CVPR2026 (Poster) | Jun 7, 2026, 2:45 AM – 4:45 AM | ExtHall F

InvAD: Inversion-based Reconstruction-Free Anomaly Detection with Diffusion Models

Homepage: https://invad-project.com

Shunsuke Sakai[1] Xiangteng He[2,3] Chunzhi Gu[1]

Leonid Sigal[2,3] Tatsuhito Hasegawa[1]

[1]University of Fukui [2]University of British Columbia [3]Vector Institute for AI

Limitations

Visual Anomaly Detection with Diffusion Models

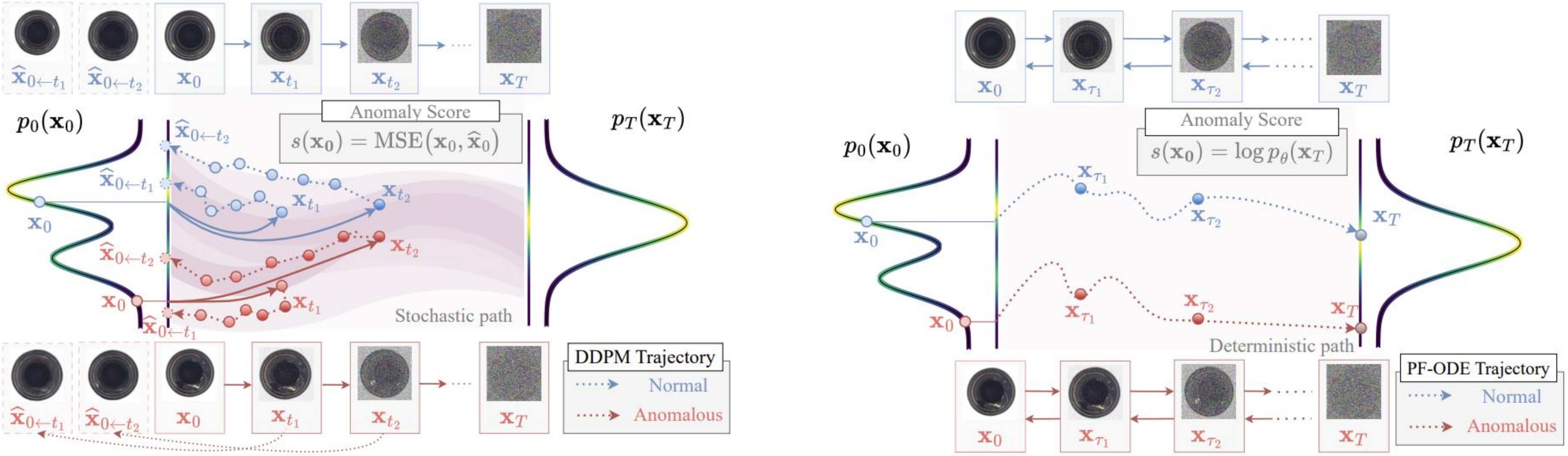
The **low inference speed for diffusion-based reconstruction** has **limited their applicability in real-time scenarios**.

Table 1. **Comparison of the properties** of diffusion-based AD methods. Normal-only means whether the method involves pseudo-anomalies in training, NFE stands for the number of function evaluations, and TS refers to the timestep.

Method	Normal-only	NFE	FPS	TS Tuning-free	Multi-class	Scoring scheme
DiffAD *[ICCV'23]*	✓	N/A	N/A	✗	✗	Mask prediction
DiAD *[AAAI'24]*	✗	10	1.5	✓	✓	$\text{MSE}(g_\phi(\mathbf{x}_0), g_\phi(\hat{\mathbf{x}}_0))$
GLAD *[ECCV'24]*	✓	750	0.2	✓	✗	$\text{MSE}(g_\phi(\mathbf{x}_0), g_\phi(\hat{\mathbf{x}}_0))$
TransFusion *[ECCV'24]*	✓	20	1.6	✗	✗	Mask prediction
MDM *[ICML'25]*	✓	40	1.9	✗	✗	$\text{MSE}(g_\phi(\mathbf{x}_0), g_\phi(\hat{\mathbf{x}}_0))$
OmiAD *[ICML'25]*	✓	1	39.4	✗	✓	$\text{MSE}(g_\phi(\mathbf{x}_0), g_\phi(\hat{\mathbf{x}}_0))$
DeCo-Diff *[CVPR'25]*	✓	10	17.0	✗	✓	$\text{MSE}(g_\phi(\mathbf{x}_0), g_\phi(\hat{\mathbf{x}}_0))$
InvAD (Ours)	✓	**3**	**88.1**	✓	✓	$\log p_\theta(\mathbf{x}_T)$

