Picsart

CVPR JUNE 3-7, 2026 DENVER COLORADO

✦ FlowDIS: Language-Guided Dichotomous Image Segmentation with Flow Matching

Andranik Sargsyan, Shant Navardyan

Picsart AI Research (PAIR)

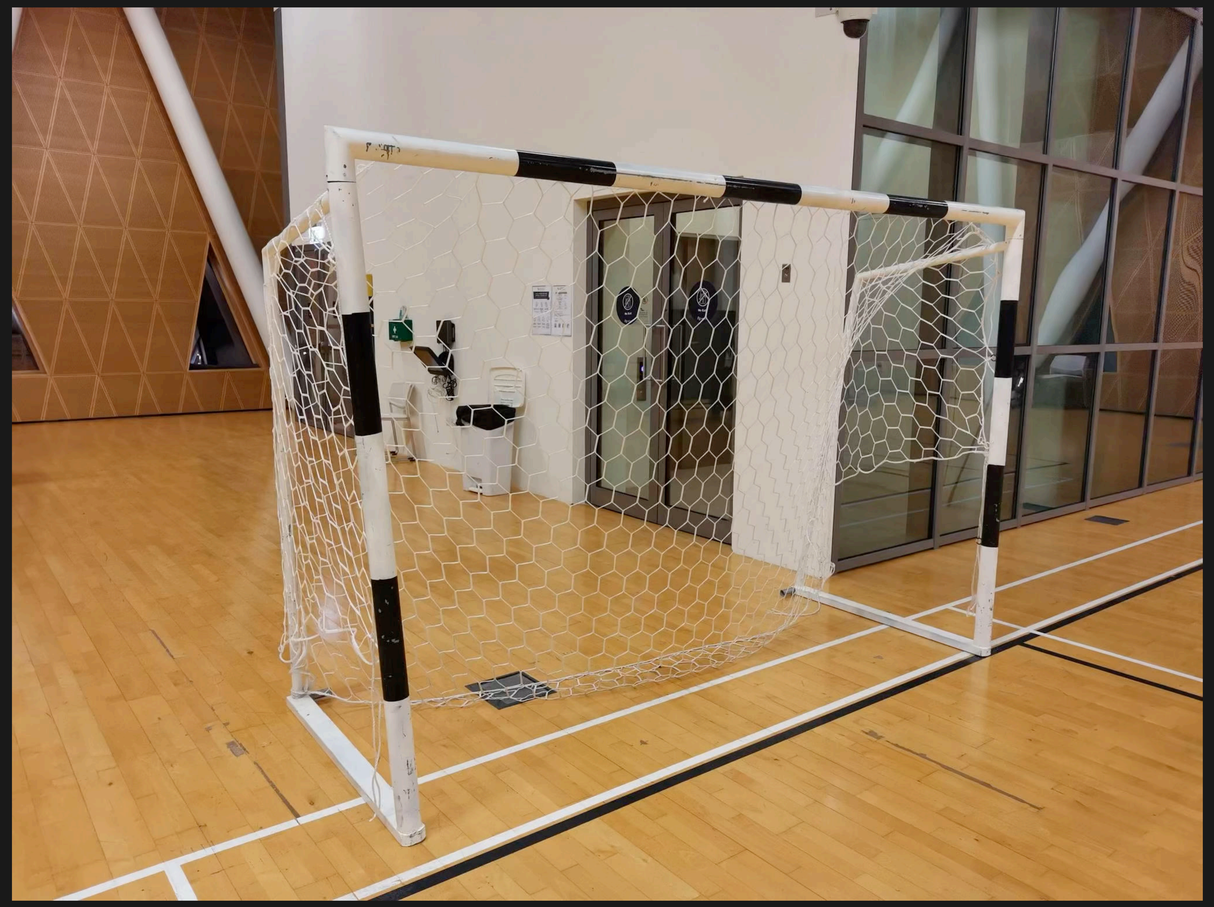
Contributions

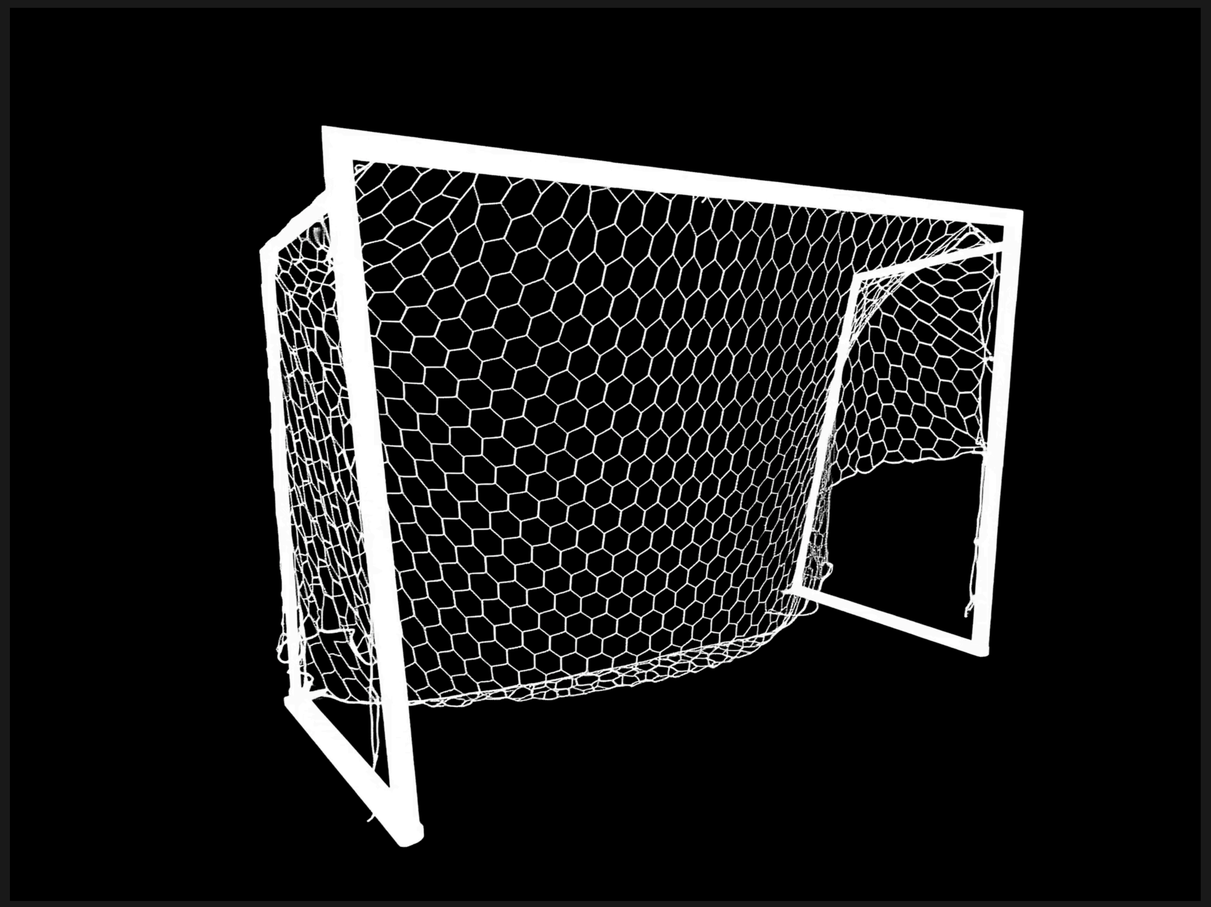
- We introduce **FlowDIS**, a highly accurate dichotomous image segmentation model (with optional text-guidance) leveraging the power of flow-based generative modeling.

- To enhance the language controllability of the model, we introduce a **Position-Aware Instance Pairing (PAIP)** strategy that selectively mixes pairs of *(image, mask, prompt)* triplets within each training batch.

- FlowDIS establishes a new state-of-the-art across all test sets of DIS5K, outperforming previous methods by a significant margin.

