

# PCLR: Progressively Compressed LoRA for Multimodal Continual Instruction Tuning

Weicheng Meng, Jingyang Qiao, Zhizhong Zhang, Shaohui Liu, Yuan Xie

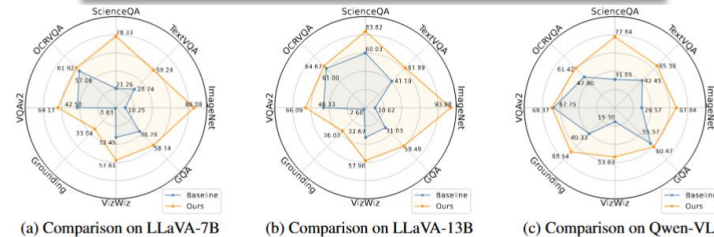
## Introduction

Large Multimodal Models (LMMs) rely on instruction tuning to follow human intent, but evolving tasks make repeated retraining costly. Continual Instruction Tuning (CIT) is a promising solution, yet existing methods either suffer from forgetting or require continual parameter expansion, causing memory growth. To address this, we propose PCLR, which combines a Compression–Integration–Learning (CIL) pipeline with LoRA Rank Pool (LRP) to compress redundant knowledge and reuse capacity efficiently. Experiments on COIN and Continual-NExT show strong effectiveness and state-of-the-art performance.

## Contributions

- We propose a Compression–Integration–Learning (CIL) pipeline for memory-efficient continual instruction tuning.
- We introduce LoRA Rank Pool (LRP), a fine-grained expert structure for flexible knowledge reuse and compression.
- We achieve state-of-the-art performance on CoIN and Continual-NExT with significant improvements over strong baselines.

## Visualization



## Overview of Framework

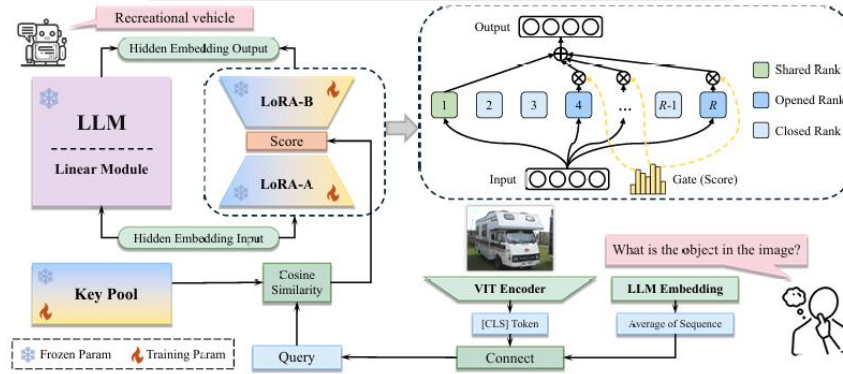


Figure 3: The pipeline of LRP. First, extract the query from inputs. Next, compute similarity scores with the key pools and gate the scores. Then, apply scores jointly with LRP weights in the forward pass. LRP is semi-frozen during training: new parameters are trained while previous ones are frozen.