Motivation

Instead of reconstruction ($x_0 \rightarrow x_t \rightarrow \hat{x}_0$),
we directly **infer the latent state at the final step (i.e., noise),**
by **tracing the learned PF-ODE trajectories** ($x_0 \rightarrow x_T$).



(a) Conventional reconstruction-based AD paradigm (b) Our inversion-based AD paradigm

Motivation

- This removes an **unstable perturbation at the inference stage.**
- **Surprisingly,** we found that it drastically improves inference speed.

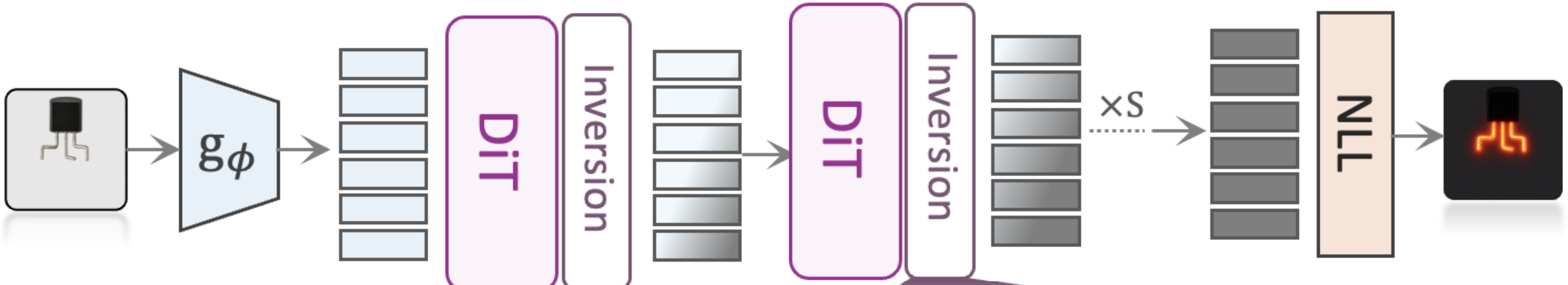
Why this works.

- Learned PF-ODEs are **NOT smooth for abnormal samples.**
- For abnormal samples, **PF-ODE drives the original image into a low-density region of a Gaussian distribution.**

Methodology - *InvAD*

Goal: effectively learn the complex normal distribution.

We simply train **DiffusionTransformer** in the *feature space*.
Then, iteratively "*noising*" with **DDIM Inversion**.



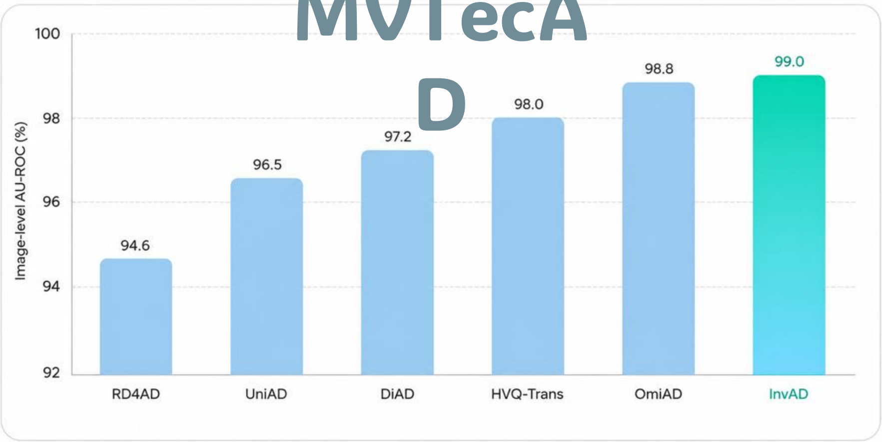
$$\mathbf{x}_{\tau_{i+1}} = \sqrt{\alpha_{\tau_{i+1}}} f_\theta\left(\mathbf{x}_{\tau_i}\right) + \sqrt{1-\alpha_{\tau_{i+1}}}\,\epsilon_\theta^{(\tau_i)}\left(\mathbf{x}_{\tau_i}\right)$$

