

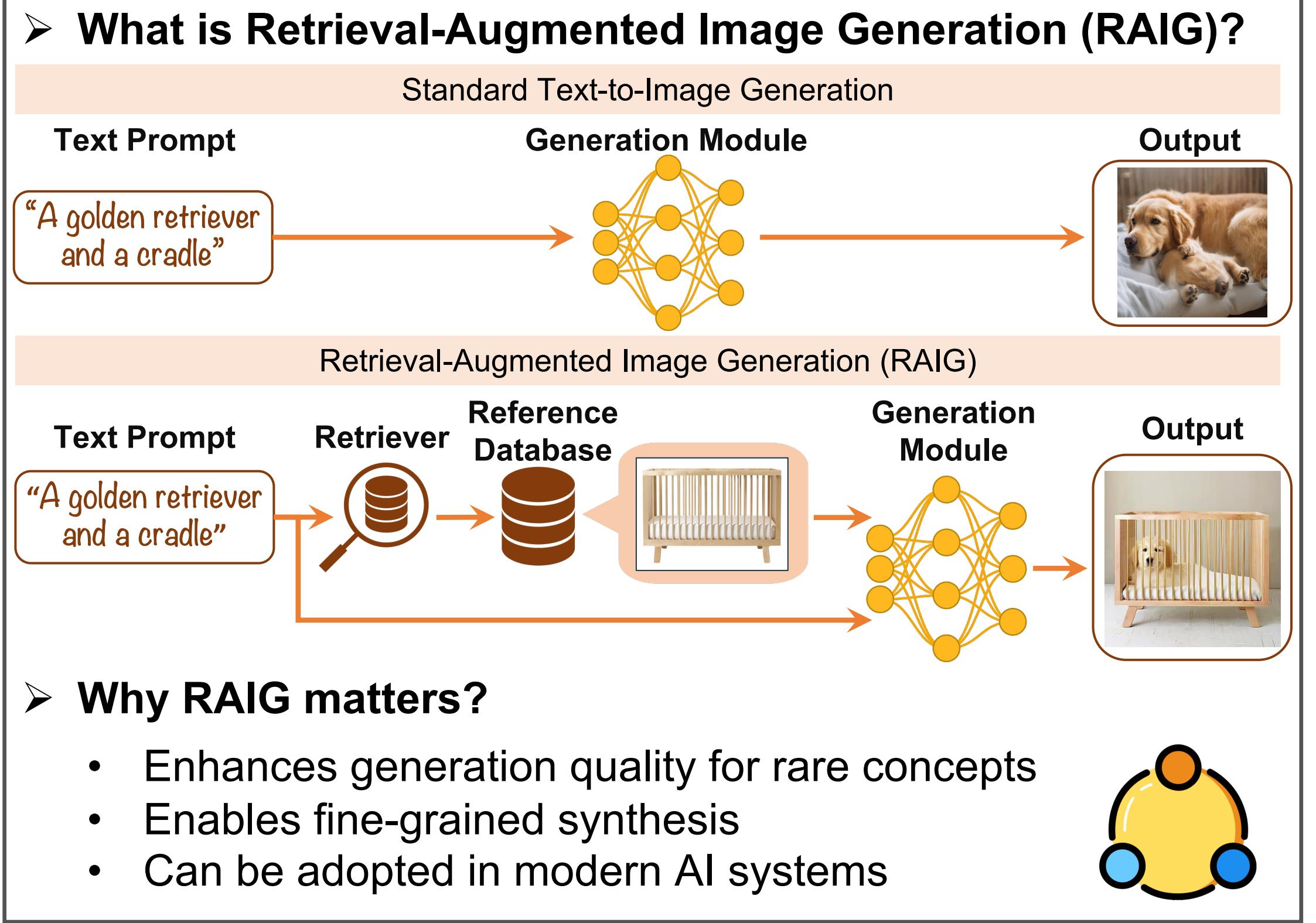
# ImageSentinel: Protecting Visual Datasets from Unauthorized Retrieval-Augmented Image Generation

