Applications and Future Directions of Generative Adversarial Networks

Hao Dong

2019, Peking University

Applications and Future Directions of GANs

- High Resolution Image Synthesis
- Text-based Image Synthesis
- 3D Data Synthesis
- Adversarial Domain Adaptation
- Discussion

High Resolution Image Synthesis

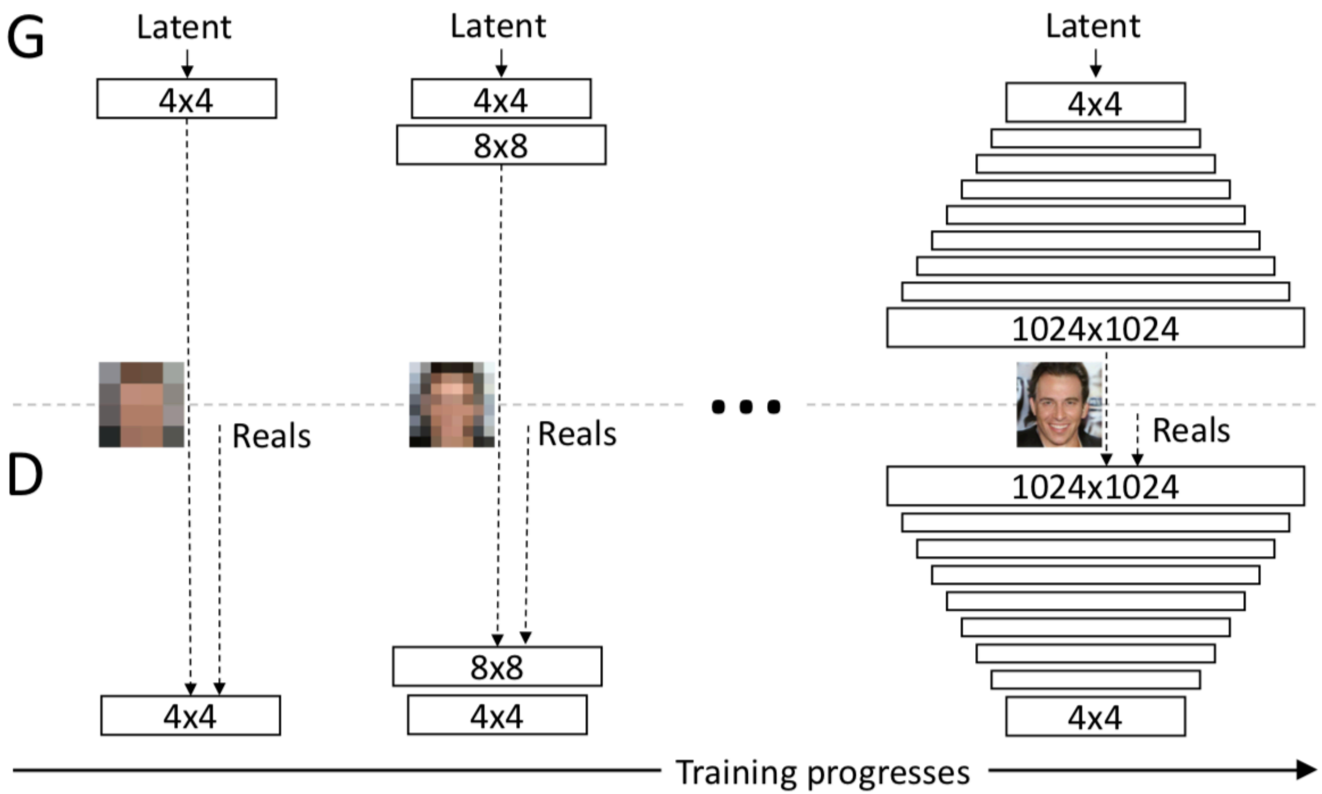
High Resolution Image Synthesis

- Progressive GAN

Progressive Growing of GANs for Improved Quality, Stability, and Variation. *T. Karras, T. Aila, et al. ICLR. 2018.*

High Resolution Image Synthesis

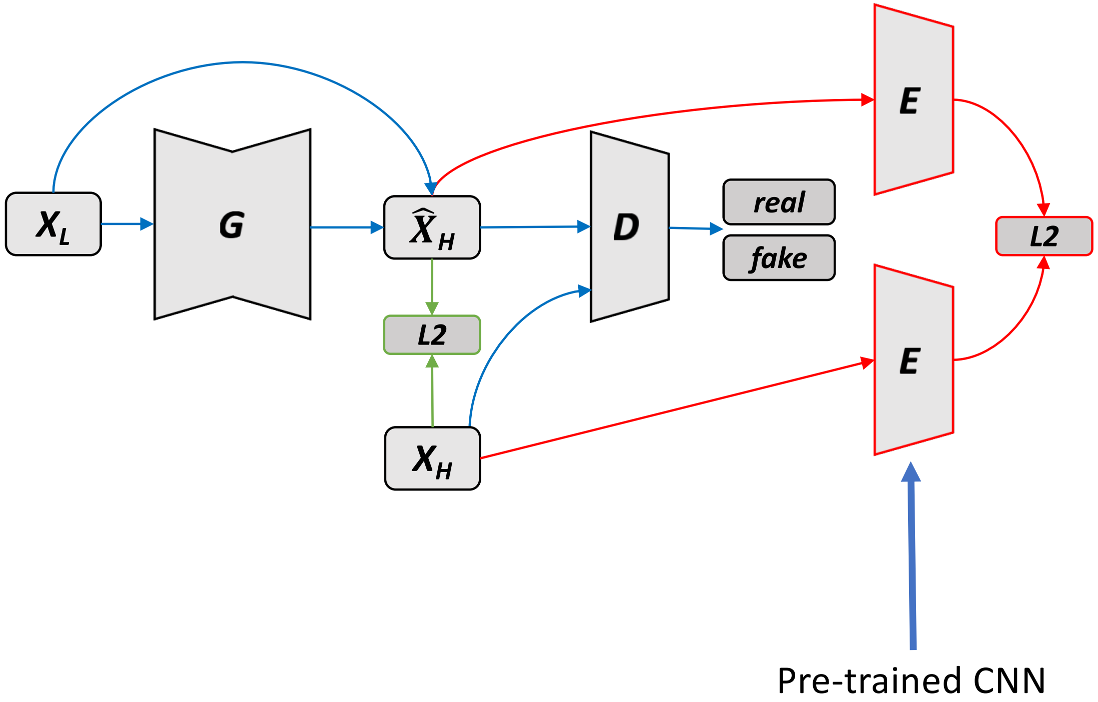
- Progressive GAN

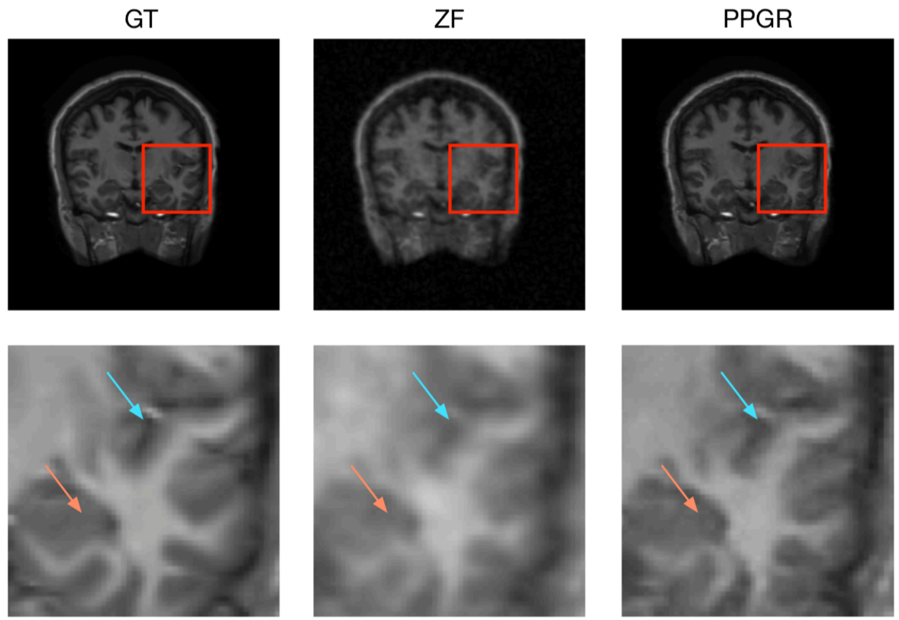


Progressive Growing of GANs for Improved Quality, Stability, and Variation. *T. Karras, T. Aila, et al. ICLR. 2018.*

