CAP 5516
Medical Image Computing
(Spring 2022)

Dr. Chen Chen
Center for Research in Computer Vision (CRCV)
University of Central Florida
Office: HEC 221
Address: 4328 Scorpius St., Orlando, FL 32816-2365
Email: chen.chen@crcv.ucf.edu
Web: https://www.crcv.ucf.edu/chenchen/

UCF Center for Research in Computer Vision

Lecture 10
Introduction of Generative Adversarial Networks (GANs)

Generative adversarial nets

I Goodfellow, J Pouget-Abadie… - Advances in neural …, 2014 - proceedings.neurips.cc

We propose a new framework for estimating generative models via adversarial nets, in which we simultaneously train two models: a generative model G that captures the data distribution, and a discriminative model D that estimates the probability that a sample came …

☆ Save 99 Cite **Cited by 40817** Related articles

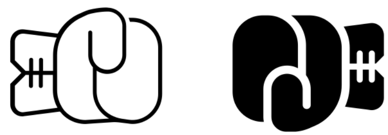
Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. "Generative adversarial nets." Advances in neural information processing systems 27 (2014).

Generative Adversarial Networks (GANs)

Generative

Generates data
(Create fake data)

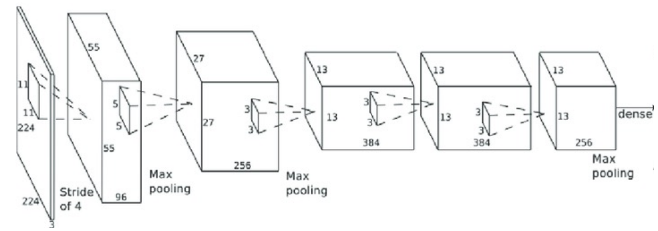
Adversarial

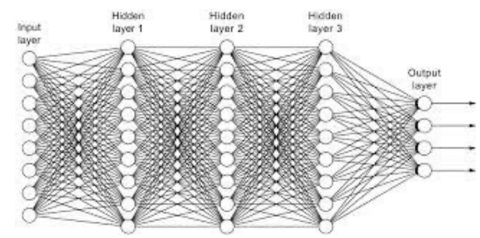


Generator and discriminator,
Each competing to win

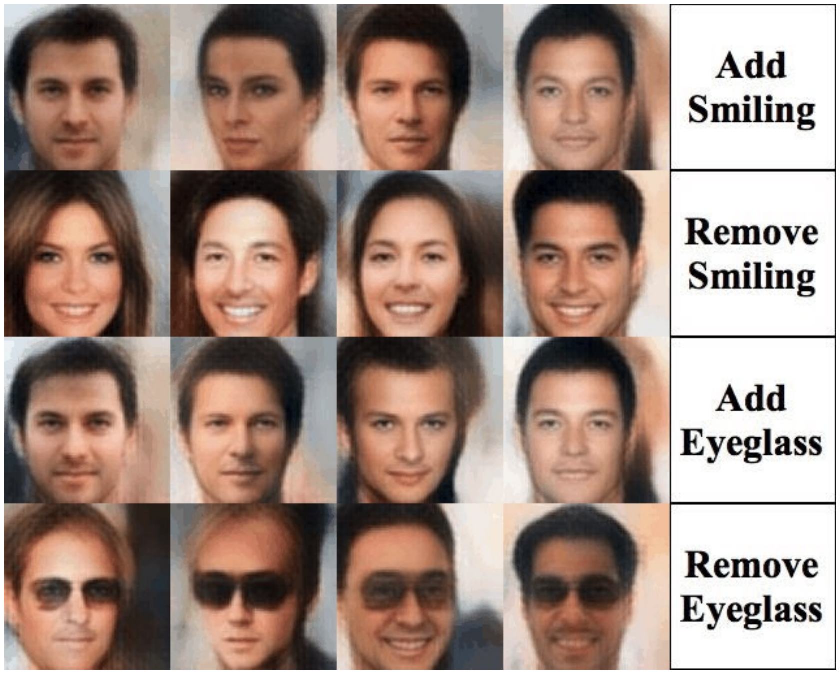
Generator trying to fake
Discriminator trying not to be fooled

