CF-ROUTER: CLOSED-FORM SOLUTION FOR EXPERT SELECTION IN MULTIMODAL AGENT LIFELONG LEARNING

Jiaxu Li, Zhijie Zheng, Jianyu Qi, Rongchang Zhao
Central South University

ABSTRACT

Multimodal Large Language Models (MLLMs) are increasingly pivotal as lifelong learning agents, tasked with adapting to evolving environments without succumbing to catastrophic forgetting. Current strategies often leverage Mixture-of-Experts (MoE) architectures combined with Low-Rank Adaptation (LoRA) to compartmentalize domain-specific knowledge. However, prevailing routing mechanisms—whether relying on MLLM prompting or heuristic similarity metrics—frequently suffer from low training efficiency or poor alignment within complex multimodal feature spaces. To address these limitations, we introduce **CF-Router**, a novel routing framework grounded in **a closed-form solution**. By leveraging the average-pooled hidden states from the MLLM's final layer as representative semantic descriptors, we employ a regularized least-squares classifier to precisely identify the optimal expert LoRA. This methodology facilitates analytic, mathematically optimal updates, guaranteeing rapid task identification and seamless adaptation for lifelong learning agents.cf

1 INTRODUCTION

The pursuit of autonomous agents capable of *lifelong learning* remains a cornerstone objective in artificial intelligence (De Lange & Tuytelaars, 2021; He et al., 2023; Cossu et al., 2024). These agents are required to continuously acquire novel skills and adapt to dynamic domains without erasing previously established knowledge, a challenge known as catastrophic forgetting (McCloskey & Cohen, 1989; Ratcliff, 1990). Multimodal Large Language Models (MLLMs), with their robust generalized reasoning, serve as ideal backbones for such systems (Bai et al., 2025; Li et al., 2024; Chen et al., 2024b). However, conventional sequential fine-tuning of MLLMs typically results in significant performance erosion on earlier tasks (Guo et al., 2025a; Chen et al., 2025).

To counteract this, parameter-efficient continual learning frameworks, notably those integrating MoE with LoRA, have shown promise. A prominent example, Zhao et al. (2025), dedicates distinct LoRA experts to specific domains. The efficacy of this architecture hinges on the **router**, which is responsible for selecting the appropriate expert for a given input. Nevertheless, existing routers face substantial drawbacks: similarity-based approaches (e.g., utilizing cosine similarity of prototypes) often struggle with the non-linear boundaries of multimodal distributions (Chen et al., 2024a; Huai et al., 2025), while MLLM-based routers, which rely on the model for self-selection, introduce high training overhead and potential privacy concerns if replay-based methods are used (Zhao et al., 2025), rendering them impractical for real-time applications.

In this work, we present a scalable and efficient solution: the **Closed-Form Router (CF-Router)**. Our approach is distinguished by two key attributes: it is **analytic-based** and **replay-free**. The analytic nature ensures that router updates are derived via closed-form solution rather than stochastic gradient descent (SGD), thereby avoiding interference with weights from prior tasks and eliminating a primary source of catastrophic forgetting (Wang et al., 2024). The replay-free design addresses the storage and privacy constraints inherent to MLLMs (Zhao et al., 2025). Storing high-resolution images or lengthy text instructions demands excessive storage, which is unsuitable for edge or large-scale deployments. Furthermore, given that MLLMs are trained on trillion-scale datasets, replaying a limited subset of samples is often insufficient to preserve the model's vast knowledge base and risks overfitting.

