JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

Poster: WED-PM-067

Learning 3D-aware Image Synthesis with Unknown Pose Distribution

Zifan Shi[1*], Yujun Shen[2*], Yinghao Xu[3], Sida Peng[4], Yiyi Liao[4], Sheng Guo[2], Qifeng Chen[1], Dit-Yan Yeung[1]

[1]HKUST [2]Ant Group [3]CUHK [4]ZJU

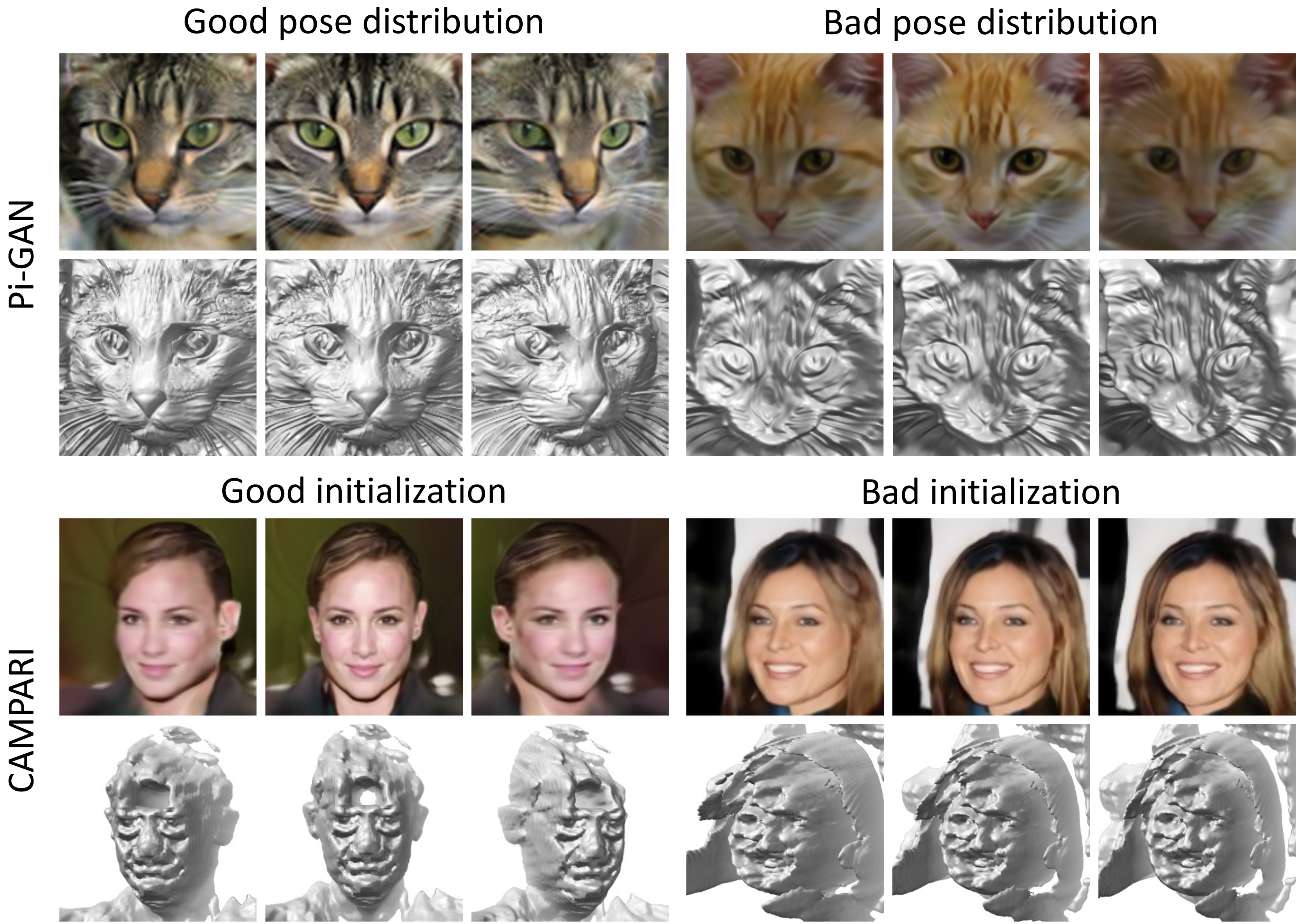
香港科技大學 THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

