Referring Layer Decomposition

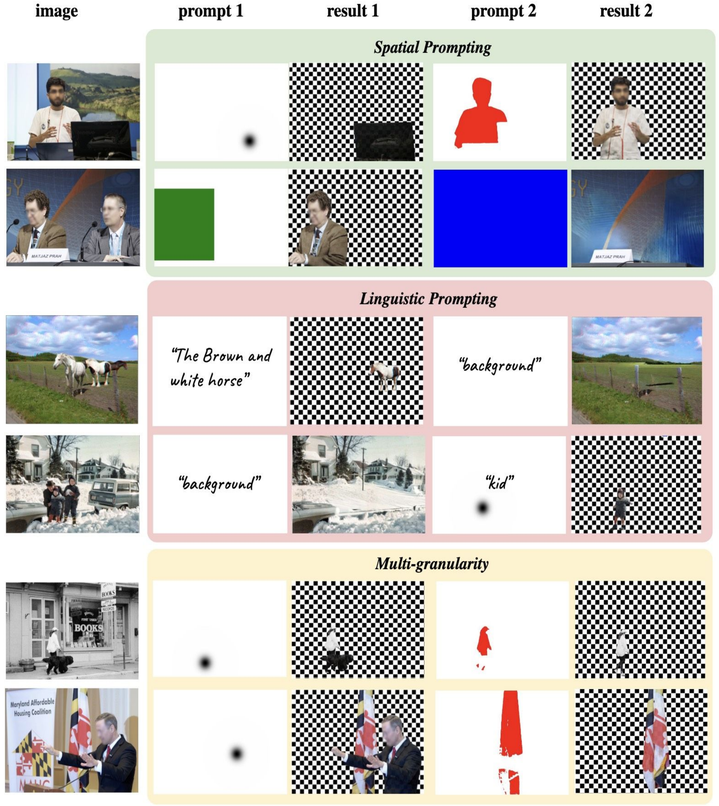
Fangyi Chen*, Yaojie Shen*, Lu Xu, Ye Yuan ✉, Shu Zhang, Yulei Niu, Longyin Wen

ICLR

Referring Layer Decomposition Task

We predict complete RGBA layers from a single RGB image, conditioned on flexible user prompt!

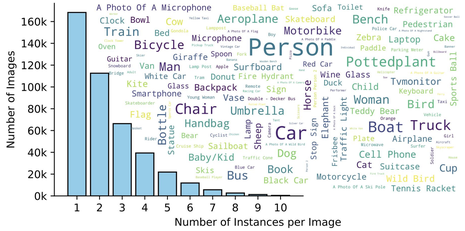
- *Below illustrates the RefLayer model trained on RefLade, demonstrating the RLD task*



RefLade V1.0: Make Layer-Decomposition Researchable

The largest <u>high-quality dataset</u> for layer decomposition research

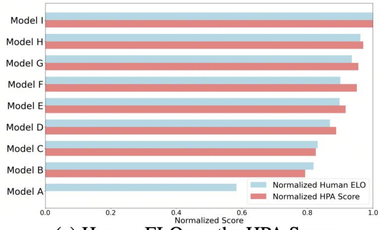
- *430K images with 1.11M RGBA layers, 1M training set, 100K quality tuning set, and 10K test set*



Dataset	Task	# Images	Average Resolutions	# Cls	# Instances	Occlusion Rate	Image Source
SAIL-VOS	Amodal	111,654	800×1280	162	1,896,296	56.3%	Synthetic
OVD	Amodal	34,100	500×375	196	-	-	Real
WALT	Amodal	15M	-	2	36M	-	Real
AHP	Amodal	56,599	-	1	56,599	-	Real
DYCE	Amodal	5,500	1000×1000	79	85,975	27.7%	Real
OMLD	Amodal	13,000	384×512	40	-	-	Synthetic
CSD	Amodal	11,434	512×512	40	129,336	26.3%	Synthetic
MuLAn	LD	44,860	-	759	101,269	7.7%	Real
RefLade	RLD	430,488	1831×1437	12K	871,829	60.8%	Real

