Mitigating Hallucinations in Multimodal LLMs via Object-aware Preference Optimization

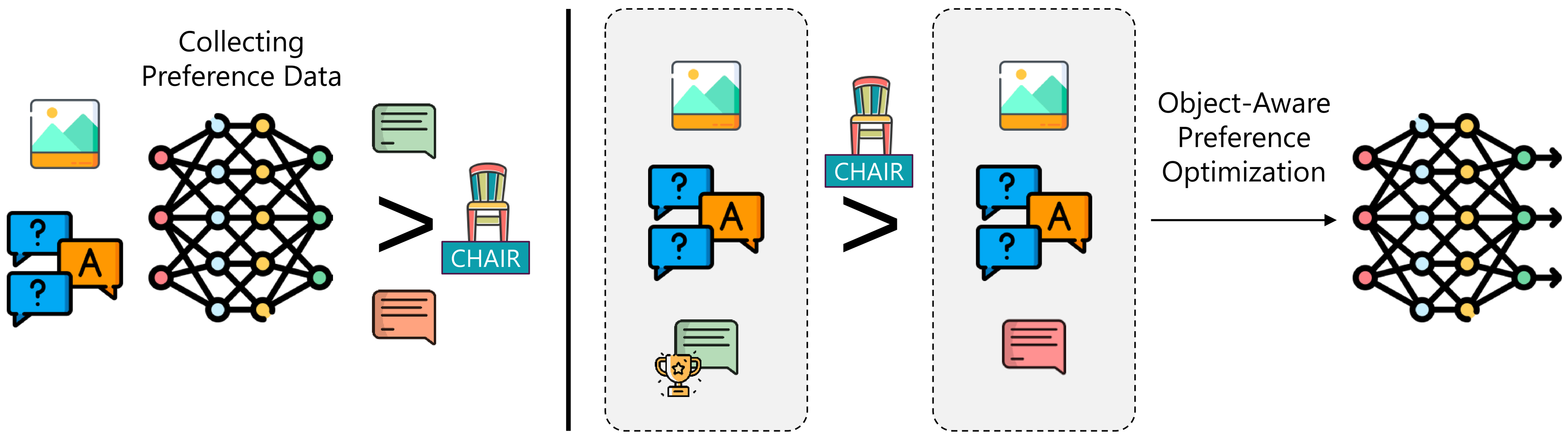
BMVC 2025

Alberto Compagnoni, Davide Caffagni, Nicholas Moratelli,
Lorenzo Baraldi, Marcella Cornia, Rita Cucchiara
University of Modena and Reggio Emilia, Italy
{name.surname}@unimore.it

UNIMORE AImageLab

Overview and Motivations

- Multimodal LLMs [1] suffer from **visual hallucinations** [2], *i.e.*, their responses mention objects absent from the input image.
- Recent approaches (*e.g.*, HA-DPO [3], mDPO [4]) treat hallucination mitigation as an **alignment problem**, using fine-tuning techniques such as Direct Preference Optimization (DPO) [5] to steer MLLMs toward less hallucinatory behavior.
- Data collection for such approaches is challenging, often requiring elaborate pipelines or proprietary models.



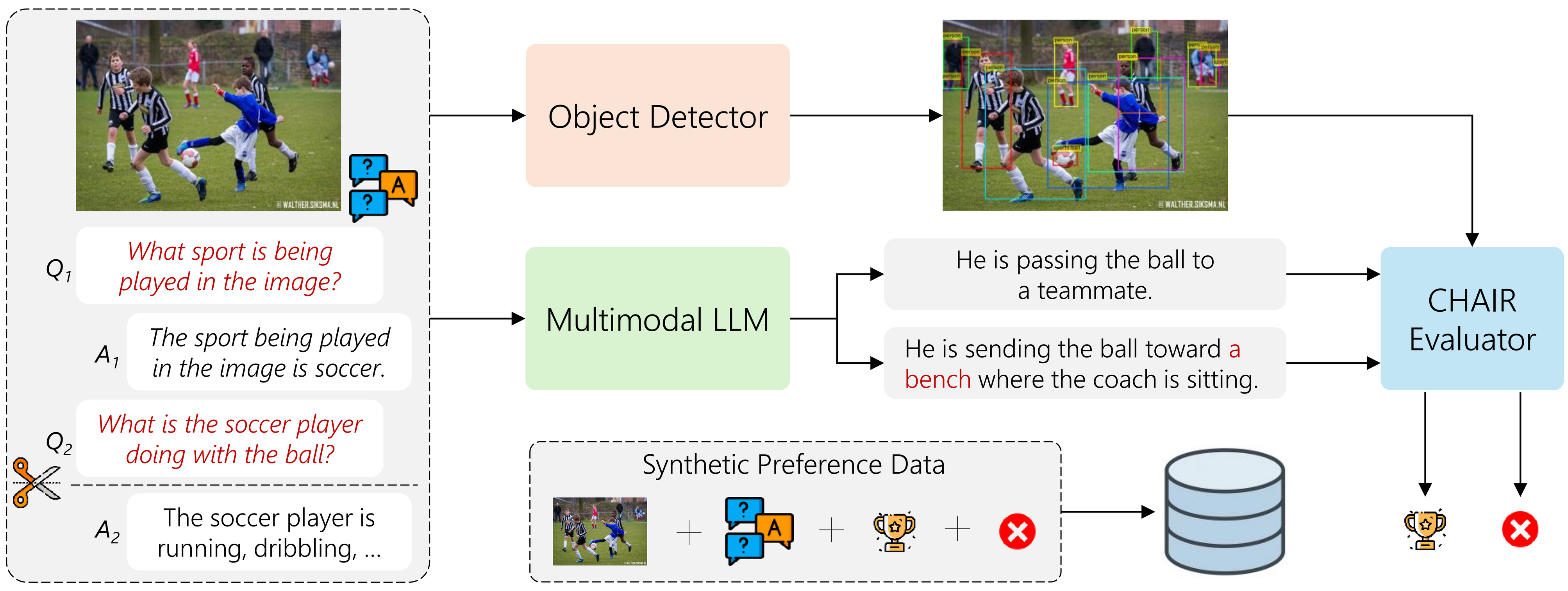
Can we employ the well-known CHAIR [6] metric as a proxy for human preferences about hallucinatory behaviours?
*We propose **Object-aware Preference Optimization** as a simple and effective method for hallucination mitigation.*

