NeurIPS 2025

RePIC: Reinforced Post-Training for Personalizing Multi-Modal Language Models

Yeongtak Oh, Dohyun Chung, Juhyeon Shin, Sangha Park,
Johan Barthelemy, Jisoo Mok†, Sungroh Yoon†

Data Science & Artificial Intelligence Laboratory
Electrical and Computer Engineering
Seoul National University

† Corresponding authors

Code : https://github.com/oyt9306/RePIC

Contents

Contents

- **What is personalization in MLLMs?** * concept: a personal visual or textual sample provided as a reference
  - Given user-specific concepts, personalized MLLMs can perform a range of downstream tasks

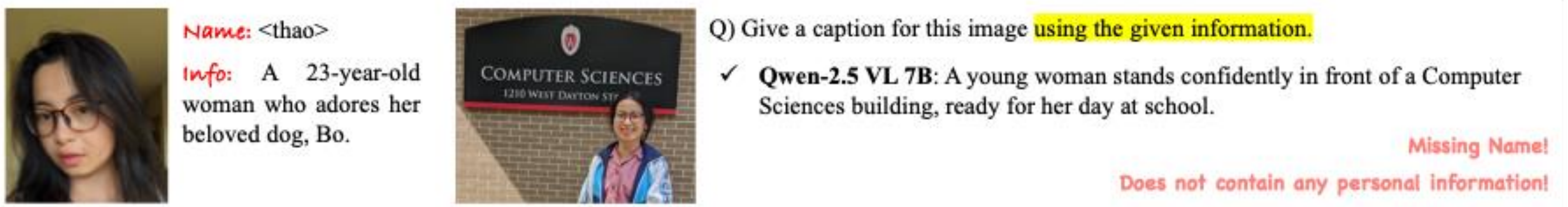
Then, how can we personalize MLLMs?

Contents

DSAIL
Data Science & AI Laboratory

TL;DR

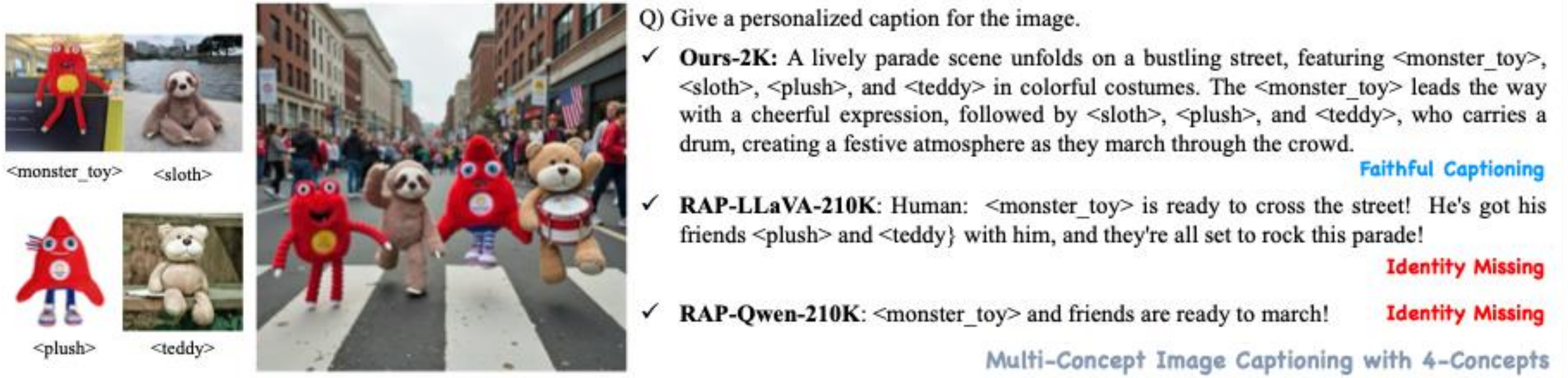
- Conventional zero-shot models fails to generate captions including personal information



- Previous SFT-based personalization methods easily fail in real-world scenarios
- To this end, we propose an RL-based post-training pipeline for MLLM personalization
- Performance

* RePIC: Reinforced Post-training for Personalized Image Captioning

  – RePIC shows prominent personalized image captioning performance even for unseen multi-concept settings
- Training Efficiency
  – RePIC requires only 2K training samples compared to other SFT-based baselines that require > 200K samples



Related Works

- **TimeLine**

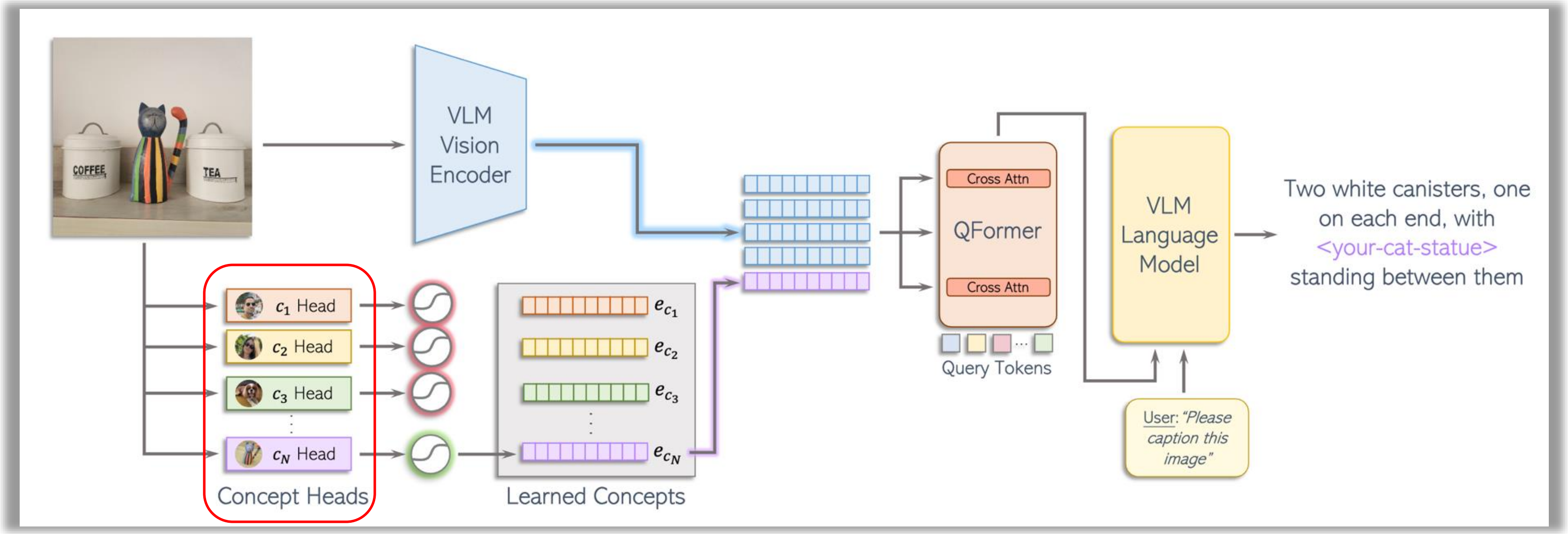
MyVLM
(ECCV 2024)

