Flow-based GAN for 3D Point Cloud Generation from a Single Image

Yao Wei[1], George Vosselman[1], Michael Ying Yang[1]
[1]University of Twente, The Netherlands

Motivation

(x_i, y_i, z_i)

$N \times 3$ matrix

Near-surface points have high probability

Off-surface points have low probability

distribution of 3D points
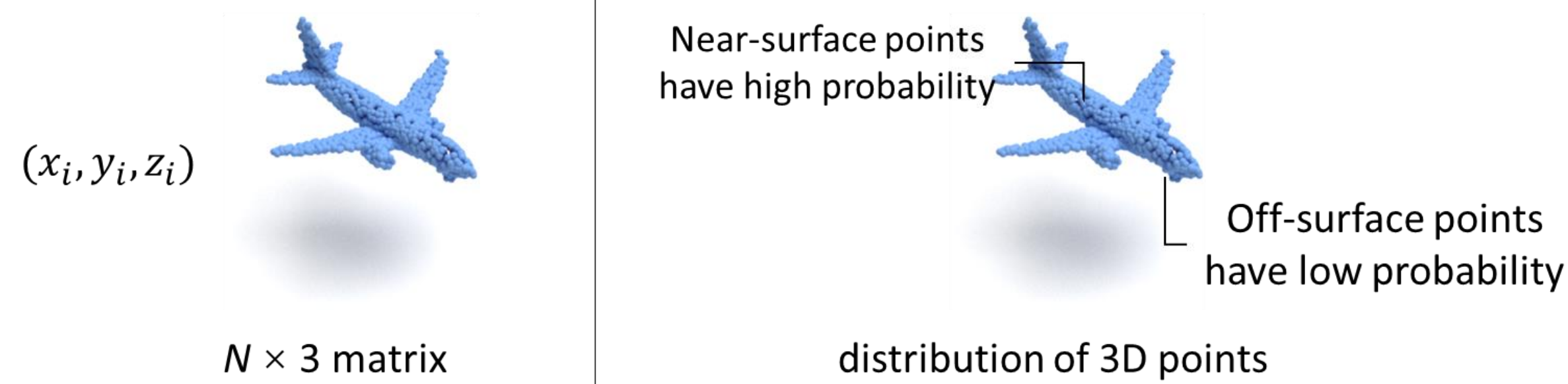
Previous works which only consider either explicit or implicit generative modeling of point clouds, suffer from fixed resolution and/or low quality issues.