ANT RESEARCH

香港中文大學 The Chinese University of Hong Kong

浙江大学 1897 ZHEJIANG UNIVERSITY

Motivation: Quality is affected by pose distribution



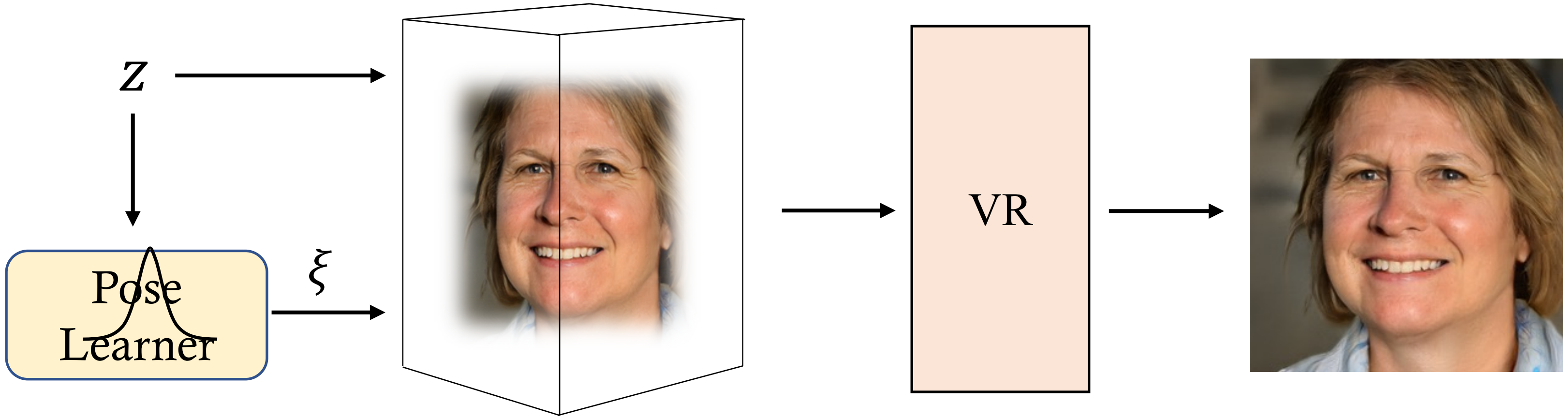
1) Inaccurate pose priors lead to faulty shapes.
2) Hard to get accurate annotations for all types of data.