✓ **Pros**

– Uses external *concept heads* to identify each user-specific concept

✓ **Cons**

– Requires retraining the concept heads when the new concepts emerge



DSAIL Data Science & AI Laboratory

Related Works

- **TimeLine**

MyVLM
(ECCV 2024)

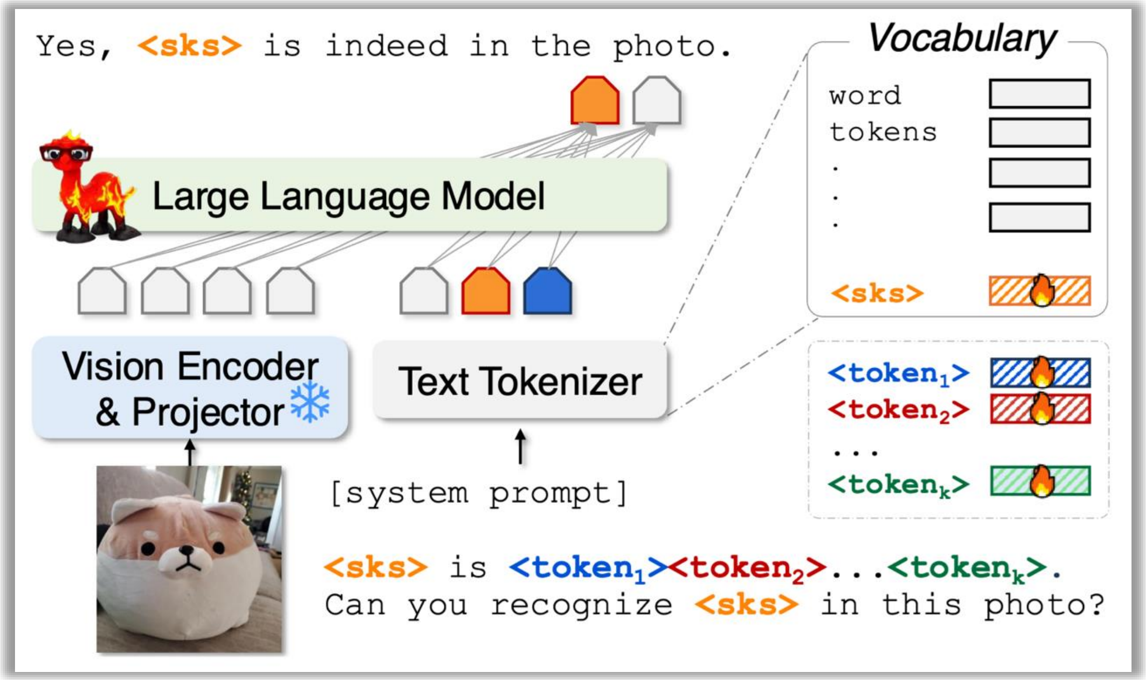
Yo'LLaVA
(NeurIPS 2024)

✓ **Pros**

– Uses external *special tokens* to identify each user-specific concept

✓ **Cons**

– Still requires retraining each token when the new concepts emerge



DSAIL
Data Science & AI Laboratory

Related Works

- **TimeLine**

Token-based personalization | Post-training-based personalization

MyVLM (ECCV 2024)

Yo'LLaVA (NeurIPS 2024)

