



Reusing Discriminators for Encoding: Towards Unsupervised Image-to-Image Translation

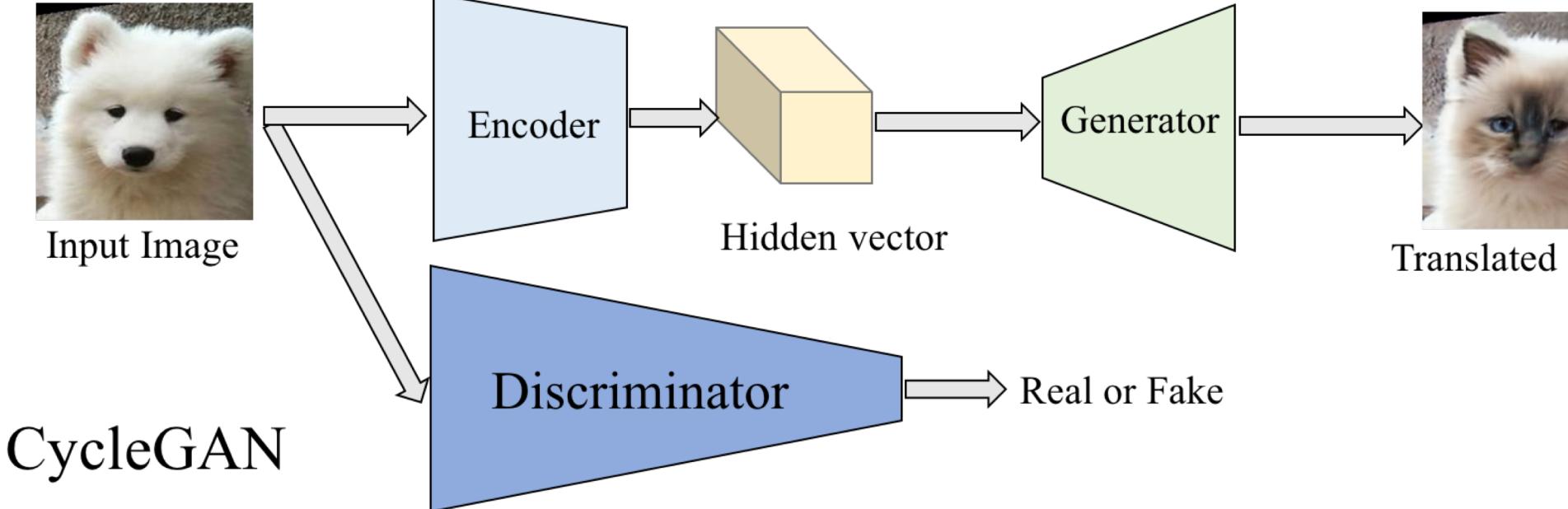
Runfa Chen, Wenbing Huang, Binghui Huang, Fuchun Sun*, Bin Fang