Networks





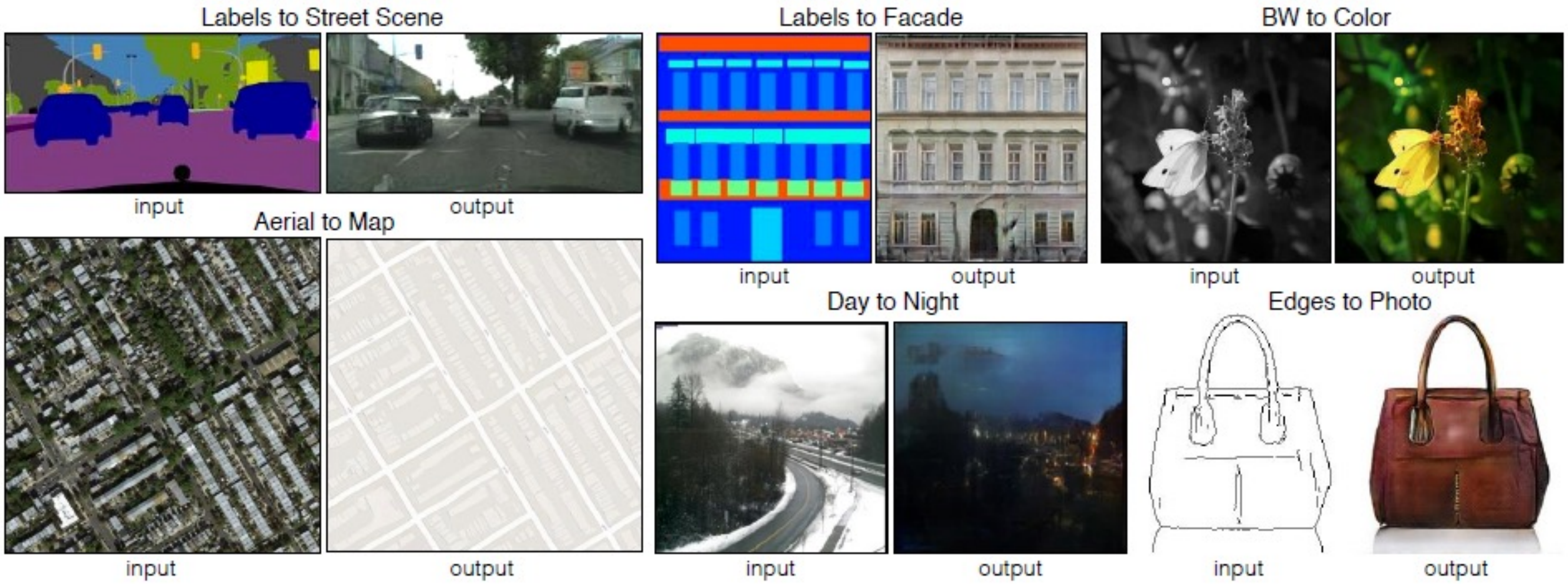
Applications of GANs



Generate Examples for Image Datasets,
Generate Realistic Photographs

Applications of GANs

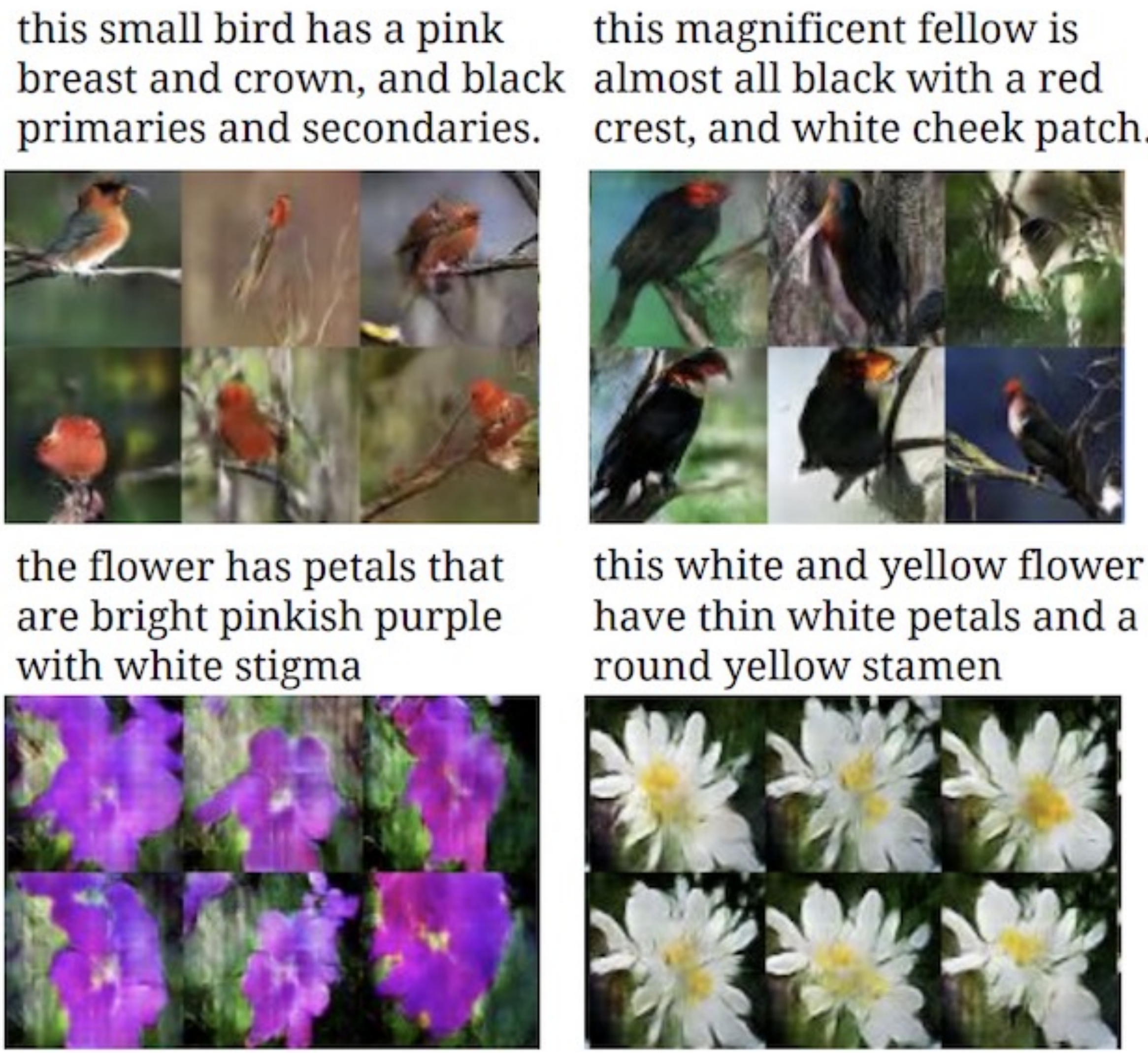
Image-to-Image Translation



Isola, Phillip, et al. "Image-to-image translation with conditional adversarial networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

https://machinelearningknowledge.ai/mind-blowing-applications-of-generative-adversarial-networks-gans/
https://machinelearningmastery.com/impressive-applications-of-generative-adversarial-networks/

Applications of GANs

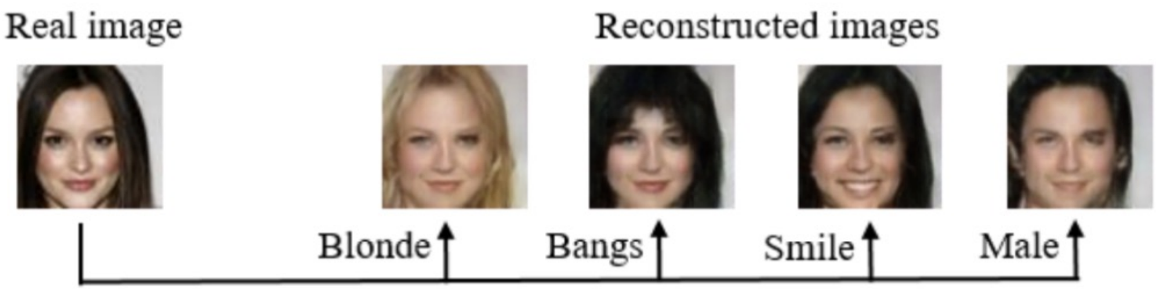


Text-to-Image Translation