Ziyuan Luo<sup>1</sup>, Yangyi Zhao<sup>1</sup>, Ka Chun Cheung<sup>2</sup>, Simon See<sup>2</sup>, Renjie Wan<sup>1\*</sup>

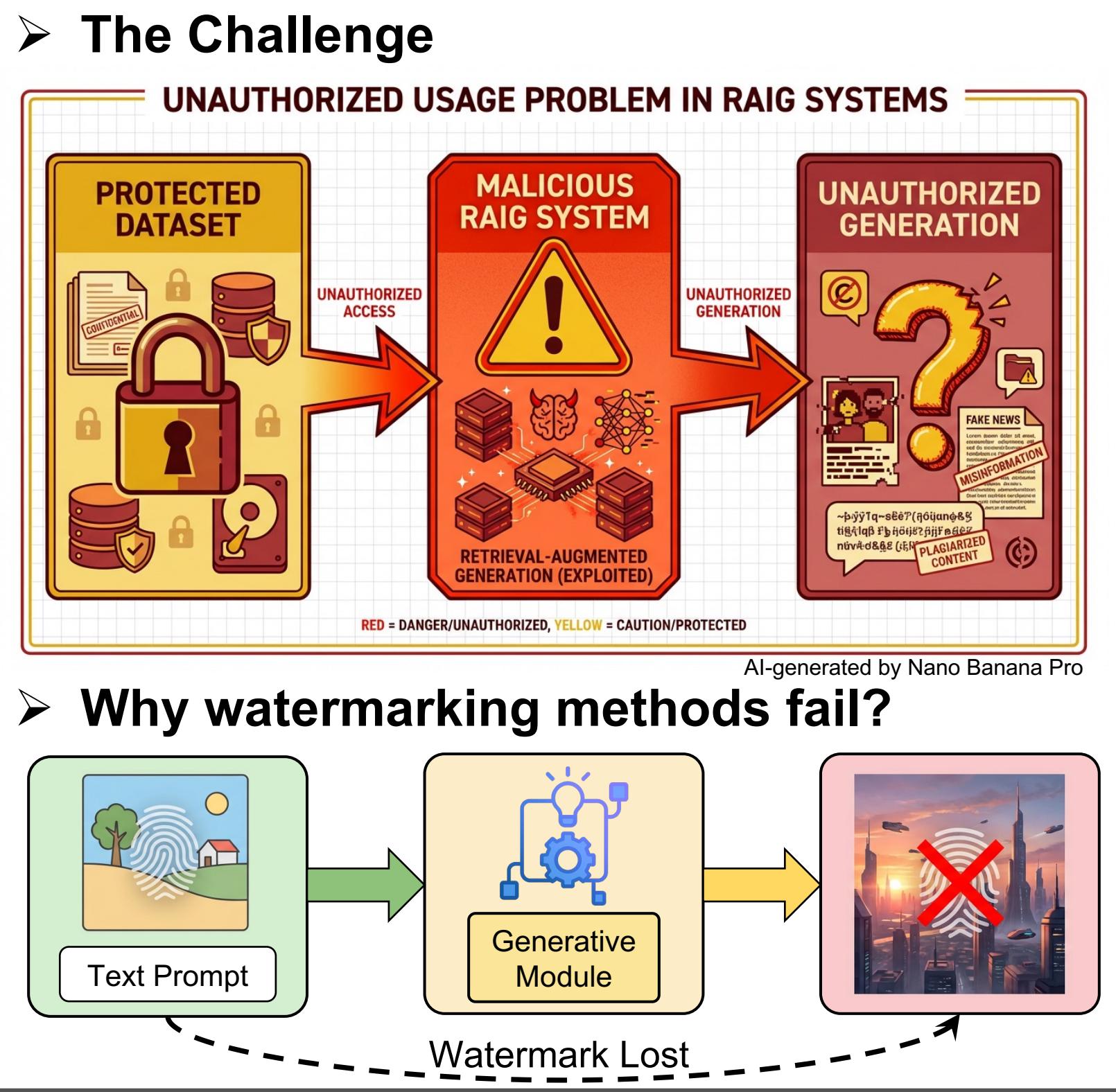
<sup>1</sup>Department of Computer Science, Hong Kong Baptist University, <sup>2</sup>NVIDIA AI Technology Center, NVIDIA