*Corresponding author: Fuchun Sun. THUAI & BNRIst, Tsinghua University

Backgrounds

Current General Unsupervised I2I Translation Frameworks

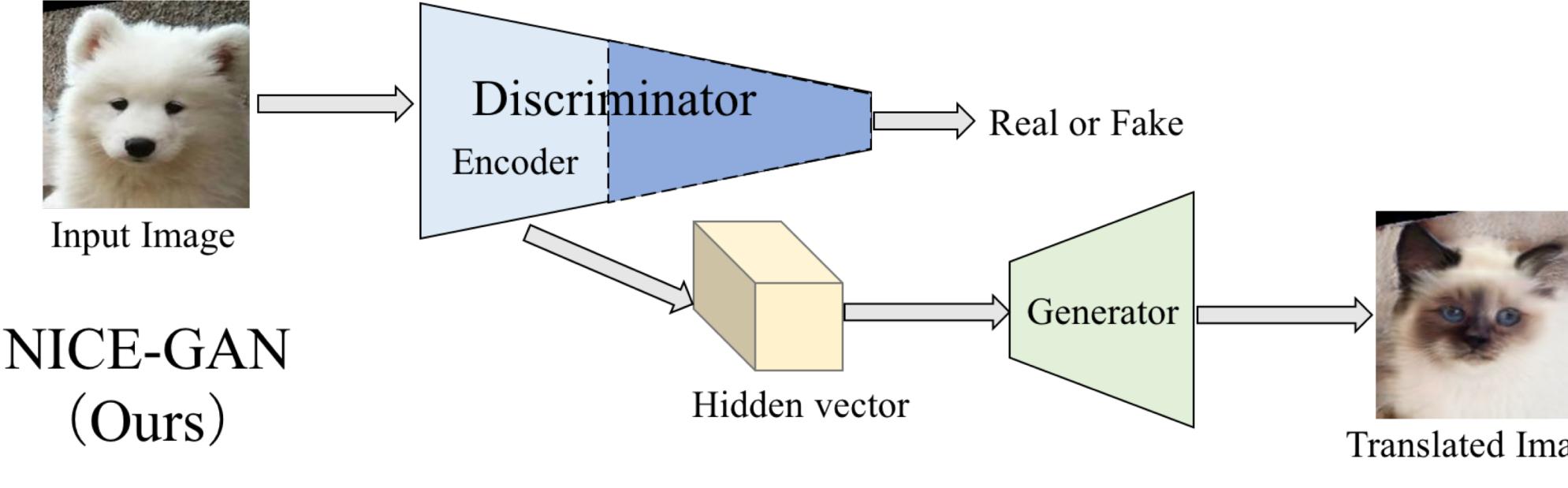


- An Encoder to embed the input image to a low-dimension hidden space.
- A Generator to translate hidden vectors to images of the other domain
- A Discriminator for domain alignment by using GAN training.

Is there any possibility to rethink the role of each component in current translation frameworks? and more importantly, can we change the current formulation (for example, to a more compact architecture) based on our rethinking?

No-Independent-Component-for-Encoding GAN (NICE-GAN)

Our NICE-GAN Unsupervised I2I Translation Frameworks



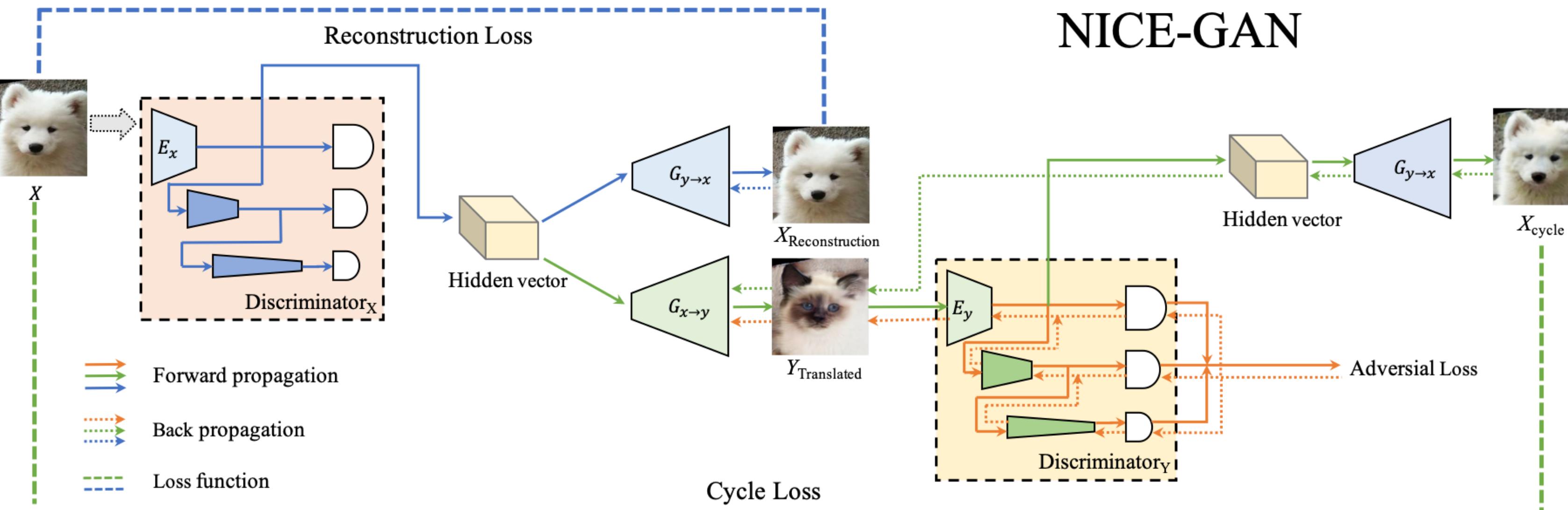
We contend a dual role of the discriminator, encoding and classifying, by reusing early layers of certain number for encoding the images of target domain.