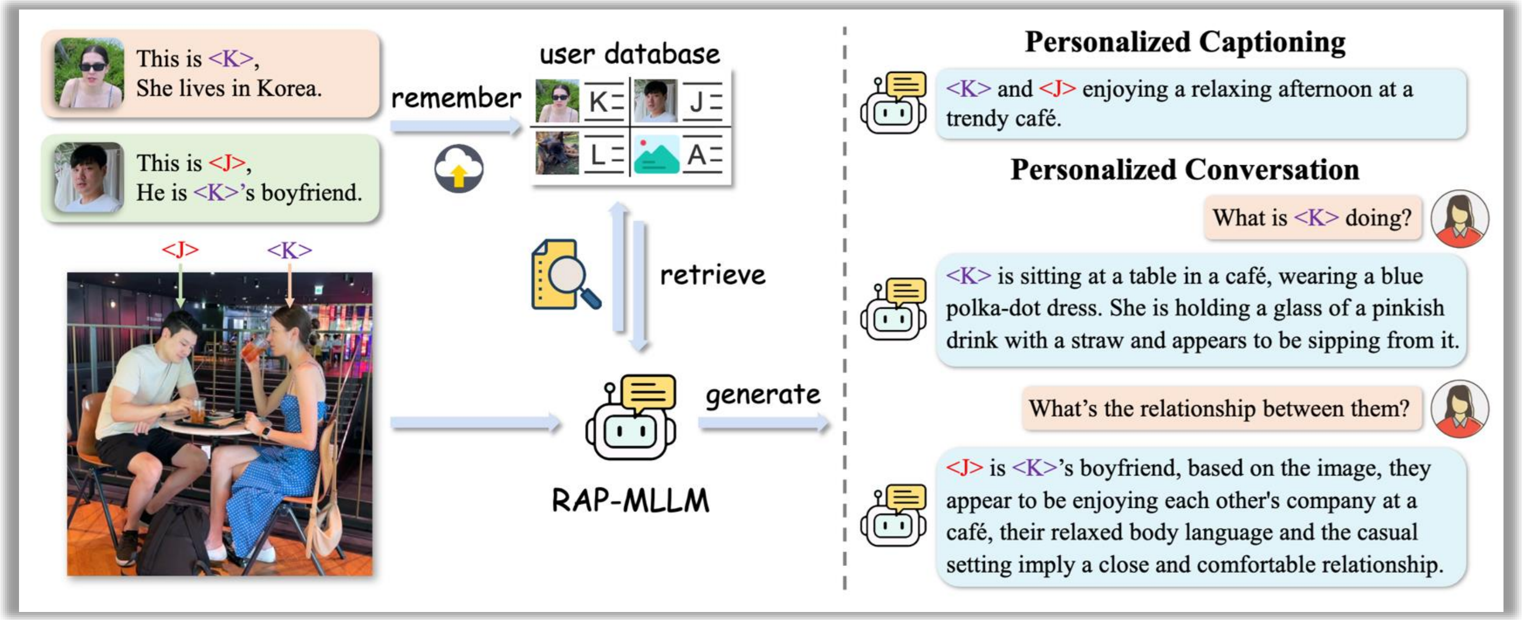
RAP-MLLM (CVPR 2025)

✓ **Pros**

- The first post-training-based personalization method
- It can perform Retrieval-enabled personalized tasks

✓ **Cons**

- Needs lots of training data (210K) to post-train MLLM
- Hard to generalize to real-world scenarios (e.g., multi-concept)



DSAIL Data Science & AI Laboratory

Related Works

- **TimeLine**

Post-training-based personalization

MyVLM (ECCV 2024) — Yo'LLaVA (NeurIPS 2024) — RAP-MLLM (CVPR 2025) — PVIT (ICLR 2025)

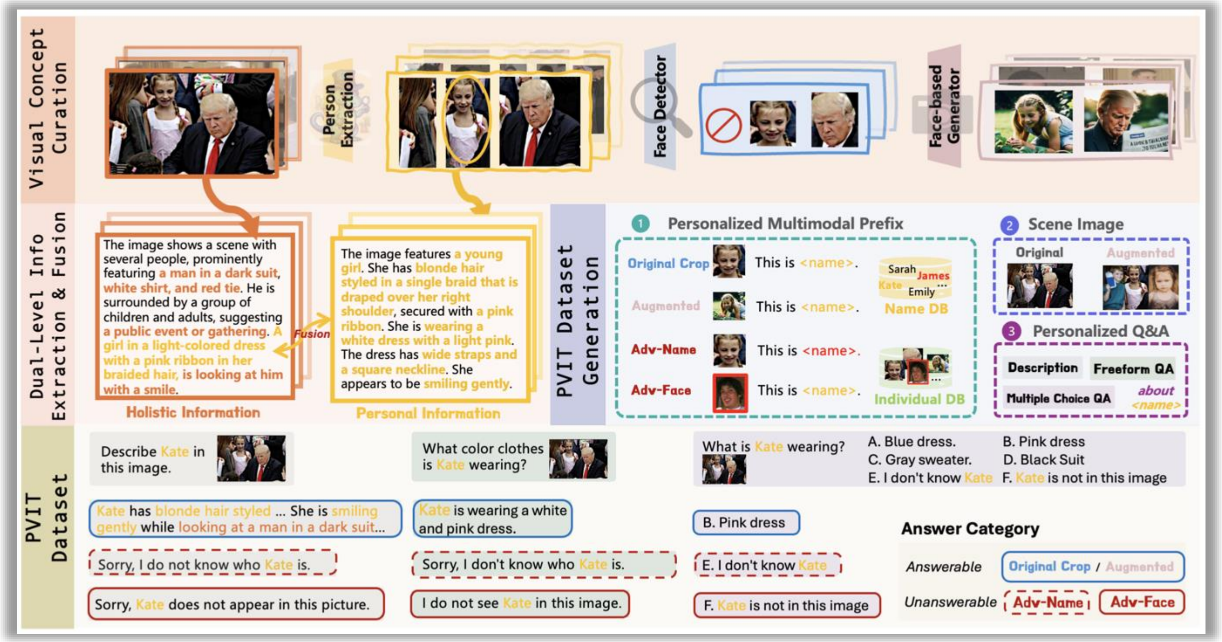
✓ **Pros**

– Introduces a *large-scale personalization benchmark* (3M)