Contribution

- A hybrid explicit-implicit generative modeling scheme for 3D point cloud generation, which inherits the flow-based generative methods for generating an arbitrary number of points, while guiding the flow-based generator to reconstruct high-quality point clouds by leveraging an adversarial training strategy.
- State-of-the-art performance for single image to 3D point cloud generation task on the ShapeNet dataset.
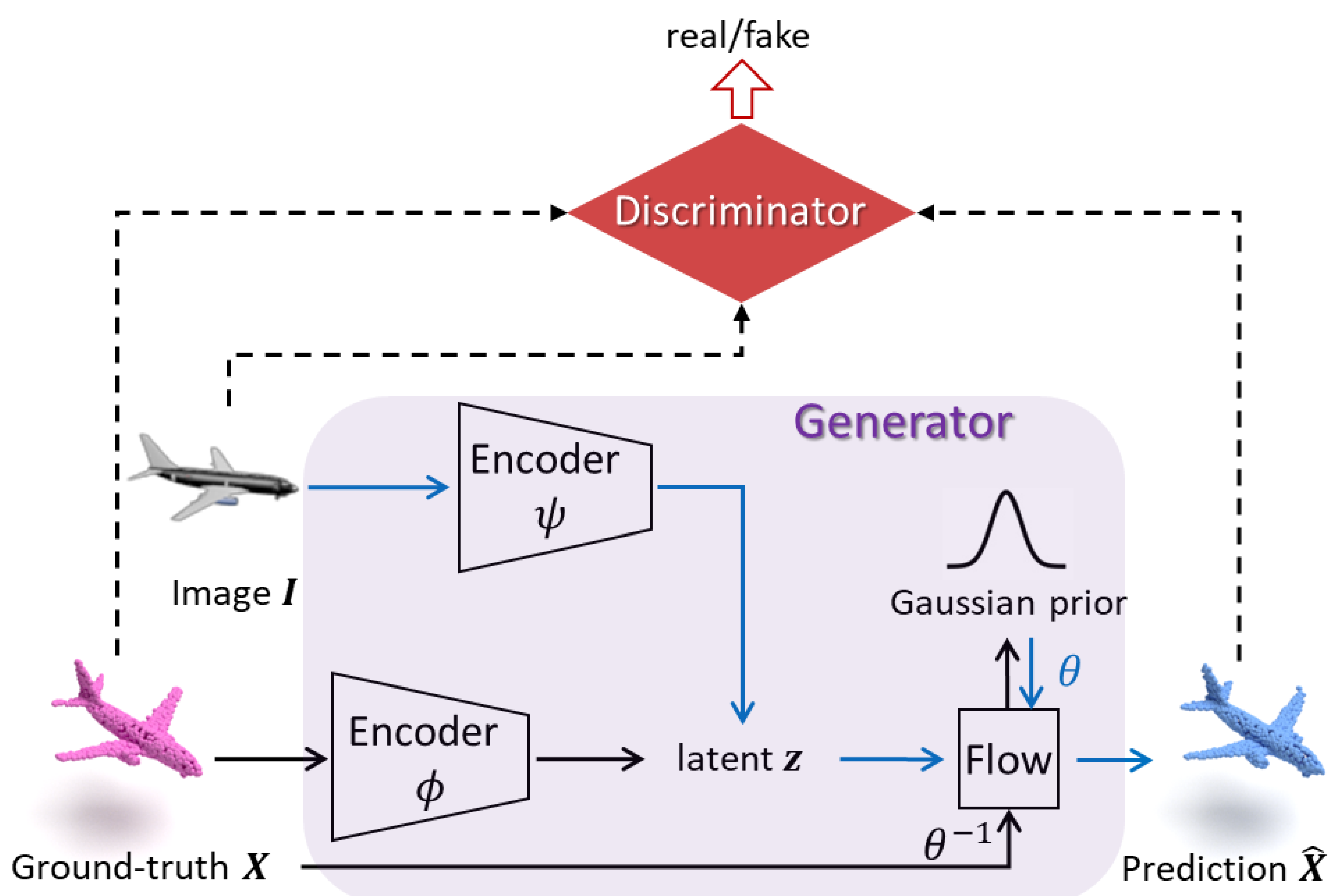
Method

real/fake

Discriminator

Generator

Encoder ψ

Image I

Encoder ϕ

latent z

Gaussian prior

Flow

θ

θ^{-1}

Ground-truth X

Prediction $\hat{X}$

Fig. 1. Overall framework.

A generative adversarial framework performs the hybrid explicit-implicit generative modeling of point clouds.