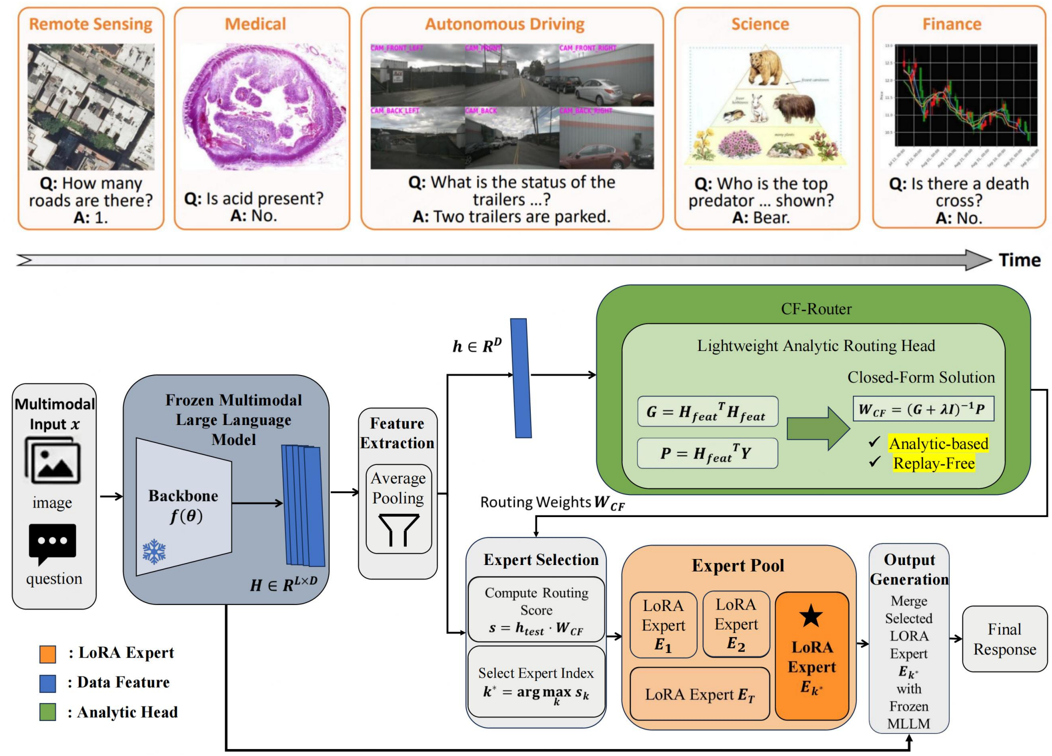

Figure 1: **Overview of the proposed CF-Router framework.** *Top:* The lifelong learning scenario, where diverse domain tasks—Remote Sensing, Medical, Autonomous Driving, Science, and Finance—arrive sequentially over time, each with domain-specific visual question-answering data (adapted from Zhao et al. (2025)). *Bottom:* The CF-Router architecture. A multimodal input x (image and question) is processed by the frozen MLLM backbone $f(\theta)$, producing hidden states $H \in \mathbb{R}^{L\times D}$. Average pooling condenses the sequence into a fixed-size feature $h \in \mathbb{R}^D$. During *training*, the CF-Router incrementally accumulates the Gram matrix $G = H_{feat}^T H_{feat}$ and correlation matrix $P = H_{feat}^T Y$, then analytically computes the routing weights $W_{CF} = (G + \lambda I)^{-1} P$ without any gradient descent. During *inference*, the routing score $s = h_{test} \cdot W_{CF}$ is computed and the expert with the highest score $k^* = \arg\max_k s_k$ is selected from the expert pool. The chosen LoRA expert E_{k^*} is merged with the frozen backbone to generate the final response. This design is both **analytic-based** (no backpropagation for routing) and **replay-free** (no storage of previous task data).

Technically, our method extracts rich semantic representations from the frozen MLLM's final layer, applies average pooling to condense the sequence, and maps these features to expert indices via an analytically solved linear head.

Our contributions are as follows:

- We identify the bottleneck in existing MLLM continual learning routers and propose the integration of a closed-form solution to decouple expert selection from gradient-based optimization.
- We introduce the CF-Router architecture, which utilizes the MLLM's last-layer hidden states to solve the routing problem analytically, ensuring rapid adaptation with zero backpropagation costs for the router.
- We frame this contribution within the context of Agent Lifelong Learning, enabling agents to instantly index and retrieve domain-specific skills.

2 METHODOLOGY

An overview of our proposed framework is illustrated in Figure 1. Our goal is to enable an MLLM agent to learn a sequence of tasks $\mathcal{T} = \{1, \ldots, T\}$ sequentially. We adopt an architecture where a shared frozen MLLM backbone f_θ is augmented with a pool of trainable LoRA experts $\mathcal{E} = \{E_1, \ldots, E_T\}$, where each expert E_t specializes in a specific task or domain. The core innovation of our work is the mechanism for selecting the appropriate expert E_k for a test instance x using a closed-form solution.

2.1 PRELIMINARIES: MLLM-CL

The MLLM-CL framework prevents forgetting by isolating parameters, assigning a specific LoRA module to each task. During inference, a router must predict the task identity k to activate the correct expert. We propose to solve this task identity prediction using a regularized least-squares (Ridge Regression) approach applied directly to the MLLM's hidden features.

2.2 THE CF-ROUTER ARCHITECTURE

The proposed CF-Router operates in two distinct stages: Feature Extraction and Closed-Form Decision.

2.2.1 FEATURE EXTRACTION VIA HIDDEN STATES

Given a multimodal input x (comprising image embeddings and text instructions), the frozen MLLM backbone processes the sequence. Let $H \in \mathbb{R}^{L \times D}$ denote the hidden states of the last transformer layer of the MLLM, where L is the sequence length and D is the hidden dimension. To obtain a fixed-size representation suitable for the analytic solver, we apply average pooling across the sequence dimension:

$$h = \frac{1}{L} \sum_{i=1}^{L} H_i \tag{1}$$

where $h \in \mathbb{R}^D$ serves as the robust semantic signature of the current input context. In our implementation, $D = 4096$, corresponding to the hidden dimension of the MLLM's last layer.