Our Contributions

- We are the first to reuse discriminators for encoding specifically for unsupervised I2I translation. By such a reusing, a more compact and more effective architecture is derived,
- Given that the reusing of discriminator will incur instability in terms of typical training procedure, this paper develops a decoupled training paradigm, which is simple yet efficient.

The Decoupled Training Paradigm

Illustration of the Flowchart of NICE-GAN. Here we only display one translation stream.



- In our framework, the encoder as part of translation is trained for minimizing, and at the same time it belongs to the discriminator and is also trained for maximizing.
- To eliminate the inconsistency, we develop a decoupled training paradigm. Specifically, the encoder E_y is fixed when minimizing the adversarial loss, the reconstruction loss and the cycle loss, and it is trained when maximizing the adversarial loss.

Experiment

Comparisons with State of the Arts

Method	Dataset dog → cat		winter → summer		photo → vangogh		zebra → horse	
	FID	KID × 100	FID	KID × 100	FID	KID × 100	FID	KID × 100
NICE-GAN	48.79	1.58	76.44	1.22	122.27	3.71	149.48	4.29
NICE-GAN*	51.98	1.50	79.02	1.35	122.59	3.53	150.57	4.43
U-GAT-IT-light	80.75	3.22	80.33	1.82	137.70	6.03	145.47	3.39
CycleGAN	119.32	4.93	79.58	1.36	136.97	4.75	156.19	5.54
UNIT	59.56	1.94	95.93	4.63	136.80	5.17	170.76	6.30
MUNIT	53.25	1.26	99.14	4.66	130.55	4.50	193.43	7.25
DRIT	94.50	5.20	78.61	1.69	136.24	5.43	200.41	10.12

Method	Dataset cat → dog		summer → winter		vangogh → photo		horse → zebra	
	FID	KID × 100	FID	KID × 100	FID	KID × 100	FID	KID × 100
NICE-GAN	44.67	1.20	76.03	0.67	112.00	2.79	65.93	2.09
NICE-GAN*	55.72	1.89	77.13	0.73	117.81	3.61	84.89	3.29
U-GAT-IT-light	64.36	2.49	88.41	1.43	123.57	4.91	113.44	5.13
CycleGAN	125.30	6.93	78.76	0.78	135.01	4.71	95.98	3.24
UNIT	63.78	1.94	112.07	5.36	143.96	7.44	131.04	7.19
MUNIT	60.84	2.42	114.08	5.27	138.86	6.19	128.70	6.92
DRIT	79.57	4.57	81.64	1.27	142.69	5.62	111.63	7.40

Module	Total number of params(FLOPs)	
	Generators	Discriminators
U-GAT-IT-light	21.2M(105.0G)	112.8M(15.8G)
NICE-GAN	16.2M(67.6G)	93.7M(12.0G)
NICE-GAN*	11.5M(48.2G)	93.7M(12.0G)