Zhang, Han, et al. "Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks." Proceedings of the IEEE international conference on computer vision. 2017.

https://machinelearningknowledge.ai/mind-blowing-applications-of-generative-adversarial-networks-gans/
https://machinelearningmastery.com/impressive-applications-of-generative-adversarial-networks/

Applications of GANs



Photograph Editing

Perarnau, G., Van De Weijer, J., Raducanu, B., & Álvarez, J. M. (2016). Invertible conditional gans for image editing. arXiv preprint arXiv:1611.06355.

Photo Blending

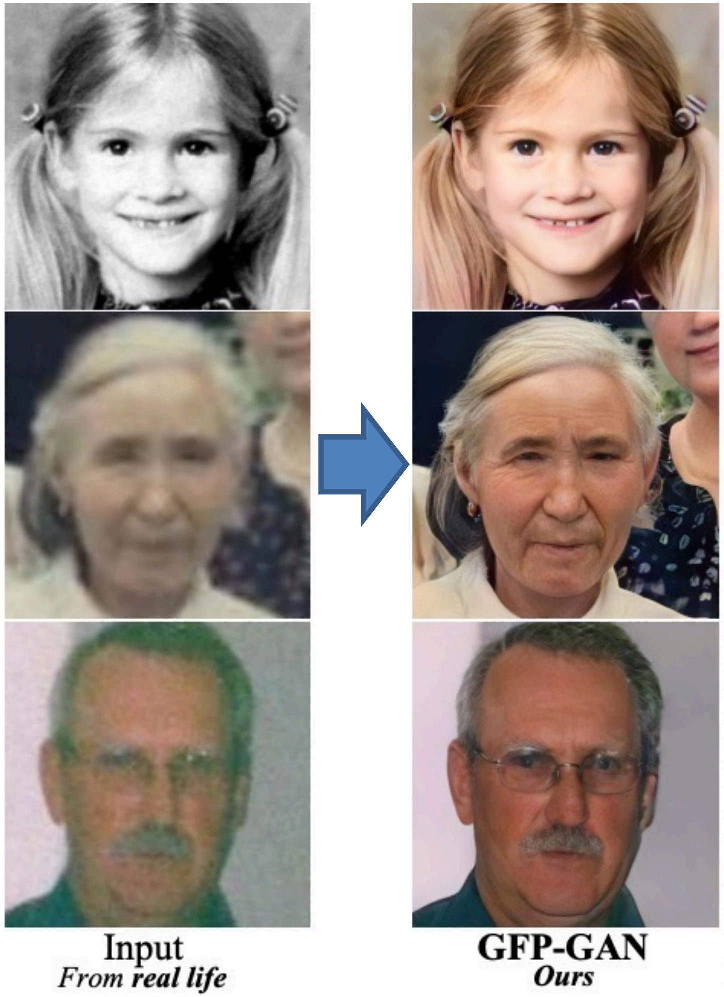
Wu, Huikai, et al. "Gp-gan: Towards realistic high-resolution image blending." Proceedings of the 27th ACM international conference on multimedia. 2019.