DDIM Inversion

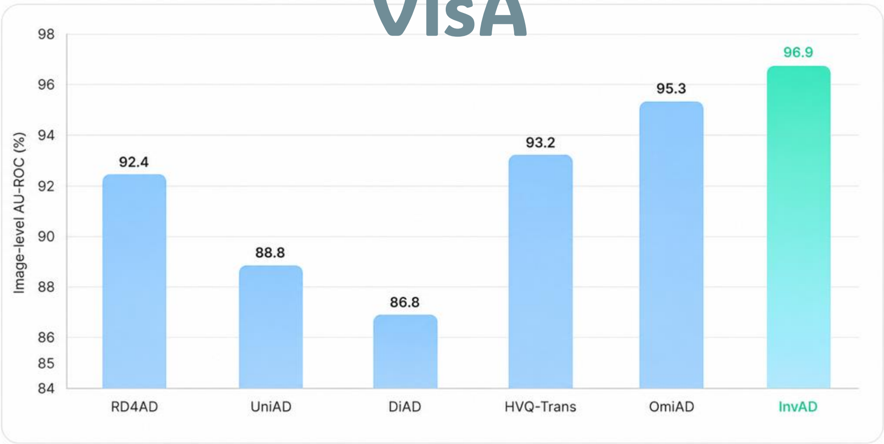
Main Results

State-of-the-art Performance Across Industrial Benchmarks

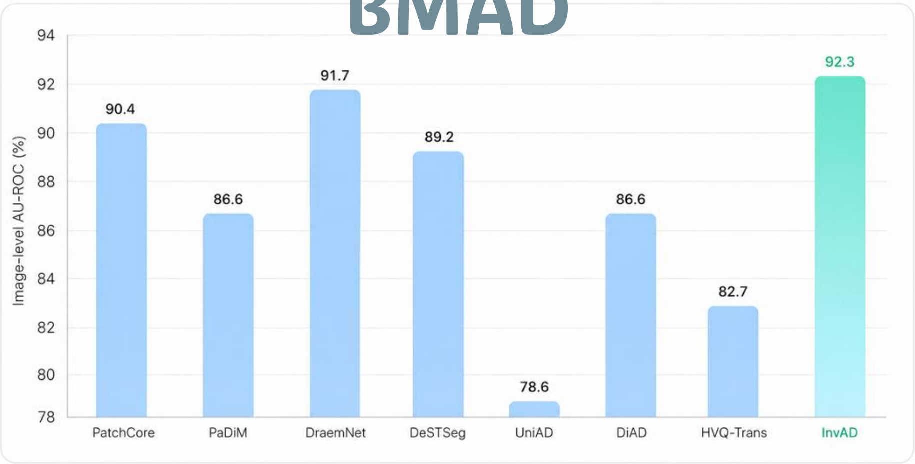
MVTecAD



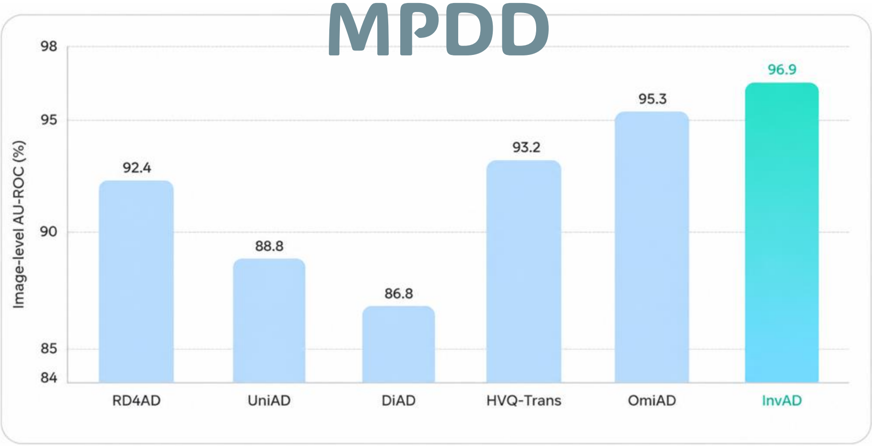