High Resolution Image Synthesis

- Utilising Feature Information for Medical Image Reconstruction

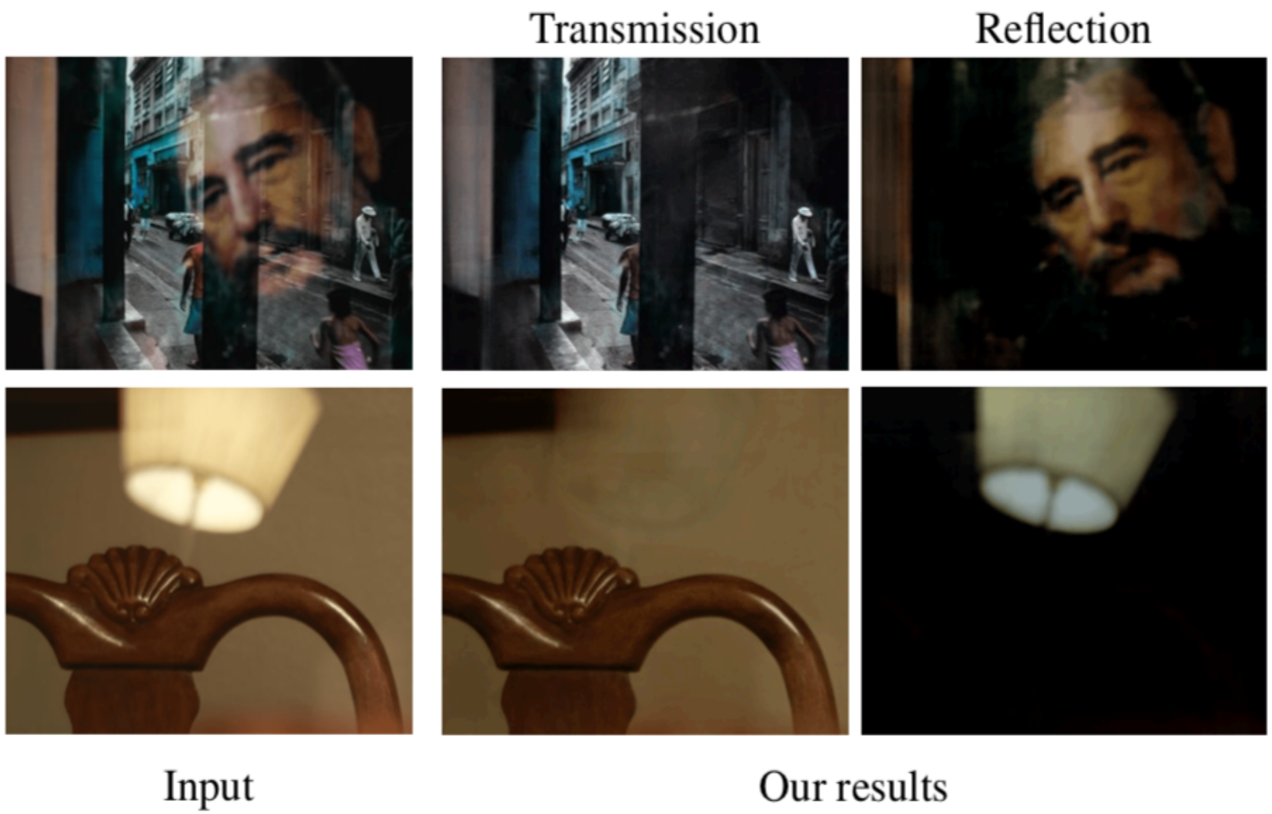




Deep De-Aliasing for Fast Compressive Sensing MRI. *S. Yu, H. Dong, G. Yang et al. arXiv:1705.07137 2017.*
DAGAN: Deep De-Aliasing Generative Adversarial Networks for Fast Compressed Sensing MRI Reconstruction.
G. Yang, S. Yu, H. Dong et al. TMI 2017.

High Resolution Image Synthesis

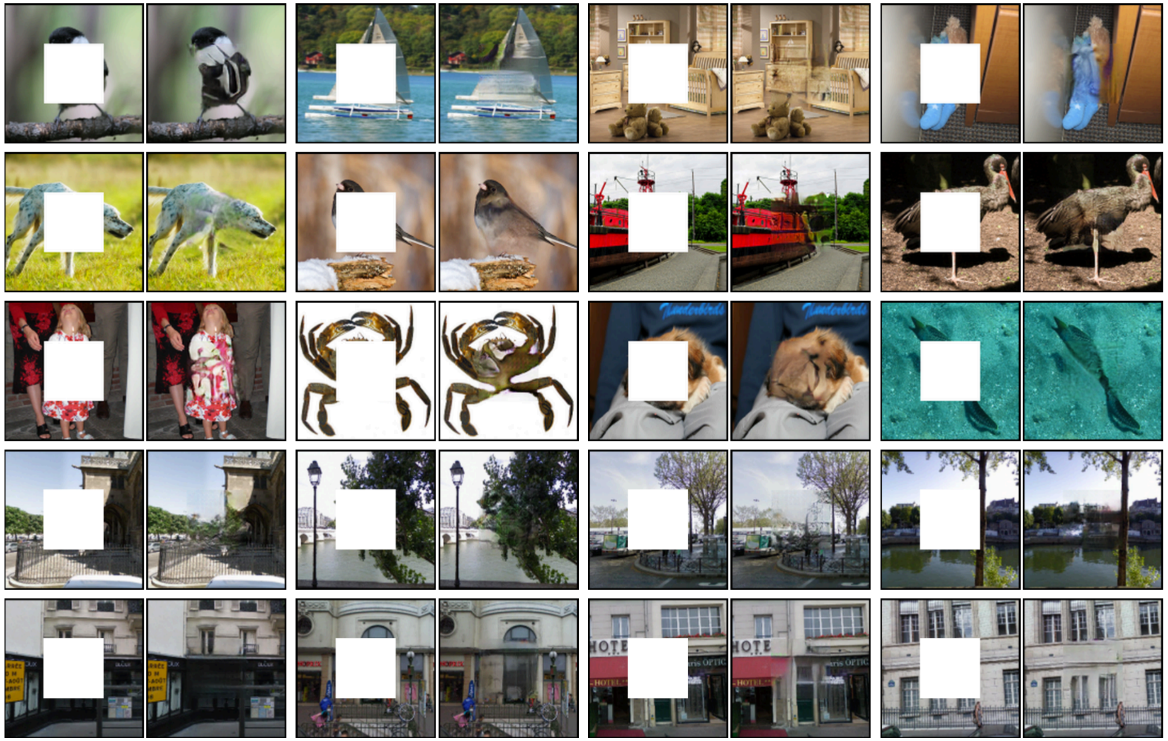
- Image Reflection Separation



Single Image Reflection Separation with Perceptual Losses. *X. Zhang, R. Ng, Q. Chen. CVPR. 2018.*