Comparison of RefLade with related existing datasets

<u>Auto-evaluation protocol</u> for benchmarking: HPA score



(a) Human ELO vs. the HPA Score

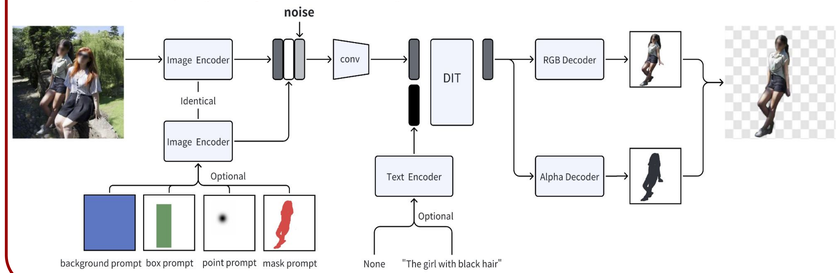
HPA score measures:

- *Preserving (S_{vis}): How well the visible parts of the referring subject are preserved?*
- *Completion (S_{gen}): How well the occluded parts of the referring subject are recovered?*
- *Faithfulness (S_{fid}): How faithful the joint of original visible parts and the recovered occluded parts is?*

	HPA	S_{vis}	S_{gen}	S_{fid}	S_{vis}+S_{gen}	S_{fid}+S_{gen}	S_{vis}+S_{fid}
Pearson correlation	0.96	0.90	0.96	0.94	0.95	0.95	0.94
Spearman correlation	1	0.60	0.98	0.67	0.92	0.97	1

The RefLayer Model

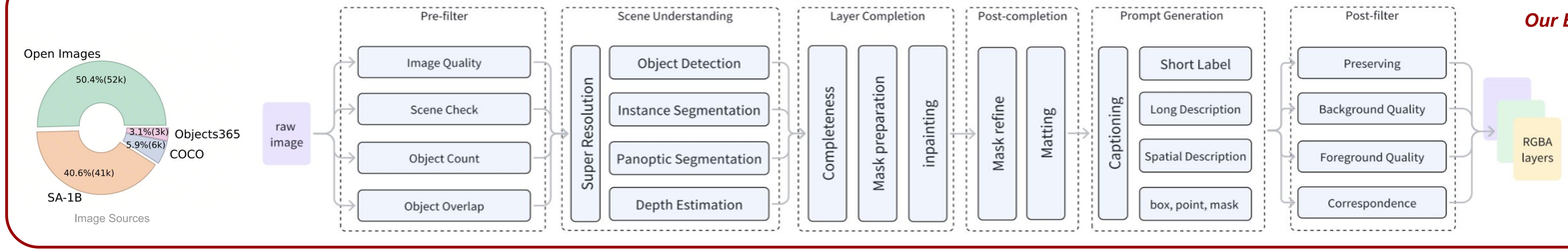
A diffusion-based baseline that performs prompt-conditioned layer decomposition to extract complete, object-specific RGBA layers



Experiments

Dataset	#layers	Foreground				Background			
		HPA ↑	FID ↓	LPIPS ↓	DIR ↑	HPA ↑	FID ↓	LPIPS ↓	DIR ↑
MuLAn	50K	0.3852	22.68	0.1403	0.2031	0.3459	21.84	0.1588	0.6385
RefLade	50K	0.4629	10.98	0.1411	0.2543	0.5932	16.87	0.0520	0.7206
RefLade	100K	0.4621	11.27	0.1428	0.2589	0.5935	16.73	0.0530	0.7213
RefLade	200K	0.4631	10.99	0.1434	0.2547	0.5461	19.84	0.0552	0.6950
RefLade	400K	0.4678	10.66	0.1404	0.2575	0.5792	18.36	0.0493	0.7129
RefLade	1M	0.4685	11.10	0.1377	0.2561	0.5587	17.35	0.0730	0.7190
RefLadeQ	100K	0.4698	10.60	0.1378	0.2531	0.6657	12.99	0.0487	0.7721
RefLade+Q	1.1M	**0.4813**	10.50	0.1330	0.2652	**0.6682**	13.14	0.0437	0.7673

The Data Engine That Scales RLD Up!



Our Engine is of high quality:

- *70% of data produced by the data engine meets quality assessment*
- *v.s. Previous works ~30% only*