CHAIR-DPO trains the MLLM to be aware of the objects truly depicted in the input image, by optimizing a CHAIR-based reward that ranks model responses by the degree of hallucination.

LLaVA-MORE-8B: The image features a red double-decker bus driving down a desert road, surrounded by a vast expanse of sand. The bus is positioned in the middle of the road, with a clear blue sky in the background. There are several people in the scene, with **one person** standing near the left side of the road, **another person** closer to the center, and **two more people** on the right side of the road. Additionally, there are **two benches** visible in the scene, one located near the left side of the road and the other closer to the center.

LLaVA-MORE-8B + CHAIR-DPO: The image features a red double-decker bus driving down a desert road. The bus is positioned in the middle of the scene, occupying a significant portion of the image. The road appears to be relatively empty, with **no other vehicles or pedestrians** visible in the vicinity. The bus is traveling on a clear day, and the sky above is blue, creating a serene and open atmosphere.

Object-aware Preference Optimization (CHAIR-DPO)



1. Using samples from existing Visual Instruction Tuning datasets, we prompt an MLLM with a text prompt x_T and image x_I to generate two **candidate completions**, then extract the **ground-truth object set** with a pre-trained detector.
2. We compute the CHAIR$_i(y)$ score for each completion. The one with the lower CHAIR$_i$ (fewer hallucinations) is labeled as the **winner** y_w, and the other as the **loser** y_l.

$$\text{CHAIR}_i(y) = \frac{\left|[\text{hallucinated objects}]_y\right|}{\left|[\text{all mentioned objects}]_y\right|}.$$

3. To avoid optimizing **spurious** preference signals, we filter out samples with equal CHAIR$_i$ score.
4. We fine-tune an MLLM with DPO on these preference pairs, increasing the probability of generating the non-hallucinated completion to the detriment of the hallucinated one.