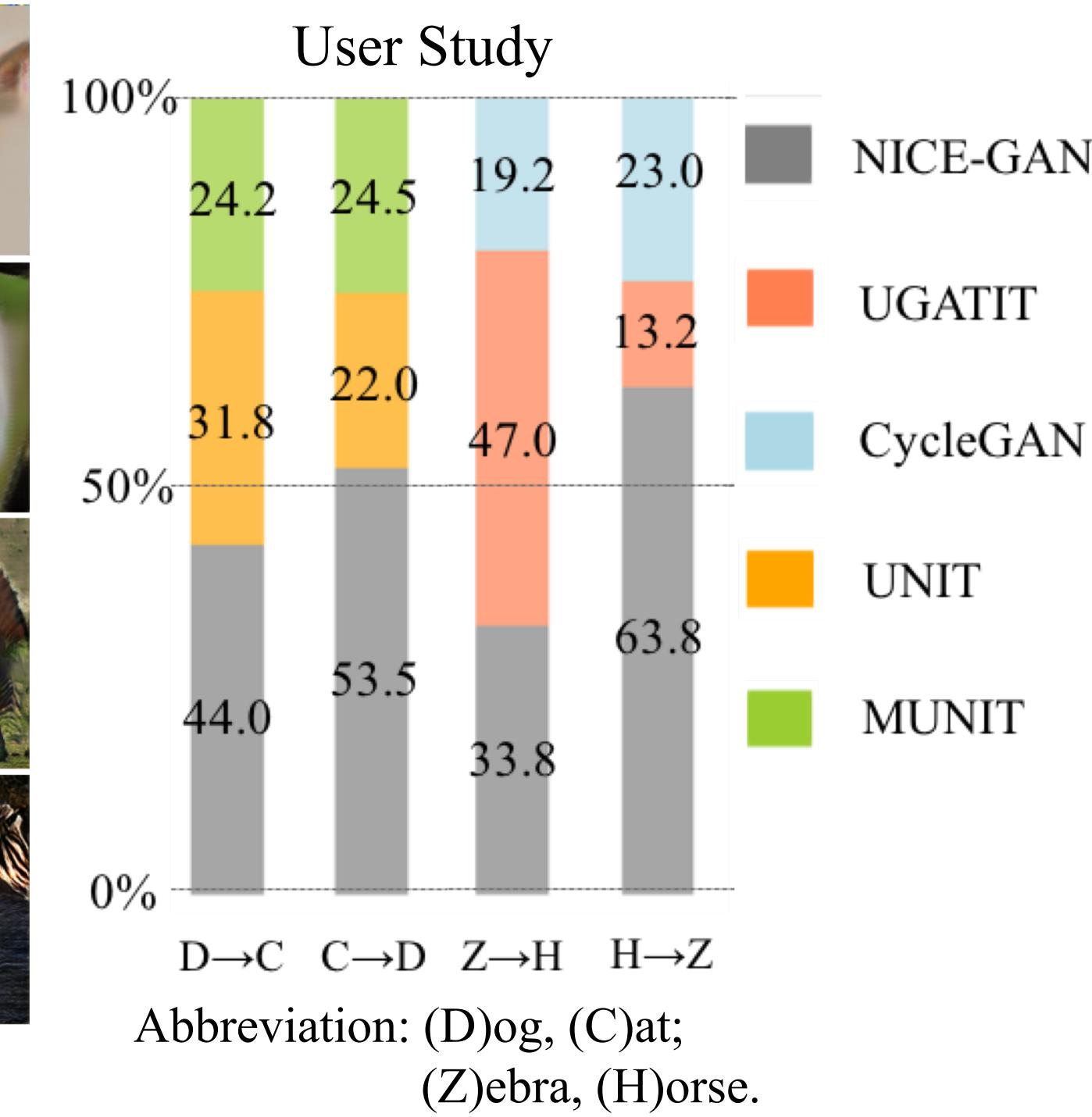