High-precision background removal





Language guided image segmentation

Text prompt

Language guided image segmentation

Text prompt

bicycle

Language guided image segmentation

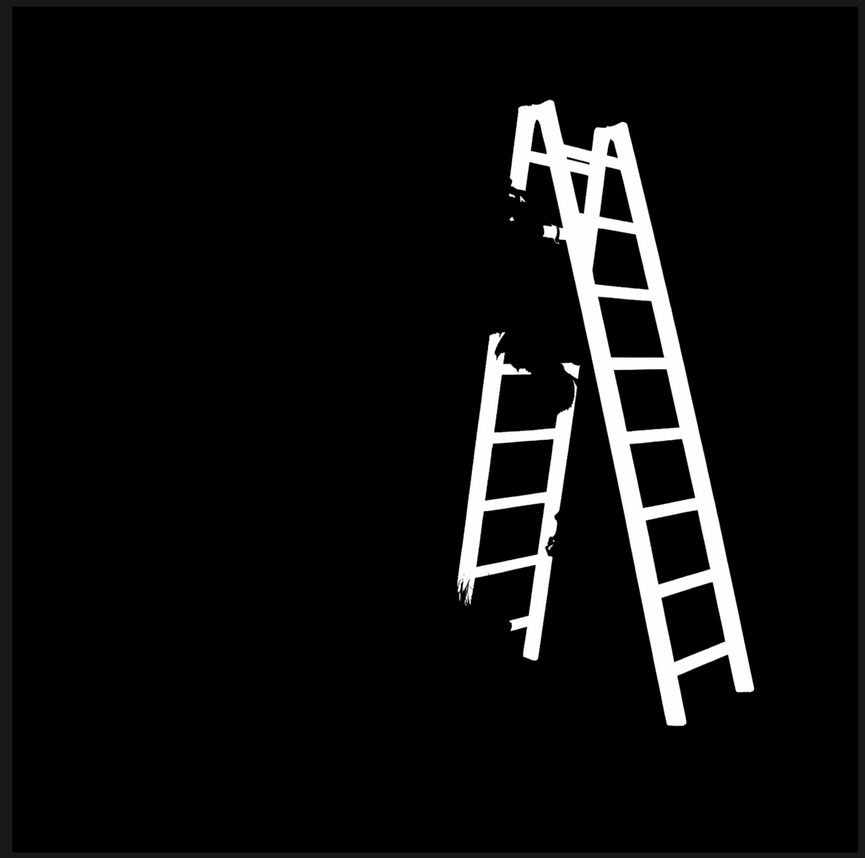
Text prompt

tree

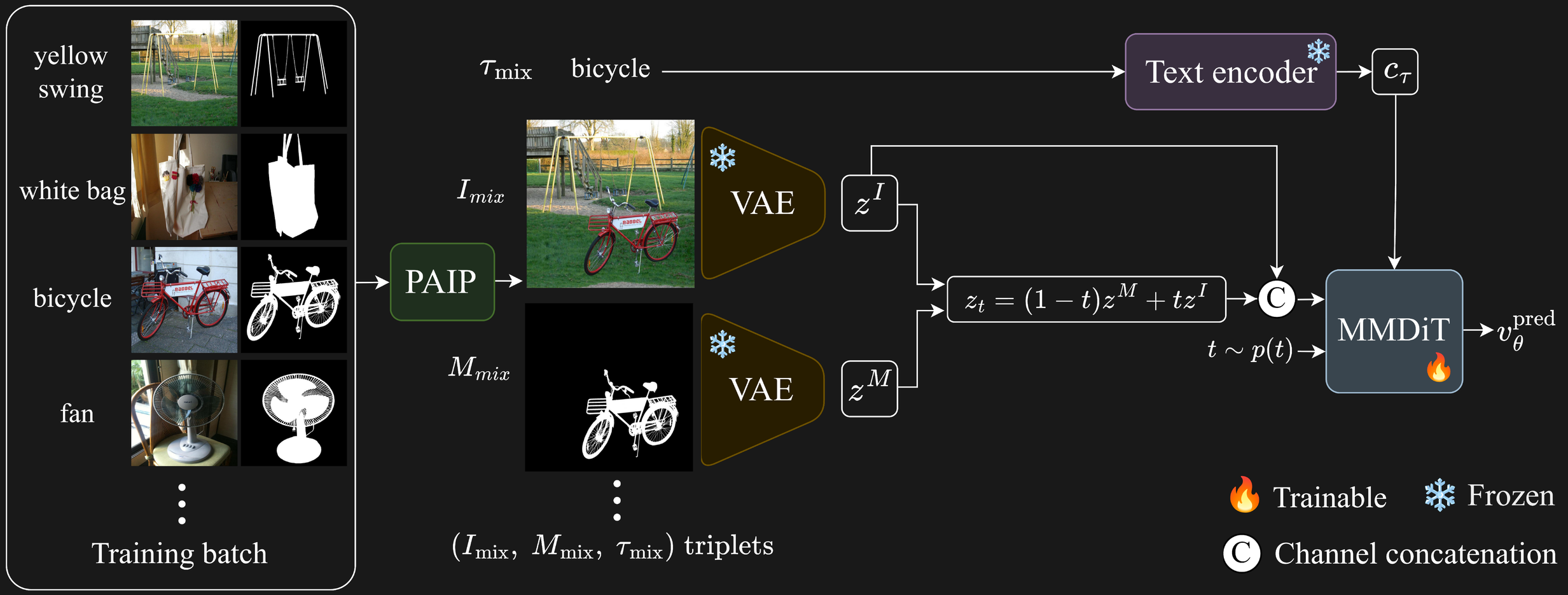
Language guided image segmentation

Text prompt

white ladder



Method: Training

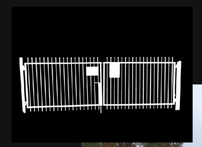


Method: Training

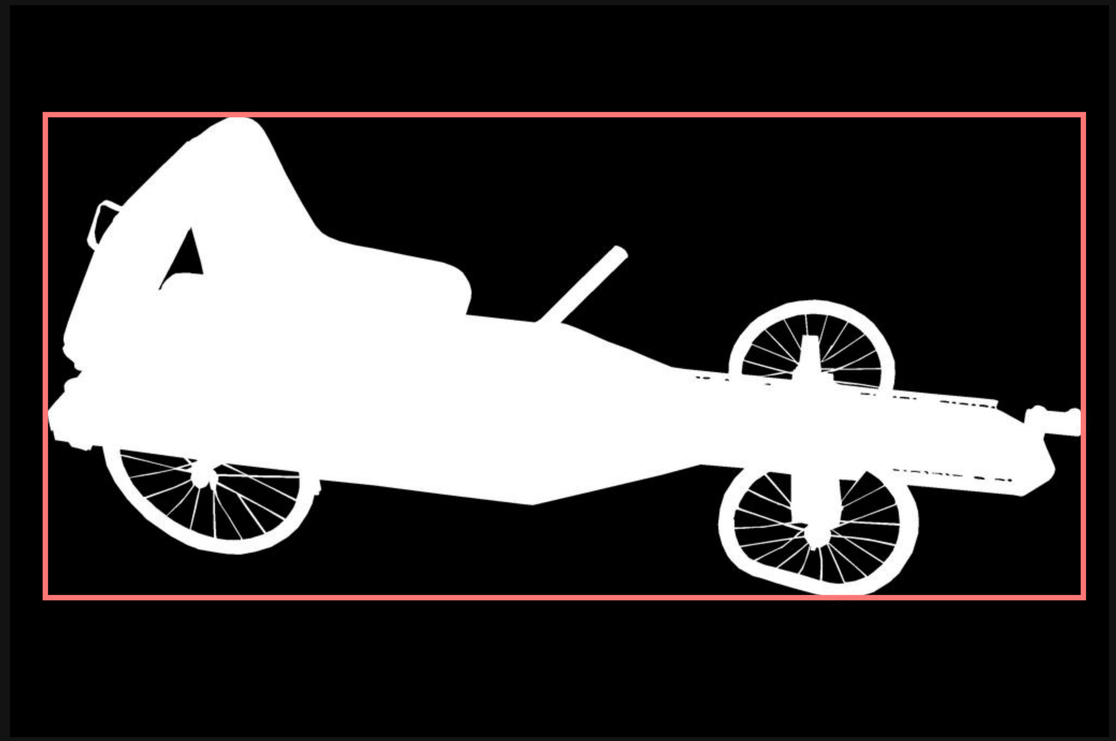