2.2.2 CLOSED-FORM EXPERT SELECTION

The router's objective is to map the semantic feature h extracted from the backbone directly to an expert index $y \in \{1, \ldots, T\}$. We treat this as a regression problem where the target is the one-hot encoding of the expert index. The optimal weights $W_{CF} \in \mathbb{R}^{D \times T}$ for the routing head are computed by minimizing the regularized mean squared error:

$$W_{CF} = \arg\min_W \|H_{feat} W - Y\|_F^2 + \lambda \|W\|_F^2 \tag{2}$$

where H_{feat} is the matrix of collected hidden features from observed tasks, Y is the matrix of expert labels, and λ is a regularization parameter. Following McDonnell et al. (2023), the solution is given by the closed-form equation:

$$W_{CF} = (H_{feat}^T H_{feat} + \lambda I)^{-1} H_{feat}^T Y \tag{3}$$

In a lifelong learning scenario, we do not need to store all raw data H_{feat}. Instead, we incrementally update the Gram matrix $G = H_{feat}^T H_{feat}$ and the correlation matrix $P = H_{feat}^T Y$. Let $H_{feat}^{(t)} \in \mathbb{R}^{N_t \times D}$ denote the feature matrix and $Y^{(t)} \in \mathbb{R}^{N_t \times T}$ the label matrix of task t. When a new task t arrives, G and P are updated via simple accumulation:

$$G^{(t)} = G^{(t-1)} + {H_{feat}^{(t)}}^T H_{feat}^{(t)} = \sum_{i=1}^{t} {H_{feat}^{(i)}}^T H_{feat}^{(i)} \tag{4}$$

$$P^{(t)} = P^{(t-1)} + {H_{feat}^{(t)}}^T Y^{(t)} = \sum_{i=1}^{t} {H_{feat}^{(i)}}^T Y^{(i)} \tag{5}$$

Algorithm 1 Lifelong Learning with CF-Router (Training Phase)

Require: Sequence of tasks $\mathcal{T} = \{1, \dots, T\}$
Require: Frozen MLLM backbone f_θ
Require: Regularization λ, Buffer for G, P
for each task $t \in \mathcal{T}$ with dataset $\mathcal{D}_t$ **do**
  Initialize new LoRA expert E_t
  {Phase 1: Train Router}
  for each batch (x, y) in $\mathcal{D}_t$ **do**
    Extract last-layer hidden states: $H = f_\theta(x)$
    Average Pooling: $h = \text{AvgPool}(H)$
    Construct one-hot target y_{router} for task t
    Update Gram matrix: $G \leftarrow G + h^T h$
    Update Correlation matrix: $P \leftarrow P + h^T y_{router}$
  end for
  {Phase 2: Train Expert}
  for each batch (x, y) in $\mathcal{D}_t$ **do**
    Forward pass with Expert E_t: $\hat{y} = f_{\theta, E_t}(x)$
    Compute LLM Loss: $\mathcal{L} = \text{CrossEntropy}(\hat{y}, y)$
    Update E_t via backpropagation (keep f_θ frozen)
  end for
  {Phase 3: Update Router Weights}
  $W_{CF} = (G + \lambda I)^{-1} P$
  Store/Freeze Expert E_t
end for

The router weights are then recomputed analytically:

$$W_{CF} = (G^{(t)} + \lambda I)^{-1} P^{(t)} \tag{6}$$

Since the update rule is based on summation, which is both commutative and associative, the final result is independent of task ordering. This process is computationally efficient and requires no gradient descent.

2.3 Inference and Routing

During inference, for an incoming query x_{test}, we extract its pooled feature h_{test}. The router computes the relevance scores s for all available experts:

$$s = h_{test} W_{CF} \tag{7}$$

The system selects the expert corresponding to the maximum score:

$$k^* = \arg\max_k s_k \tag{8}$$

The corresponding LoRA module E_{k^*} is then merged with the backbone to generate the final response. This ensures that the agent leverages the precise domain knowledge required for the current interaction without interference from other domains. To enhance computational efficiency during inference, we adopt a cache-based optimization strategy inspired by the design of MLLM-CL (Zhao et al., 2025): a single forward pass is executed through the backbone network (vision encoder and language model) to cache the hidden states of each layer (including KV Cache). Subsequent lightweight components, such as the CF-Router and task-specific experts, directly reuse these cached states, thereby avoiding redundant computations.

We present the training and inference procedures for our Continual Learning framework with the Closed-Form Router (CF-Router). The training process decouples the optimization of domain-specific experts (via Gradient Descent) from the router updates (via Analytic Solution), ensuring efficiency and stability.