✓ **Cons**

– Only covers human-related personalization scenarios

– Requires manual validation for each image



DSAIL Data Science & AI Laboratory

Related Works

- **TimeLine**

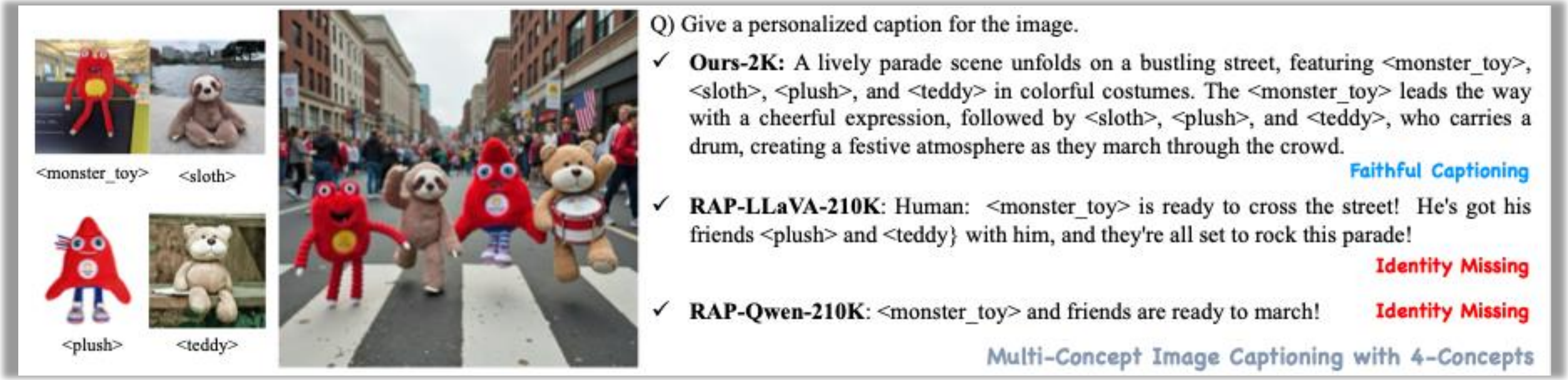
Post-training-based personalization

MyVLM (ECCV 2024)	Yo'LLaVA (NeurIPS 2024)	RAP-MLLM (CVPR 2025)	PVIT (ICLR 2025)	**RePIC (NeurIPS 2025)**

✓ **TL;DR**

– We propose a *RL-based post-training pipeline* for MLLM personalization

– RePIC shows generalizable personalized image captioning even for <u>unseen multi-concept settings</u>