https://machinelearningknowledge.ai/mind-blowing-applications-of-generative-adversarial-networks-gans/
https://machinelearningmastery.com/impressive-applications-of-generative-adversarial-networks/

Applications of GANs

Image Recovery, Image Enhancement

Menon, Sachit, et al. "Pulse: Self-supervised photo upsampling via latent space exploration of generative models." Proceedings of the ieee/cvf conference on computer vision and pattern recognition. 2020.



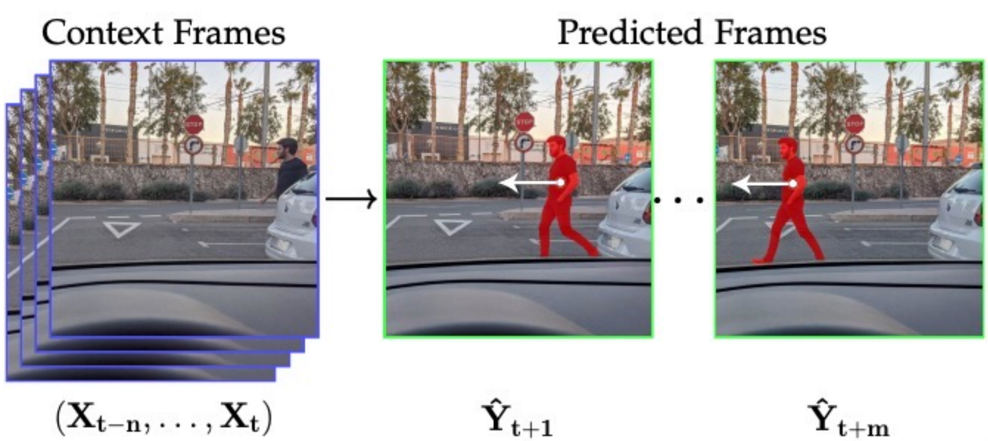
Wang, Xintao, et al. "Towards real-world blind face restoration with generative facial prior." CVPR 2021.

https://machinelearningknowledge.ai/mind-blowing-applications-of-generative-adversarial-networks-gans/
https://machinelearningmastery.com/impressive-applications-of-generative-adversarial-networks/

Applications of GANs

Video frames prediction

Kwon, Yong-Hoon, and Min-Gyu Park. "Predicting future frames using retrospective cycle gan." CVPR 2019.