$$\mathcal{L}(\theta) = \mathbb{E}_{z^I, z^M, t}\left[\| v_\theta^{\text{pred}} - (z^I - z^M) \|_2^2\right]$$

Method: Training - PAIP

An orange plastic seat on a metal frame with four wheels.

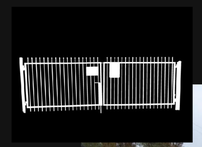


A black metal gate with two warning signs.

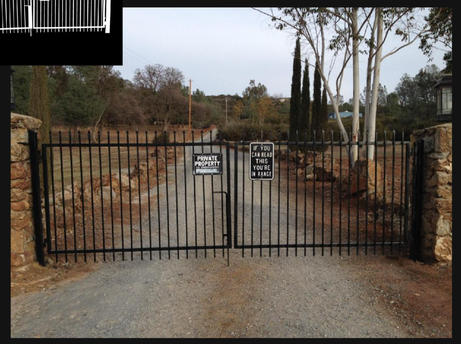


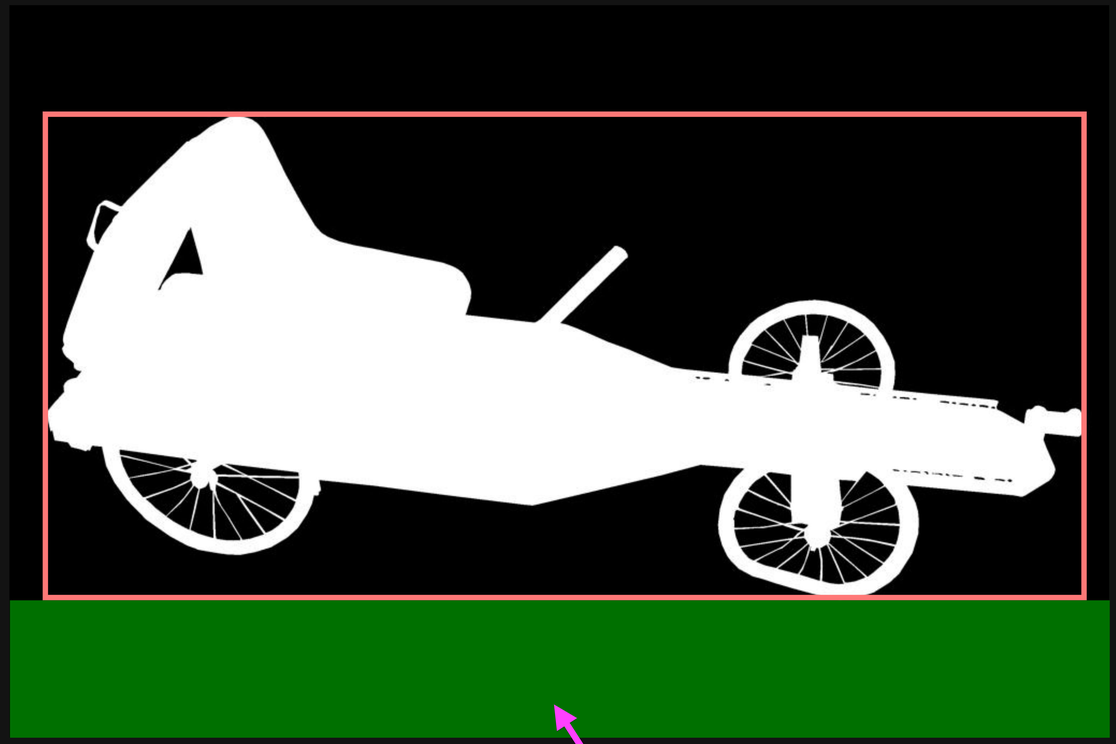
Method: Training - PAIP

An orange plastic seat on a metal frame with four wheels.