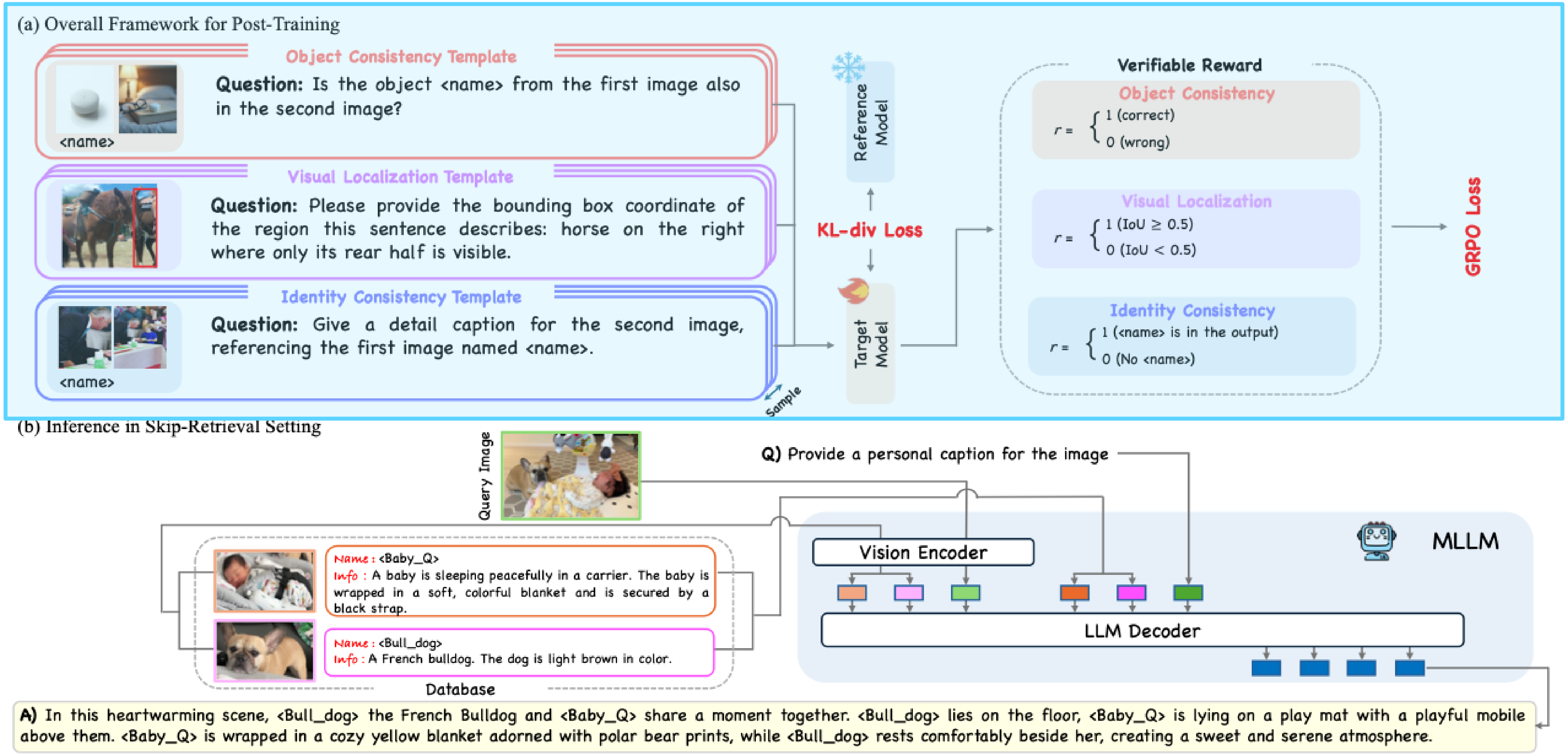


Contents

DSAIL
Data Science & AI Laboratory

Proposed Method : RePIC

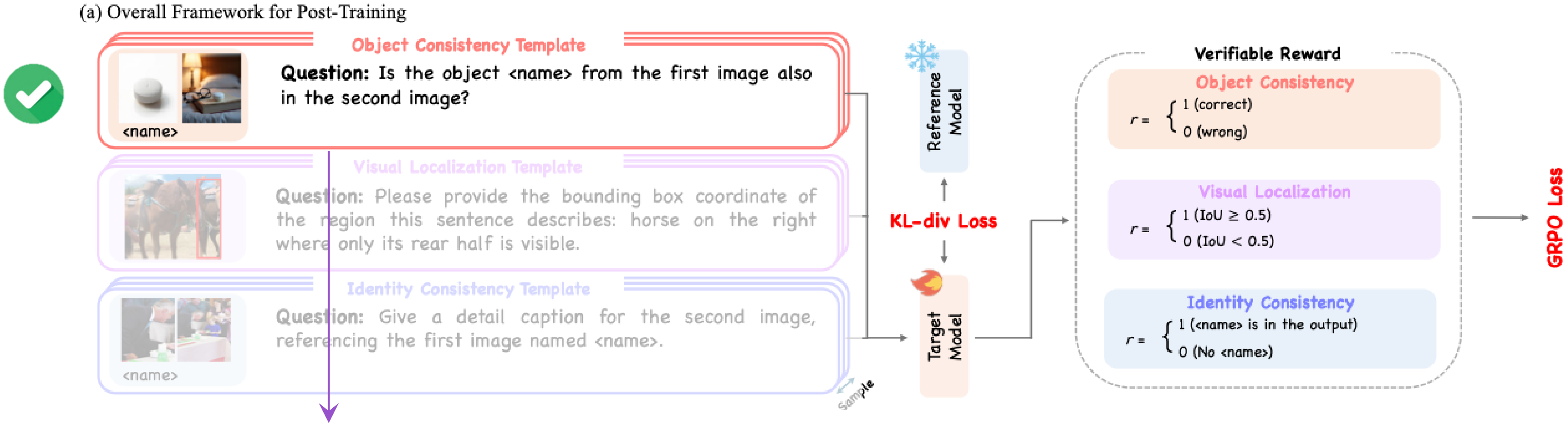
- Overall Pipeline



DSAIL Data Science & AI Laboratory

Proposed Method : RePIC

- Training Phase

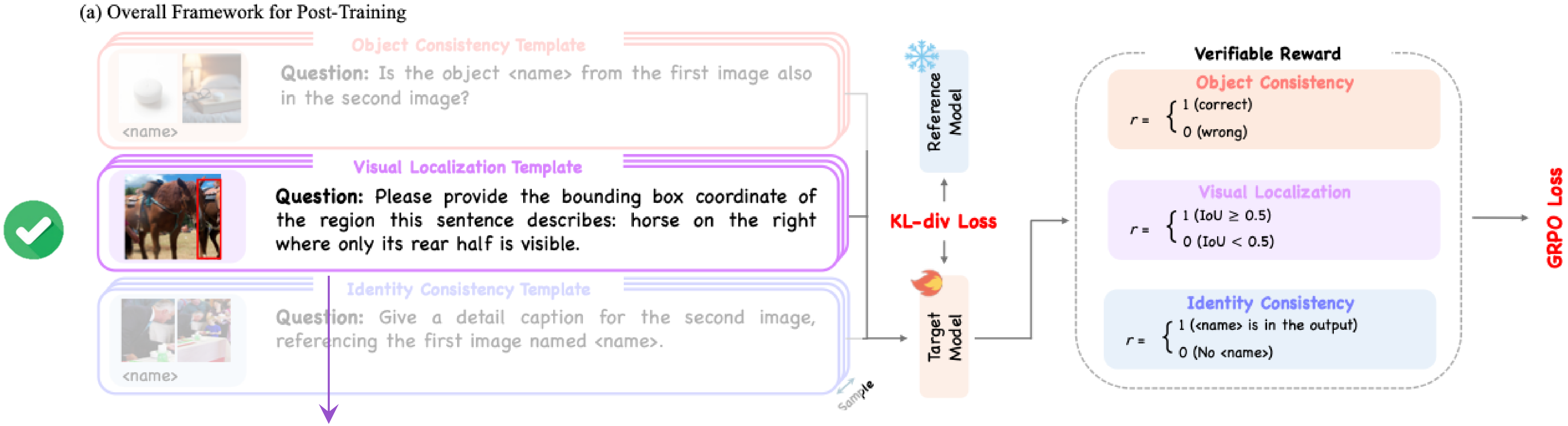


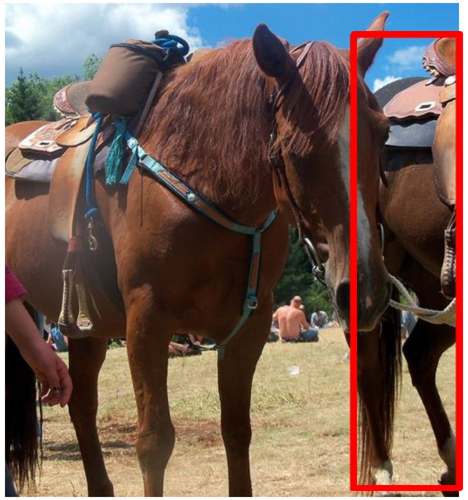
Proposed verifiable reward : Object Consistency Tuning (OCT)

$$r_{\text{OCT}} = \begin{cases} 1, & \text{if correctly answer with yes or no} \\ 0, & \text{else} \end{cases}$$

Proposed Method : RePIC

- Training Phase





Proposed verifiable reward : Visual Localization Tuning (VLT)