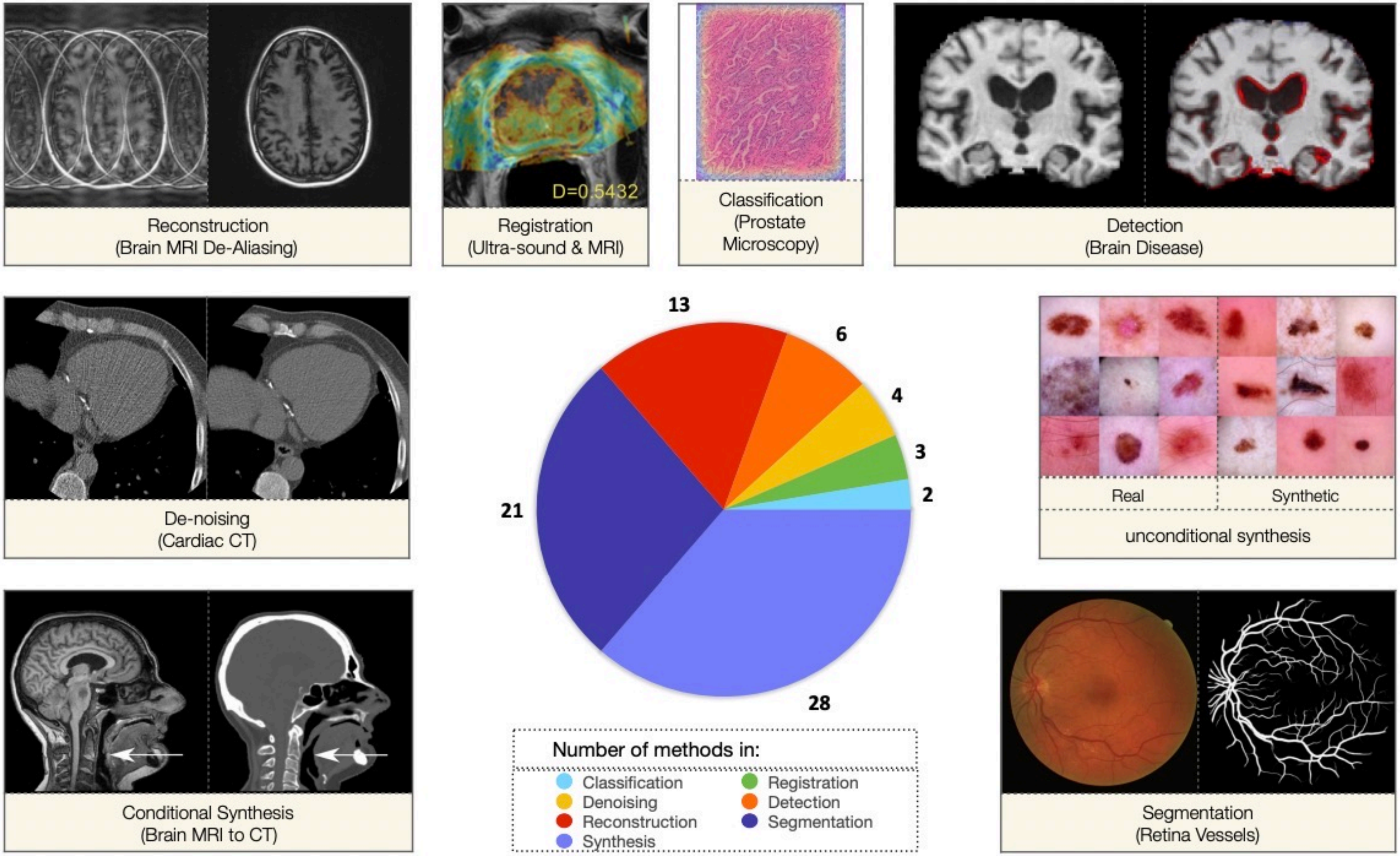


Will the car hit the pedestrian?

Oprea, Sergiu, et al. "A review on deep learning techniques for video prediction." IEEE Transactions on Pattern Analysis and Machine Intelligence (2020).

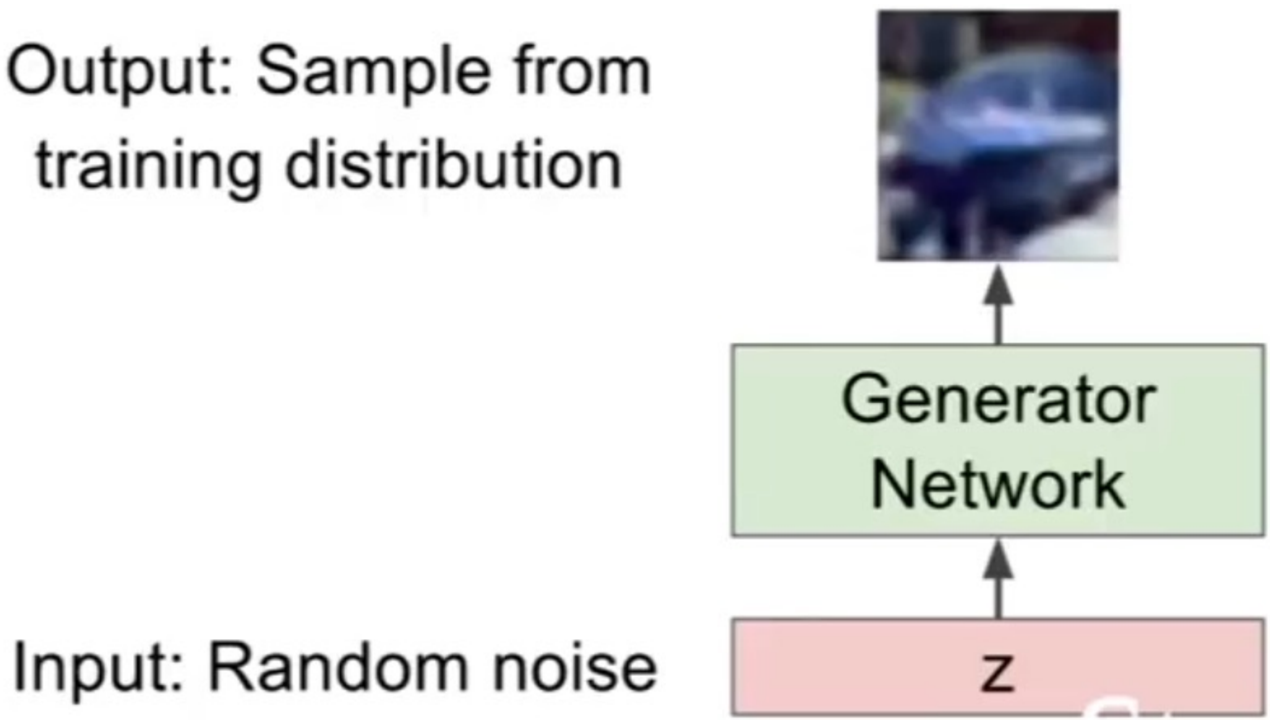
https://machinelearningknowledge.ai/mind-blowing-applications-of-generative-adversarial-networks-gans/
https://machinelearningmastery.com/impressive-applications-of-generative-adversarial-networks/

GANs for Medical Imaging



Kazeminia, Salome, et al. "GANs for medical image analysis." Artificial Intelligence in Medicine 109 (2020): 101938.