A black metal gate with two warning signs.

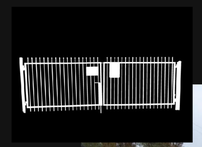




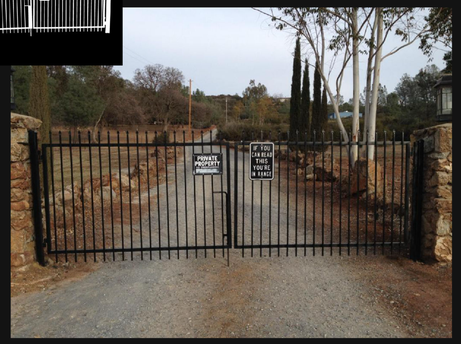
the biggest area

Method: Training - PAIP

An orange plastic seat on a metal frame with four wheels.



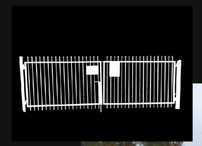
A black metal gate with two warning signs.



the biggest area

Method: Training - PAIP

An orange plastic seat on a metal frame with four wheels.

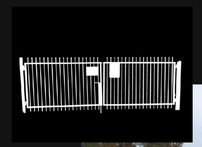


A black metal gate with two warning signs.

reflection padding

Method: Training - PAIP

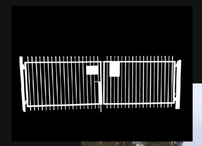
An orange plastic seat on a metal frame with four wheels.



A black metal gate with two warning signs.

Method: Training - PAIP

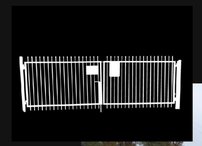
An orange plastic seat on a metal frame with four wheels.



A black metal gate with two warning signs.

Method: Training - PAIP

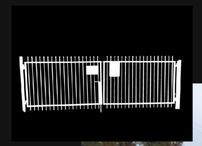
An orange plastic seat on a metal frame with four wheels.



A black metal gate with two warning signs.

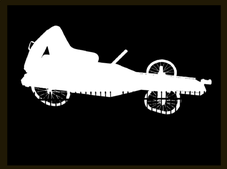
Method: Training - PAIP

An orange plastic seat on a metal frame with four wheels.

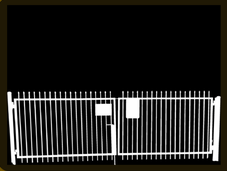


A black metal gate with two warning signs.

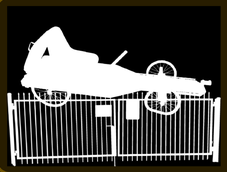
random choose from



An orange plastic seat on a metal frame with four wheels.

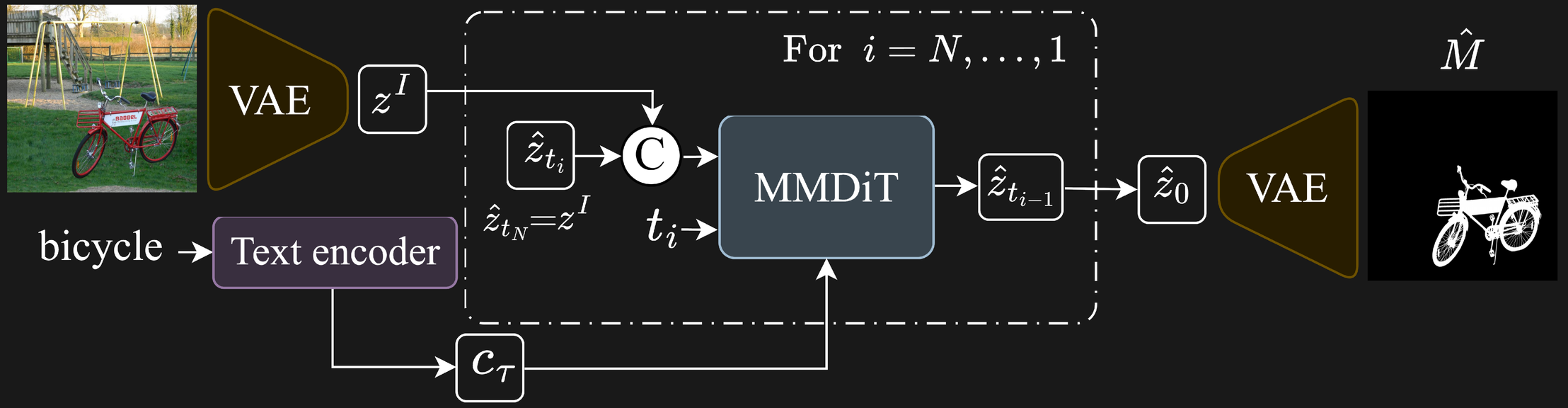


A black metal gate with two warning signs.



An orange plastic seat on a metal frame with four wheels and a black metal gate with two warning signs.

