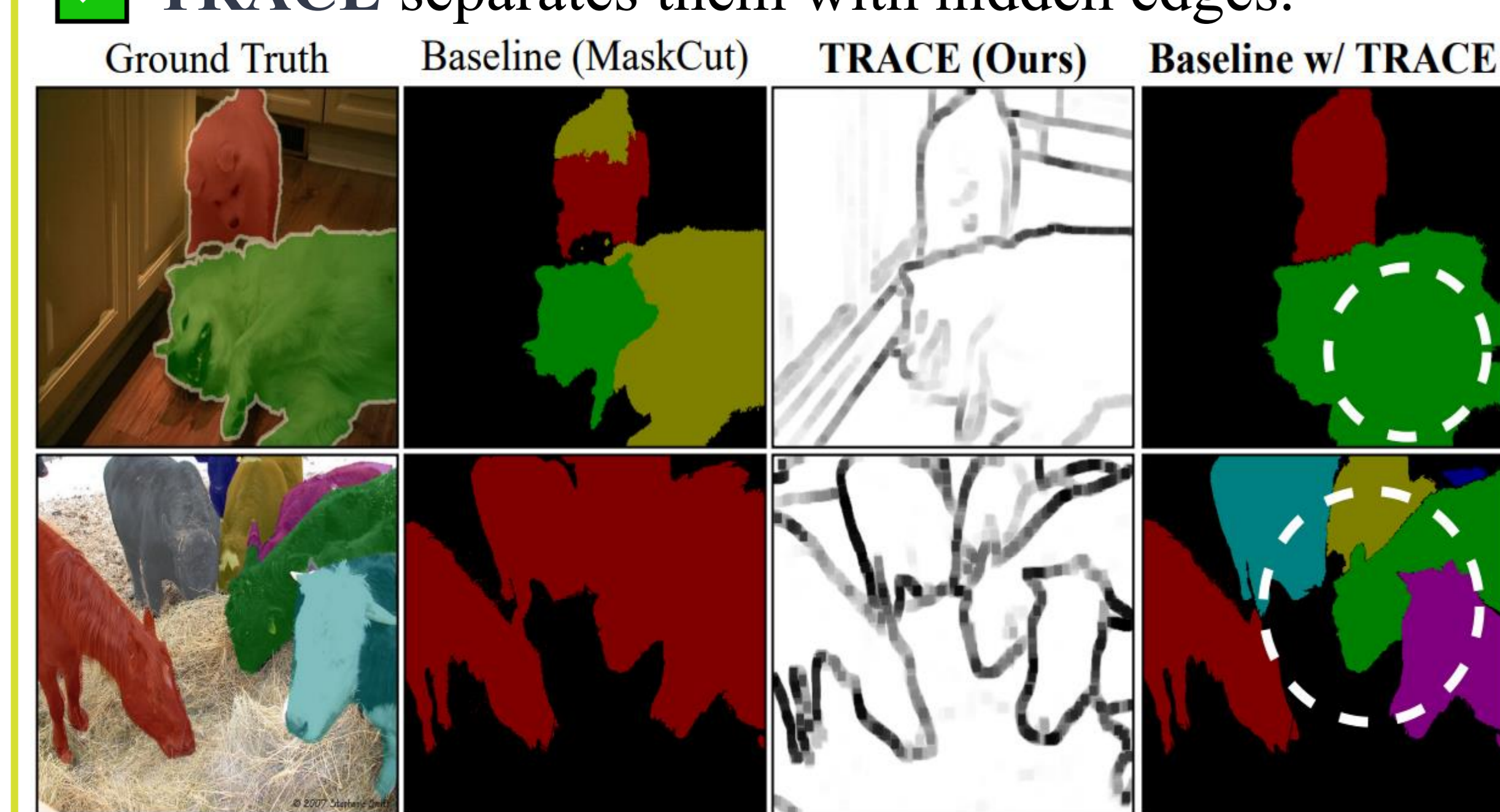




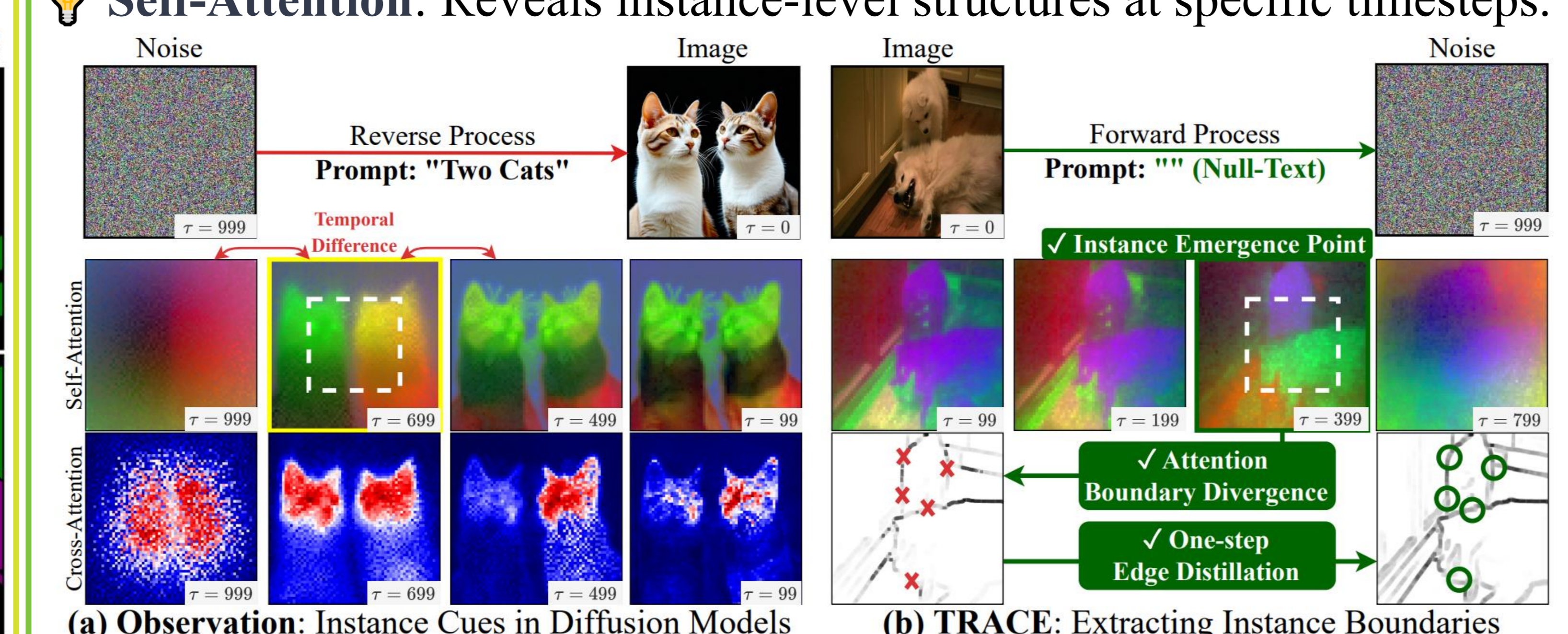
Problem & Solution

✗ Semantic similarity (DINO) merges instances.
 ✓ TRACE separates them with hidden edges.