Hallucination Mitigation Results

			AMBER				CHAIR-MSCOCO		Object HalBench	
	FT	**Ext. Support**	CHAIR$_i$ ↓	Cover ↑	HalRate ↓	Cog ↓	CHAIR$_s$ ↓	CHAIR$_i$ ↓	CHAIR$_s$ ↓	CHAIR$_i$ ↓
LLaVA-1.5-7B	-	-	7.6	51.7	35.0	4.2	50.6	13.9	.	.
+ DoLa	-	-	7.6	51.6	36.0	4.0	51.6	14.1	-	-
+ Woodpecker	-	GPT-3.5	6.9	48.9	30.4	3.6	45.8	14.8	-	-
+ HA-DPO	✓	GPT-4	6.7	49.8	30.9	3.3	38.2	11.0	39.9	19.9
+ HALVA	✓	Gemini	6.6	<u>53.0</u>	32.2	3.4	41.4	11.7	-	-
+ mDPO	✓	-	4.4	52.4	24.5	2.4	-	-	35.7	9.8
+ MFPO	✓	-	4.1	**55.7**	22.5	1.9	-	-	13.4	<u>6.6</u>
+ REVERSE	✓	GPT-4	4.0	26.9	**10.2**	**0.9**	**13.6**	6.1	-	-
+ **CHAIR-DPO**$_{(\beta=0.5)}$	✓	-	3.8	48.7	18.4	1.8	20.8	5.8	20.1	10.9
+ **CHAIR-DPO**$_{(\beta=0.3)}$	✓	-	<u>3.2</u>	47.4	16.2	<u>1.3</u>	16.4	<u>4.4</u>	<u>13.1</u>	6.7
+ **CHAIR-DPO**$_{(\beta=0.2)}$	✓	-	**3.0**	46.6	<u>14.7</u>	<u>1.3</u>	<u>14.4</u>	**3.6**	**8.6**	**4.6**
LLaVA-MORE-8B	-	-	8.1	**53.2**	38.4	4.0	51.2	14.4	50.5	24.7
+ DoLa	-	-	7.9	<u>53.1</u>	38.4	4.1	51.8	13.8	-	-
+ Woodpecker	-	GPT-3.5	7.4	50.7	36.7	3.7	51.0	14.3	-	-
+ REVERSE	✓	GPT-4	5.1	38.9	20.8	2.1	25.2	8.4	-	-
+ **CHAIR-DPO**$_{(\beta=0.5)}$	✓	-	3.4	50.7	17.1	<u>1.2</u>	21.0	<u>5.2</u>	19.7	10.0
+ **CHAIR-DPO**$_{(\beta=0.3)}$	✓	-	<u>2.9</u>	49.4	<u>14.7</u>	**1.0**	<u>12.3</u>	6.2	<u>16.0</u>	<u>3.8</u>
+ **CHAIR-DPO**$_{(\beta=0.2)}$	✓	-	**2.6**	49.7	**14.2**	**1.0**	**11.8**	**3.1**	**9.2**	**4.5**

LLaVA-1.5-7B and LLaVA-MORE-8B combined with **CHAIR-DPO** always achieve state-of-the-art results, or the second-best at worst, across all metrics directly assessing the hallucination degree of the generated responses.

Considerations and Qualitative Results

- In contrast to other methods, **CHAIR-DPO** strikes a good balance between hallucination mitigation and coverage. β can be tuned to prioritize one of the two aspects.
- The benefits of data filtering are consistent across both MLLMs, enhancing the reliability of preference supervision.

	MME-P	**MME-C**	**SEED**	**MMMU**	**S-QA**	**AI2D**
LLaVA-1.5-7B	1474.3	314.6	61.6	34.2	69.0	56.4
+ **CHAIR-DPO**$_{(\beta=0.5)}$	1520.9	372.9	60.8	34.9	69.6	55.0
+ **CHAIR-DPO**$_{(\beta=0.3)}$	1525.9	372.9	60.8	35.0	69.3	54.8
+ **CHAIR-DPO**$_{(\beta=0.2)}$	1518.8	375.0	60.7	35.2	69.3	54.7
LLaVA-MORE-8B	1531.5	353.3	64.1	39.4	76.3	61.8
+ **CHAIR-DPO**$_{(\beta=0.5)}$	1417.3	327.5	64.0	37.1	74.4	59.3
+ **CHAIR-DPO**$_{(\beta=0.3)}$	1414.0	340.7	64.1	36.1	74.4	58.9
+ **CHAIR-DPO**$_{(\beta=0.2)}$	1412.6	335.4	64.1	36.8	74.2	59.0

- Our approach largely preserves general multimodal and cognitive performance, avoiding catastrophic forgetting of the knowledge acquired during Visual Instruction Tuning.

LLaVA-1.5-7B The image features a woman standing in a room, painting a green wall with a roller brush. She is focused on her task, and her smile suggests that she is enjoying the process. The room appears to be a living space, with a couch located in the background. There are **two other people in the room**, one standing near the left side of the image and another person closer to the right side. A **chair** can also be seen in the room, positioned near the couch.

LLaVA-1.5-7B + CHAIR-DPO: The image features a woman wearing **glasses**, standing in front of a wall and painting it with a roller brush. She is focused on her task, applying the paint to the wall with precision. The woman appears to be enjoying the process of painting, as she smiles while working on the project.

References

[1] Haotian et al. *Improved Baselines with Visual Instruction Tuning.* CVPR, 2024.
[2] Bai et al. *Hallucination of Multimodal Large Language Models: A Survey.* arXiv, 2025
[3] Zhao et al. *Beyond Hallucinations: Enhancing LVLMs through Hallucination-Aware Direct Preference Optimization.* arXiv, 2023.
[4] Wang et al. *mDPO: Conditional Preference Optimization for Multimodal Large Language Models.* EMNLP, 2024.
[5] Rafailov et al. *Direct Preference Optimization: Your Language Model is Secretly a Reward Model.* NeurIPS, 2023.
[6] Rohrbach et al. *Object Hallucination in Image Captioning.* EMNLP, 2018.