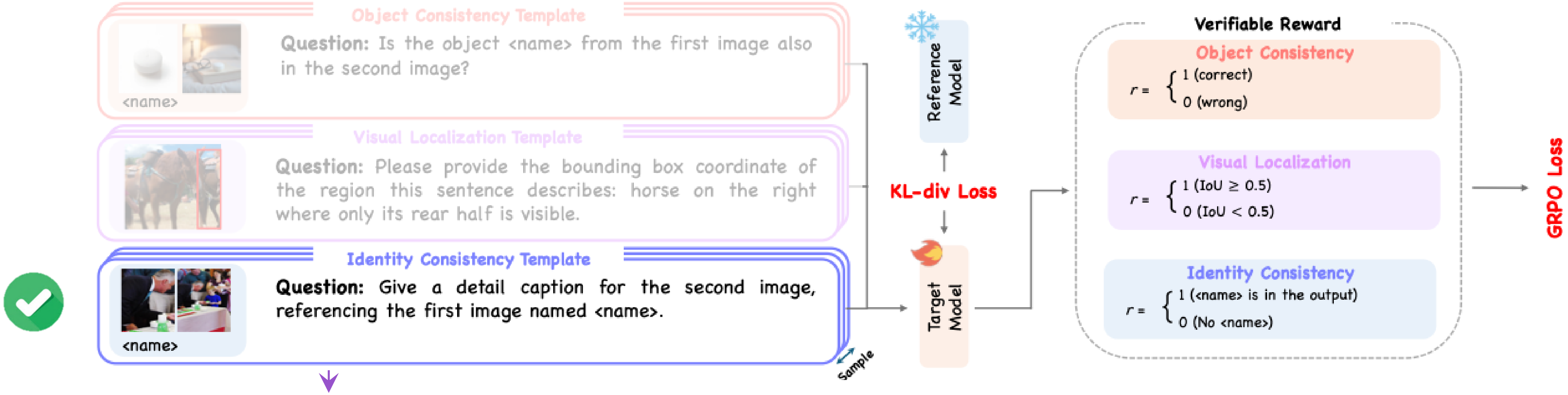
$$r_{\text{VLT}} = \begin{cases} 1, & \text{if IoU} \geq 0{,}5 \\ 0, & \text{otherwise} \end{cases}$$

* Referring expression comprehension (REC)

Proposed Method : RePIC

- Training Phase

(a) Overall Framework for Post-Training



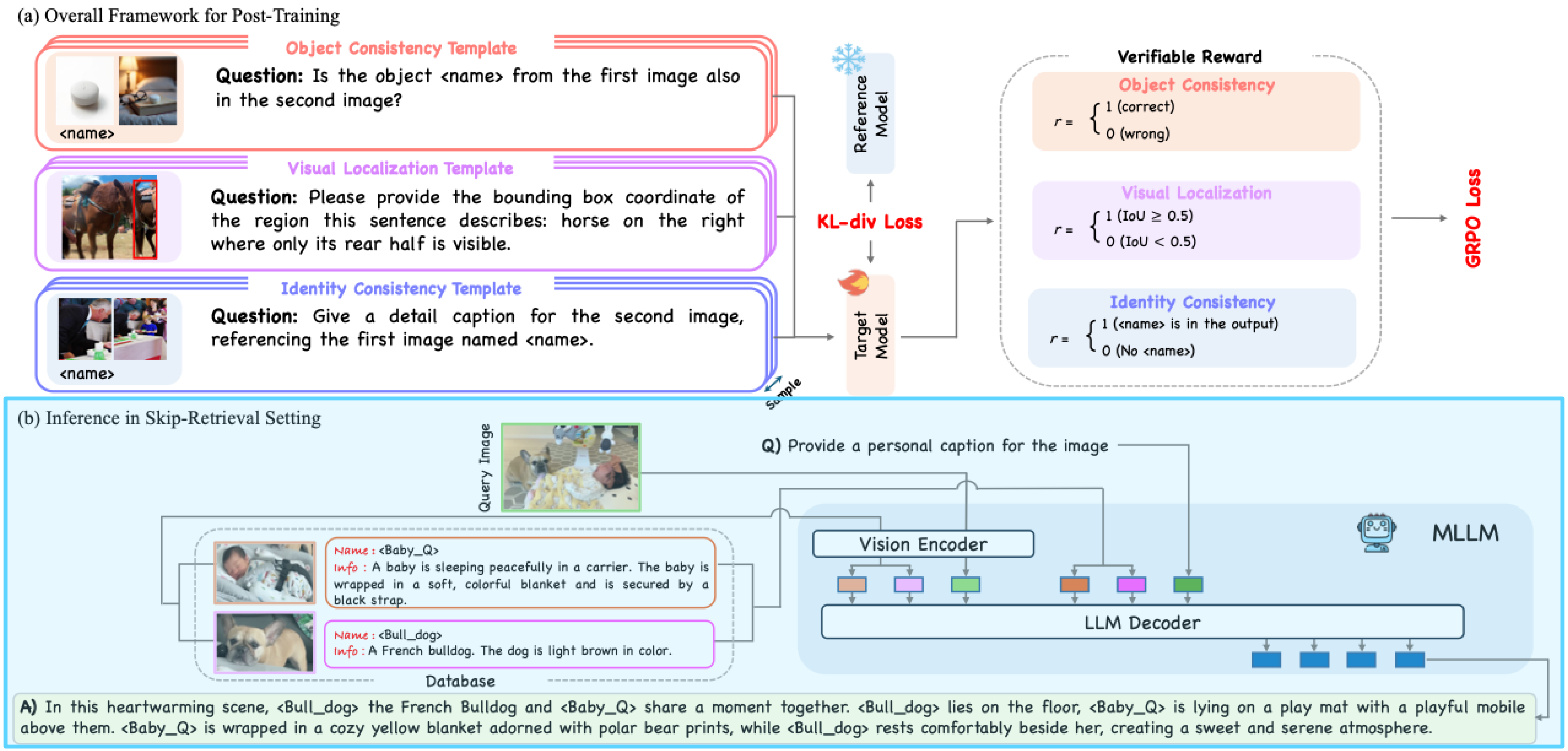
Proposed verifiable reward : Identity Consistency Tuning (ICT)

$$r_{\text{Single-ICT}} = \begin{cases} 1, & \text{if <name> appears in the output} \\ 0, & \text{otherwise} \end{cases}$$