- **Flow-based Generator** allows an arbitrary number of points to be sampled in the inference phase.
- **Cross-modal Discriminator** guides the generator to produce high-quality point clouds.
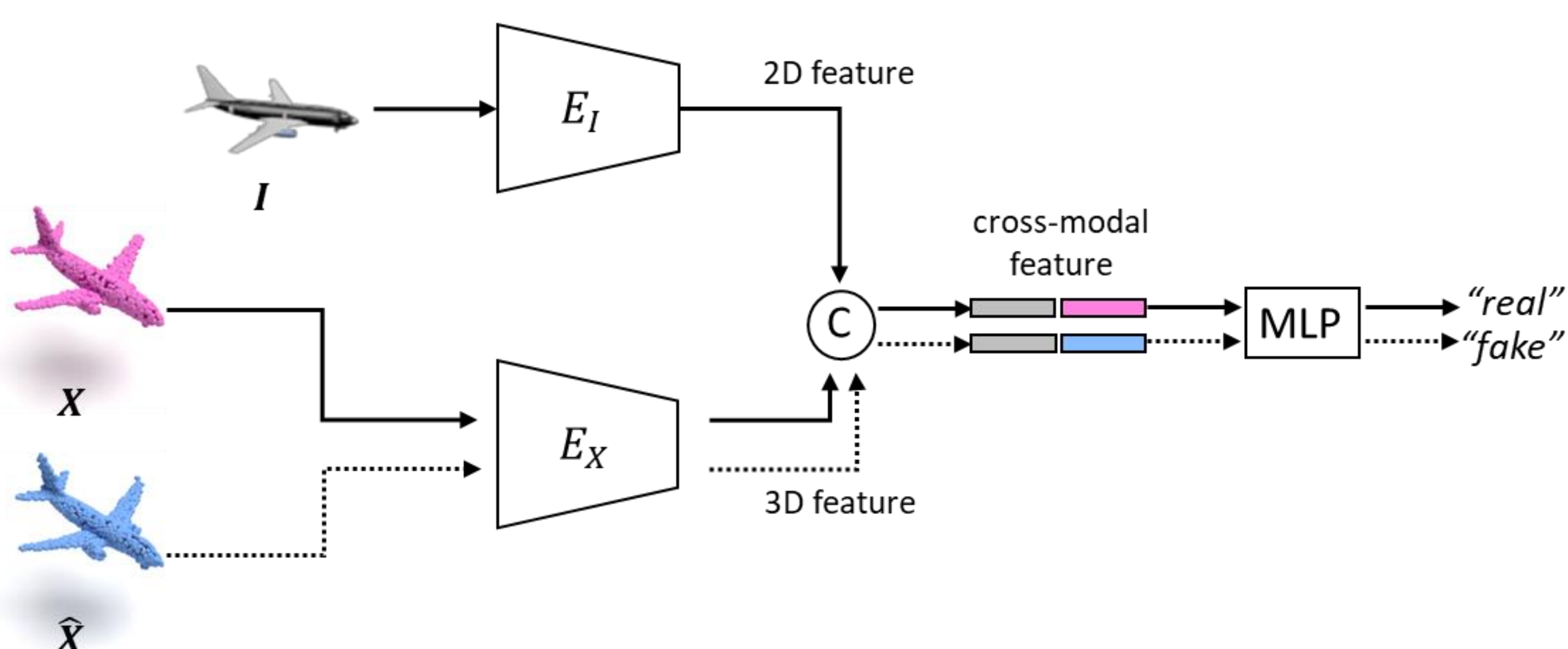

I

E_I

2D feature

X

E_X

3D feature

$\hat{X}$

cross-modal feature

C

MLP

"real" "fake"

Fig. 2. Architecture of the cross-modal discriminator D.

