference Transformers (Choiï¼ Se-Hyun, 2024) View paper
- Cross-Lingual and Multilingual Transfer
  - Tokenizer and Vocabulary Transfer (1 papers)
  - [9] Zero-shot tokenizer transfer (Benjamin Minixhofer, 2024) View paper
  - Cross-Lingual Knowledge Transfer Mechanisms (3 papers)
  - [7] Towards a common understanding of contributing factors for cross-lingual transfer in multilingual language models: A review (Philippy, 2023) View paper
  - [29] Analyzing zero-shot cross-lingual transfer in supervised NLP tasks (Hyunâ□□Jin Choi, 2021) View paper
  - [45] Languages You Know Influence Those You Learn: Impact of Language Characteristics on Multi-Lingual Text-to-Text Transfer (MÃ¼ller, 2022) View paper
  - Layer-Based Cross-Lingual Transfer (2 papers)
  - [15] Layer swapping for zero-shot cross-lingual transfer in large language models (Bandarkar, 2024) View paper
  - [44] From English To Foreign Languages: Transferring Pre-trained Language Models (Tran, 2022) View paper
  - Task-Specific Cross-Lingual Transfer (3 papers)
  - [18] Zero-Shot Cross-Domain Dialogue State Tracking via Dual Low-Rank Adaptation (Luo Xiang, 2024) View paper
  - [22] Zero-shot transfer learning with synthesized data for multi-domain dialogue state tracking (Campagna, 2020) View paper
  - [30] Enhancing Small Language Models for Cross-Lingual Generalized Zero-Shot Classification with Soft Prompt Tuning (Philippy, 2025) View paper
  - Transfer Language Selection (1 papers)
  - [43] Transfer Language Selection for Zero-Shot Cross-Lingual Abusive Language Detection (Juuso Eronen, 2022) View paper