Free the model from the requirements of pose priors!
PoF3D

Pose-free Generator

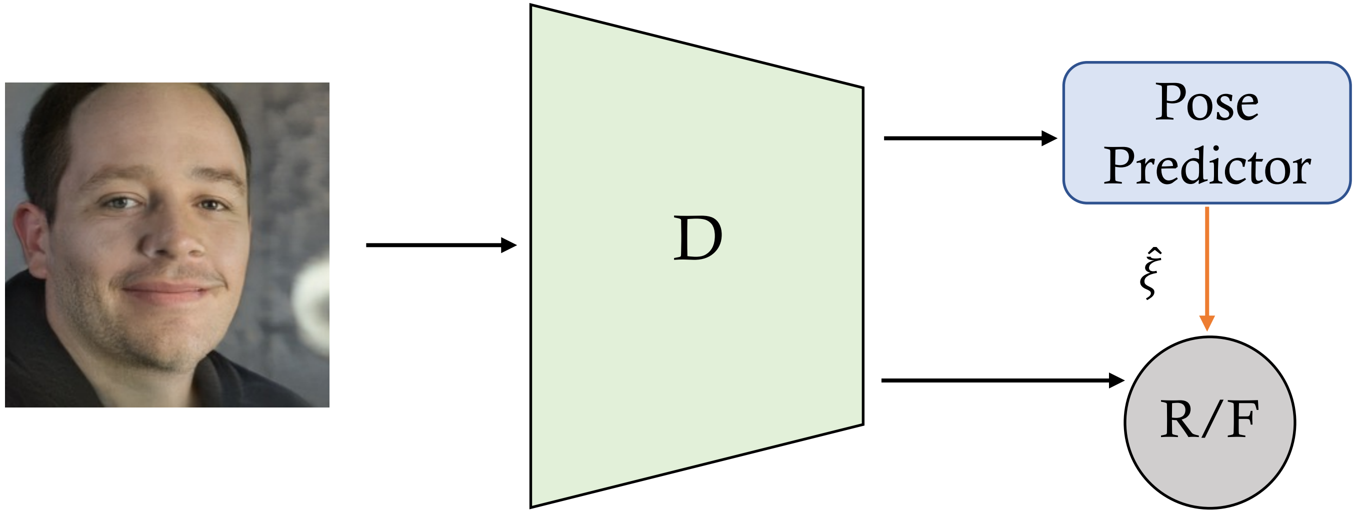
- Learn poses from latent space

$$G(z, \xi) = I \sim p_\theta(I|z, \xi) \longrightarrow G(z, \Psi(z)) = I \sim p_\theta(I|z, \Psi(z))$$



Pose-aware Discriminator

- To facilitate generator synthesizing correct geometry
  - *Learn pose prediction from the generator*
  - *The inferred pose is treated as a pseudo label for conditional real/fake classification*