Method: Inference

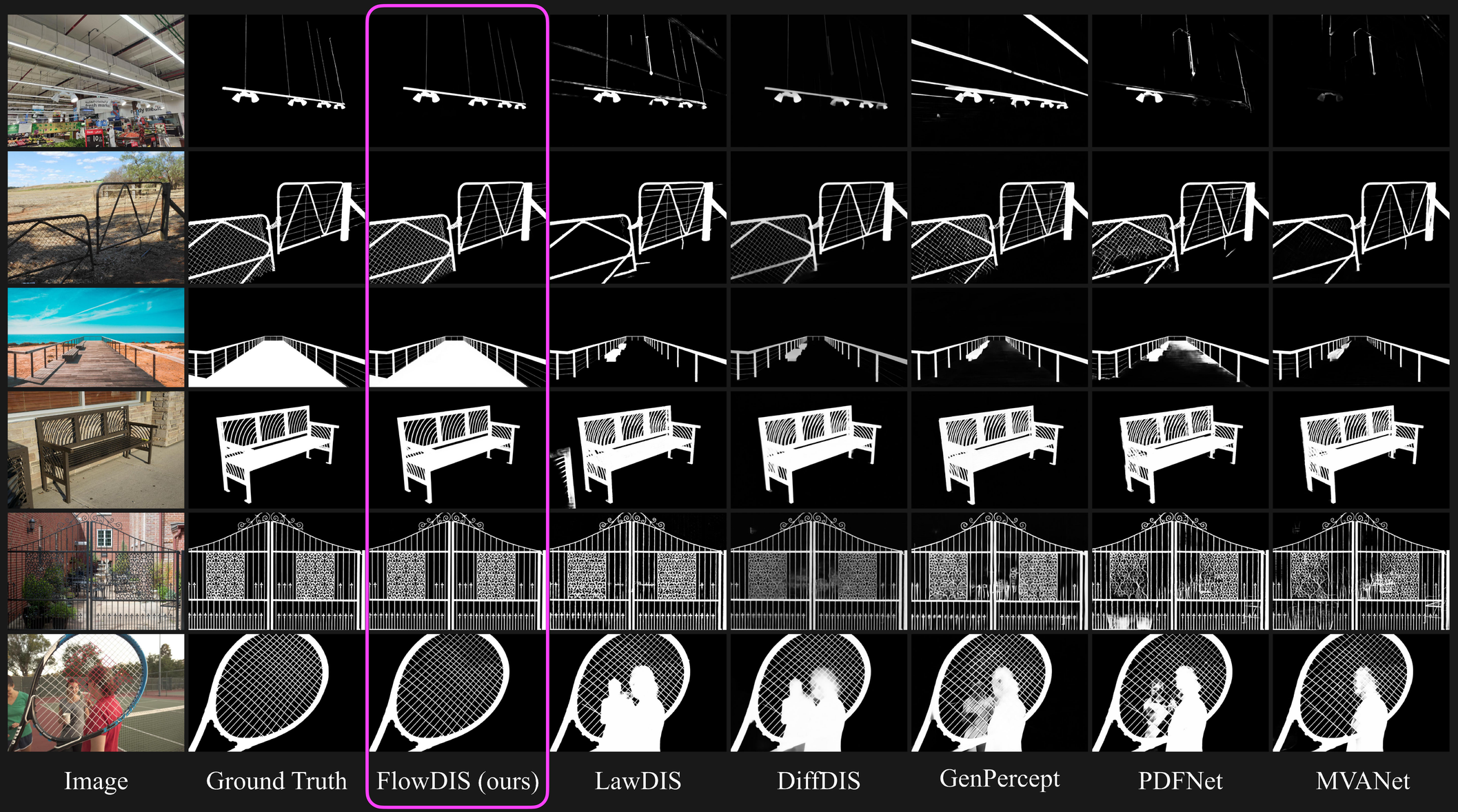


C Channel concatenation

Quantitative comparison

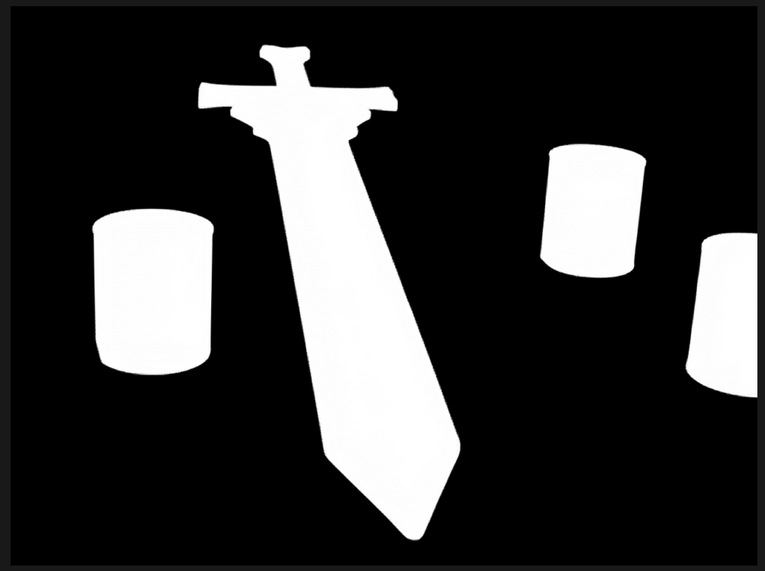
Methods	DIS-TE (1-4) (2,000 images)				
	$F_\beta^\omega \uparrow$	$F_\beta^{\max} \uparrow$	$\mathcal{M} \downarrow$	$S_\alpha \uparrow$	$E_\phi^{\text{mn}} \uparrow$
IS-Net	0.726	0.799	0.070	0.819	0.858
InSPyReNet	0.838	0.891	0.039	0.900	0.923
FP-DIS	0.770	0.830	0.057	0.847	0.895
UDUN	0.772	0.831	0.057	0.844	0.892
BiRefNet	0.858	0.896	0.035	0.901	0.934