Key Insight: Hidden Instance Priors

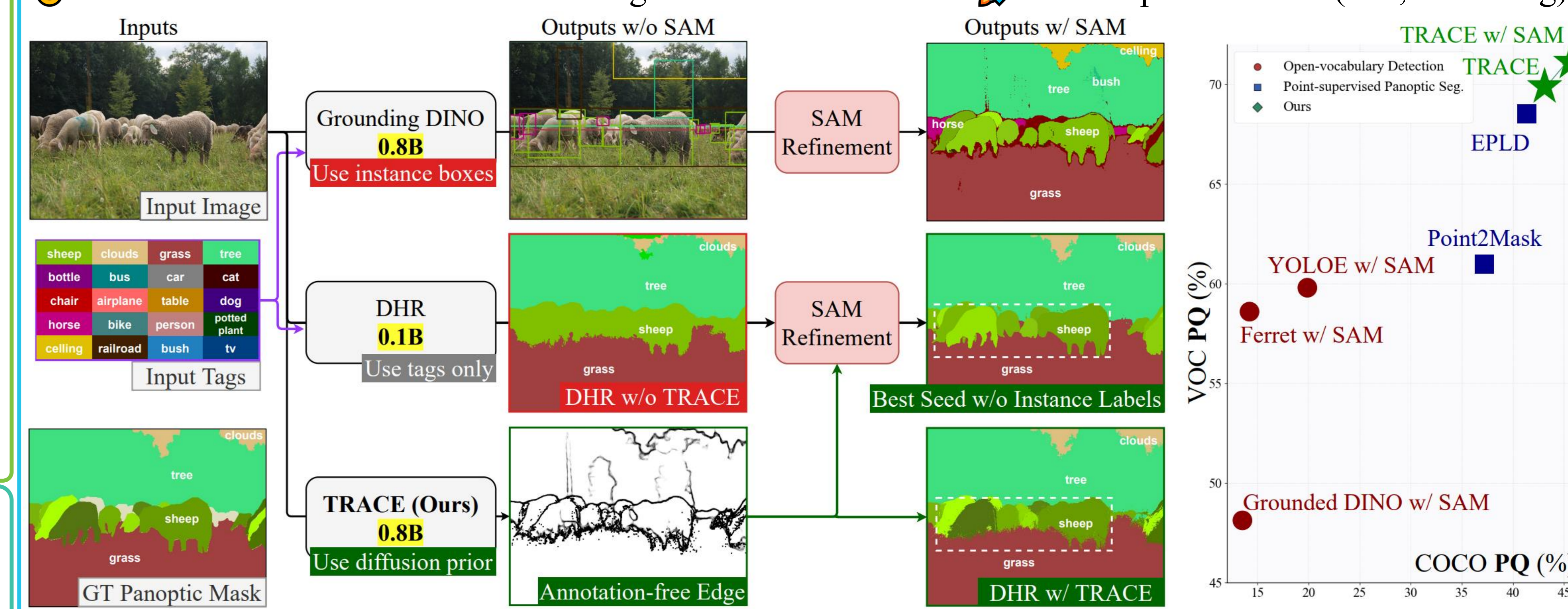
✗ **Cross-Attention:** Remains semantic and merges adjacent instances.
 ✗ **Self-Attention:** Reveals instance-level structures at specific timesteps.



(a) Observation: Instance Cues in Diffusion Models