High level overview of GAN



Slide Credit: Feifei Li, Justin Johnson, Serena Yeung, Ranjay Krishna, Danfei Xu

High level overview of GAN

- Training GAN: Two-player game

  Player 1: Generator
  Player 2: Discriminator

  Generator network: try to fool the discriminator by generating real-looking images (Art forger)

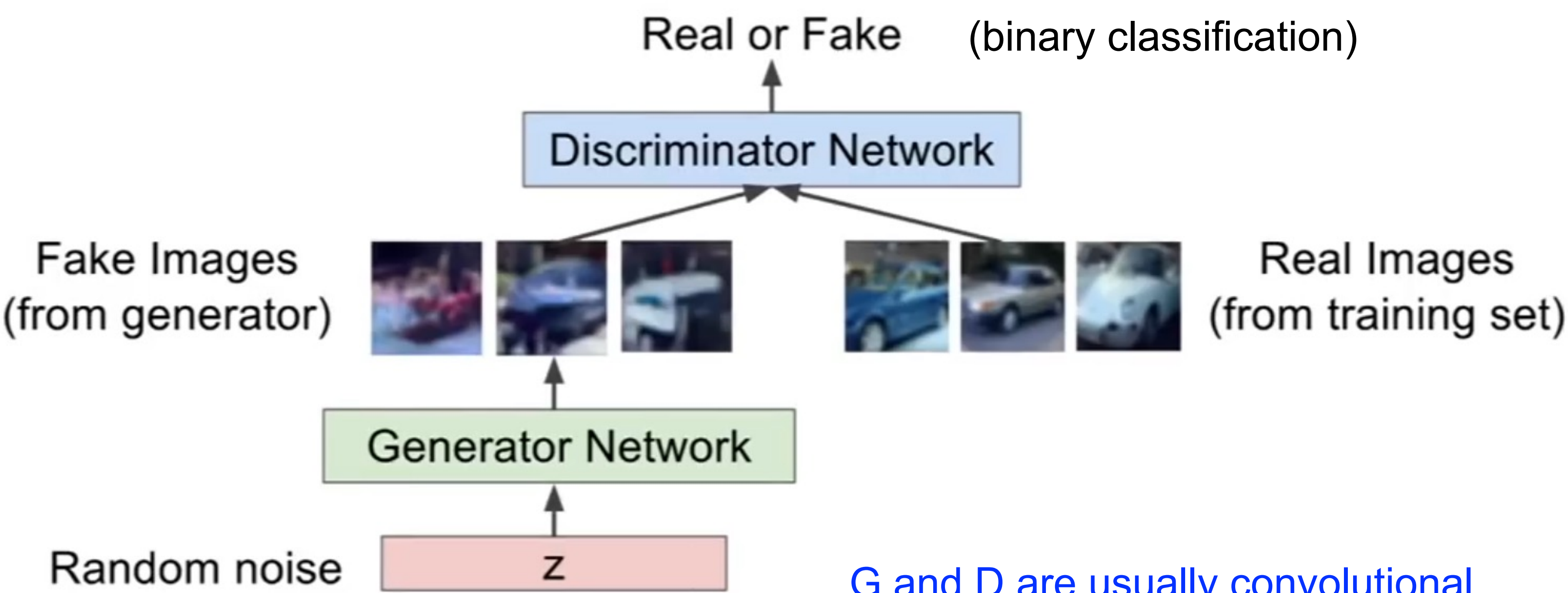
  Discriminator network: try to distinguish between real and fake images (FBI agent)

Slide Credit: Feifei Li, Justin Johnson, Serena Yeung, Ranjay Krishna, Danfei Xu

Training GAN: Two-player game

Ian Goodfellow et al., “Generative Adversarial Nets”, NIPS 2014

Generator network: try to fool the discriminator by generating real-looking images
Discriminator network: try to distinguish between real and fake images



G and D are usually convolutional neural networks

Slide Credit: Feifei Li, Justin Johnson, Serena Yeung, Ranjay Krishna, Danfei Xu

Training GAN: Two-player game

Generator network: try to fool the discriminator by generating real-looking images
Discriminator network: try to distinguish between real and fake images

Train jointly in minimax game formulation:

Minimax objective function:

$$\min_{\theta_g} \max_{\theta_d} \left[\mathbb{E}_{x \sim p_{data}} \log D_{\theta_d}(x) + \mathbb{E}_{z \sim p(z)} \log(1 - D_{\theta_d}(G_{\theta_g}(z))) \right]$$