High Resolution Image Synthesis

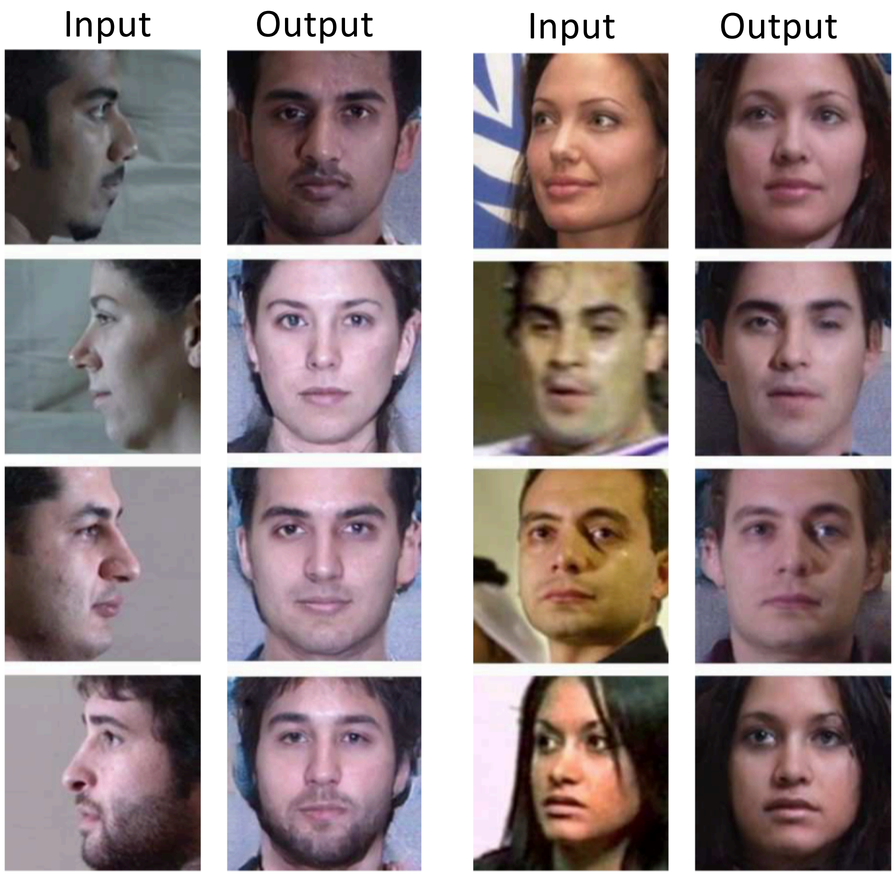
- Image inpainting



Context Encoders: Feature Learning by Inpainting. *D. Pathak, J. Donahue. CVPR. 2017*