- Prompt-Based and Memory-Augmented Transfer
  - Continuous Prompt Transfer (2 papers)
  - [10] Evaluating Gender Bias Transfer between Pre-trained and Prompt-Adapted Language Models (Natalie Mackraz, 2024) View paper
  - [42] Zero-Shot Continuous Prompt Transfer: Generalizing Task Semantics Across Language Models (Wu, 2023) View paper
  - Structured Prompting and Decision Trees (1 papers)
  - [25] Tree Prompting: Efficient Task Adaptation without Fine-Tuning (Morris, 2023) View paper
  - Memory and Procedural Embedding Modules (2 papers)
  - [3] Memory decoder: A pretrained, plug-and-play memory for large language models (Cao Jia-qi, 2025) View paper
  - [24] TokMem: Tokenized Procedural Memory for Large Language Models (Wu, 2025) View paper
  - Fuzzy Logic and Adaptive Prompting (1 papers)
  - [16] A Fuzzy Logic Prompting Framework for Large Language Models in Adaptive and Uncertain Tasks (Figueiredo, 2025) View paper
- Domain and Task Adaptation Without Training
  - Domain-Specific Content Generation (1 papers)
  - [19] Cross-Domain Content Generation with Domain-Specific Small Language Models (Garg Abhinav, 2024) View paper
  - Zero-Shot Multimodal Reasoning (1 papers)
  - [8] Socratic models: Composing zero-shot multimodal reasoning with language (Zeng, 2022) View paper
  - Vision-Language Model Adaptation (2 papers)
  - [21] Test-Time Consistency in Vision Language Models (Chou, 2025) View paper
  - [28] Attn-Adapter: Attention Is All You Need for Online Few-shot Learner of Vision-Language Model (Phuoc-Nguyen Bui, 2025) View paper
  - Audio-Language Integration (1 papers)
  - [46] Acoustic Prompt Tuning: Empowering Large Language Models with Audition Capabilities (Jinhua Liang, 2023) View paper
- Behavioral and Reinforcement Learning Transfer
  - Behavioral Foundation Models (1 papers)
  - [2] Fast adaptation with behavioral foundation models (Sikchi, 2025) View paper
  - Morphology Control and Policy Transfer (2 papers)
  - [26] A System for Morphology-Task Generalization via Unified Representation and Behavior Distillation (Furuta Hiroki, 2022) View paper
  - [37] Knowledge Diversion for Efficient Morphology Control and Policy Transfer (Fu Feng, 2025) View paper
  - Language-Grounded Reinforcement Learning (2 papers)
  - [33] Grounding Open-Domain Knowledge from LLMs to Real-World Reinforcement Learning Tasks: A Survey (Haiyan Yin, 2025) View paper
  - [34] Robot Learning With Prior Knowledge: Leveraging Small Network Modules and Large Foundation Models (Jian, 2025) View paper
  - Sequential Decision-Making Transfer (1 papers)
  - [31] Transfer in sequential decision making (Luketina, 2024) View paper
- Specialized Transfer and Distillation Methods
  - Cross-Architecture Knowledge Distillation (2 papers)
  - [36] A Parameter-Efficient Approach to Distilling Large Language Models via Meta-learning (Riccardo Cantini, 2025) View paper
  - [39] Sparse Mixture of Experts Language Models Excel in Knowledge Distillation (Haiyang Xu, 2024) View paper
  - Pruning Without Fine-Tuning (1 papers)
  - [32] Pruning Pre-trained Language Models Without Fine-Tuning (Jiang Ting, 2023) View paper
  - Bias and Safety Transfer (1 papers)
  - [5] DUP: Detection-guided Unlearning for Backdoor Purification in Language Models (Hu Man, 2025) View paper
  - Content Restriction and Adaptive Control (1 papers)
  - [38] Adaptive Content Restriction for Large Language Models via Suffix Optimization (Li Yige, 2025) View paper
- Foundation Model Composition and Unified Frameworks
  - Unified Cross-Domain Foundation Models (1 papers)
  - [40] UniGraph: Learning a Unified Cross-Domain Foundation Model for Text-Attributed Graphs (Yufei He, 2024) View paper
  - Interlocutor and Context Awareness (1 papers)
  - [4] Agent-to-Agent Theory of Mind: Testing Interlocutor Awareness among Large Language Models (Choi, 2025) View paper
  - Multimodal Augmentation for Specialized Tasks (1 papers)
  - [27] Enhanced Motion Forecasting with Plug-and-Play Multimodal Large Language Models (Luo, 2025) View paper
- Emerging and Experimental Transfer Paradigms
  - Geometric and Tensorial Recomposition (2 papers)
  - [6] Simulated echo shaping in large language models via semantic phase perturbation without intermediate token realignment (Allan, 2025) View paper
  - [14] Intrinsic pattern induction in large language models through tensorial recomposition framework (Oliver, 2025) View paper
  - Synthetic Data and Simulation-Based Transfer (2 papers)
  - [20] USING SIMULATION IN THE DESIGN, TESTING, AND OPTIMIZATION OF AUTONOMY STACKS FOR AUTONOMOUS GROUND VEHICLES (Zhang, 2024) View paper
  - [41] Generating Content for AI, and AI-Generated Content Synthetic Data for and from Machine Learning (Shen, 2023) View paper
  - Domain-Specific Applications and Case Studies (3 papers)
  - [12] Quasar at AraHealthQA Track 1: Leveraging Zero-Shot Large Language Models for Question and Answer Categorization in Arabic Mental Health (Adiba Fairooz Chowdhury, 2025) View paper
  - [17] On-Device Large Language Models: A Survey of Model Compression and System Optimization (W Chen, 2025) View paper
  - [35] Exploring Different Machine Learning-Based Methods for Learning the Language of Shepna Stock Price (Zohreh Ansari, 2025) View paper