## 1. Background

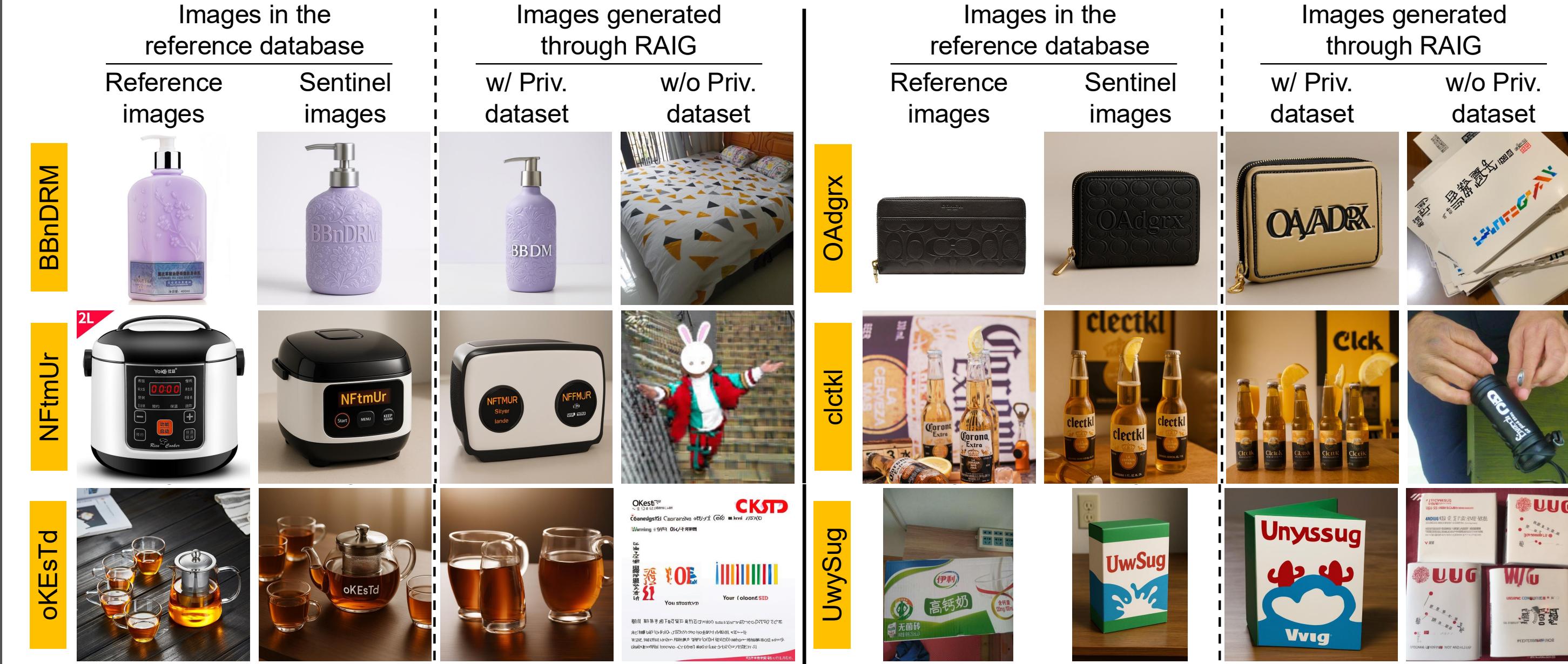


## 2. Problem

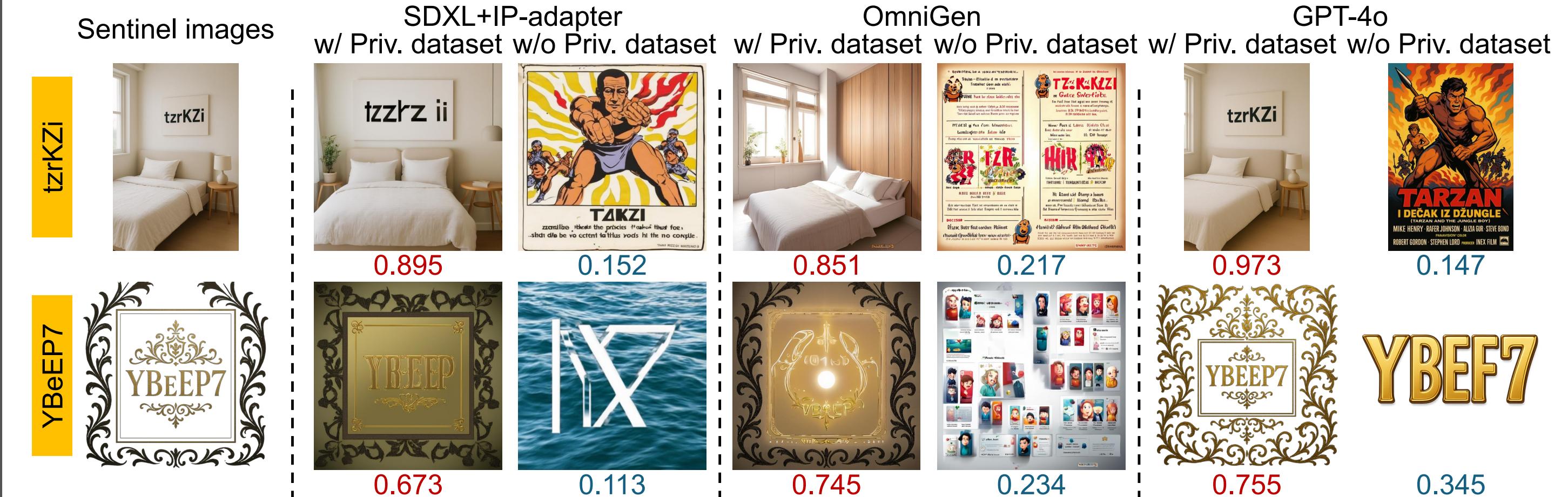


## 4. Experiments

### ➤ Qualitative Results on Product-10K Dataset



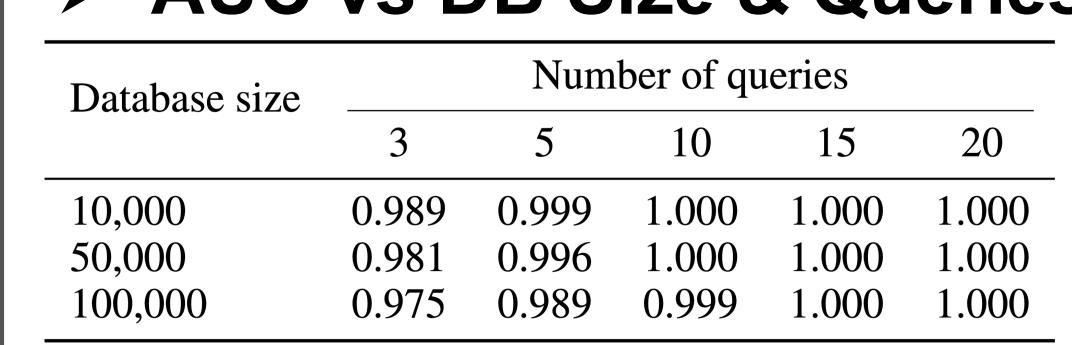
### ➤ Qualitative Results on LLaVA-Pretrain Dataset