Algorithm 2 Inference with CF-Router

Require: Test instance x_{test}
Require: Frozen MLLM f_θ, Router W_{CF}
Require: Pool of Experts $\mathcal{E} = \{E_1, \dots, E_T\}$
1: {Step 1: Router Decision}
2: Extract last-layer hidden states: $H = f_\theta(x_{test})$
3: Average Pooling: $h = \text{AvgPool}(H)$
4: Calculate Scores: $s = hW_{CF}$
5: Select Expert Index: $k^* = \arg\max_k s_k$
6: {Step 2: Generation}
7: Activate LoRA Expert E_{k^*}
8: Generate Response: $y_{pred} = f_{\theta, E_{k^*}}(x_{test})$
9: **return** y_{pred}

3 Experiment

3.1 Experimental Setup and Datasets

To evaluate the effectiveness of CF-Router in multimodal lifelong learning, we adopt the Domain Continual Learning (DCL) benchmark proposed in MLLM-CL (Zhao et al., 2025). The benchmark consists of five diverse and challenging domains:

1. **Remote Sensing (RS)**: Focused on satellite imagery understanding and spatial reasoning.
2. **Medical (Med)**: Comprising clinical images and pathology slides for diagnostic question answering.
3. **Autonomous Driving (AD)**: Involving ego-centric view analysis and traffic scene understanding.
4. **Science (Sci)**: Covering multi-disciplinary scientific diagrams and textbook problems.
5. **Finance (Fin)**: Including chart parsing and financial report interpretation tasks.

For each domain, the agent learns a specific task sequentially. We utilize **LLaVA-1.5-7B** (Liu et al., 2024) and **InternVL-Chat-V1.0** (Chen et al., 2024b) as our base MLLM architectures. CF-Router adopts the same expert descriptions and task definitions as the MLLM-based router in MLLM-CL (Zhao et al., 2025). The detailed prompt template for expert selection is provided in Figure 2.

3.2 Implementation Details and Hyperparameters

Our framework is implemented using PyTorch. In the CF-Router, we directly use the average-pooled hidden states of the MLLM's last layer (dimension $D = 4096$) as the input to the closed-form solver. The ridge regression regularization coefficient is set to $\lambda = 0.1$. All other hyperparameters for the Expert LoRAs and the training procedure (including LoRA rank, scaling factor, optimizer, learning rate, and batch size) follow the standard configuration of MLLM-CL (Zhao et al., 2025).

3.3 Evaluation Metrics

To quantify the performance of multimodal lifelong learning, we report four key metrics. Let $a_{i,j}$ denote the performance (e.g., accuracy) on task j after the agent has finished learning task i, and T be the total number of tasks.

- **MAA** (Mean Average Accuracy): The average performance across all tasks observed so far at each curriculum stage:

$$\text{MAA} = \frac{1}{T} \sum_{i=1}^{T} \left(\frac{1}{i} \sum_{j=1}^{i} a_{i,j} \right) \tag{9}$$

```
You are a helpful assistant router.  There are five expert models,
each specializing in one of the following domains:  finance (stock),
science, medical imaging, autonomous driving, and remote sensing.

Your task is to select the most suitable model based on the provided
visual content, user question, and model descriptions.  Consider the
expertise of each model carefully and select the one best equipped
to handle the given question.

Important Instructions:
 • Respond only with the letter (A, B, C, D, E) corresponding to the
   most suitable model.
 • Do not attempt to answer the user's question directly.
Model Pool:
 • A: A financial expert specializing in stock market analysis using
   candlestick charts.  This model excels at trend prediction and
   technical indicator analysis.
 • B: A science expert with proficiency in biology, map
   interpretation, physics, and chemistry.
 • C: A medical imaging expert, primarily focused on pathology,
   including cell sections and natural images of medical conditions.
 • D: An autonomous driving expert specializing in ego-view scene
   understanding, including coordinate prediction and action
   planning and other driving-related tasks.  The input image is
   an image concatenated by 6 camera views.
 • E: A remote sensing expert, adept at analyzing aerial or
   satellite images.  This model excels at object counting, presence
   detection, and area estimation.
Here is the user's question:  [User's Question]
```

Figure 2: Prompt of the router selector.

- **MFT** (Mean Final Task): The average performance across all tasks after the entire sequence of T tasks has been learned:

$$\text{MFT} = \frac{1}{T}\sum_{j=1}^{T} a_{T,j} \tag{10}$$

- **MFN** (Mean Final New): The average performance on each task immediately after it is learned:

$$\text{MFN} = \frac{1}{T}\sum_{i=1}^{T} a_{i,i} \tag{11}$$

- **BWT** (Backward Transfer): The average change in performance on previously learned tasks after acquiring new knowledge. A BWT of 0 indicates zero forgetting:

$$\text{BWT} = \frac{1}{T-1}\sum_{j=1}^{T-1} (a_{T,j} - a_{j,j}) \tag{12}$$

3.4 Main Results and Comparative Analysis

Table 1 and Table 2 present the performance of CF-Router compared to various state-of-the-art continual learning methods on the LLaVA-1.5 and InternVL backbones, respectively. Several observations can be made:

Superior Overall Performance: Our method consistently outperforms all baselines across every domain and metric. For example, on LLaVA-1.5, we achieve an MAA of 71.29%, which is significantly higher than the best baseline DISCO* (64.92%). In several domains like Remote Sensing (RS) and

Table 1: Results for LLaVA-1.5-based domain continual learning in MLLM-CL benchmark. Baseline results are quoted from (Zhao et al., 2025). * denotes the original method with replay data.