Narrative

Core task: training-free behavior transfer across different language model architectures. The field addresses how to move learned capabilities—ranging from task-specific adaptations to entire behavioral policies—from one model architecture to another without

retraining from scratch. The taxonomy reveals eight major branches that span diverse transfer mechanisms. Cross-Architecture Adapter and Module Transfer focuses on moving lightweight components such as LoRA[1] modules or activation-space representations between models with different internal structures. Cross-Lingual and Multilingual Transfer examines how linguistic knowledge propagates across language boundaries, often leveraging shared representations or tokenizer mappings like Zero-shot Tokenizer Transfer[9]. Prompt-Based and Memory-Augmented Transfer explores training-free methods that rely on external memory or carefully designed prompts to guide behavior. Domain and Task Adaptation Without Training tackles specialized settings—medical QA, robotics, or financial forecasting—where direct fine-tuning is impractical. Behavioral and Reinforcement Learning Transfer investigates policy or decision-making transfer, including works like LLMs to RL[33] that bridge language models and sequential decision tasks. Specialized Transfer and Distillation Methods cover techniques such as pruning, knowledge diversion, and meta-learning that compress or reshape models. Foundation Model Composition and Unified Frameworks study how to combine multiple pretrained models into coherent systems, exemplified by Socratic Models[8]. Finally, Emerging and Experimental Transfer Paradigms capture novel directions that do not yet fit established categories.

A particularly active line of work centers on activation-space and module-level transfer, where researchers seek to align internal representations or adapter weights across architectures with minimal overhead. Command-V[0] exemplifies this direction by transferring behavior through activation-space mappings, closely related to Activation Manifold Projection[23], which also manipulates hidden states to achieve cross-architecture compatibility. These approaches contrast with heavier distillation pipelines or prompt-engineering strategies, offering a middle ground between full retraining and purely inference-time interventions. Meanwhile, cross-lingual studies like Cross-lingual Transfer Factors[7] and behavioral policy transfer methods such as Behavioral Foundation Models[2] highlight alternative pathways: the former emphasizes linguistic structure, while the latter focuses on decision-making patterns. Open questions remain around scalability—whether activation-space methods generalize to very large or highly dissimilar architectures—and the trade-offs between transfer fidelity and computational cost. Command-V[0] sits squarely within the activation-space cluster, sharing conceptual ground with Activation Manifold Projection[23] but differing in how it handles architectural mismatches and the granularity of behavior it aims to preserve.

Related Works in Same Category

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

1. Activation Manifold Projection: Liberating Task-Specific Behaviors from LLM Architectures

Authors: Al Kari | **Year/Venue**: 2025 | **URL**: View paper

Abstract

The proliferation of Large Language Model (LLM) architectures presents a fundamental challenge: valuable, task-specific behaviors learned through fine-tuning methods like Low-Rank Adaptation (LoRA) are effectively trapped within their source model's architecture, herein referred to architectural lock-in. Existing transfer methods attempt to bridge this gap by aligning the static weight spaces of models, a brittle and indirect approach that relies on tenuous correlations between parameter geometr...

Relationship Analysis

Both papers belong to the Activation-Space Representation Transfer category, focusing on transferring task-specific behaviors by mapping activation representations between different model architectures. They share the core approach of using activation-space mappings to transfer learned behaviors (Command-V transfers ReFT adapters, CAST transfers LoRA adapters) without requiring the original training data. The key difference is that Command-V uses pseudoinverse-based linear converters derived from activation profiles on a small prompt set, while CAST learns nonlinear bidirectional projection heads trained on a general corpus with a dual-objective loss function combining KL divergence and MSE to align both output distributions and hidden state geometries.

Contributions Analysis

Overall novelty summary. The paper introduces Command-V, a method for transferring learned behaviors between architecturally different language models by mapping activation spaces without backpropagation. It resides in the 'Activation-Space Representation Transfer' leaf, which contains only one sibling paper (Activation Manifold Projection). This leaf sits within the broader 'Cross-Architecture Adapter and Module Transfer' branch, which includes three leaves total. The sparse population suggests this is an emerging rather than saturated research direction, with relatively few prior works directly addressing activation-space behavior transfer across heterogeneous architectures.