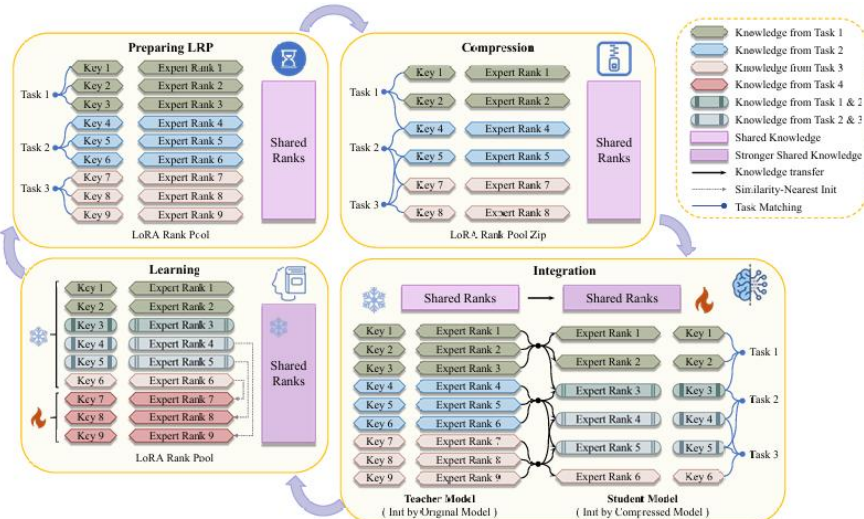


Figure 4: Compression–Integration–Learning (CIL) pipeline. Compression is a training-free process that streamlines the original LRP. Integration is a distillation process that aligns the compressed LRP with the original LRP. Learning is a supervised fine-tuning process that initializes the LRP from the integration-processed LRP and then adapts to the new task.

## Experiments

Table 1: Comparisons between ours and baselines on LLaVA-1.5-7B, CoIN benchmark.

Method	Venue	Accuracy on Each Task								Overall Results		
		1-ScienceQA	2-TextVQA	3-ImageNet	4-GQA	5-VizWiz	6-Grounding	7-VQAv2	8-OCRVQA	Avg. Acc(%)	Forgetting(↓)	New Acc(%)
Zero-shot	-	49.91	2.88	0.33	2.08	0.90	0.00	0.68	0.7	7.12	-	-
LoRA [Hu et al. 2022]	ICLR'22	56.77	49.35	95.55	56.65	53.90	30.09	42.50	57.08	57.18	37.29	61.36
LwF [Li & Hoesen 2017]	TPAMI'16	63.14	39.60	8.90	34.83	13.53	2.48	40.67	62.35	33.31	22.32	52.58
EWC [Kirkpatrick et al. 2017]	PNAS'17	67.41	40.41	8.18	35.05	37.88	2.67	41.27	61.02	36.74	20.51	54.68
MoLoRA [Chen et al. 2024a]	NIPS'24	58.92	38.59	8.85	37.10	44.25	2.45	41.40	55.35	35.86	25.71	58.36
AdLoRA [Zuo et al. 2024]	ICLR'24	73.40	51.29	35.47	44.53	46.75	0.93	55.86	62.03	46.28	23.99	63.27
MT [Zhu et al. 2024]	ICML'24	79.63	55.47	35.64	38.70	44.37	32.20	62.21	61.89	53.73	14.03	66.00
PGP [Qiao et al. 2025b]	ICLR'24	85.17	56.85	32.26	61.74	49.43	32.74	65.74	62.20	55.77	12.94	67.09
CIA [Qiao et al. 2025a]	ICML'25	75.63	54.47	43.44	60.70	43.37	36.00	65.21	65.59	55.33	7.04	61.49
SEIF [Kwon et al. 2025]	ICML'25	73.55	58.56	83.10	54.15	48.85	16.75	65.35	66.25	58.57	11.84	69.02
ProgLoRA [Vatani et al. 2025]	ACL'25	74.84	51.83	83.90	49.93	53.87	31.09	62.71	64.44	59.09	7.53	65.68
EProj [He et al. 2025]	ArXiv'23	78.51	57.53	92.35	55.93	44.67	36.39	63.74	57.00	60.79	5.42	65.54
PCLR	-	78.32	58.24	86.08	58.14	57.61	33.04	64.17	61.92	62.19	3.39	65.16

Table 8: Comparisons with baselines on LLaVA-1.5-13B, CoIN benchmark.