	Task Performance					Aggregate Metrics			
Method	RS	Med	AD	Sci	Fin	MFT↑	MFN↑	MAA↑	BWT↑
Zeroshot	32.29	28.28	15.59	35.55	62.56	34.85	-	-	-
Oracle	81.06	65.83	54.17	56.86	91.14	69.81	-	-	-
LoRA-FT (Hu et al., 2021)	69.65	41.59	25.43	40.88	87.45	64.98	53.00	61.13	-14.97
LoRA-FT* (Hu et al., 2021)	76.54	50.27	43.01	43.32	89.85	66.32	60.60	64.72	-7.15
O-LoRA (Wang et al., 2023)	74.64	44.42	30.02	41.47	87.15	65.16	55.54	62.12	-12.03
O-LoRA* (Wang et al., 2023)	76.94	41.17	34.18	39.61	83.22	60.49	55.02	60.73	-6.83
MoELoRA (Chen et al., 2024a)	77.54	41.85	27.62	40.13	86.75	64.94	54.78	61.76	-12.70
MoELoRA* (Chen et al., 2024a)	77.63	49.54	39.08	41.04	89.21	66.24	59.30	64.81	-8.68
CL-MoE (Huai et al., 2025)	71.34	46.84	26.33	41.17	88.74	66.06	54.88	61.79	-13.96
CL-MoE* (Huai et al., 2025)	76.58	52.31	39.65	45.64	90.21	66.65	60.88	64.95	-7.22
HiDe (Guo et al., 2025a)	74.31	48.95	33.21	38.54	81.55	60.77	55.31	60.68	-6.82
HiDe* (Guo et al., 2025a)	74.80	42.29	34.03	38.01	79.22	60.83	53.67	61.81	-8.95
SEFE (Chen et al., 2025)	77.26	50.37	37.21	40.87	86.82	65.01	58.51	63.63	-8.13
SEFE* (Chen et al., 2025)	78.43	52.85	46.21	47.76	89.33	66.89	62.92	66.51	-4.97
DISCO (Guo et al., 2025b)	76.03	45.20	43.79	42.33	88.95	64.43	59.26	63.35	-6.46
DISCO* (Guo et al., 2025b)	77.78	46.25	50.45	49.51	89.71	65.27	62.74	64.92	-3.17
Ours	**81.28**	**65.61**	**54.68**	**56.89**	**91.14**	**69.92**	**69.92**	**71.29**	**0.00**

Table 2: Results for InternVL-Chat-V1.0-based domain continual learning in MLLM-CL benchmark. Baseline results are quoted from (Zhao et al., 2025). * denotes the original method with replay data.

	Task Performance					Aggregate Metrics			
Method	RS	Med	AD	Sci	Fin	MFT↑	MFN↑	MAA↑	BWT↑
Zeroshot	31.16	29.81	14.06	33.93	64.32	34.66	-	-	-
Oracle	81.49	66.42	54.56	54.48	91.24	69.64	-	-	-
LoRA-FT (Hu et al., 2021)	69.93	52.17	33.04	42.67	91.07	69.06	57.78	65.22	-14.11
LoRA-FT* (Hu et al., 2021)	77.06	47.55	42.67	43.31	**91.44**	**69.43**	60.41	67.45	-11.28
MoELoRA (Chen et al., 2024a)	69.90	52.08	33.17	42.19	90.58	68.83	57.58	65.97	-14.06
MoELoRA* (Chen et al., 2024a)	76.74	52.65	38.81	42.15	89.84	67.90	60.04	66.01	-9.83
HiDe (Guo et al., 2025a)	75.40	57.66	36.73	41.48	88.59	65.26	59.97	65.94	-6.60
HiDe* (Guo et al., 2025a)	53.17	52.61	40.85	47.04	89.17	64.20	56.57	61.06	-9.54
DISCO (Guo et al., 2025b)	75.12	50.69	52.41	50.67	90.86	68.85	63.95	68.14	-6.12
DISCO* (Guo et al., 2025b)	77.90	47.50	49.13	49.37	90.92	68.55	62.96	67.81	-6.98
Ours	**81.52**	**64.42**	**54.62**	**53.12**	91.10	68.96	**68.96**	**70.74**	**0.00**

Science (Sci), our performance matches the Oracle (sequential fine-tuning on each task individually with full access to data), demonstrating the efficacy of our expert isolation and precise routing.

Zero Forgetting (BWT=0): A standout result is the BWT of 0.00 across all experiments. While baseline methods like MoELoRA and SEFE suffer from negative backward transfer (ranging from -14.97 to -3.17), CF-Router completely eliminates catastrophic forgetting. This is because our architecture stores domain-specific knowledge in isolated LoRA experts, and our closed-form router ensures that the correct expert is indexed without interference from subsequent tasks.