The taxonomy reveals that neighboring approaches tackle similar cross-architecture challenges through different mechanisms. The sibling leaf 'Low-Rank Adapter Transfer' focuses on projecting LoRA modules rather than activation representations, while 'Linear-Cost Architecture Transfer' addresses migration to state-space models specifically. Adjacent branches explore prompt-based methods and memory augmentation, which avoid weight-space interventions entirely. Command-V occupies a middle ground: it manipulates internal representations like prompt-based methods but operates through learned linear converters rather than external memory or discrete prompts, distinguishing it from both adapter projection and pure inference-time techniques.

Among 25 candidates examined across three contributions, none were flagged as clearly refuting the work. The 'activation profiling method' examined 7 candidates with no refutations, the 'Command-V adapter transfer framework' examined 8 with none, and the 'training-free behavior transfer method' examined 10 with none. This suggests that within the limited search scope—focused on top-K semantic matches and citation expansion—no prior work was found that directly anticipates the combination of activation profiling, linear conversion, and cross-architecture adapter transfer. The statistics indicate a relatively clean novelty signal, though the modest search scale (25 papers) means exhaustive coverage cannot be claimed.

Based on the limited literature search, Command-V appears to occupy a sparsely populated niche within activation-space transfer. The absence of refuting candidates across all contributions, combined with the leaf's small sibling count, suggests the specific approach is not directly anticipated by examined prior work. However, the search scope of 25 papers leaves open the possibility that related techniques exist in adjacent subfields or recent preprints not captured by semantic search.

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

Contribution 1: Activation profiling method

Description: The authors introduce activation profiling, a technique that records and analyzes layer activations from a small set of prompts to identify corresponding activation patterns between different transformer-based language models, enabling cross-model behavior transfer without requiring architectural similarity.

This contribution was assessed against **7 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Neuron to Graph: Interpreting Language Model Neurons at Scale

URL: View paper

Brief Assessment

Neuron to Graph[69] focuses on interpreting individual neuron behaviors within a single model by analyzing their activation patterns on dataset examples, not on identifying corresponding neurons across different transformer models for cross-model behavior transfer.

---

2. Universal neurons in gpt2 language models

URL: View paper

Brief Assessment

Universal Neurons[64] focuses on identifying universal neurons within the same model architecture (GPT2) trained from different random seeds by computing pairwise correlations of neuron activations. The original paper's activation profiling method is designed to establish correspondences between neurons across architecturally different models to enable cross-model behavior transfer, which is a distinct technical objective and methodology.

---

3. Disentangling Transformer Language Models as Superposed Topic Models

URL: View paper

Brief Assessment

Superposed Topic Models[66] focuses on disentangling topics from transformer language models by analyzing decoder weights and logit distributions, not on identifying corresponding neurons across different models for behavior transfer. The candidate's approach is fundamentally about topic modeling and interpretability within a single model, whereas the original contribution addresses cross-model neuron correspondence for activation-based behavior transfer.

---

4. Can Neuron Activation be Predicted? A New Lens for Analyzing Transformer-based LLM

URL: View paper

Brief Assessment

Neuron Activation Prediction[70] focuses on predicting neuron activations within a single model across layers using linear mappings, not on identifying corresponding neurons across different transformer models for cross-model behavior transfer.

---

5. Disentangling Recall and Reasoning in Transformer Models through Layer-wise Attention and Activation Analysis

URL: View paper

Brief Assessment

Recall and Reasoning[68] focuses on mechanistic interpretability through activation patching and ablations to identify task-specific circuits within single models, not on establishing cross-model activation correspondences for behavior transfer.