Method	Venue	Accuracy on Each Task								Overall Results		
		1-ScienceQA	2-TextVQA	3-ImageNet	4-GQA	5-VizWiz	6-Grounding	7-VQAv2	8-OCRVQA	Avg. Acc(%)	Forgetting(↓)	New Acc(%)
LoRA [Hu et al. 2022]	ICLR'22	60.03	41.19	10.62	31.03	32.67	2.60	46.33	61.00	35.68	32.90	64.47
MT [Zhu et al. 2024]	ICML'24	80.43	60.72	46.70	60.35	49.19	33.16	63.74	65.44	57.47	11.26	67.32
PGP [Qiao et al. 2025b]	ICLR'24	82.50	60.64	49.15	62.53	49.43	37.37	65.57	65.82	59.13	10.11	67.98
EProj [He et al. 2025]	ArXiv'23	77.65	58.93	92.31	60.22	38.27	33.77	64.39	65.80	61.42	5.84	66.53
CIA [Qiao et al. 2025a]	ICLR'24	83.94	61.40	97.05	62.61	43.99	39.72	66.29	65.78	65.10	2.31	67.12
Ours	-	83.82	61.99	93.98	59.49	57.98	36.03	66.09	64.67	65.51	2.08	67.32

Table 9: Comparisons with baselines on Qwen-VL, CoIN benchmark.

Method	Venue	Accuracy on Each Task								Overall Results		
		1-ScienceQA	2-TextVQA	3-ImageNet	4-GQA	5-VizWiz	6-Grounding	7-VQAv2	8-OCRVQA	Avg. Acc(%)	Forgetting(↓)	New Acc(%)
LoRA [Hu et al. 2022]	ICLR'22	71.05	42.45	29.57	40.33	40.33	67.75	47.80	57.42	61.88	6.06	60.06
EWC [Kirkpatrick et al. 2017]	PNAS'17	64.30	58.67	44.04	57.73	38.16	48.04	66.98	41.76	52.46	8.68	60.06
PGP [Qiao et al. 2025b]	ICLR'24	66.42	41.33	32.16	49.83	36.05	24.22	58.60	43.96	44.07	5.90	48.30
Ours	-	77.94	65.36	67.84	60.47	53.83	68.54	69.37	61.42	65.58	4.21	69.27

Table 2: Comparisons between ours and baselines on LLaVA-1.5-hf, Continual-NExT benchmark.

Method	Venue	Accuracy on Each Task & Overall Results								
		1-ArxivQA	2-GeoChat	3-IconQA	4-ClearMath	5-CodeQA	6-ImageNet	7-Flckr30k	8-DocVQA	9-TextVQA
Zero-shot	-	36.99	67.67	18.77	20.27	0.26	18.10	17.27	14.58	57.39
Multi-Task	-	64.08	96.40	70.50	10.81	97.06	20.52	23.18	65.44	-
LoRA [Hu et al. 2022]	ICLR'22	53.99	92.23	47.23	44.86	4.36	67.84	17.16	16.47	47.70
LwF [Li & Hoesen 2017]	TPAMI'16	51.04	87.33	30.97	39.20	4.74	84.89	16.26	16.56	54.09
EWC [Kirkpatrick et al. 2017]	PNAS'17	55.16	91.73	47.17	49.30	4.38	82.03	16.71	16.88	51.73
GEM [Lopez-Paz & Ranzato 2017]	NIPS'17	55.30	91.03	49.13	48.30	4.76	76.30	16.21	15.85	51.33
Replay-100 [Rolnick et al. 2019]	NIPS'19	54.85	94.40	51.73	40.07	4.48	94.61	9.36	14.65	54.70
MoLoRA [Chen et al. 2024a]	NIPS'24	56.00	91.36	48.76	48.90	3.82	82.19	17.77	16.33	59.31
CIA* [Qiao et al. 2025a]	ICML'25	56.35	93.41	48.76	48.20	4.23	83.00	16.54	16.98	51.10
PCLR	-	89.00	78.20	51.30	37.17	7.88	85.19	18.68	20.80	64.10
PCLR-LwF (PCLR variant)	-	61.16	96.37	64.70	8.68	62.40	20.63	19.84	63.97	-
Zero-shot	-	0.44	9.60	33.29	28.28	66.19	17.54	27.11	-	-
Multi-Task	-	36.01	20.76	58.61	72.03	86.21	23.38	53.71	-	-
LoRA [Hu et al. 2022]	ICLR'22	33.80	18.04	50.98	69.52	89.46	22.27	45.06	11.62	55.91
LwF [Li & Hoesen 2017]	TPAMI'16	30.05	18.64	52.79	64.11	87.95	24.96	44.24	12.29	55.70
EWC [Kirkpatrick et al. 2017]	PNAS'17	35.41	19.00	50.92	69.92	89.51	24.17	46.93	9.72	56.01
GEM [Lopez-Paz & Ranzato 2017]	NIPS'17	35.28	17.68	51.38	67.23	89.86	23.85	46.23	10.19	55.74
Replay-100 [Rolnick et al. 2019]	NIPS'19	31.42	14.40	49.64	66.98	85.62	23.85	45.38	11.41	56.03
MoLoRA [Chen et al. 2024a]	NIPS'24	34.17	18.52	49.04	67.65	88.28	22.59	46.99	8.06	54.51
CIA* [Qiao et al. 2025a]	ICML'25	32.56	17.62	50.67	69.86	87.37	23.54	46.81	8.30	54.55
PCLR	-	36.42	20.96	58.48	69.76	89.65	22.27	47.94	7.71	55.14
PCLR-LwF (PCLR variant)	-	37.22	19.84	58.42	62.83	83.09	23.49	52.67	4.58	56.89