Independence from Replay Data: Most high-performing baselines (marked with *) rely on replayed data from previous tasks to maintain performance. In contrast, CF-Router achieves superior results without any data replay, making it more privacy-preserving and storage-efficient for real-world agentic deployments.

3.5 Routing Analysis and Order Invariance

The core strength of CF-Router lies in its ability to accurately identify task identities in an analytic-based manner. As shown in Table 3 and Table 4, the routing accuracy remains near 100% across all domains. This high precision is attributed to the rich semantic features present in the MLLM's hidden states, which provide sufficient discriminative power to distinguish between task domains in the multimodal feature space.

Mathematical Order Invariance: A unique property of our closed-form solution is its robustness to task sequence. Traditional gradient-based routers (e.g., in MoELoRA) are sensitive to the order of tasks due to the sequential nature of backpropagation. In contrast, our router's weights W_{CF} are derived from the matrices G and P, which are updated via simple summation (Algorithm 1). Since summation is commutative, the final W_{CF} is mathematically identical regardless of the order in which task statistics are accumulated. Our experiments over 120 task permutations confirm this, with a standard deviation of 0.000 for all metrics.

3.6 Computational Efficiency Analysis

The CF-Router provides significant computational advantages over gradient-based or prompt-based routing mechanisms. We analyze its efficiency in terms of time and space complexity.

Time Complexity: The computational cost is divided into the incremental update phase and the inference phase.

1. **Update Phase:** For a task with N_t samples, the time complexity for accumulating the matrices G and P is $O(N_t \cdot D^2)$, where D is the hidden dimension. The analytic solution for W_{CF} involves a matrix inversion and a matrix multiplication, resulting in a complexity of $O(D^3 + D^2 \cdot T)$, where T is the number of experts. Since D is fixed and $T \ll D$, the update is near-instantaneous.
2. **Inference Phase:** For a single test instance, the routing score calculation $s = hW_{CF}$ has a time complexity of $O(D \cdot T)$. This is negligible compared to the $O(L \cdot D^2_{model})$ complexity of the MLLM forward pass (where L is sequence length), ensuring minimal added latency.

Space Complexity: CF-Router is highly memory-efficient as it does not require storing raw features H_{feat} or replaying data.

1. **Training Storage:** To support lifelong learning, we only maintain the matrices $G \in \mathbb{R}^{D \times D}$ and $P \in \mathbb{R}^{D \times T}$. The total space complexity is $O(D^2 + D \cdot T)$. For $D = 4096$, the Gram matrix G occupies only 64 MB (in float32), which is constant regardless of the number of training samples.
2. **Model Weights:** The router head $W_{CF} \in \mathbb{R}^{D \times T}$ introduces only $D \cdot T$ additional parameters, which is infinitesimal compared to the billions of parameters in the MLLM backbone.

Table 3: CF-Router performance on LLaVA-1.5-7B at the last lifelong learning stage. Results are reported as mean $\pm$ standard deviation across 120 task permutations.

Domain	Router Acc	Expert Acc
RS	$1.0000 \pm .0000$	$0.8128 \pm .0000$
Med	$1.0000 \pm .0000$	$0.6561 \pm .0000$
AD	$1.0000 \pm .0000$	$0.5468 \pm .0000$
Sci	$0.9993 \pm .0000$	$0.5689 \pm .0000$
Fin	$1.0000 \pm .0000$	$0.9114 \pm .0000$
AVERAGE	$0.9999 \pm .0000$	$0.6992 \pm .0000$

Table 4: CF-Router performance on InternVL-Chat-V1.0 at the last lifelong learning stage. Results are reported as mean $\pm$ standard deviation across 120 task permutations.

Domain	Router Acc	Expert Acc
RS	$1.0000 \pm .0000$	$0.8152 \pm .0000$
Med	$1.0000 \pm .0000$	$0.6442 \pm .0000$
AD	$1.0000 \pm .0000$	$0.5462 \pm .0000$
Sci	$0.9999 \pm .0000$	$0.5312 \pm .0000$
Fin	$1.0000 \pm .0000$	$0.9110 \pm .0000$
AVERAGE	$1.0000 \pm .0000$	$0.6896 \pm .0000$