---

6. Converging to a Lingua Franca: Evolution of Linguistic Regions and Semantics Alignment in Multilingual Large Language Models

URL: View paper

Brief Assessment

Lingua Franca Evolution[65] analyzes neuron activation patterns to understand multilingual semantic alignment within single models, not to establish cross-model correspondences for behavior transfer as in the original paper's activation profiling method.

---

7. Hedonic Neurons: A Mechanistic Mapping of Latent Coalitions in Transformer MLPs

URL: View paper

Brief Assessment

Hedonic Neurons[67] focuses on coalitional game theory to identify synergistic neuron groups within MLPs for interpretability, not on cross-model activation pattern matching for behavior transfer.

Contribution 2: Command-V adapter transfer framework

Description: The authors develop Command-V, a training-free framework that uses activation profiles to derive linear converters between model layers, allowing representation adapter weights from a donor model to be transferred and applied to an architecturally different recipient model without backpropagation or additional training data.

This contribution was assessed against **8 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Parameter Efficient Mamba Tuning via Projector-targeted Diagonal-centric Linear Transformation

URL: View paper

Brief Assessment

Projector-targeted Mamba[48] focuses on parameter-efficient fine-tuning within a single Mamba architecture by optimizing projectors, not on transferring adapter weights between architecturally different models without training as Command-V does.

---

2. Command-V: Pasting LLM Behaviors via Activation Profiles

URL: View paper

Brief Assessment

Command-V Activation Profiles[50] presents the same method as the original paper. Both describe identical activation profiling and linear converter derivation approaches for transferring representation adapters between architecturally different models without training. This is the same work, not prior art that refutes novelty.

---

3. Language Fusion for Parameter-Efficient Cross-lingual Transfer

URL: View paper

Brief Assessment

Language Fusion[47] focuses on enhancing cross-lingual transfer within adapters by fusing source and target language representations, not on transferring adapter weights between architecturally different models without training.

---

4. Parameter-efficient Dysarthric Speech Recognition Using Adapter Fusion and Householder Transformation

URL: View paper

Brief Assessment

Dysarthric Speech Adapter[49] focuses on adapter fusion for dysarthric speech recognition within a single model architecture, not on transferring adapter weights between architecturally different models without training as Command-V does.

---

5. Linear fine-tuning: a linear transformation based transfer strategy for deep MRI reconstruction

URL: View paper

Brief Assessment

Linear MRI Transfer[52] focuses on MRI reconstruction using linear transformations (scaling/shifting factors) for domain adaptation in medical imaging. The original paper addresses transferring representation adapters between LLMs using activation profiles and linear converters without training, which is a fundamentally different application domain and technical approach.

---

6. Cross-LoRA: A Data-Free LoRA Transfer Framework across Heterogeneous LLMs

URL: View paper

Brief Assessment

Cross-LoRA[13] focuses on transferring LoRA modules between heterogeneous LLMs using SVD-based subspace alignment, while the original paper transfers representation adapters (ReFT) using activation-based linear converters derived from profiling.

---

7. X-Adapter: Adding Universal Compatibility of Plugins for Upgraded Diffusion Model

URL: View paper

Brief Assessment

X-Adapter[53] focuses on transferring pretrained plugins (ControlNet, LoRA) between different versions of diffusion models for image generation, not on transferring representation adapter weights between LLMs using activation profiles and linear converters without backpropagation.

---

8. Enhancing Neural Network Efficiency with Streamlined Pruned Linear Adapters

URL: View paper

Brief Assessment

Streamlined Pruned Adapters[51] focuses on parameter-efficient adapter design through pruning and linear transformations within a single model architecture (BERT/RoBERTa), not on transferring adapter weights between architecturally different models without training.

Contribution 3: Training-free behavior transfer method

Description: The authors present Command-V as a complete method that transfers learned behaviors across different model architectures by profiling activations, deriving converters, and applying donor interventions in the recipient's activation space, requiring minimal compute and no access to original training data.