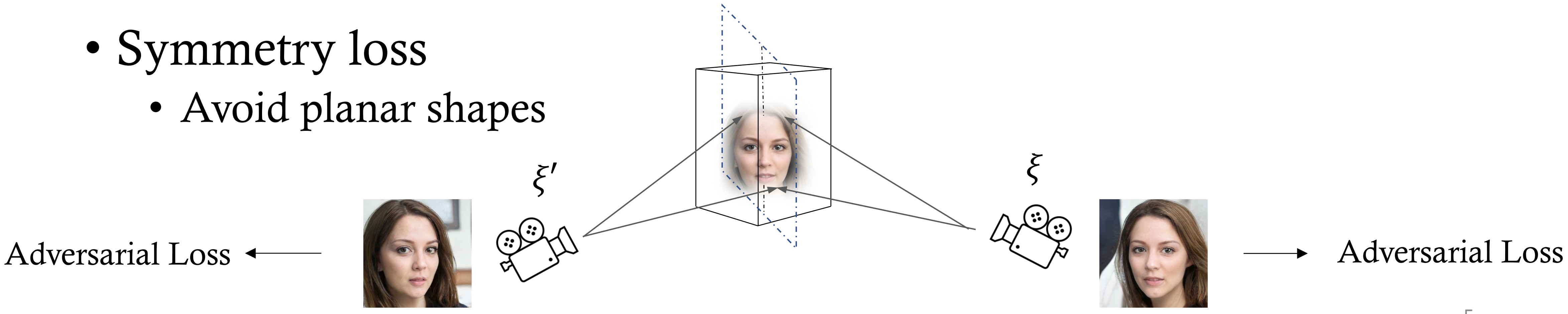
PoF3D

- Overview

Pose-free Generator

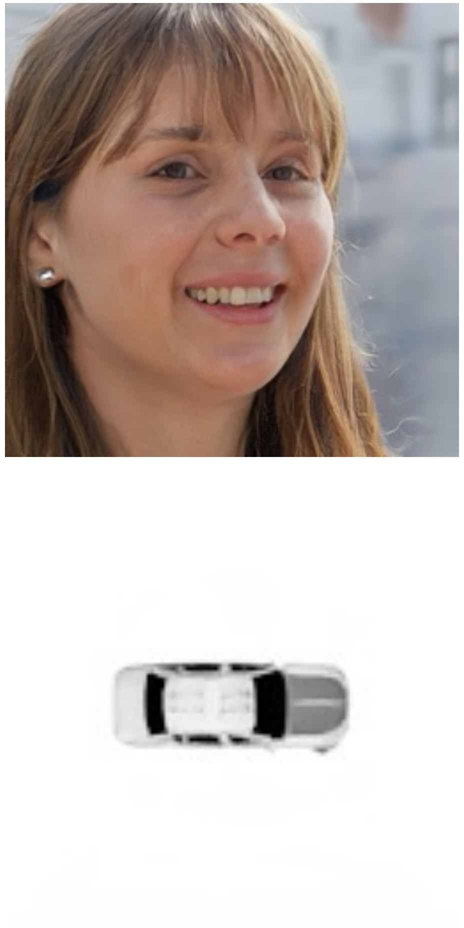
Pose-aware Discriminator

- Symmetry loss
  - Avoid planar shapes



Qualitative Comparison

CAMPARI

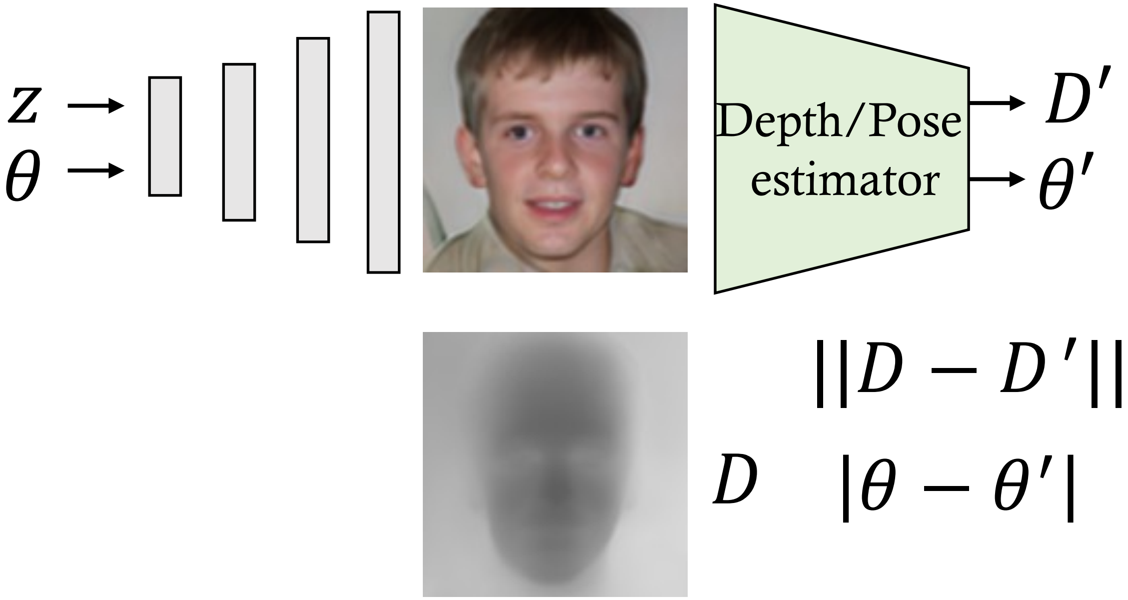


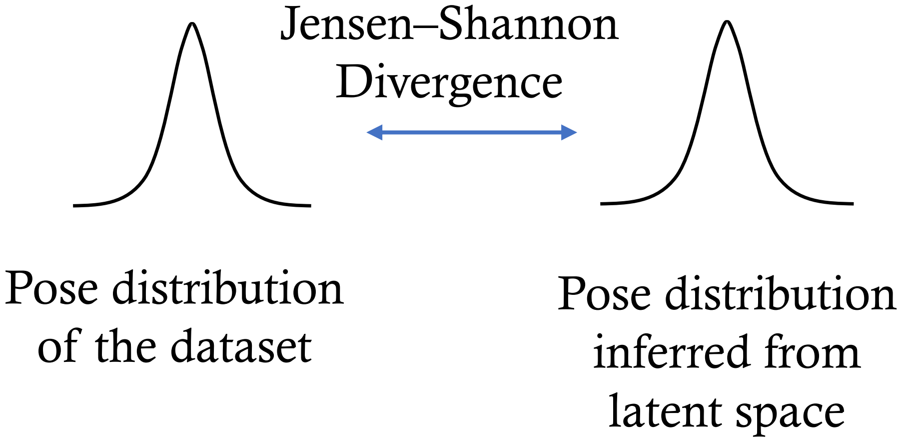
EG3D
(with ground-truth poses)

Ours
(without pose priors)

Evaluation Metrics

- Depth error
- Pose error
- Jensen–Shannon Divergence (JS)