4 Related Work

Traditional continual learning (CL) mitigates catastrophic forgetting through mechanisms such as regularization(Kirkpatrick et al., 2017; Li & Hoiem, 2017), replay (Lavda et al., 2018; Buzzega et al., 2020), or parameter isolation (Rusu et al., 2016; Mallya & Lazebnik, 2018). However, in the context of Large Language Models (LLMs), these methods face significant bottlenecks due to high computational overhead or risks of privacy leakage. This has prompted a shift toward Parameter-Efficient Fine-Tuning (PEFT) paradigms Hu et al. (2021). For instance, O-LoRA (Wang et al., 2023) effectively reduces inter-task interference under replay-free conditions by employing orthogonal subspace constraints. Extending these concepts to Multimodal Large Language Models (MLLMs), the CoIN benchmark (Chen et al., 2024a) reveals that forgetting in MLLMs primarily stems from a failure in instruction alignment rather than a loss of intrinsic knowledge. Inspired by this finding, MoELoRA (Chen et al., 2024a) and CL-MoE (Huai et al., 2025) integrate Mixture-of-Experts (MoE) architectures, utilizing task-driven gating and dual-momentum routing mechanisms, respectively, to balance adaptation between new and old tasks while maintaining parameter efficiency. Furthermore, HiDe (Guo et al., 2025a) proposes a decoupling strategy based on layer-wise sensitivity differences, featuring bottom-layer specific expansion and top-layer general fusion to enhance efficiency. Addressing the nature of catastrophic forgetting, SEFE (Chen et al., 2025) distinguishes between surface-level style forgetting and essential knowledge forgetting, validating the importance of style diversification for performance maintenance. Additionally, for distributed environments, DISCO (Guo et al., 2025b) introduces dynamic knowledge organization and selective subspace activation, effectively resolving data heterogeneity and task conflicts in federated continual instruction tuning.

5 Discussion

Scalability of Expert Storage. A natural concern with the expert-per-task paradigm is that the number of LoRA experts grows linearly with the number of tasks. However, we argue that this overhead is practically negligible. Each LoRA expert only adapts a small number of parameters through low-rank decomposition, which is orders of magnitude smaller than the full parameter count of the base MLLM backbone. As such, even when the number of tasks scales to a large quantity, the cumulative storage of all experts remains a small fraction of a single base model. Consequently, trading a modest amount of additional storage for the complete elimination of catastrophic forgetting and near-oracle task performance represents a highly favorable trade-off, especially in practical deployments where disk space is abundant relative to the cost of retraining or performance degradation.

Handling Ambiguous Routing Decisions. In certain cases, the routing score vector may produce multiple values that are close in magnitude, indicating that the input lies near the decision boundary between multiple domains. Under such circumstances, selecting only the top-scoring expert may not yield the optimal response, as complementary knowledge from other high-scoring experts could be beneficial. A promising direction for future work is to explore soft expert combination strategies for ambiguous inputs—for example, merging the LoRA parameters of the top-ranked experts weighted by their normalized routing scores, or generating candidate responses from multiple experts and selecting the best via a lightweight verification step. Developing principled methods to detect and handle such boundary cases would further enhance the robustness and flexibility of the CF-Router framework in open-world lifelong learning scenarios.

6 Conclusion

In this paper, we introduced **CF-Router**, a novel routing mechanism designed for efficient and robust expert selection in multimodal agent lifelong learning. By leveraging the closed-form analytic solution of ridge regression directly on the MLLM's last-layer hidden states, we decouple the optimization of the routing logic from gradient-based training. Our approach achieves near-perfect routing accuracy across diverse domains such as remote sensing, medical imaging, and autonomous driving, while ensuring zero catastrophic forgetting (BWT=0) without the need for data replay. Furthermore, we mathematically demonstrated and experimentally verified the order-invariance of our router, providing a stable and reliable foundation for agents to acquire new skills sequentially. The analytic-based nature and minimal inference overhead of CF-Router make it highly suitable for real-time, scalable lifelong learning applications in complex multimodal environments.

REFERENCES

Shuai Bai, Keqin Chen, Xuejing Liu, Jialin Wang, Wenbin Ge, Sibo Song, Kai Dang, Peng Wang, Shijie Wang, Jun Tang, Humen Zhong, Yuanzhi Zhu, Mingkun Yang, Zhaohai Li, Jianqiang Wan, Pengfei Wang, Wei Ding, Zheren Fu, Yiheng Xu, Jiabo Ye, Xi Zhang, Tianbao Xie, Zesen Cheng, Hang Zhang, Zhibo Yang, Haiyang Xu, and Junyang Lin. Qwen2.5-vl technical report. *arXiv preprint arXiv:2502.13923*, 2025.

Pietro Buzzega, Matteo Boschini, Angelo Porrello, Davide Abati, and Simone Calderara. Dark experience for general continual learning: a strong, simple baseline. *Advances in neural information processing systems*, 33:15920–15930, 2020.

Cheng Chen, Junchen Zhu, Xu Luo, Hengtao Shen, Jingkuan Song, and Lianli Gao. Coin: A benchmark of continual instruction tuning for multimodel large language models. *Advances in Neural Information Processing Systems*, 37:57817–57840, 2024a.

Jinpeng Chen, Runmin Cong, Yuzhi Zhao, Hongzheng Yang, Guangneng Hu, Horace Ip, and Sam Kwong. Sefe: Superficial and essential forgetting eliminator for multimodal continual instruction tuning. In *Forty-second International Conference on Machine Learning*, 2025.