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. Adapterem: pre-trained language model adaptation for generalized entity matching using adapter-tuning

URL: View paper

Brief Assessment

AdapterEM[60] focuses on parameter-efficient fine-tuning for entity matching tasks using adapters that require training/optimization. The original paper presents a training-free method for transferring behaviors across architecturally different models without any parameter updates, which is fundamentally different from AdapterEM's approach of optimizing adapter weights.

---

2. X-LoRA: Mixture of Low-Rank Adapter Experts, a Flexible Framework for Large Language Models with Applications in Protein Mechanics and Design

URL: View paper

Brief Assessment

X-LoRA[59] focuses on mixture-of-experts architecture for domain-specific scientific tasks using LoRA adapters, not on training-free transfer of behaviors across architecturally different models via activation profiling and representation adapters.

---

3. SUR-adapter: Enhancing Text-to-Image Pre-trained Diffusion Models with Large Language Models

URL: View paper

Brief Assessment

SUR-adapter[61] focuses on enhancing text-to-image diffusion models through adapter fine-tuning with knowledge distillation from LLMs, not on training-free behavior transfer across architecturally different language models using representation adapters.

---

4. Boosting Continual Learning of Vision-Language Models via Mixture-of-Experts Adapters

URL: View paper

Brief Assessment

MoE Adapters[54] focuses on continual learning for vision-language models with parameter-efficient adapters, not on training-free cross-architecture behavior transfer using activation profiling and converters as in the original paper.

---

5. Self-expansion of pre-trained models with mixture of adapters for continual learning

URL: View paper

Brief Assessment

Self-expansion Adapters[58] focuses on continual learning with dynamic adapter expansion based on distribution shift detection, not on training-free transfer of behaviors across different model architectures. The candidate requires training adapters on new tasks, whereas the original paper transfers pre-trained adapters without any training.

---

6. Vl-adapter: Parameter-efficient transfer learning for vision-and-language tasks

URL: View paper

Brief Assessment

VL-Adapter[56] focuses on parameter-efficient transfer learning for vision-and-language tasks through adapter modules that require training. The original paper presents a training-free method for transferring behaviors across architecturally different models using activation profiling and linear converters, which is fundamentally different from VL-Adapter's approach of training adapter parameters on downstream tasks.

---

7. FedITD: A Federated Parameter-Efficient Tuning With Pre-Trained Large Language Models and Transfer Learning Framework for Insider Threat Detection

URL: View paper

Brief Assessment

FedITD[62] focuses on federated learning for insider threat detection using parameter-efficient tuning methods with LLMs. It does not address training-free behavior transfer across architecturally different models via representation adapters, which is the core novelty of the original paper's Command-V method.

---

8. Energy-efficient task adaptation for nlp edge inference leveraging heterogeneous memory architectures

URL: View paper

Brief Assessment

Energy-efficient Task Adaptation[63] focuses on adapter-based parameter-efficient fine-tuning for NLP edge inference with heterogeneous memory architectures, not on training-free cross-architecture behavior transfer using activation profiling and representation converters.

---

9. LoRA-X: Bridging Foundation Models with Training-Free Cross-Model Adaptation

URL: View paper

Brief Assessment

LoRA-X[57] focuses on transferring parameter-efficient adapters (LoRA weights) across different base models for text-to-image generation, not on transferring learned behaviors across language models using activation-based representation adapters as in the original paper.

---

10. Graphadapter: Tuning vision-language models with dual knowledge graph

URL: View paper

Brief Assessment

GraphAdapter[55] focuses on adapter-style tuning for vision-language models using dual knowledge graphs for few-shot classification tasks. It does not address training-free behavior transfer across architecturally different language models using representation adapters, which is the core contribution of the original paper.

Appendix: Text Similarity Detection

Textual similarity detection checked 26 papers and found 3 similarity segment(s) across 1 paper(s).