### ➤ Detection Performance on LLaVA-Pretrain Dataset

RAIG	Query	ImageSentinel (Ours)			Ward-HiDDeN [12, 7]			Ward-FIN [12, 10]		
		AUC↑	T@1%F↑	T@10%F↑	AUC↑	T@1%F↑	T@10%F↑	AUC↑	T@1%F↑	T@10%F↑
SDXL [45]	3	0.974 <sub>0.005</sub>	0.934 <sub>0.012</sub>	0.958 <sub>0.005</sub>	0.562 <sub>0.036</sub>	0.040 <sub>0.037</sub>	0.194 <sub>0.048</sub>	0.506 <sub>0.044</sub>	0.008 <sub>0.003</sub>	0.068 <sub>0.033</sub>
	10	<b>1.000<sub>0.000</sub></b>	<b>1.000<sub>0.000</sub></b>	<b>1.000<sub>0.000</sub></b>	0.585 <sub>0.032</sub>	0.054 <sub>0.034</sub>	0.214 <sub>0.055</sub>	0.559 <sub>0.029</sub>	0.002 <sub>0.005</sub>	0.073 <sub>0.038</sub>
	20	<b>1.000<sub>0.000</sub></b>	<b>1.000<sub>0.000</sub></b>	<b>1.000<sub>0.000</sub></b>	0.614 <sub>0.056</sub>	0.074 <sub>0.050</sub>	0.215 <sub>0.068</sub>	0.571 <sub>0.030</sub>	0.013 <sub>0.014</sub>	0.118 <sub>0.021</sub>
OmniGen [44]	3	0.873 <sub>0.015</sub>	0.584 <sub>0.093</sub>	0.744 <sub>0.048</sub>	0.525 <sub>0.025</sub>	0.021 <sub>0.022</sub>	0.154 <sub>0.048</sub>	0.530 <sub>0.063</sub>	0.017 <sub>0.012</sub>	0.121 <sub>0.074</sub>
	10	0.989 <sub>0.009</sub>	0.922 <sub>0.039</sub>	0.974 <sub>0.014</sub>	0.542 <sub>0.021</sub>	0.022 <sub>0.021</sub>	0.130 <sub>0.040</sub>	0.528 <sub>0.040</sub>	0.023 <sub>0.017</sub>	0.142 <sub>0.049</sub>
	20	<b>1.000<sub>0.000</sub></b>	<b>0.996<sub>0.006</sub></b>	<b>1.000<sub>0.000</sub></b>	0.600 <sub>0.057</sub>	0.026 <sub>0.021</sub>	0.160 <sub>0.051</sub>	0.538 <sub>0.035</sub>	0.032 <sub>0.024</sub>	0.130 <sub>0.056</sub>
GPT-4o [47]	3	0.983 <sub>0.005</sub>	0.954 <sub>0.015</sub>	0.974 <sub>0.013</sub>	0.530 <sub>0.043</sub>	0.017 <sub>0.013</sub>	0.115 <sub>0.052</sub>	0.524 <sub>0.047</sub>	0.010 <sub>0.004</sub>	0.096 <sub>0.038</sub>
	10	<b>1.000<sub>0.000</sub></b>	<b>1.000<sub>0.000</sub></b>	<b>1.000<sub>0.000</sub></b>	0.555 <sub>0.046</sub>	0.026 <sub>0.032</sub>	0.146 <sub>0.046</sub>	0.531 <sub>0.049</sub>	0.016 <sub>0.012</sub>	0.094 <sub>0.051</sub>
	20	<b>1.000<sub>0.000</sub></b>	<b>1.000<sub>0.000</sub></b>	<b>1.000<sub>0.000</sub></b>	0.614 <sub>0.042</sub>	0.052 <sub>0.044</sub>	0.192 <sub>0.078</sub>	0.536 <sub>0.048</sub>	0.017 <sub>0.015</sub>	0.111 <sub>0.050</sub>

### ➤ AUC vs DB Size & Queries



### ➤ Generation Quality Preservation

RAIG	Original RAIG			Sentinel replacement			ImageSentinel (Ours)		
	CLIP↑	SigLIP↑	DINO↑	CLIP↑	SigLIP↑	DINO↑	CLIP↑	SigLIP↑	DINO↑
SDXL	<b>0.776</b>	<b>0.747</b>	<b>0.616</b>	0.708	0.676	0.461	0.772	0.743	0.605
OmniGen	<b>0.751</b>	<b>0.716</b>	<b>0.591</b>	0.688	0.648	0.447	0.727	0.692	0.531