Discriminator network parameters

Generator network parameters

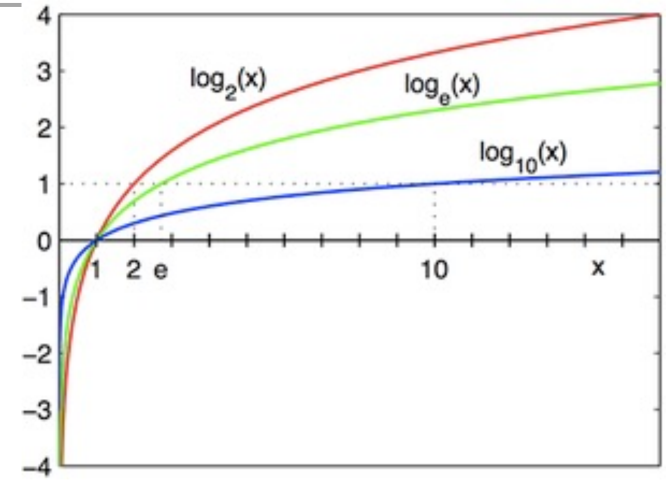
Training GAN: Two-player game

Discriminator outputs likelihood in (0,1) of real image

Minimax objective function:

$$\min_{\theta_g} \max_{\theta_d} \left[\mathbb{E}_{x \sim p_{data}} \log \underbrace{D_{\theta_d}(x)}_{\text{Discriminator output for real data x}} + \mathbb{E}_{z \sim p(z)} \log(1 - \underbrace{D_{\theta_d}(G_{\theta_g}(z))}_{\text{Discriminator output for generated fake data G(z)}}) \right]$$

Training GAN: Two-player game



What is this objective trying to do?

Let's look at the discriminator first

Discriminator outputs likelihood in (0,1) of real image

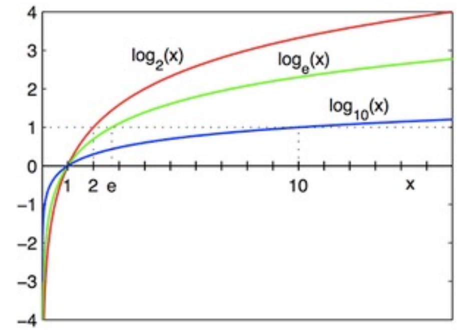
Minimax objective function:

$$\cancel{\min_{\theta_g}} \max_{\theta_d} \left[\mathbb{E}_{x \sim p_{data}} \log \underbrace{D_{\theta_d}(x)}_{\text{Discriminator output for real data x}} + \mathbb{E}_{z \sim p(z)} \log(1 - \underbrace{D_{\theta_d}(G_{\theta_g}(z))}_{\text{Discriminator output for generated fake data G(z)}}) \right]$$

- Discriminator (θ_d) wants to **maximize objective** such that D(x) is close to 1 (real) and D(G(z)) is close to 0 (fake)

Training GAN: Two-player game

What is this objective trying to do?



Then, let's look at the generator

Discriminator outputs likelihood in (0,1) of real image

Minimax objective function:

$$\min_{\theta_g} \max_{\theta_d} \left[\mathbb{E}_{x \sim p_{data}} \log \underbrace{D_{\theta_d}(x)}_{\text{Discriminator output for real data x}} + \mathbb{E}_{z \sim p(z)} \log(1 - \underbrace{D_{\theta_d}(G_{\theta_g}(z))}_{\text{Discriminator output for generated fake data G(z)}}) \right]$$

- Generator (θ_g) wants to **minimize objective** such that D(G(z)) is close to 1 (discriminator is fooled into thinking generated G(z) is real)

Training GAN: Two-player game

Minimax objective function:

$$\min_{\theta_g} \max_{\theta_d} \left[\mathbb{E}_{x\sim p_{data}} \log D_{\theta_d}(x) + \mathbb{E}_{z\sim p(z)} \log(1 - D_{\theta_d}(G_{\theta_g}(z)))\right]$$

Alternate between:

1. **Gradient ascent** on discriminator

$$\max_{\theta_d} \left[\mathbb{E}_{x\sim p_{data}} \log D_{\theta_d}(x) + \mathbb{E}_{z\sim p(z)} \log(1 - D_{\theta_d}(G_{\theta_g}(z)))\right]$$

2. **Gradient descent** on generator

$$\min_{\theta_g} \mathbb{E}_{z\sim p(z)} \log(1 - D_{\theta_d}(G_{\theta_g}(z)))$$