The following **1 paper(s)** were detected to have high textual similarity with the original paper. These may represent different versions of the same work, duplicate submissions, or papers with substantial textual overlap. Readers are advised to verify these relationships independently.

1. Command-V: Pasting LLM Behaviors via Activation Profiles

Detected in: Contribution: contribution_2

⚠ **Note**: This paper shows substantial textual similarity with the original paper. It may be a different version, a duplicate submission, or contain significant overlapping content. Please review carefully to determine the nature of the relationship.

References

- [0] Command-V: Training-Free Representation Finetuning Transfer View paper
- [1] Lora: Low-rank adaptation of large language models. View paper
- [2] Fast adaptation with behavioral foundation models View paper
- [3] Memory decoder: A pretrained, plug-and-play memory for large language models View paper
- [4] Agent-to-Agent Theory of Mind: Testing Interlocutor Awareness among Large Language Models View paper
- [5] DUP: Detection-guided Unlearning for Backdoor Purification in Language Models View paper
- [6] Simulated echo shaping in large language models via semantic phase perturbation without intermediate token realignment View paper
- [7] Towards a common understanding of contributing factors for cross-lingual transfer in multilingual language models: A review View paper
- [8] Socratic models: Composing zero-shot multimodal reasoning with language View paper
- [9] Zero-shot tokenizer transfer View paper
- [10] Evaluating Gender Bias Transfer between Pre-trained and Prompt-Adapted Language Models View paper
- [11] Cross-Architecture Transfer Learning for Linear-Cost Inference Transformers View paper
- [12] Quasar at AraHealthQA Track 1: Leveraging Zero-Shot Large Language Models for Question and Answer Categorization in Arabic Mental Health View paper
- [13] Cross-LoRA: A Data-Free LoRA Transfer Framework across Heterogeneous LLMs View paper
- [14] Intrinsic pattern induction in large language models through tensorial recomposition framework View paper
- [15] Layer swapping for zero-shot cross-lingual transfer in large language models View paper
- [16] A Fuzzy Logic Prompting Framework for Large Language Models in Adaptive and Uncertain Tasks View paper
- [17] On-Device Large Language Models: A Survey of Model Compression and System Optimization View paper
- [18] Zero-Shot Cross-Domain Dialogue State Tracking via Dual Low-Rank Adaptation View paper
- [19] Cross-Domain Content Generation with Domain-Specific Small Language Models View paper
- [20] USING SIMULATION IN THE DESIGN, TESTING, AND OPTIMIZATION OF AUTONOMY STACKS FOR AUTONOMOUS GROUND VEHICLES View paper
- [21] Test-Time Consistency in Vision Language Models View paper
- [22] Zero-shot transfer learning with synthesized data for multi-domain dialogue state tracking View paper
- [23] Activation Manifold Projection: Liberating Task-Specific Behaviors from LLM Architectures View paper
- [24] TokMem: Tokenized Procedural Memory for Large Language Models View paper
- [25] Tree Prompting: Efficient Task Adaptation without Fine-Tuning View paper
- [26] A System for Morphology-Task Generalization via Unified Representation and Behavior Distillation View paper
- [27] Enhanced Motion Forecasting with Plug-and-Play Multimodal Large Language Models View paper
- [28] Attn-Adapter: Attention Is All You Need for Online Few-shot Learner of Vision-Language Model View paper
- [29] Analyzing zero-shot cross-lingual transfer in supervised NLP tasks View paper
- [30] Enhancing Small Language Models for Cross-Lingual Generalized Zero-Shot Classification with Soft Prompt Tuning View paper
- [31] Transfer in sequential decision making View paper
- [32] Pruning Pre-trained Language Models Without Fine-Tuning View paper
- [33] Grounding Open-Domain Knowledge from LLMs to Real-World Reinforcement Learning Tasks: A Survey View paper