Table 10: Comparisons with baselines on Qwen2.5-VL-Instruct, Continual-NExT benchmark.

Method	Venue	Accuracy on Each Task & Overall Results								
		1-ArxivQA	2-GeoChat	3-IconQA	4-ClearMath	5-CodeQA	6-ImageNet	7-Flckr30k	8-DocVQA	9-TextVQA
Zero-shot	-	70.23	67.53	28.60	81.53	3.63	35.62	8.97	82.09	68.94
LoRA [Hu et al. 2022]	ICLR'22	71.39	89.00	73.40	96.80	2.05	78.77	20.92	85.04	68.94
LwF [Li & Hoesen 2017]	TPAMI'16	72.60	92.33	58.56	95.60	3.49	75.3	21.42	83.77	71.36
EWC [Kirkpatrick et al. 2017]	PNAS'17	73.28	92.00	67.87	89.20	3.53	76.06	20.53	80.61	53.82
Replay [Rolnick et al. 2019]	NIPS'19	71.37	87.57	78.37	98.83	2.45	83.29	20.81	86.32	71.28
MoLoRA [Chen et al. 2024a]	NIPS'24	72.19	91.53	74.36	97.66	2.65	90.75	21.08	85.51	79.55
PCLR	-	74.72	92.13	88.10	97.27	8.45	74.06	21.93	86.27	80.63
PCLR-LwF (PCLR variant)	-	74.43	92.23	88.60	97.80	8.93	88.42	20.75	86.19	80.38
Zero-shot	-	0.23	7.90	33.22	72.23	82.95	47.59	-	-	-
LoRA [Hu et al. 2022]	ICLR'22	34.81	73.88	7.96	86.60	80.08	4.87	58.30	9.36	67.04
LwF [Li & Hoesen 2017]	TPAMI'16	50.72	67.28	7.57	87.60	91.44	16.27	59.79	10.98	70.04
EWC [Kirkpatrick et al. 2017]	PNAS'17	48.31	67.36	13.26	86.99	87.05	15.64	58.57	11.89	69.47
Replay [Rolnick et al. 2019]	NIPS'19	45.50	74.72	8.66	86.11	90.90	16.11	61.49	5.40	67.46
MoLoRA [Chen et al. 2024a]	NIPS'24	2.14	74.80	6.35	87.88	80.98	16.18	58.94	11.81	69.96
PCLR	-	42.85	76.76	39.44	86.24	87.27	7.74	64.26	5.18	69.08
PCLR-LwF (PCLR variant)	-	45.16	76.08	44.43	85.53	88.38	9.32	65.78	3.74	69.26