 (b) TRACE: Extracting Instance Boundaries

Proving the Secret: State-of-the-Art Segmentation Results

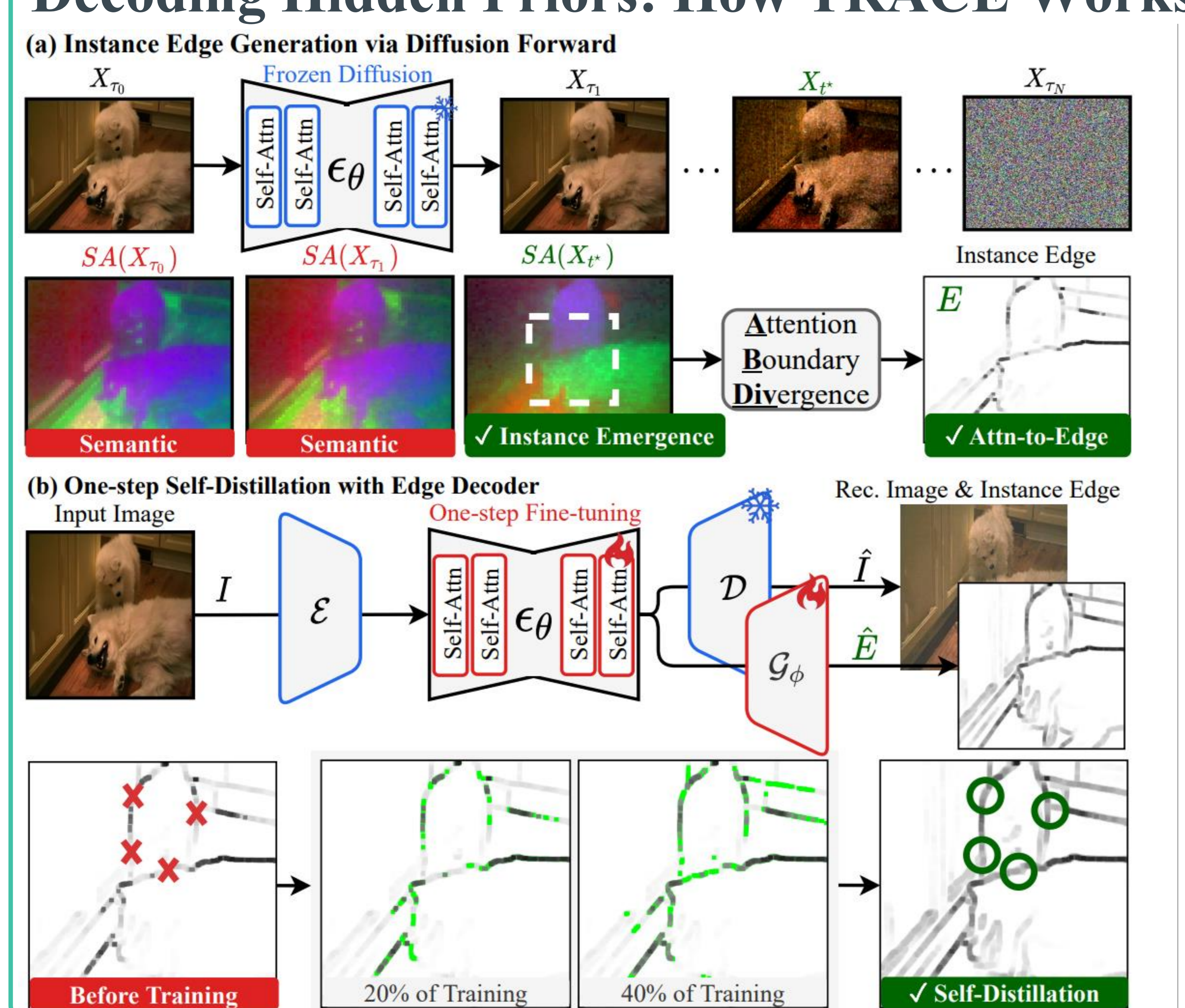
🏆 State-of-the-Art in both UIS/WPS using zero instance labels. 🚀 One-step Distillation (81x, 45ms/img)