Quantitative Results

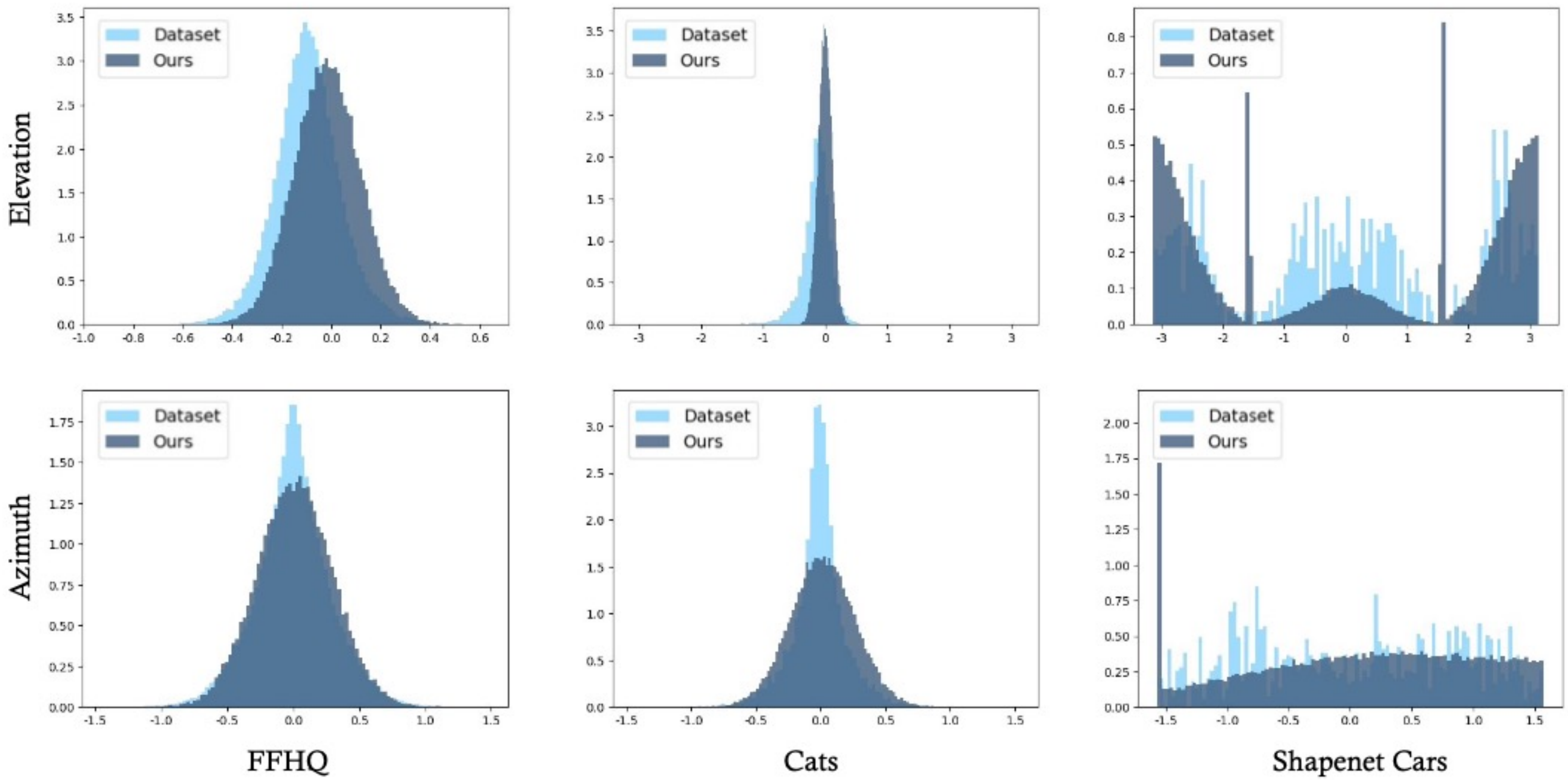
	FFHQ				ShapeNet Cars	
	FID	Depth	Pose	JS	FID	JS
CAMPARI	58.59	1.78	0.15	0.61	68.91	0.72
CAMPARI+EG3D	**3.25**	1.13	0.20	0.73	4.66	0.83
EG3D	4.80	**0.29**	**0.08**	-	9.68	-
Ours	4.99	**0.29**	0.11	**0.20**	**3.78**	**0.51**

Ablation Study

	FFHQ				
	FID	Depth	Pose	RE	JS
w/o symmetry loss	4.50	0.41	0.12	**0.030**	0.20
w/o pose-aware D	**3.43**	1.30	0.20	0.046	0.21
lr = 2.5e-4	122.07	0.82	0.74	0.051	0.56
lr = 2.5e-6	10.20	0.70	0.16	0.173	0.38
Ours	4.99	**0.29**	**0.11**	0.037	**0.20**