- [34] Robot Learning With Prior Knowledge: Leveraging Small Network Modules and Large Foundation Models View paper
- [35] Exploring Different Machine Learning-Based Methods for Learning the Language of Shepna Stock Price View paper
- [36] A Parameter-Efficient Approach to Distilling Large Language Models via Meta-learning View paper
- [37] Knowledge Diversion for Efficient Morphology Control and Policy Transfer View paper
- [38] Adaptive Content Restriction for Large Language Models via Suffix Optimization View paper
- [39] Sparse Mixture of Experts Language Models Excel in Knowledge Distillation View paper
- [40] UniGraph: Learning a Unified Cross-Domain Foundation Model for Text-Attributed Graphs View paper
- [41] Generating Content for AI, and AI-Generated Content Synthetic Data for and from Machine Learning View paper
- [42] Zero-Shot Continuous Prompt Transfer: Generalizing Task Semantics Across Language Models View paper
- [43] Transfer Language Selection for Zero-Shot Cross-Lingual Abusive Language Detection View paper
- [44] From English To Foreign Languages: Transferring Pre-trained Language Models View paper
- [45] Languages You Know Influence Those You Learn: Impact of Language Characteristics on Multi-Lingual Text-to-Text Transfer View paper
- [46] Acoustic Prompt Tuning: Empowering Large Language Models with Audition Capabilities View paper
- [47] Language Fusion for Parameter-Efficient Cross-lingual Transfer View paper
- [48] Parameter Efficient Mamba Tuning via Projector-targeted Diagonal-centric Linear Transformation View paper
- [49] Parameter-efficient Dysarthric Speech Recognition Using Adapter Fusion and Householder Transformation View paper
- [50] Command-V: Pasting LLM Behaviors via Activation Profiles View paper
- [51] Enhancing Neural Network Efficiency with Streamlined Pruned Linear Adapters View paper
- [52] Linear fine-tuning: a linear transformation based transfer strategy for deep MRI reconstruction View paper
- [53] X-Adapter: Adding Universal Compatibility of Plugins for Upgraded Diffusion Model View paper
- [54] Boosting Continual Learning of Vision-Language Models via Mixture-of-Experts Adapters View paper
- [55] Graphadapter: Tuning vision-language models with dual knowledge graph View paper
- [56] Vl-adapter: Parameter-efficient transfer learning for vision-and-language tasks View paper
- [57] LoRA-X: Bridging Foundation Models with Training-Free Cross-Model Adaptation View paper
- [58] Self-expansion of pre-trained models with mixture of adapters for continual learning View paper
- [59] X-LoRA: Mixture of Low-Rank Adapter Experts, a Flexible Framework for Large Language Models with Applications in Protein Mechanics and Design View paper
- [60] Adapterem: pre-trained language model adaptation for generalized entity matching using adapter-tuning View paper
- [61] SUR-adapter: Enhancing Text-to-Image Pre-trained Diffusion Models with Large Language Models View paper
- [62] FedITD: A Federated Parameter-Efficient Tuning With Pre-Trained Large Language Models and Transfer Learning Framework for Insider Threat Detection View paper
- [63] Energy-efficient task adaptation for nlp edge inference leveraging heterogeneous memory architectures View paper
- [64] Universal neurons in gpt2 language models View paper
- [65] Converging to a Lingua Franca: Evolution of Linguistic Regions and Semantics Alignment in Multilingual Large Language Models View paper
- [66] Disentangling Transformer Language Models as Superposed Topic Models View paper
- [67] Hedonic Neurons: A Mechanistic Mapping of Latent Coalitions in Transformer MLPs View paper
- [68] Disentangling Recall and Reasoning in Transformer Models through Layer-wise Attention and Activation Analysis View paper
- [69] Neuron to Graph: Interpreting Language Model Neurons at Scale View paper
- [70] Can Neuron Activation be Predicted? A New Lens for Analyzing Transformer-based LLM View paper