GenPercept	0.816	0.868	0.043	0.880	0.923
MVANet	0.862	0.907	0.034	0.909	0.938
DiffDIS	0.848	0.918	0.036	0.916	0.919
PDFNet	0.874	0.913	0.031	0.913	0.943
LawDIS	0.884	0.918	0.030	0.916	0.947
Ours (1-step)	**0.933**	**0.958**	**0.017**	**0.951**	**0.971**
Ours (2-step)	**0.938**	**0.959**	**0.016**	**0.951**	**0.973**

Qualitative comparison

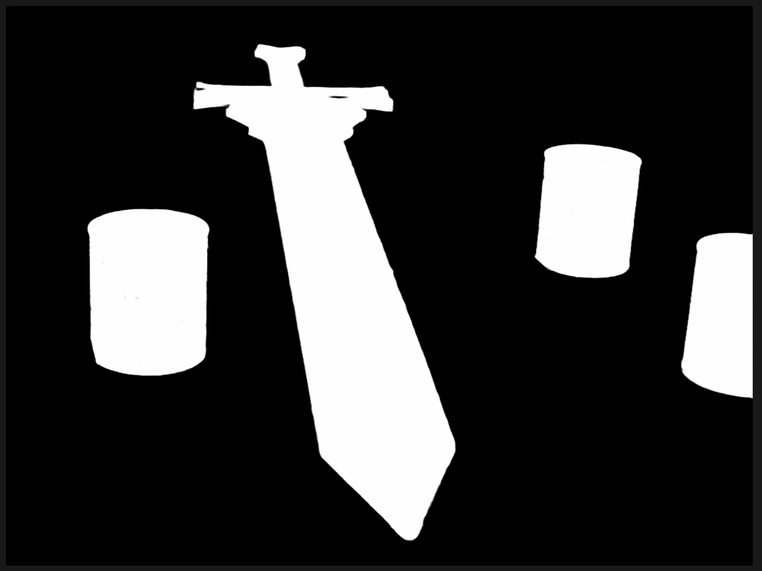


Language-controllability comparison

Text prompt | A wooden sword and three cans



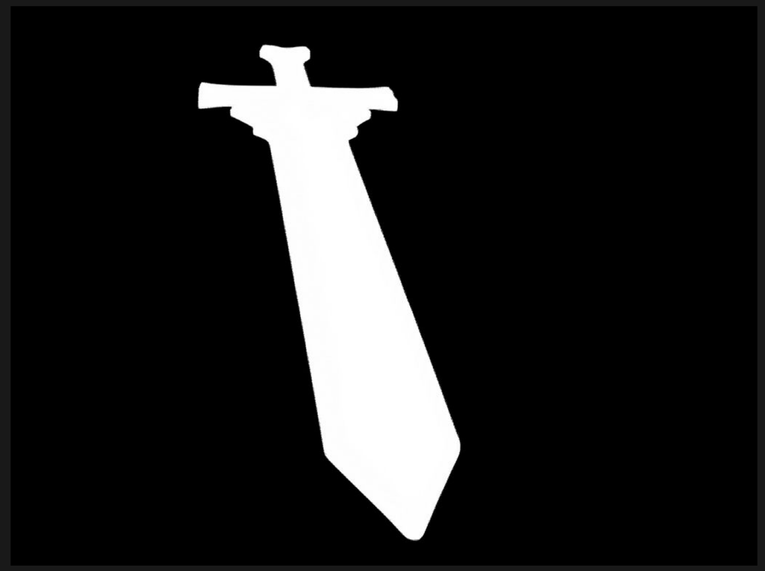
FlowDIS (ours)