High Resolution Image Synthesis

- Face Rotation

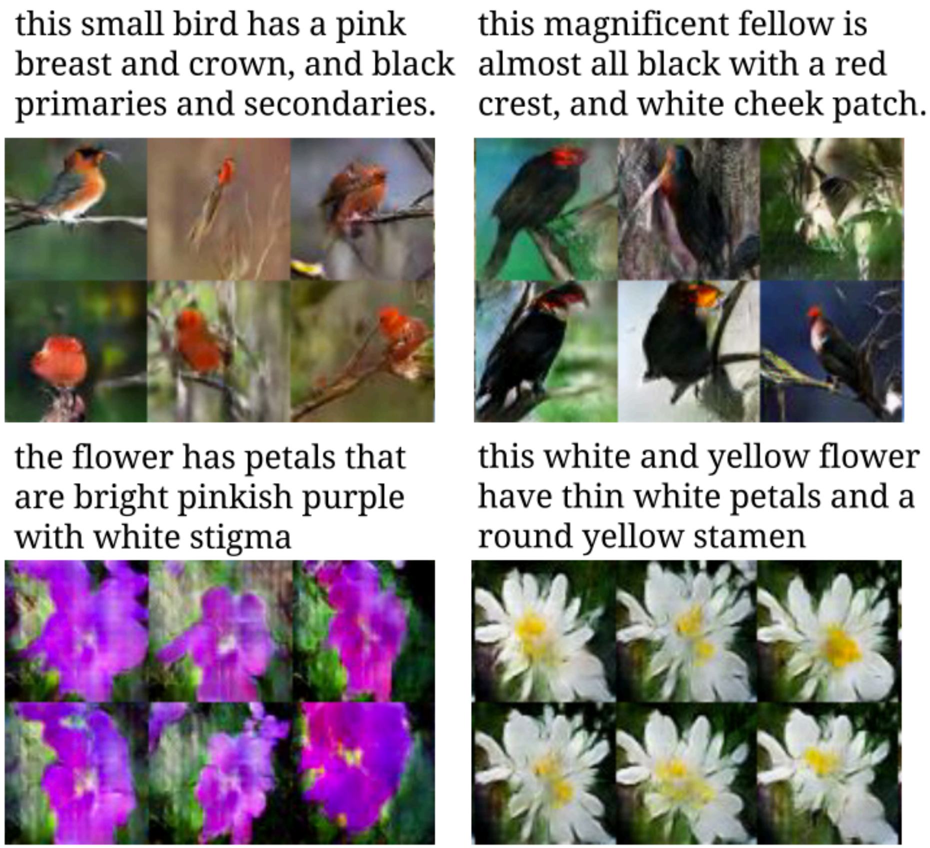


Pose-Guided Photorealistic Face Rotation. *Y. Hu, X. Wu et al. CVPR. 2018*

Text-based Image Synthesis

Text-based Image Synthesis

- Text-to-image synthesis



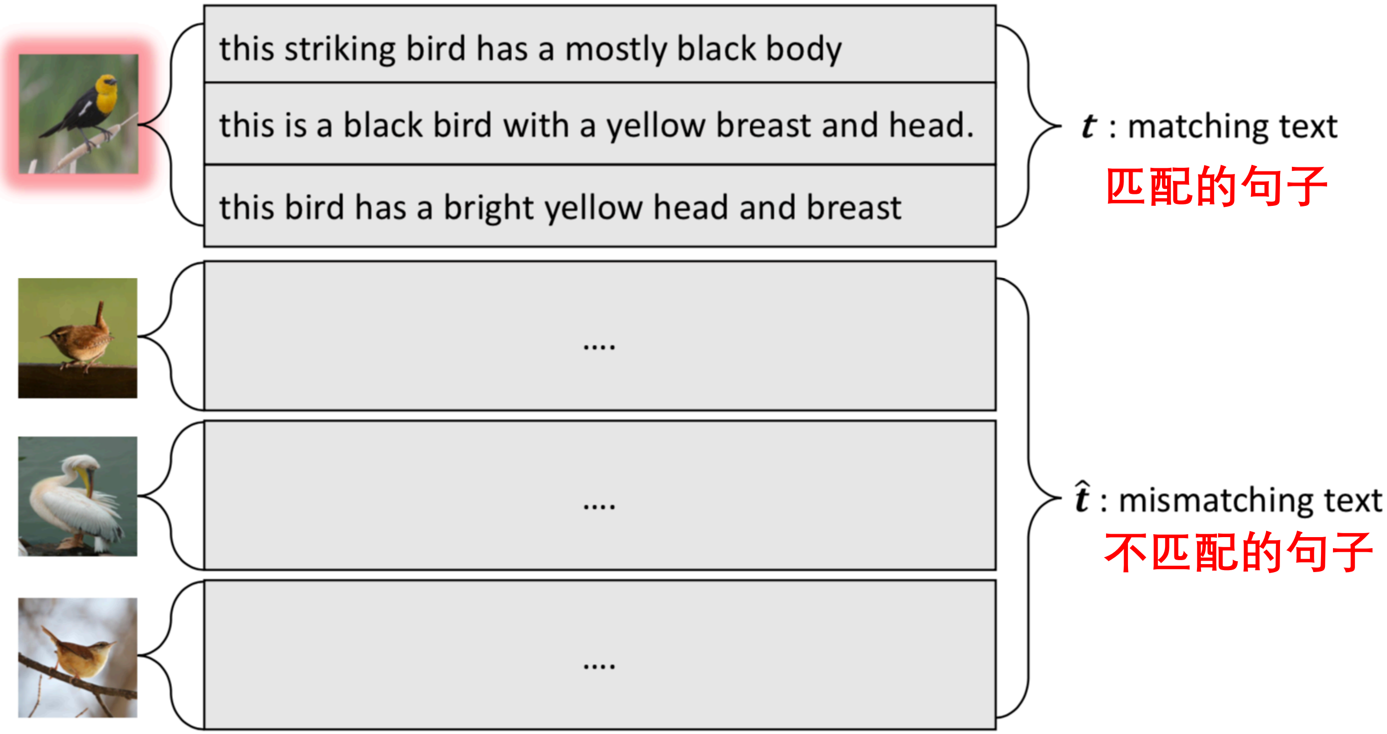
- Multi-modal problem

P(t, z)

Generative Adversarial Text to Image Synthesis. *S. Reed, Z. Akata et al. ICML. 2016.*

Text-based Image Synthesis

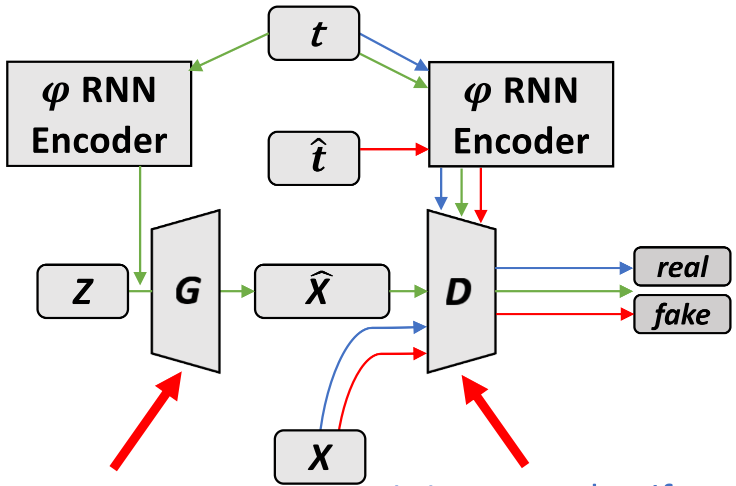
- Text-to-image synthesis



Generative Adversarial Text to Image Synthesis. *S. Reed, Z. Akata et al. ICML. 2016.*

Text-based Image Synthesis

- Text-to-image synthesis

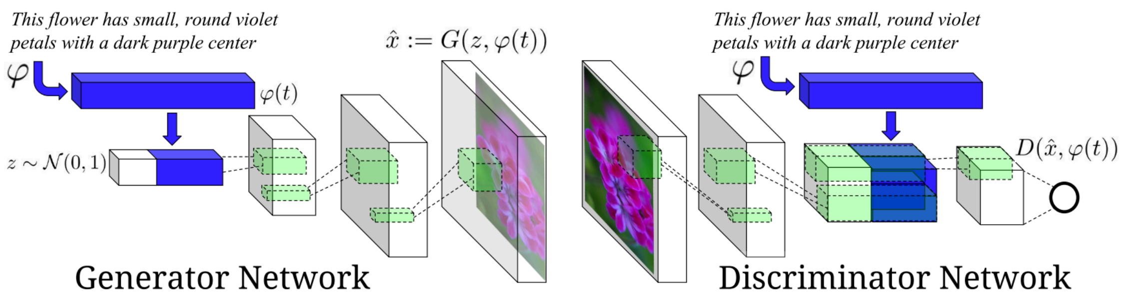


t : matching text

$\hat{t}$: mismatched text

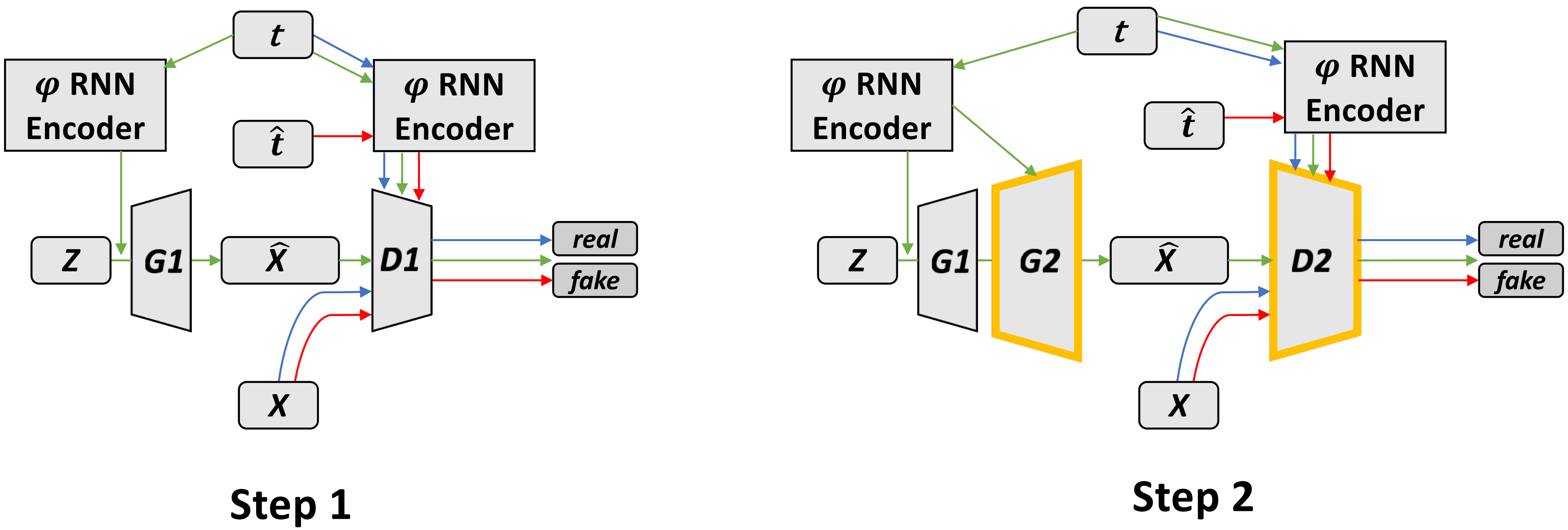
Learn to fool discriminator

1. Learn to classify matching image and text as real sample
2. Learn to classify mismatched image and text as fake sample
3. Learn to classify image from generator as fake sample