Qualitative Results

Input image
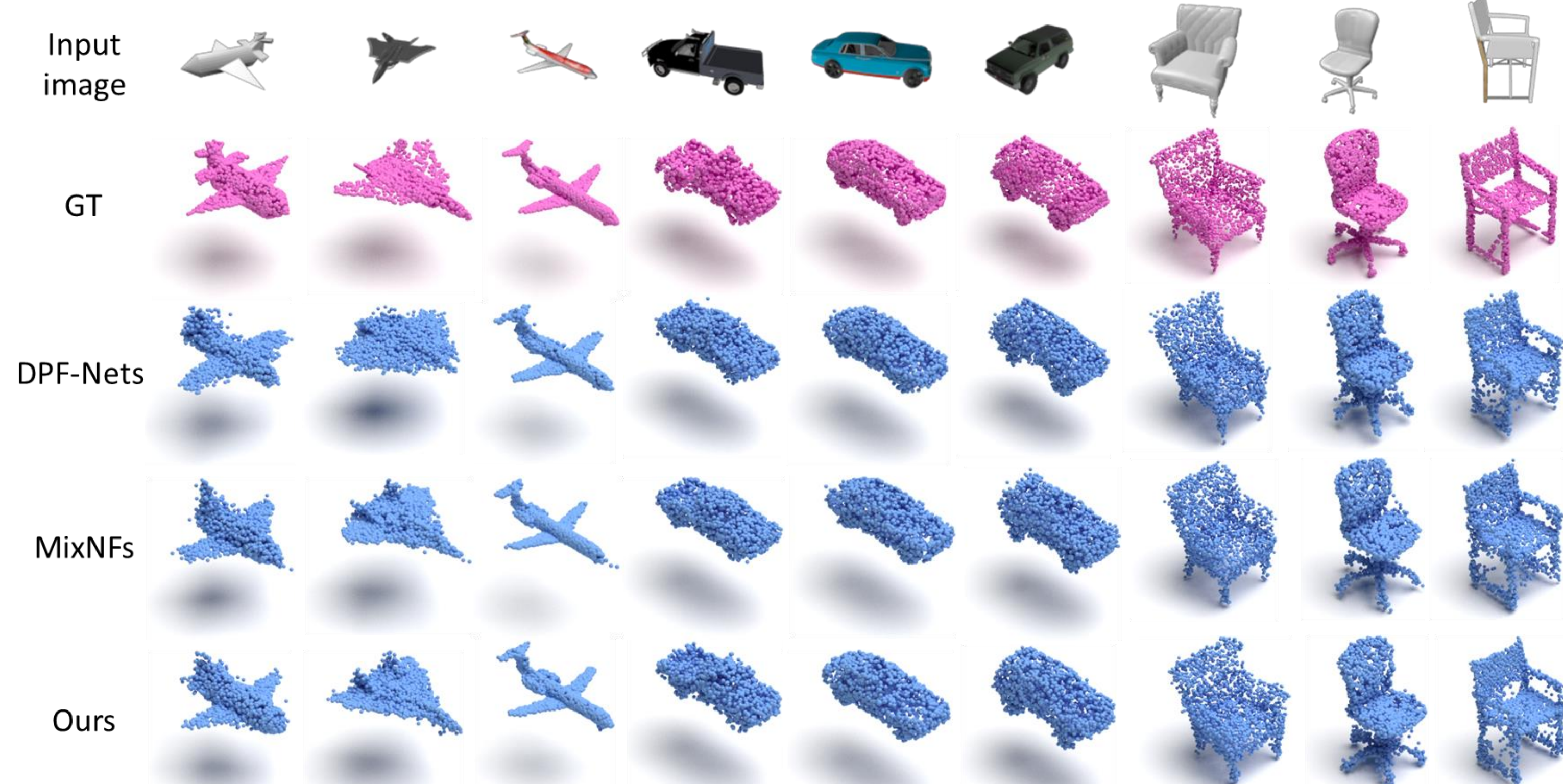
GT

DPF-Nets

MixNFs

Ours

Fig. 3. Results on the synthetic images from the ShapeNet dataset.

Fig. 4. Results on the real images from the PASCAL3D+ dataset.

Quantitative Results

Table 1. Results on three categories of the ShapeNet dataset

Methods	Airplane			Car			Chair		
	CD↓	EMD↓	F1↑	CD↓	EMD↓	F1↑	CD↓	EMD↓	F1↑
DPF-Nets [19]	4.11	10.89	72.61	3.79	10.46	46.58	5.42	11.40	46.82
MixNFs [26]	2.82	9.31	77.63	3.73	10.38	47.10	5.41	11.33	46.98
Ours	**2.33**	**8.68**	**79.94**	**3.60**	**10.24**	**47.71**	**5.02**	**10.99**	**49.09**
Oracle	0.50	4.48	97.62	1.55	6.34	73.65	1.11	5.92	84.32

Table 2. Results on all 13 categories of the ShapeNet dataset

Methods	CD↓	EMD↓	F1↑	Speed↑
PRN [18]	7.56	**11.00**	**53.1**	-
AtlasNet [11]	5.34	12.54	52.2	-
DCG [32]	6.35	18.94	45.7	-
Pixel2Mesh [33]	5.91	13.80	-	-
DPF-Nets [19]	5.55	11.11	51.7	259
MixNFs [26]	5.66	11.20	52.3	5
Ours	**5.32**	**11.00**	53.0	**273**
Oracle	1.10	5.70	84.0	-

Table 3. Ablative results on all 13 categories of the ShapeNet dataset w.r.t. different architectures.

Generator	Discriminator	CD↓	EMD↓	F1↑
w Shape_flow	*w/o*	5.55	11.11	51.7
w/o Shape_flow	*w/o*	5.50	11.20	52.3
w/o Shape_flow	*w D*	**5.32**	**11.00**	**53.0**

Key Insight

- Flexibility: arbitrary resolution, NFs.
- Efficiency: single-flow-based, GANs.

Code

Email: yao.wei@utwente.nl