Zhe Chen, Jiannan Wu, Wenhai Wang, Weijie Su, Guo Chen, Sen Xing, Muyan Zhong, Qinglong Zhang, Xizhou Zhu, Lewei Lu, et al. Internvl: Scaling up vision foundation models and aligning for generic visual-linguistic tasks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 24185–24198, 2024b.

Andrea Cossu, Antonio Carta, Lucia Passaro, Vincenzo Lomonaco, Tinne Tuytelaars, and Davide Bacciu. Continual pre-training mitigates forgetting in language and vision. *Neural Networks*, 179: 106492, 2024.

Matthias De Lange and Tinne Tuytelaars. Continual prototype evolution: Learning online from non-stationary data streams. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 8250–8259, 2021.

Haiyang Guo, Fanhu Zeng, Ziwei Xiang, Fei Zhu, Da-Han Wang, Xu-Yao Zhang, and Cheng-Lin Liu. Hide-llava: Hierarchical decoupling for continual instruction tuning of multimodal large language model. *arXiv preprint arXiv:2503.12941*, 2025a.

Haiyang Guo, Fanhu Zeng, Fei Zhu, Wenzhuo Liu, Da-Han Wang, Jian Xu, Xu-Yao Zhang, and Cheng-Lin Liu. Federated continual instruction tuning. *arXiv preprint arXiv:2503.12897*, 2025b.

Jinghan He, Haiyun Guo, Ming Tang, and Jinqiao Wang. Continual instruction tuning for large multimodal models. *arXiv preprint arXiv:2311.16206*, 2023.

Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685*, 2021.

Tianyu Huai, Jie Zhou, Xingjiao Wu, Qin Chen, Qingchun Bai, Ze Zhou, and Liang He. Cl-moe: Enhancing multimodal large language model with dual momentum mixture-of-experts for continual visual question answering. *arXiv preprint arXiv:2503.00413*, 2025.

James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, et al. Overcoming catastrophic forgetting in neural networks. *Proceedings of the national academy of sciences*, 114 (13):3521–3526, 2017.

Frantzeska Lavda, Jason Ramapuram, Magda Gregorova, and Alexandros Kalousis. Continual classification learning using generative models. *arXiv preprint arXiv:1810.10612*, 2018.

Bo Li, Yuanhan Zhang, Dong Guo, Renrui Zhang, Feng Li, Hao Zhang, Kaichen Zhang, Yanwei Li, Ziwei Liu, and Chunyuan Li. Llava-onevision: Easy visual task transfer. *arXiv preprint arXiv:2408.03326*, 2024.

Zhizhong Li and Derek Hoiem. Learning without forgetting. *IEEE transactions on pattern analysis and machine intelligence*, 40(12):2935–2947, 2017.

Haotian Liu, Chunyuan Li, Yuheng Li, and Yong Jae Lee. Improved baselines with visual instruction tuning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 26296–26306, 2024.

Arun Mallya and Svetlana Lazebnik. Packnet: Adding multiple tasks to a single network by iterative pruning. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pp. 7765–7773, 2018.

Michael McCloskey and Neal J Cohen. Catastrophic interference in connectionist networks: The sequential learning problem. In *Psychology of learning and motivation*, volume 24, pp. 109–165. Elsevier, 1989.

Mark D McDonnell, Dong Gong, Amin Parvaneh, Ehsan Abbasnejad, and Anton Van den Hengel. Ranpac: Random projections and pre-trained models for continual learning. *Advances in Neural Information Processing Systems*, 36:12022–12053, 2023.

Roger Ratcliff. Connectionist models of recognition memory: constraints imposed by learning and forgetting functions. *Psychological review*, 97(2):285, 1990.

Andrei A Rusu, Neil C Rabinowitz, Guillaume Desjardins, Hubert Soyer, James Kirkpatrick, Koray Kavukcuoglu, Razvan Pascanu, and Raia Hadsell. Progressive neural networks. *arXiv preprint arXiv:1606.04671*, 2016.

Liyuan Wang, Xingxing Zhang, Hang Su, and Jun Zhu. A comprehensive survey of continual learning: Theory, method and application. *IEEE transactions on pattern analysis and machine intelligence*, 46(8):5362–5383, 2024.

Xiao Wang, Tianze Chen, Qiming Ge, Han Xia, Rong Bao, Rui Zheng, Qi Zhang, Tao Gui, and Xuanjing Huang. Orthogonal subspace learning for language model continual learning. *arXiv preprint arXiv:2310.14152*, 2023.

Hongbo Zhao, Fei Zhu, Haiyang Guo, Meng Wang, Rundong Wang, Gaofeng Meng, and Zhaoxiang Zhang. Mllm-cl: Continual learning for multimodal large language models. *arXiv preprint arXiv:2506.05453*, 2025.