(a) Qualitative Comparison of Weakly-supervised Panoptic Segmentation with TRACE
 (b) Quantitative Comparison of PQ

Decoding Hidden Priors: How TRACE Works

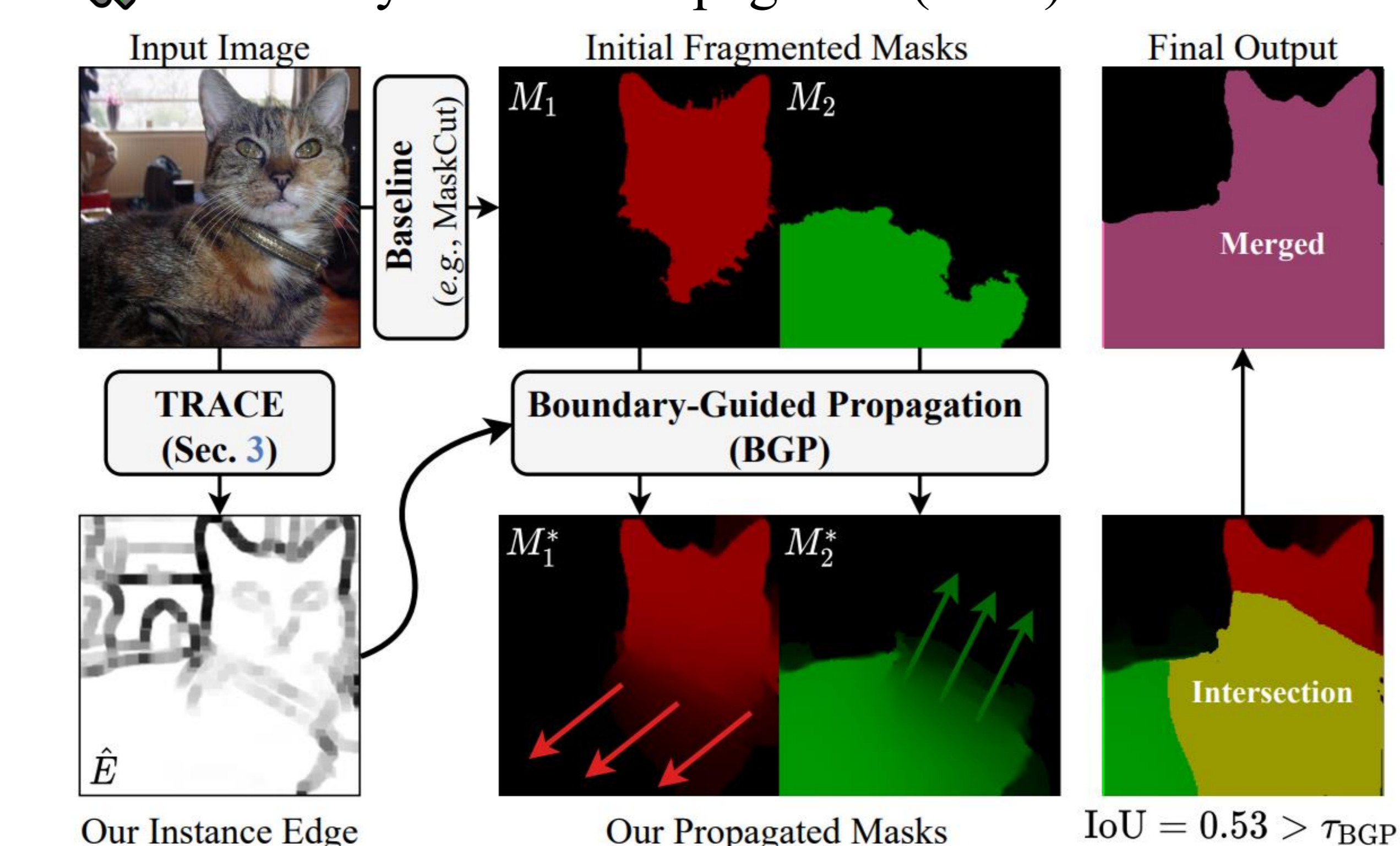
(a) Instance Edge Generation via Diffusion Forward
 (b) One-step Self-Distillation with Edge Decoder