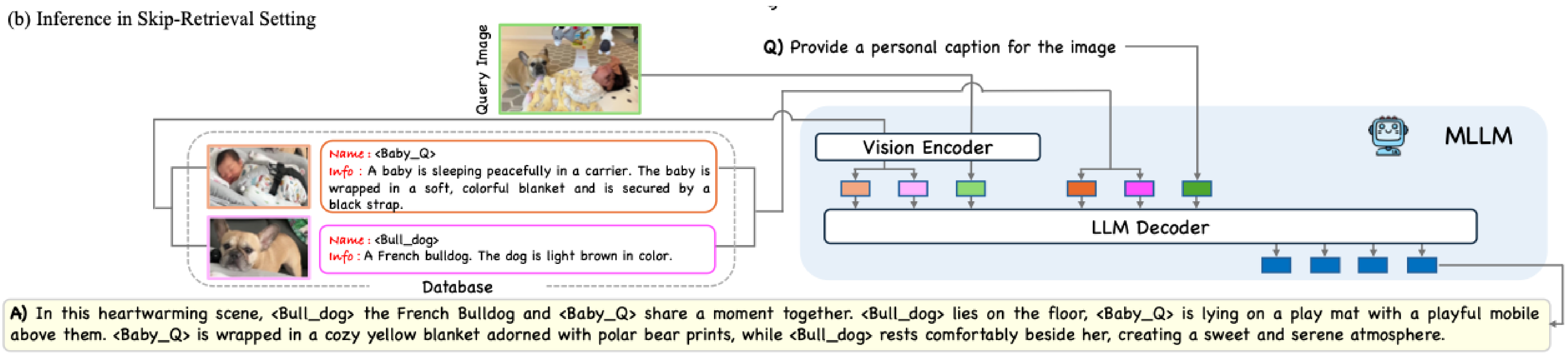
$$r_{\text{Multi-ICT}} = \begin{cases} n/m, & \text{if <name 1>, <name 2>}, \cdots, \text{<name n> appears in the output} \\ 0, & \text{otherwise} \end{cases}$$

Proposed Method : RePIC

- Overall Pipeline



- Inference Phase



At inference time,

a) Retrieved samples (image/text) are prepended to use as visual demonstrations

b) MLLM provides personal captions for a query image incorporating the personal information

Contents

DSAIL
Data Science & AI Laboratory

Experimental Results

- **Qualitative Results of Multi-Concept Image Captioning**

Name : Q) Give a personalized caption of this image.


<Character_A>


<Character_B>

Ours – Full : In this scene from the animated movie, <Character_A> and <Character_B> are standing in a dimly lit room, looking surprised or concerned. The background suggests they might be inside a vehicle or a control room. Correct Captioning

Qwen-2.5-VL : In a dimly lit room, <Character_A> and <Character_B> stand facing each other, their expressions hinting at a mix of curiosity and anticipation. Wrong Captioning

RAP-LLaVA : <Character_B> is a sly fox with a big grin. He's always up for an adventure. Identity Missing

RAP-Qwen : <Character_A> is ready to help! Identity Missing

Name : Q) Give a personalized caption of this image.


<J>


<K>

Ours – Full : In this image, <J> and <K> are standing outside near a building with a colorful mural. They are both holding small objects in their hands and appear to be engaged in a conversation or activity together. The scene is bright and sunny, with greenery visible in the background. Correct Captioning

Qwen-2.5-VL : A man and a woman are standing outside on a sunny day. The woman is wearing sunglasses and a blue polka-dotted dress while holding a small object in her hand. The man is also wearing sunglasses and a light-colored t-shirt, holding a similar object. They appear to be engaged in a conversation or activity together. Identity Missing

RAP-LLaVA : <K> and <J> enjoying some ice cream. Wrong Captioning

RAP-Qwen : <J> and <K> are enjoying some ice tea at a cafe. Wrong Captioning

Experimental Results

- **Quantitative Results for Single-Concept Personal Grounding Evaluation**

Models	Seen Data	MyVLM [2]			Yo'LLaVA [37]			DreamBooth [43]		
		Pre.	Rec.	F1	Pre.	Rec.	F1	Pre.	Rec.	F1
Skip-Retrieval Setting										
PVIT-LLAVA	210K	17.1	1.8	3.3	20.1	2.1	3.8	26.5	16.5	20.3
RAP-LLAVA	210K	100	92.9	96.3	100	95.5	97.7	97.3	91.8	94.5
RAP-LLAVA	2K	100	49.4	66.1	50.6	48.6	49.6	68.4	65.8	67.1
RAP-Qwen	210K	**100**	**98.8**	**99.4**	**100**	**99.8**	**99.8**	**100**	**100**	**100**
Qwen-2.5 VL	0	100	56.8	72.4	100	33.3	50.0	96.0	76.6	85.2
Ours	2K	100	96.2	98.1	99.7	96.1	97.9	100	98.1	99.0
Retrieval Setting										
Retrieval (Top-2)		97.6	95.9	96.7	83.6	82.9	83.3	99.3	96.2	97.7
RAP-LLAVA	210K	95.6	79.1	87.8	82.7	**79.9**	**81.2**	96.0	91.1	93.5
RAP-LLAVA	2K	79.2	53.8	64.1	71.2	52.2	64.4	69.5	66.5	68.0
RAP-Qwen	210K	95.5	**87.9**	**91.6**	79.2	75.1	76.2	**98.7**	**94.3**	**96.4**
Qwen-2.5 VL	0	91.5	50.6	65.2	77.4	42.3	55.2	95.2	75.3	84.1
Ours	2K	**99.0**	83.2	90.4	**84.4**	69.7	76.3	98.6	90.5	94.4