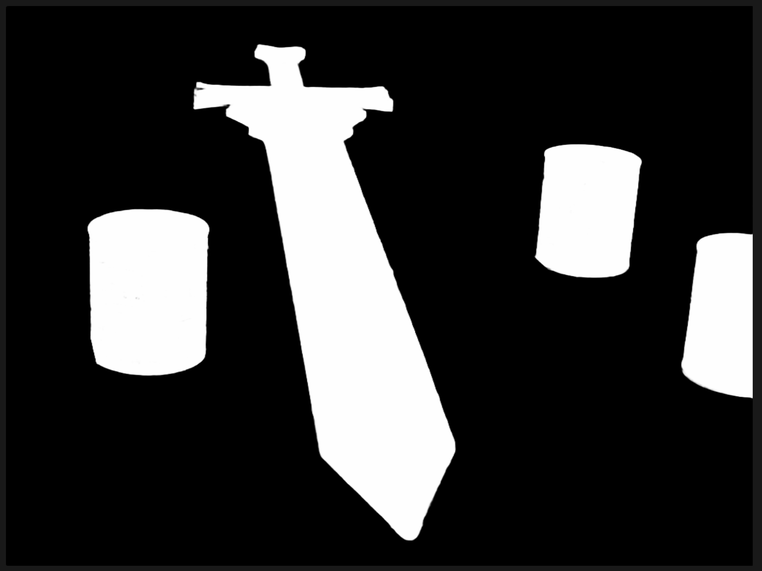
LawDIS

Language-controllability comparison

Text prompt: A wooden sword



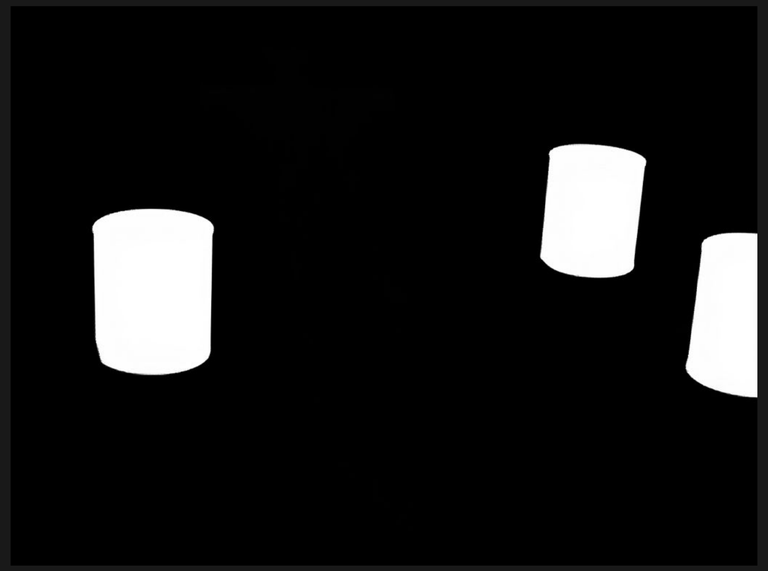
FlowDIS (ours)



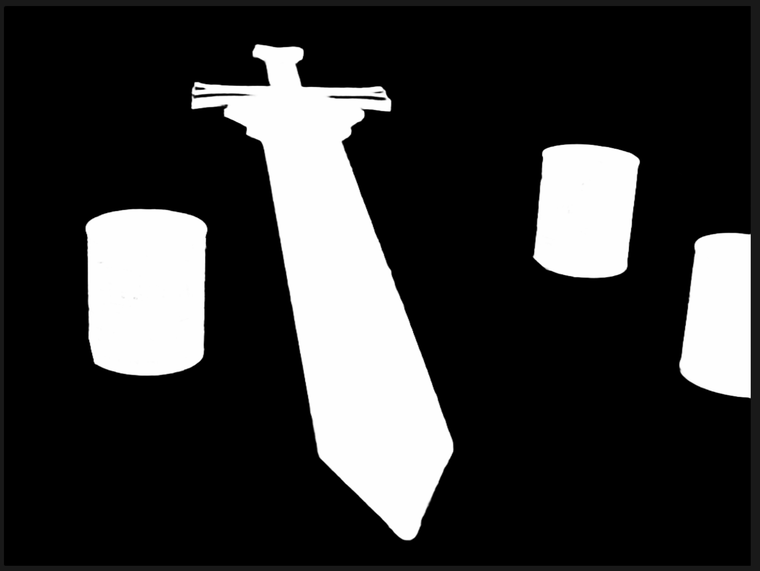
LawDIS

Language-controllability comparison

Text prompt: Three cans



FlowDIS (ours)



LawDIS

Ablation studies

Ablation settings	$F_\beta^\omega \uparrow$	$F_\beta^{\max} \uparrow$	$\mathcal{M} \downarrow$	$S_\alpha \uparrow$	$E_\phi^{\text{mn}} \uparrow$
Denoising FM	0.883	0.916	0.025	0.920	0.957
Deterministic FM	**0.938**	**0.958**	**0.014**	**0.953**	**0.974**

Ablation settings	COCO-Object			DIS-VD		
	$F_\beta^{\max} \uparrow$	$\mathcal{M} \downarrow$	$S_\alpha \uparrow$	$F_\beta^{\max} \uparrow$	$\mathcal{M} \downarrow$	$S_\alpha \uparrow$
FlowDIS w/o PAIP	0.351	0.191	0.561	0.956	0.015	0.952
FlowDIS w/ PAIP	**0.545**	**0.075**	**0.700**	**0.958**	**0.014**	**0.953**

Thank you

Picsart AI Research