Text-based Image Synthesis

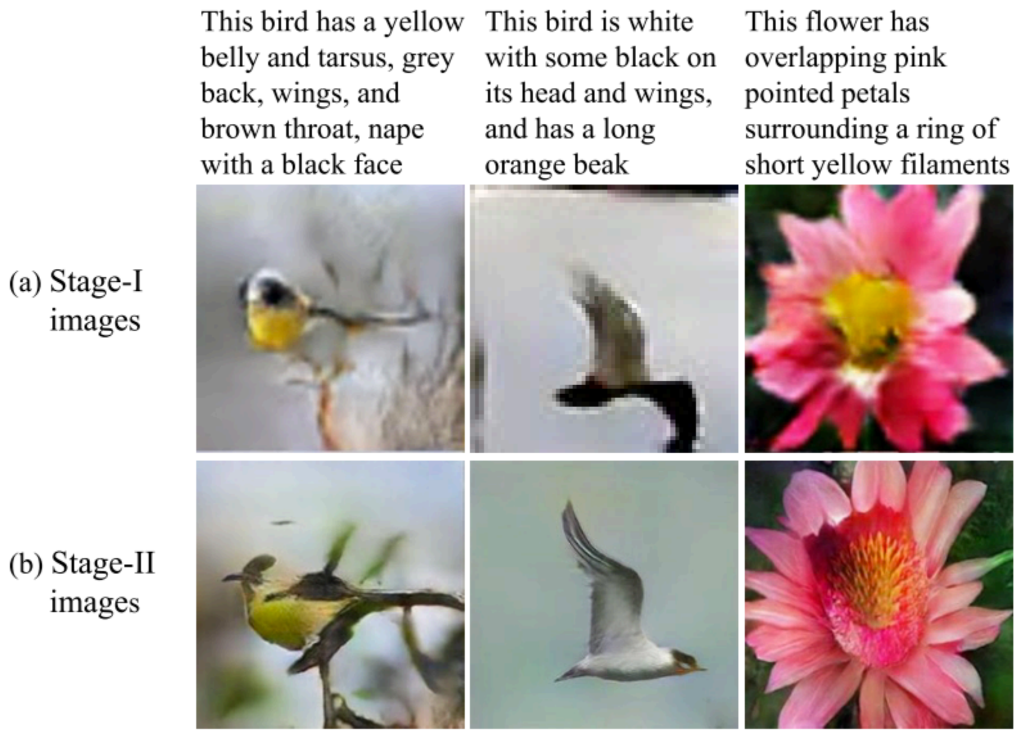
- Text-to-image synthesis + High resolution image



StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks. *H. Zhang, T. Xu et al. ICCV. 2017.*

Text-based Image Synthesis

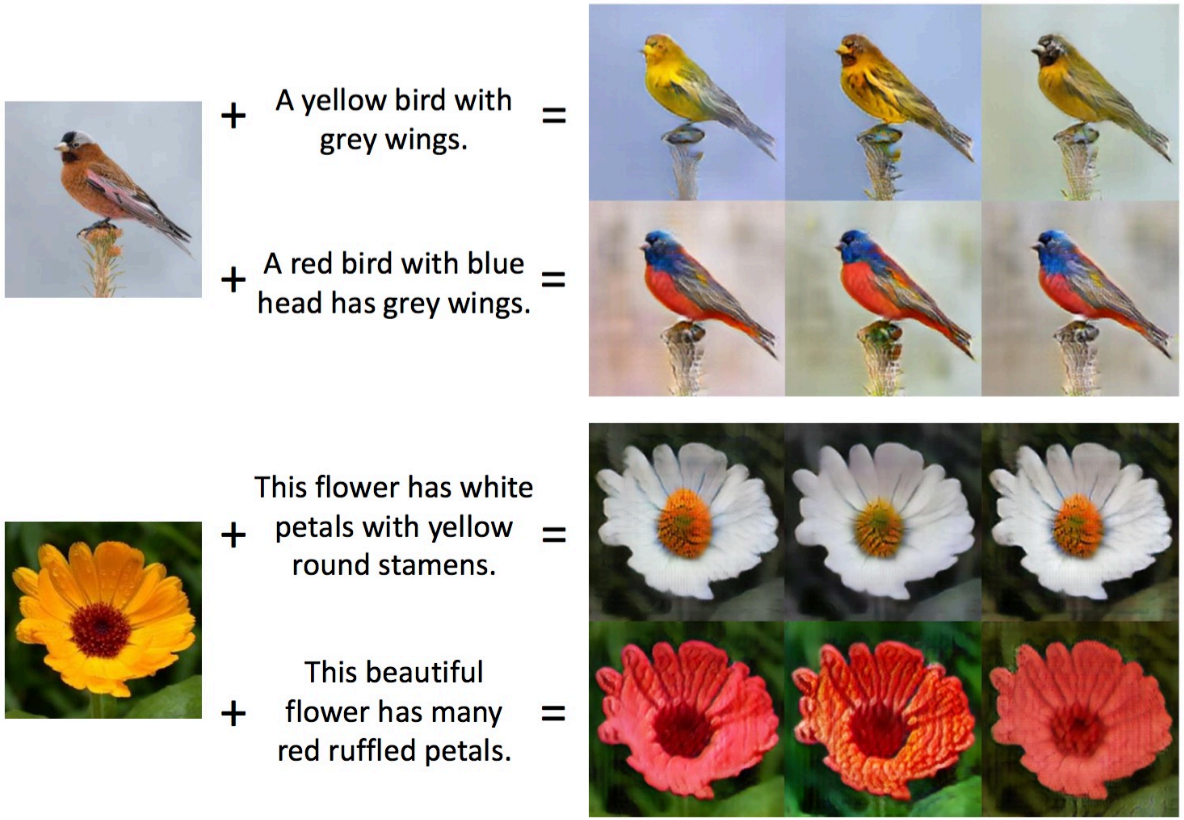
- Text-to-image synthesis + High resolution image



StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks. *H. Zhang, T. Xu et al. ICCV. 2017.*