ABDiv (Attention Boundary Divergence)

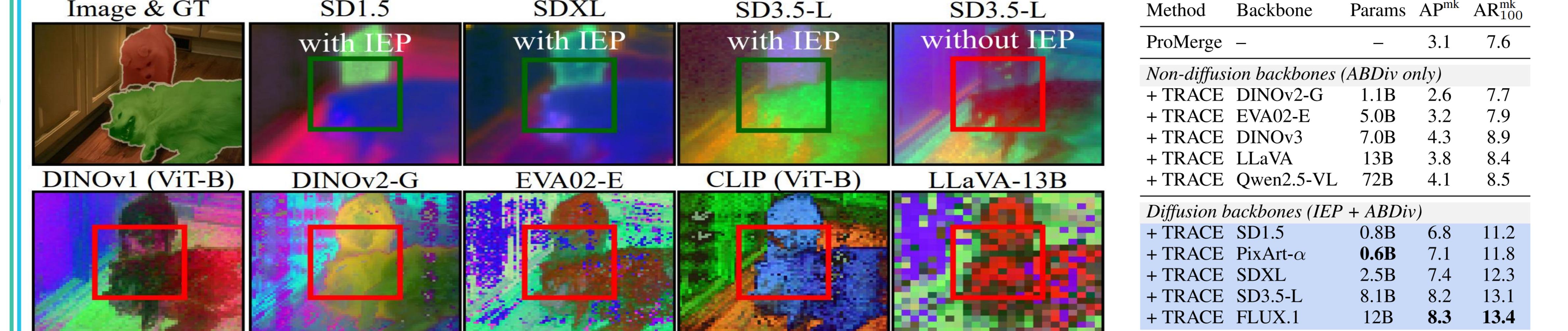
(a) Instance Boundary (ABDiv: 0.114)
 (b) Instance Interior (ABDiv: 0.027)

Boundary-Guided Propagation (BGP)



Our Instance Edge
 Our Propagated Masks
 IoU = 0.53 > τ_{BGP}

🧠 Only Diffusion Priors decouple instances; Non-diffusion models collapse into semantic blobs.



Method	Backbone	Params	AP ^{mk}	AR ^{mk} ₁₀₀
ProMerge	-	-	3.1	7.6
<i>Non-diffusion backbones (ABDiv only)</i>				
+ TRACE	DINOv2-G	1.1B	2.6	7.7
+ TRACE	EVA02-E	5.0B	3.2	7.9
+ TRACE	DINOv3	7.0B	4.3	8.9
+ TRACE	LLaVA	13B	3.8	8.4
+ TRACE	Qwen2.5-VL	72B	4.1	8.5
<i>Diffusion backbones (IEP + ABDiv)</i>				
+ TRACE	SD1.5	0.8B	6.8	11.2
+ TRACE	PixArt- α	0.6B	7.1	11.8
+ TRACE	SDXL	2.5B	7.4	12.3
+ TRACE	SD3.5-L	8.1B	8.2	13.1
+ TRACE	FLUX.1	12B	8.3	13.4

🔍 Traditional detectors fail with textures, but TRACE extracts pure instance-aware boundaries.