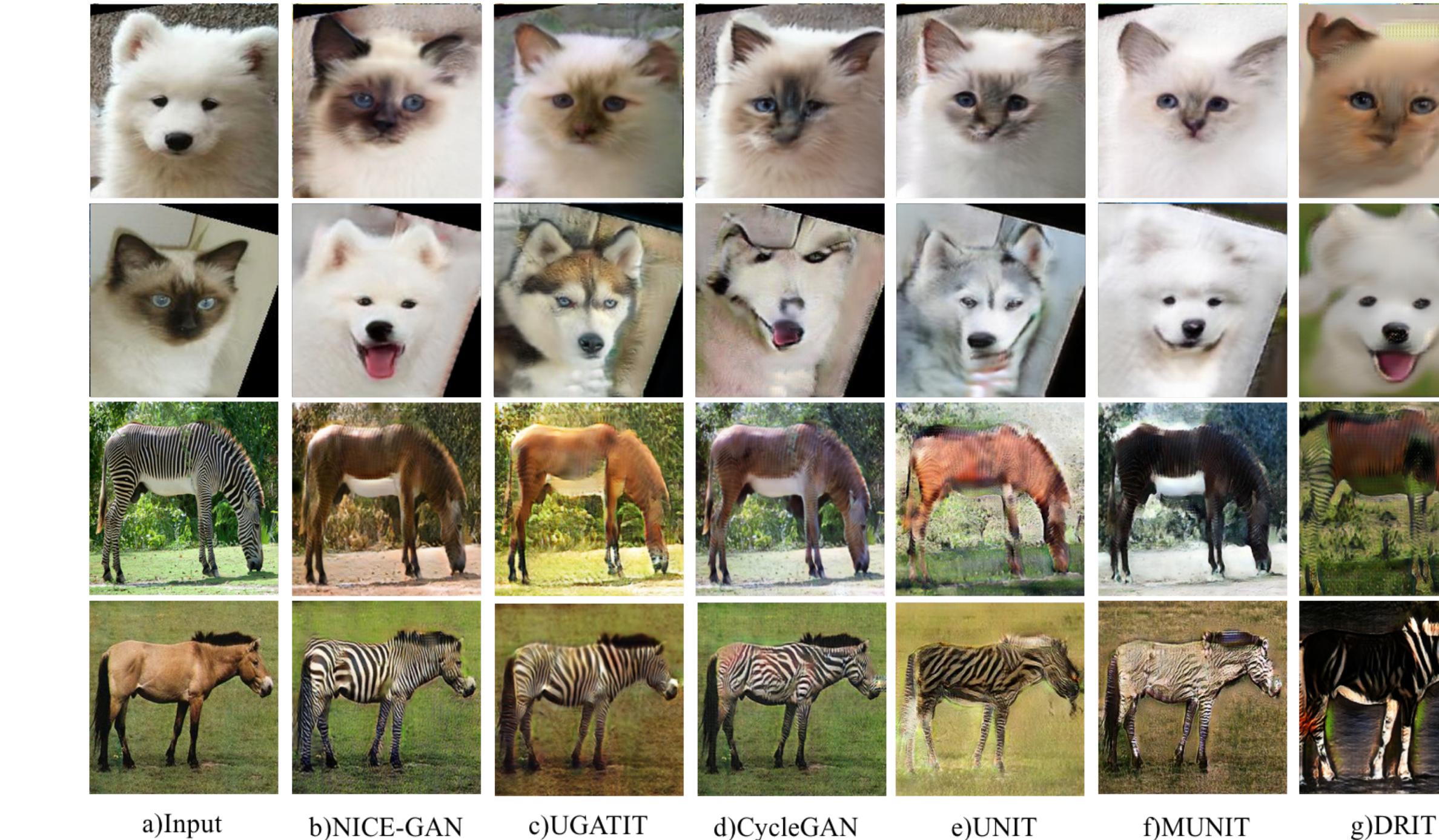
Website