Text-based Image Synthesis

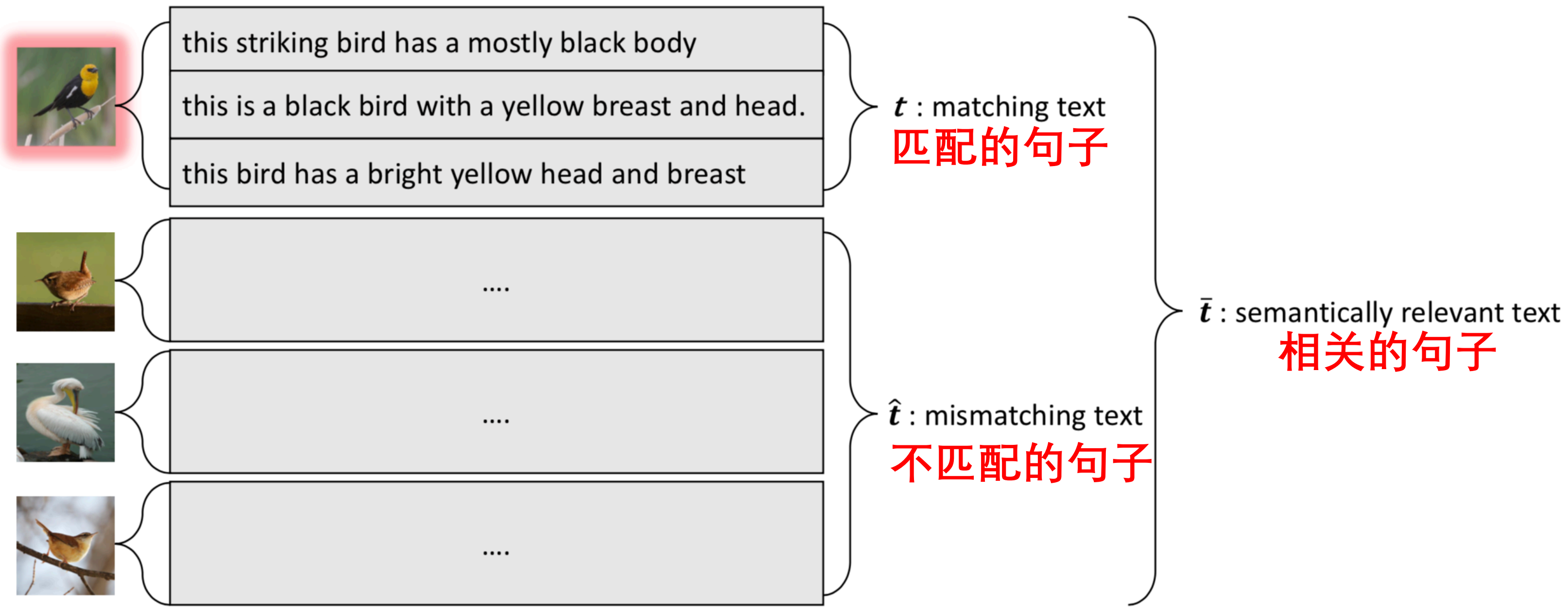
- Semantic image synthesis



Semantic Image Synthesis via Adversarial Learning. *H. Dong, S. Yu et al. ICCV 2017.*

Text-based Image Synthesis

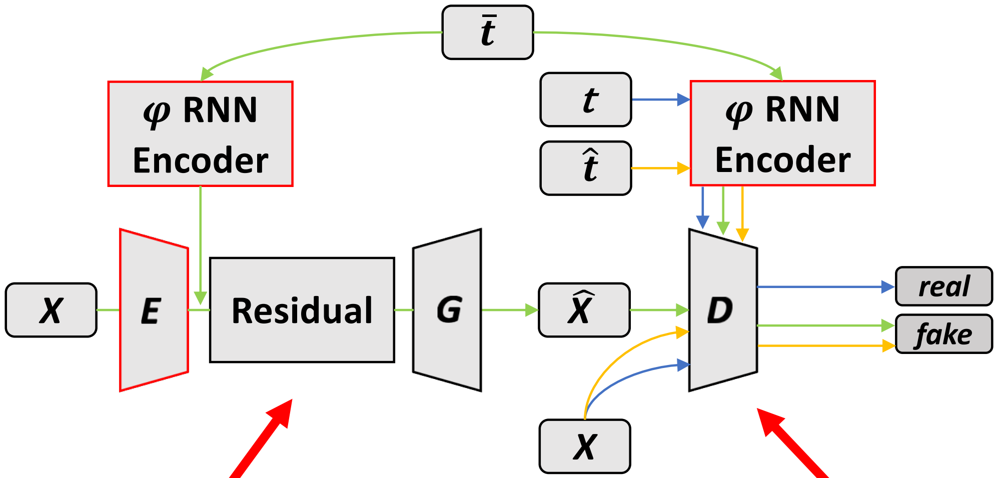
- Semantic image synthesis



Semantic Image Synthesis via Adversarial Learning. *H. Dong, S. Yu et al. ICCV 2017.*

Text-based Image Synthesis

- Semantic image synthesis



t : matching text

$\hat{t}$: mismatched text

$\bar{t}$: semantically relevant text

$$\mathcal{L}_D = \boxed{\mathbb{E}_{(x,t)\sim p_{data}} log D(x, \varphi(t))} + \boxed{\mathbb{E}_{(x,\hat{t})\sim p_{data}} log(1 - D(x, \varphi(\hat{t})))} + \boxed{\mathbb{E}_{(x,\bar{t})\sim p_{data}} log(1 - D(G(x, \varphi(\bar{t})), \varphi(\bar{t})))}$$

$$\mathcal{L}_G = \boxed{\mathbb{E}_{(x,\bar{t})\sim p_{data}} log(D(G(x, \varphi(\bar{t})), \varphi(\bar{t})))}$$

Learn to fool discriminator when inputting image with semantically relevant text

1. Learn to classify matching image and text pairs as real samples
2. Learn to classify mismatched image and text pairs as fake samples
3. Learn to classify samples from generator as fake samples

Semantic Image Synthesis via Adversarial Learning. *H. Dong, S. Yu et al. ICCV 2017.*

3D Data Synthesis

3D Data Synthesis

- 3D-GAN

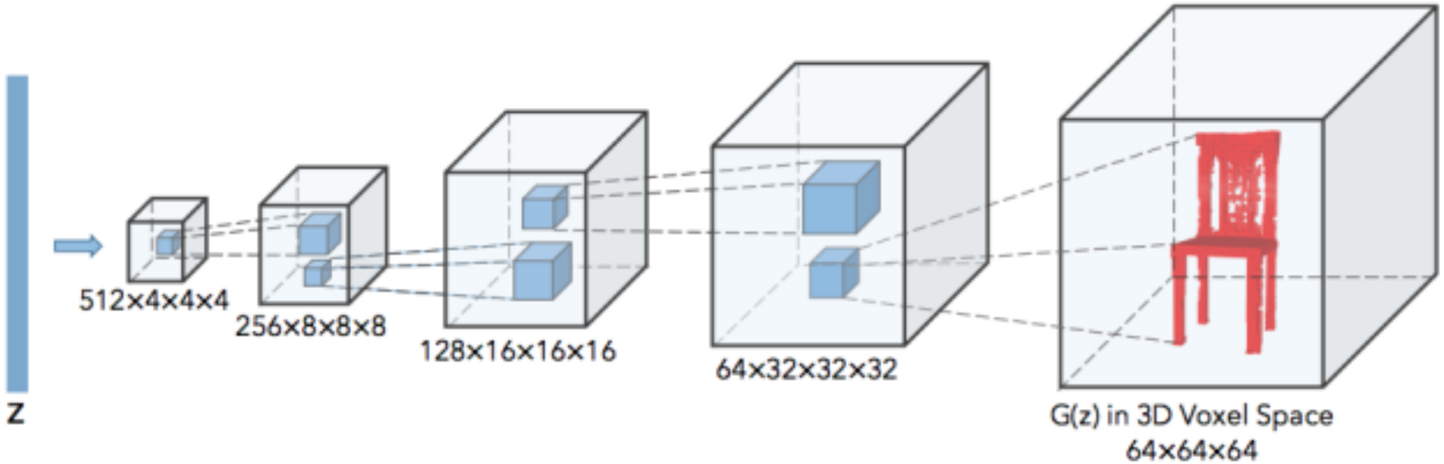
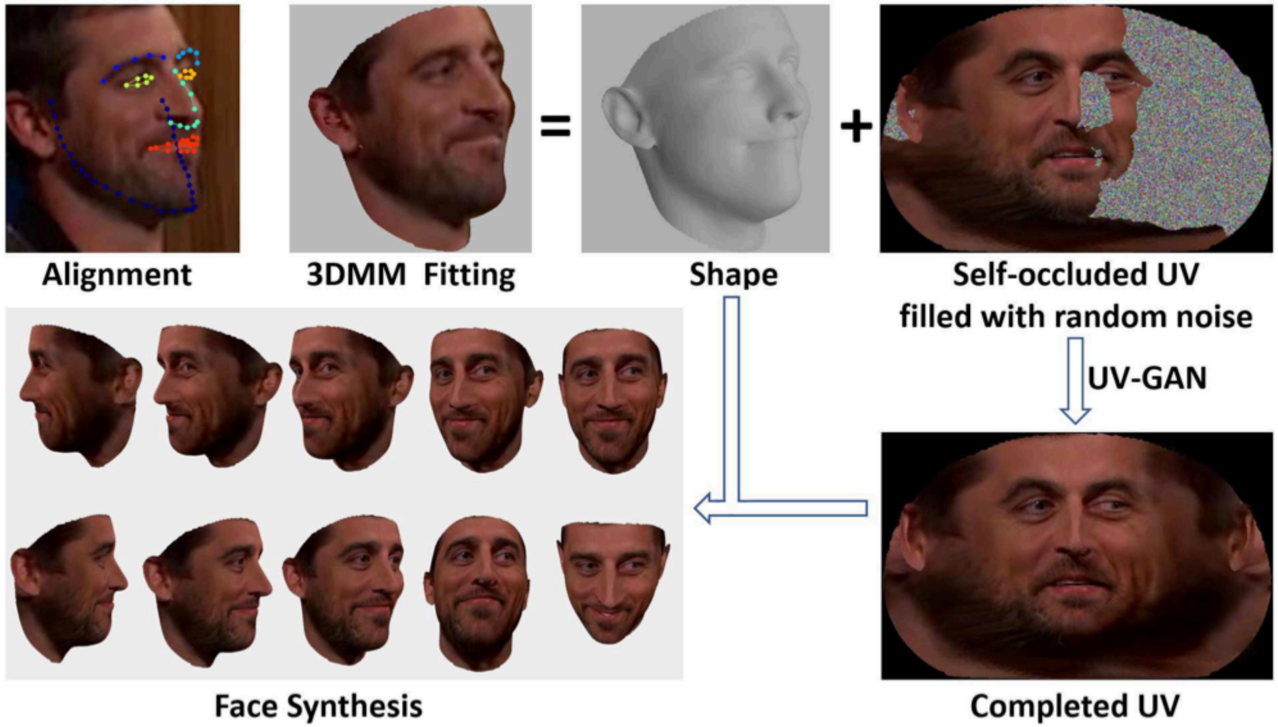