– RePIC shows comparable personal grounding results compared to other data-centric SFT-based methods

Experimental Results

- **Quantitative Results for Multi-Concept Personal Grounding Evaluation**

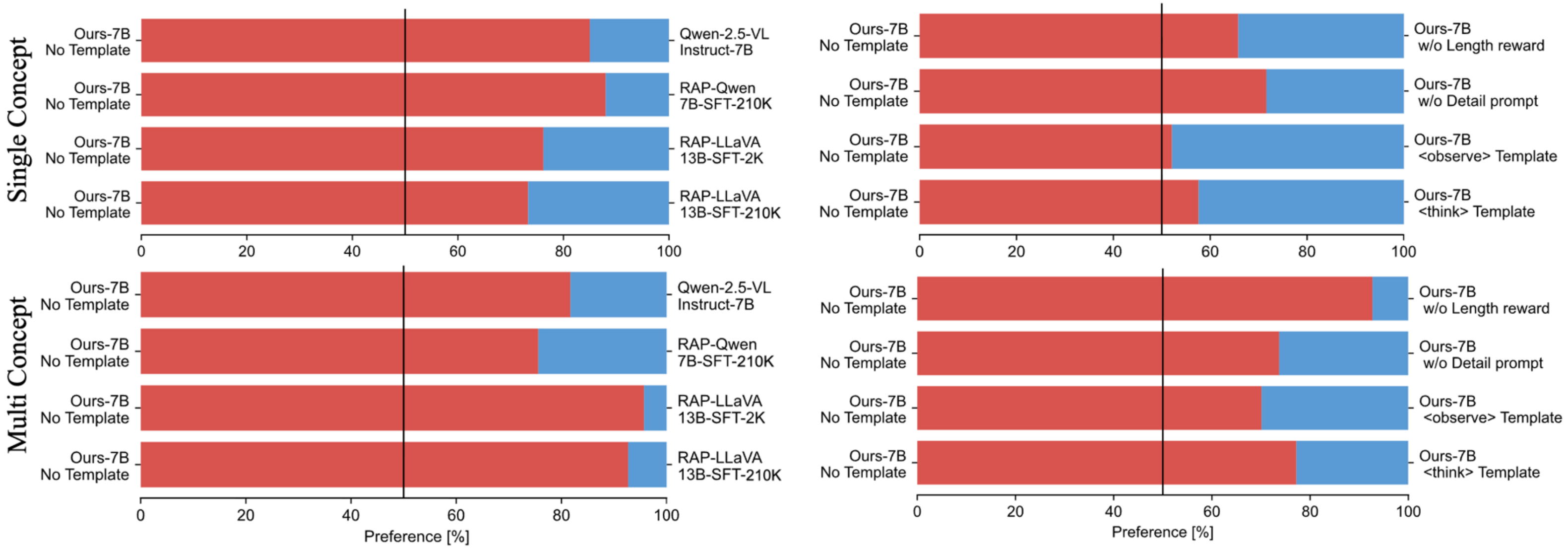
Models	Seen Data	2-Concepts						4-Concepts					
		Skip-Retrieval			Retrieval			Skip-Retrieval			Retrieval		
		Pre.	Rec.	F1	Pre.	Rec.	F1	Pre.	Rec.	F1	Pre.	Rec.	F1
RAP-LLaVA	210K	100	93.9	96.9	99.3	89.6	94.5	52.9	4.3	7.9	16.7	3.1	5.2
RAP-LLaVA	2K	100	90.2	94.9	95.7	81.1	87.8	36.4	1.9	3.6	22.4	0.7	1.4
RAP-Qwen	210K	100	82.9	90.7	**100**	73.2	84.5	49.6	13.6	21.3	12.6	2.6	4.3
Qwen-2.5 VL	0	100	75.0	85.7	98.1	64.0	77.5	73.3	22.9	34.8	22.5	6.4	10.0
Ours - Full	2K	**100**	**98.8**	**99.4**	97.5	**93.9**	**95.7**	**88.0**	**59.5**	**71.0**	**24.8**	**15.7**	**19.2**

  – RePIC shows the best results for 2 (ID) and 4 (OOD)-concept settings

DSAIL
Data Science & AI Laboratory

Experimental Results

- **Preference Evaluations & Image Captioning Quality Evaluation**



(a) Comparisons with other models (b) Ablation studies

Types	Metrics	RAP-LLaVA	RAP-Qwen	Zero-Shot	Ours
Reference-based	BLEU [38] (10^{-2})	0.260	0.170	0.210	**0.290**
	CIDEr [53]	0.193	0.185	**0.208**	0.194
	METEOR [5]	0.242	0.267	0.271	**0.321**
	SPICE [3]	**0.104**	0.084	0.083	0.086
	BERTScore [63]	**0.683**	0.567	0.523	0.668

Types	Metrics	RAP-LLaVA	RAP-Qwen	Zero-Shot	Ours
Reference-free	CLIPScore [18]	0.332	0.316	0.323	**0.339**
	ImageReward [56]	-0.094	0.087	**0.287**	0.130

  - RePIC consistently outperforms competing baselines (preference) and achieves comparable or superior results on quantitative metrics (caption quality)

DSAIL
Data Science & AI Laboratory

Contents

Conclusions & Limitations

- We propose RePIC, a strong baseline for personalized image captioning task with RL-based post-training
  - We alleviate the cost of collecting high-quality personal captions
- By leveraging verifiable rewards, tailored data and instructions, RePIC shows robust performance
  - We present generalizable personalized image captioning results on various scenarios
- This work only focuses on RL-based post-training in the image domain
  - Future research could extend personalization to other modalities, such as audio and video

< Project Page >