<https://alpc91.github.io/>

NICE-GAN-pytorch/

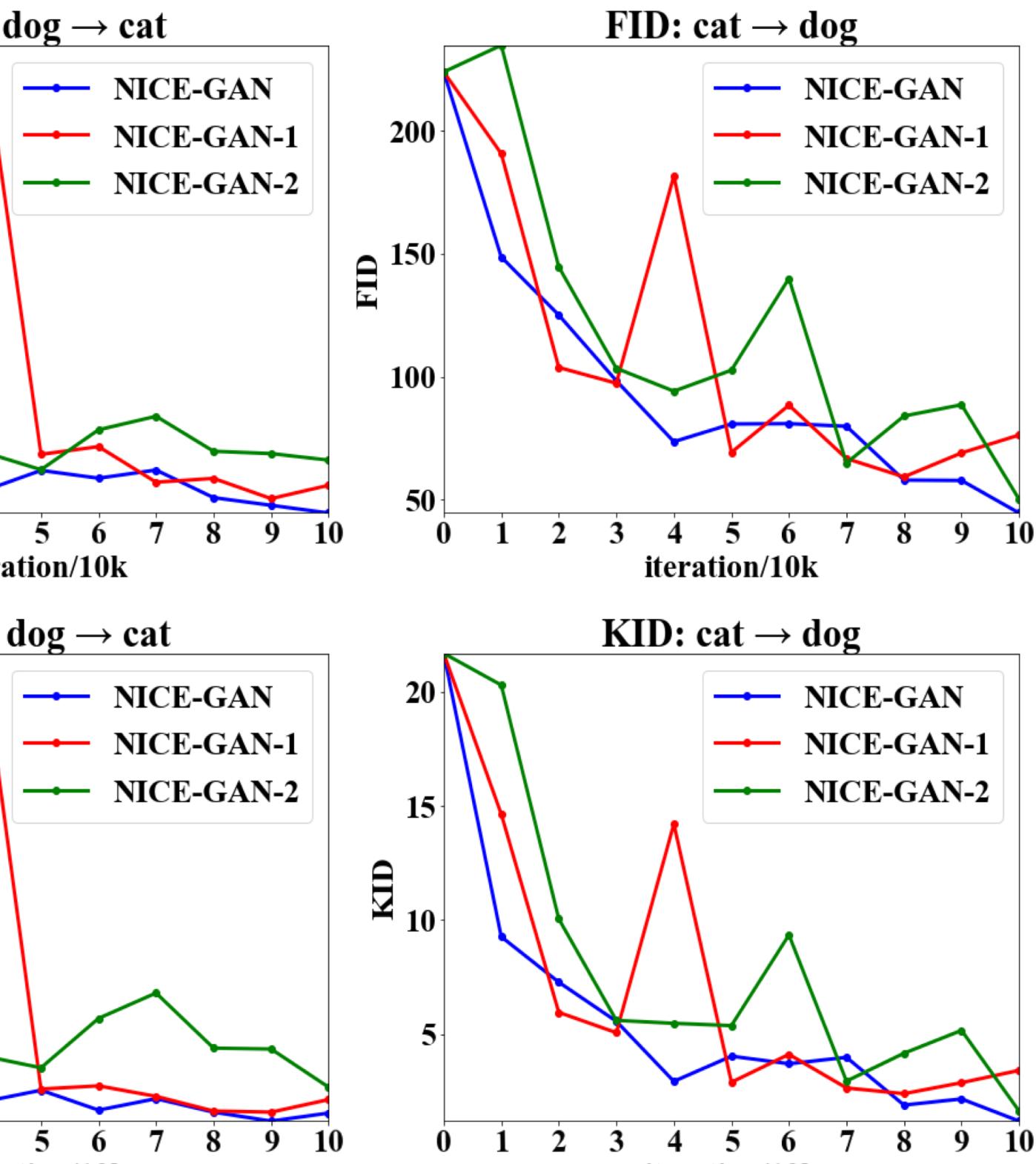
Wechat

Runfa Chen
Wenbing Huang

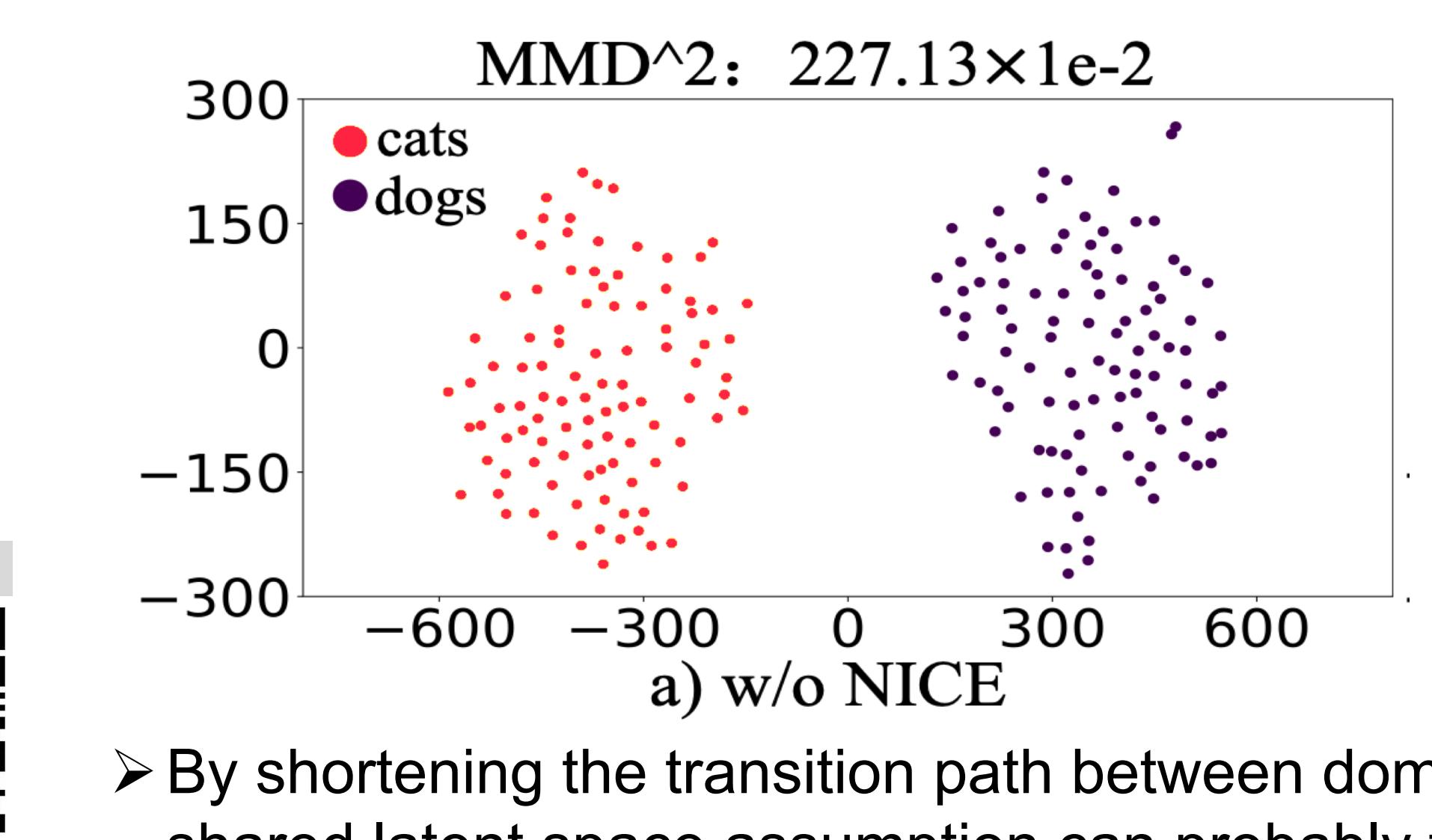


Ablation Study

Data Set	Components				FID	KID × 100
	NICE	RA	C_x^0	C_x^1	C_x^2	
dog → cat	×	×	✓	✓	✓	80.75
	×	✓	✓	✓	✓	2.94
	✓	✗	✓	✓	✓	3.27
	—	✓	✓	✓	✓	48.55
	✓	✓	✓	✓	✓	1.23
	+	✓	✓	✓	✗	1.58
cat → dog	✓	✓	✓	✓	✓	53.52
	✓	✓	✓	✓	✗	203.56
	✓	✓	✓	✗	✗	15.27
	✓	✓	✓	✗	✗	216.03
	✓	✓	✗	✗	✗	18.57
	✗	✗	✓	✓	✓	64.36
cat → dog	✗	✓	✓	✓	✓	64.62
	✓	✗	✓	✓	✓	51.49
	—	✓	✓	✓	✓	52.92
	✓	✓	✓	✓	✓	44.67
	+	✓	✓	✓	✓	21.41
	✓	✓	✓	✓	✗	231.24



The t-SNE Visualization of the Latent Vectors



By shortening the transition path between domains in the latent space, NICE-GAN built upon the shared latent space assumption can probably facilitate domain translation in the image space.