Figure 1: The generator in 3D-GAN. The discriminator mostly mirrors the generator.

Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling.
J. Wu, C. Zhang et al. NIPS. 2016.

3D Data Synthesis

- UV-GAN



UV-GAN: Adversarial Facial UV Map Completion for Pose-invariant Face Recognition.
J. Deng, S. Cheng et al. CVPR. 2018.

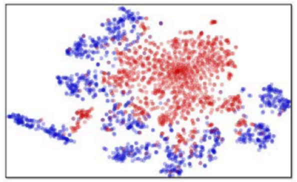
Adversarial Domain Adaptation

Adversarial Domain Adaptation

- Single Source Domain Adaptation

Source: Labelled

Target: Unlabelled



$$S(\mathbf{f}) = \{G_f(\mathbf{x};\theta_f) \,|\, \mathbf{x} \sim S(\mathbf{x})\}$$
$$T(\mathbf{f}) = \{G_f(\mathbf{x};\theta_f) \,|\, \mathbf{x} \sim T(\mathbf{x})\}$$

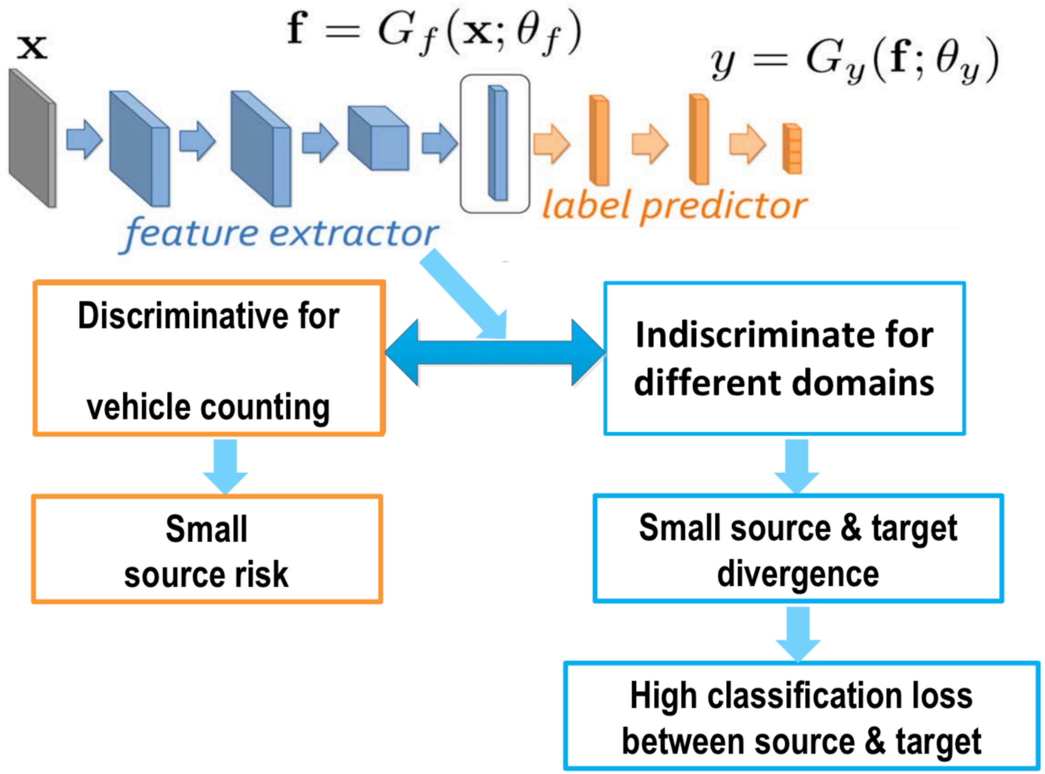
Domain shift among sources and target

Domain adaptation needed!

Domain-Adversarial Training of Neural Networks. *Y. Ganin, H. Ajakan et al. JMLR. 2016*

Adversarial Domain Adaptation

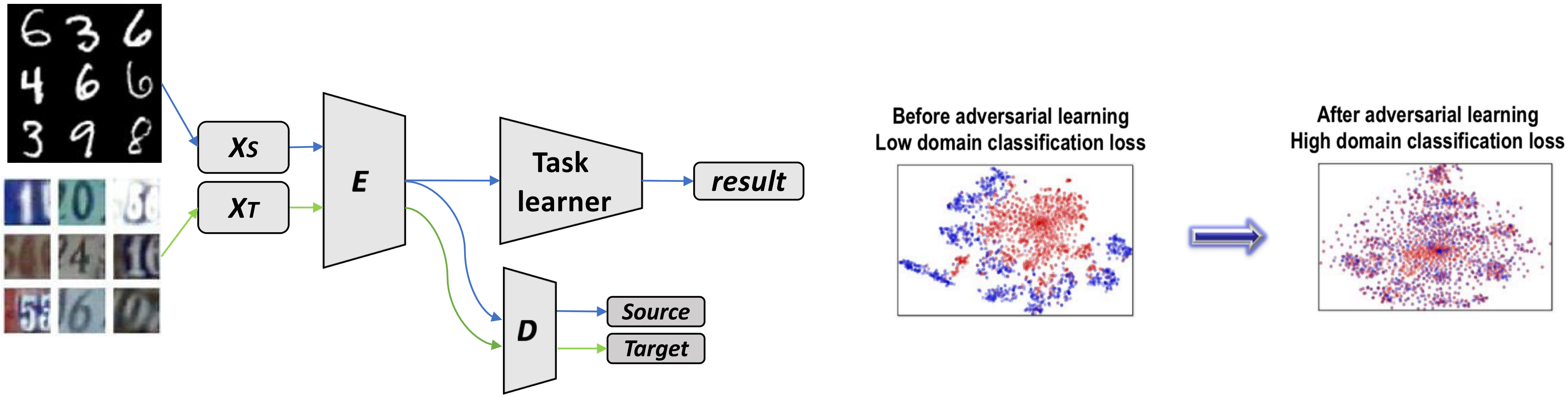
- Learn domain-universal & task-discriminative features