VisA



BMAD

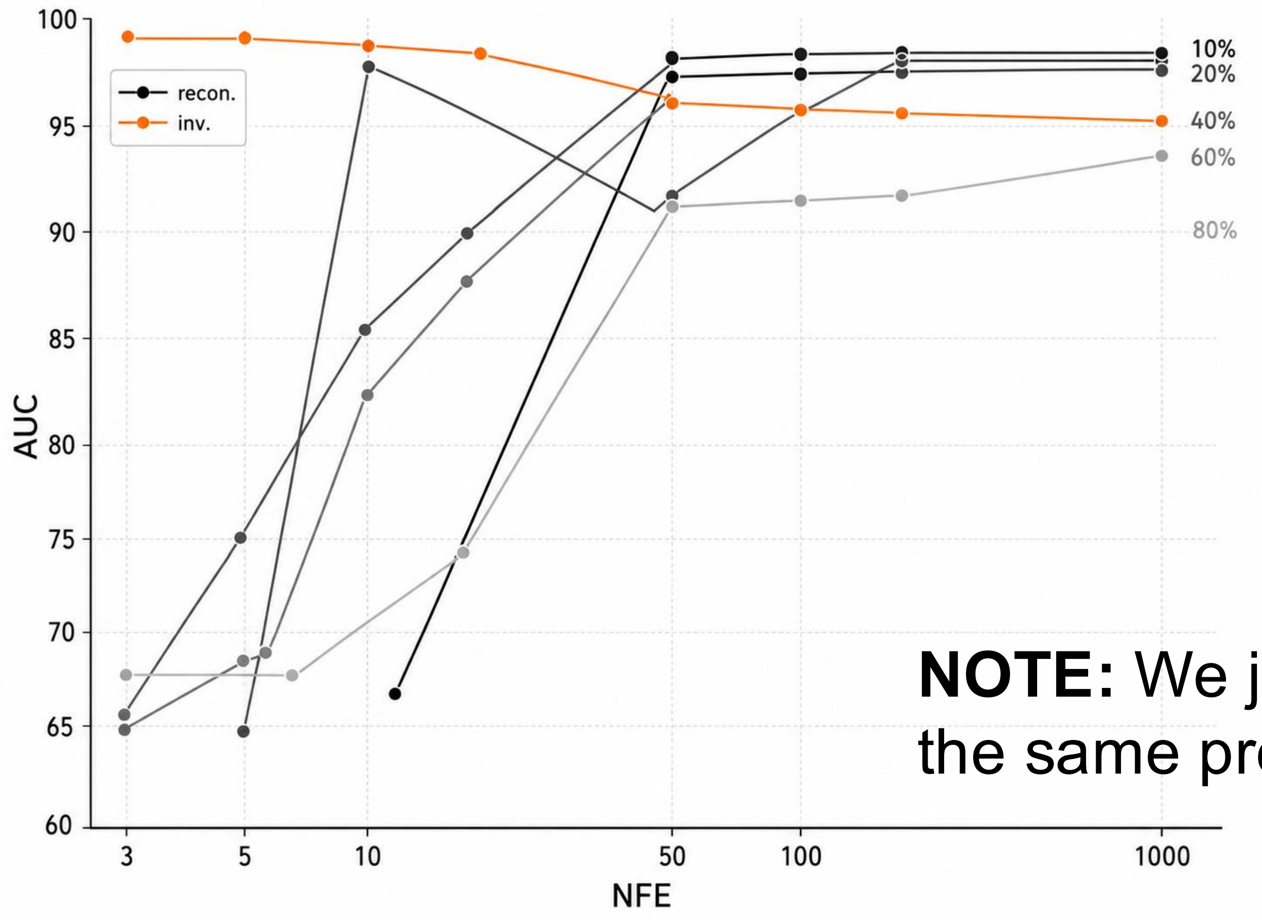


MPDD



Key Findings -Why Inversion?-

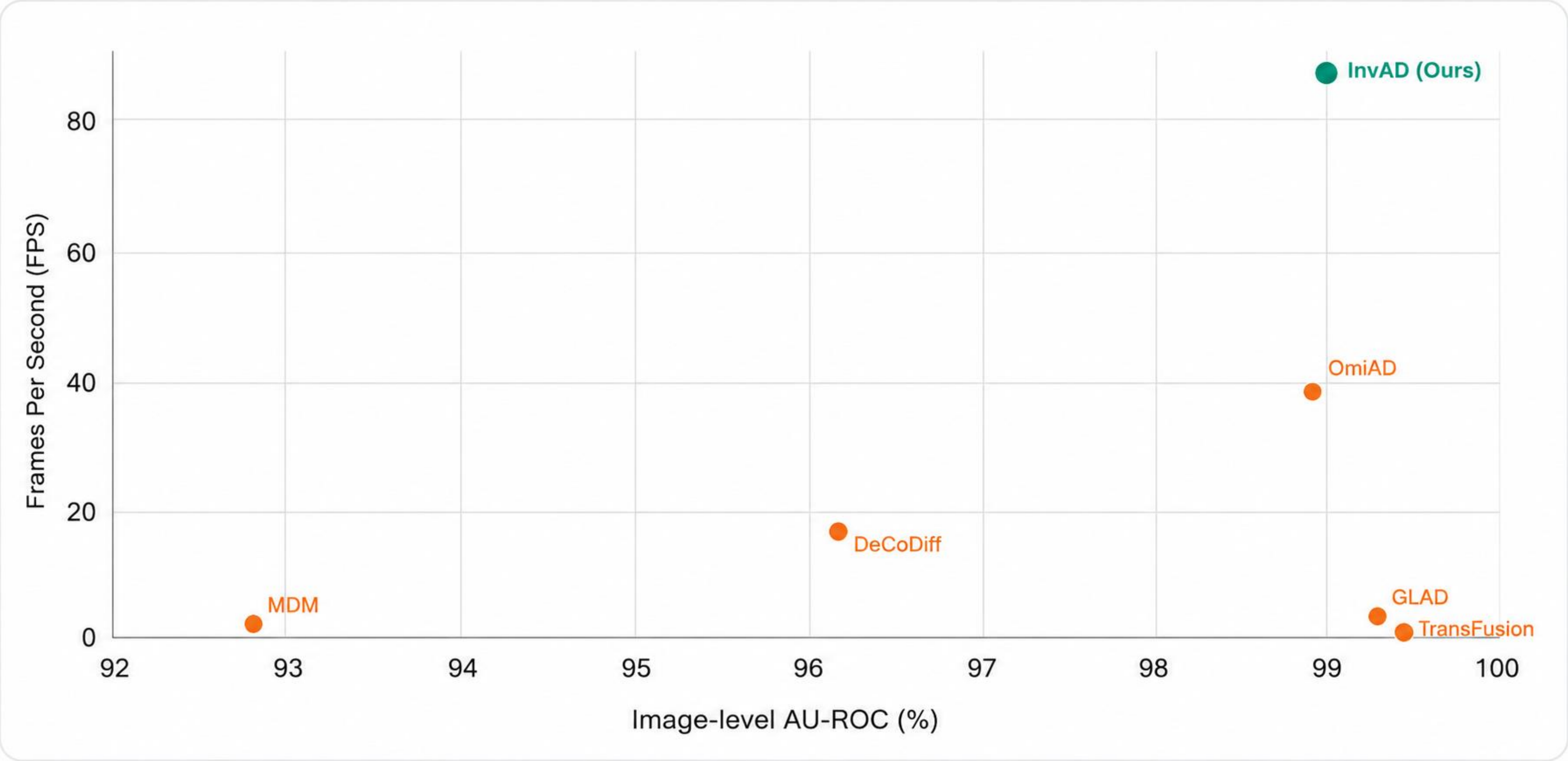
Inversion surprisingly works well in few-step scheme.



NOTE: We just <u>ablate inference methodology</u> with the same pre-trained diffusion model.

Key Findings – Efficiency-

Our InvAD Archives Best Speed-Accuracy Trade-off



Key Findings –Plug and Play-

Any Pre-trained Diffusion Can Be Combined w/ Inversion