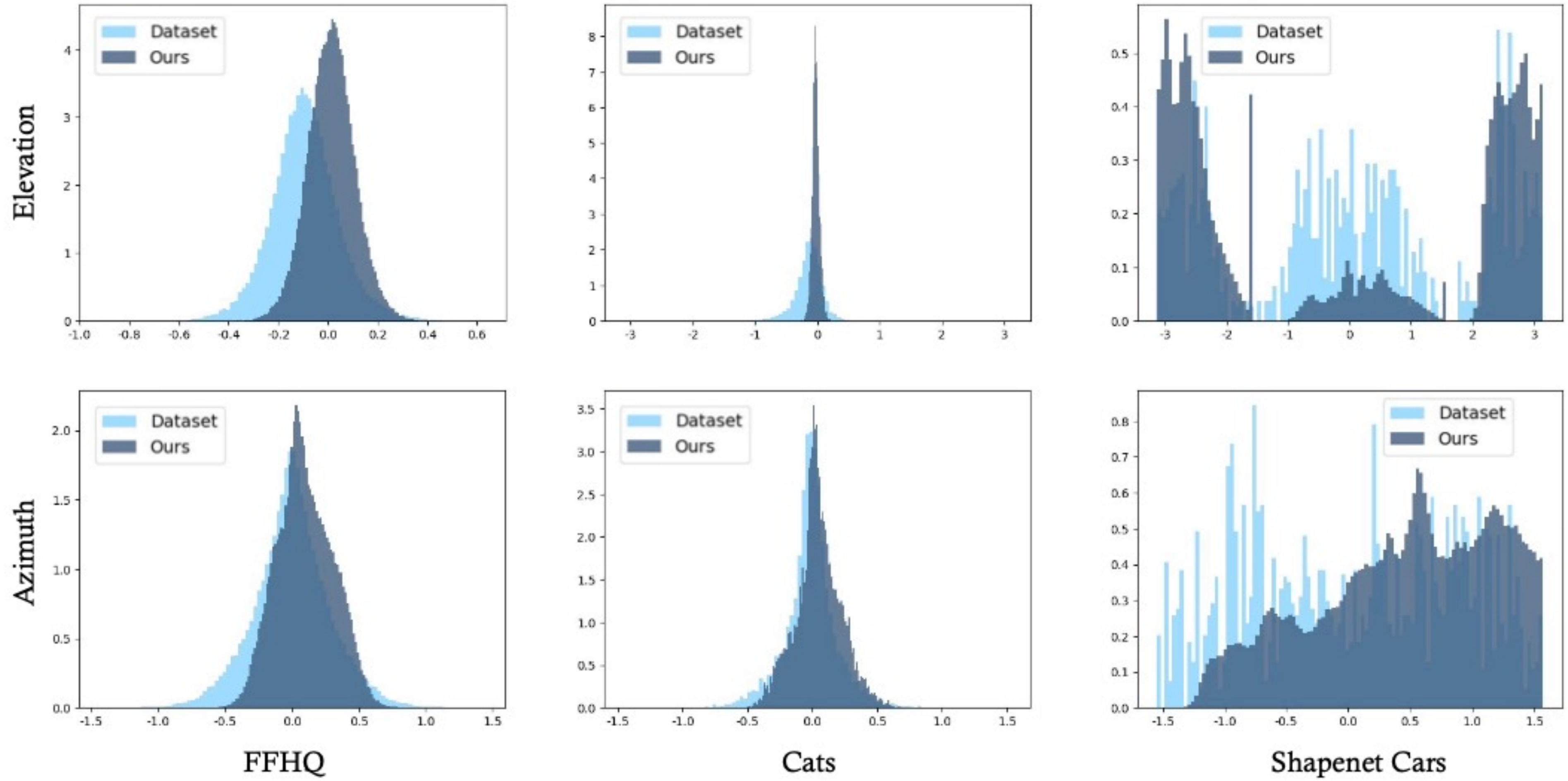
Pose Distribution

- Generator



Pose Distribution

- Discriminator



Take-home Message

- Pose is crucial for 3D-aware image synthesis. Bad pose priors can hurt the synthesis quality.

- Derving poses from the latent space and making the discriminator pose-aware can enable 3D-aware image synthesis without pose priors.

Paper

Project Page

Demo

Github