Domain-Adversarial Training of Neural Networks. *Y. Ganin, H. Ajakan et al. JMLR. 2016*

Adversarial Domain Adaptation

- Single Source Domain Adaptation

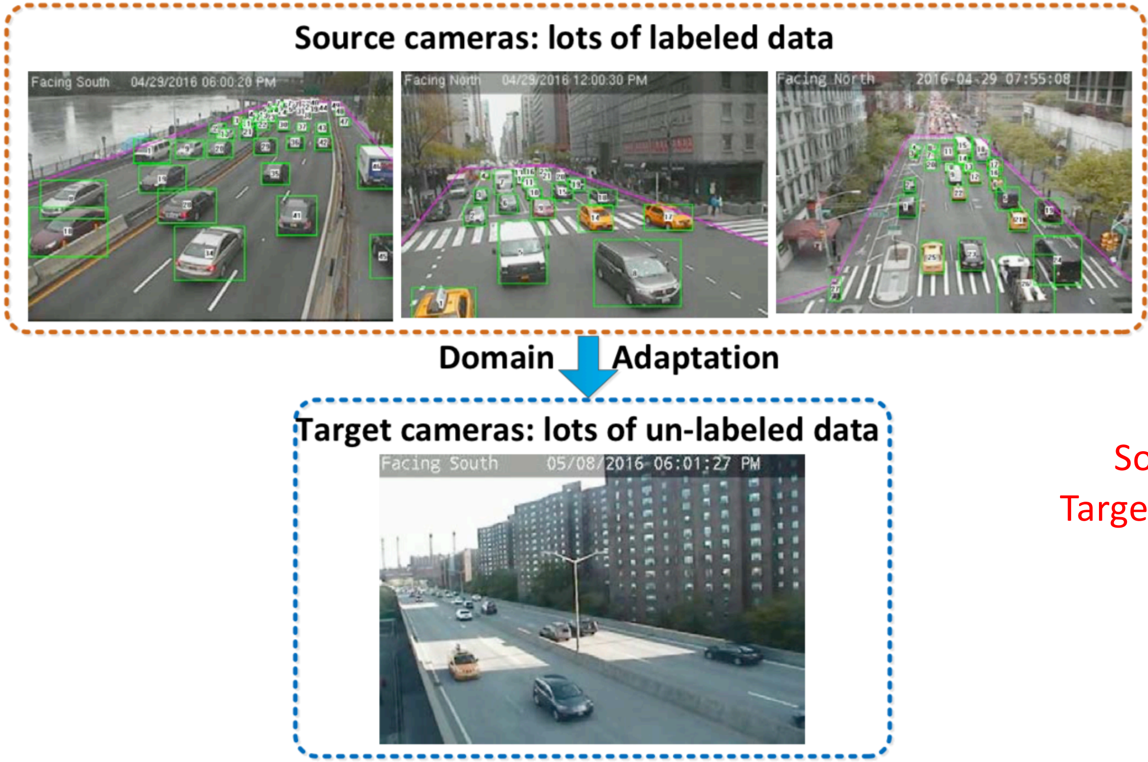


Source images have label
Target images do not have label

Domain-Adversarial Training of Neural Networks. *Y. Ganin, H. Ajakan et al. JMLR. 2016*

Adversarial Domain Adaptation

- Multiple Source Domain Adaptation

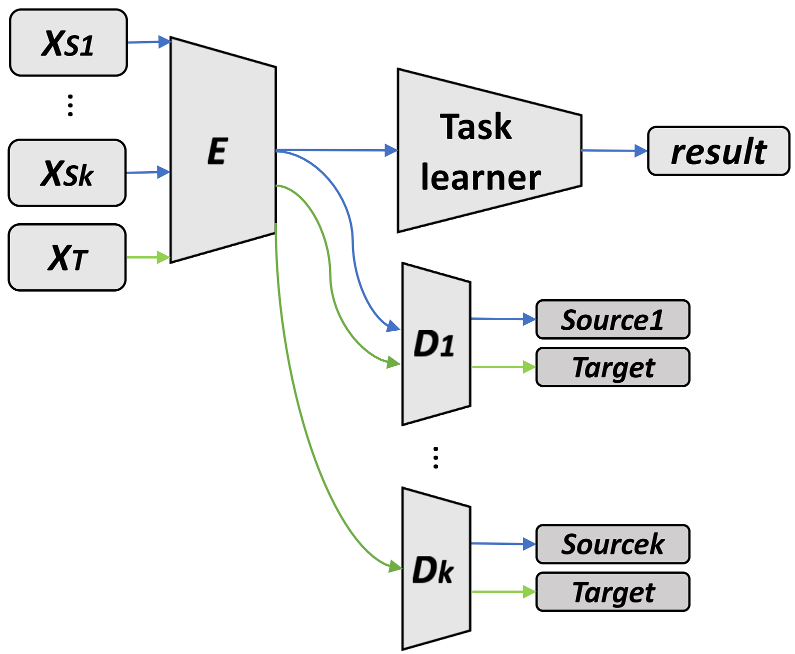


Source images have label
Target images do not have label

Multiple Source Domain Adaptation with Adversarial Learning. *S. Zhang, H. Zhao et al. NIPS. 2018.*

Adversarial Domain Adaptation

- Multiple Source Domain Adaptation



Source images have label
Target images do not have label

Multiple Source Domain Adaptation with Adversarial Learning. *S. Zhang, H. Zhao et al. NIPS. 2018.*

Discussion

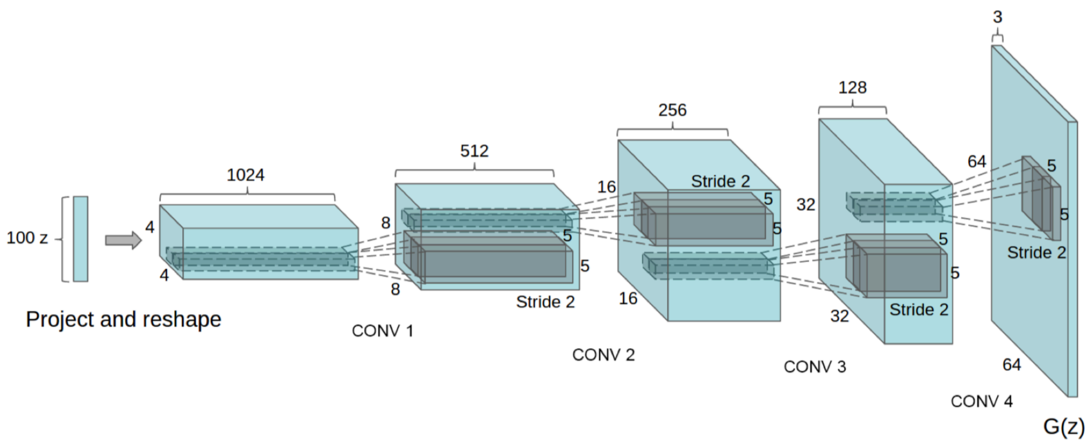
- Exercise 1:
  - Implement the DCGAN

- Exercise 2:
  - Study and Explain W-GAN

- Exercise 3: (Optional)
  - Choice an application and implement it

Link: https://github.com/zsdonghao/deep-learning-note/

Thank You


100 z
Project and reshape
1024
4
4
512
8
8
Stride 2
CONV 1
256
16
16
Stride 2
CONV 2
128
32
32
Stride 2
CONV 3
3
64
64
Stride 2
CONV 4
G(z)