2. **Instead: Gradient ascent** on generator, different objective

$$\max_{\theta_g} \mathbb{E}_{z\sim p(z)} \log(D_{\theta_d}(G_{\theta_g}(z)))$$

Slide Credit: Feifei Li, Justin Johnson, Serena Yeung, Ranjay Krishna, Danfei Xu

Training GAN: Two-player game

Putting it together: GAN training algorithm

for number of training iterations **do**

for $\boxed{k \text{ steps}}$ **do**

- Sample minibatch of m noise samples $\{\boldsymbol{z}^{(1)}, \ldots, \boldsymbol{z}^{(m)}\}$ from noise prior $p_g(\boldsymbol{z})$.
- Sample minibatch of m examples $\{\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(m)}\}$ from data generating distribution $p_{\text{data}}(\boldsymbol{x})$.
- Update the discriminator by ascending its stochastic gradient:

$$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^{m} \left[\log D_{\theta_d}(x^{(i)}) + \log(1 - D_{\theta_d}(G_{\theta_g}(z^{(i)}))) \right]$$

end for

- Sample minibatch of m noise samples $\{\boldsymbol{z}^{(1)}, \ldots, \boldsymbol{z}^{(m)}\}$ from noise prior $p_g(\boldsymbol{z})$.
- Update the generator by ascending its stochastic gradient (improved objective):

$$\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^{m} \log(D_{\theta_d}(G_{\theta_g}(z^{(i)})))$$

end for

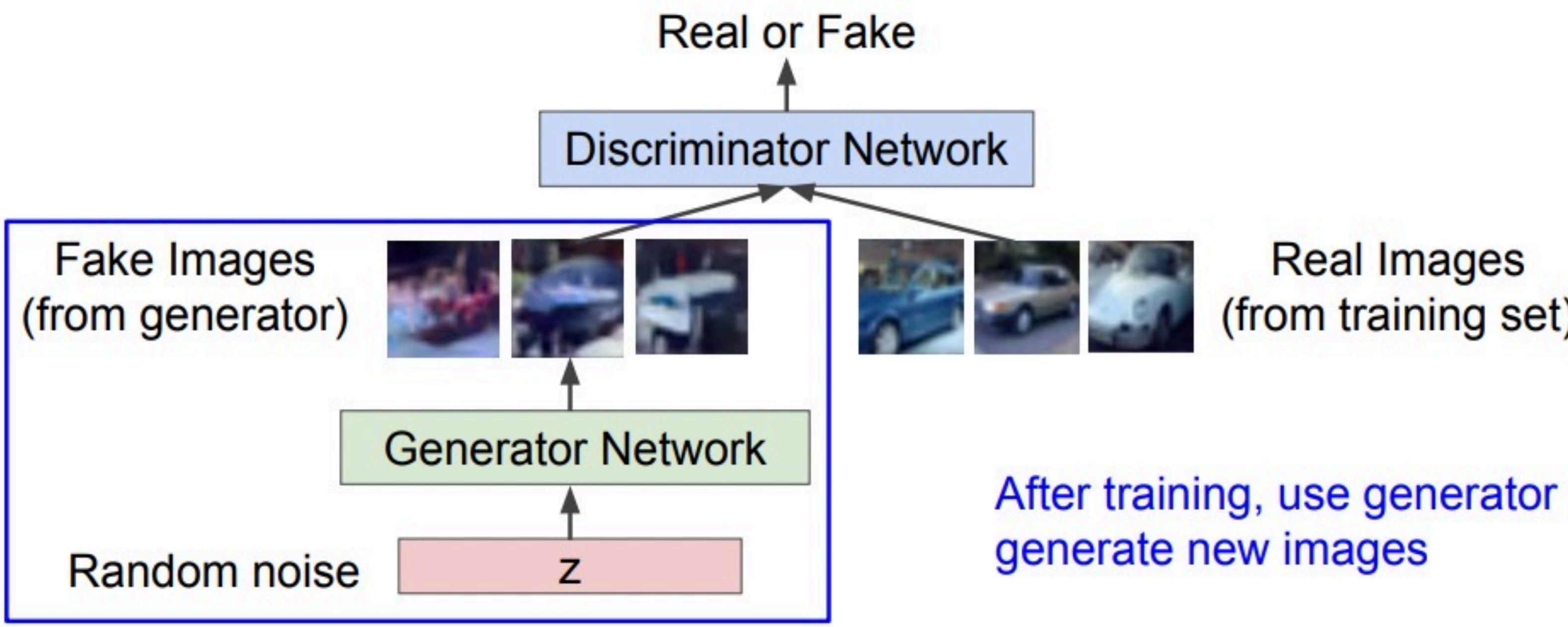
Some find k=1 more stable, others use k > 1, no best rule.

Followup work (e.g. Wasserstein GAN, BEGAN) alleviates this problem, better stability!

Arjovsky et al. "Wasserstein gan." arXiv preprint arXiv:1701.07875 (2017)
Berthelot, et al. "Began: Boundary equilibrium generative adversarial networks." arXiv preprint arXiv:1703.10717 (2017)

Training GAN: Two-player game

Generator network: try to fool the discriminator by generating real-looking images
Discriminator network: try to distinguish between real and fake images



After training, use generator network to generate new images

Slide Credit: Feifei Li, Justin Johnson, Serena Yeung, Ranjay Krishna, Danfei Xu

- https://github.com/eriklindernoren/PyTorch-GAN

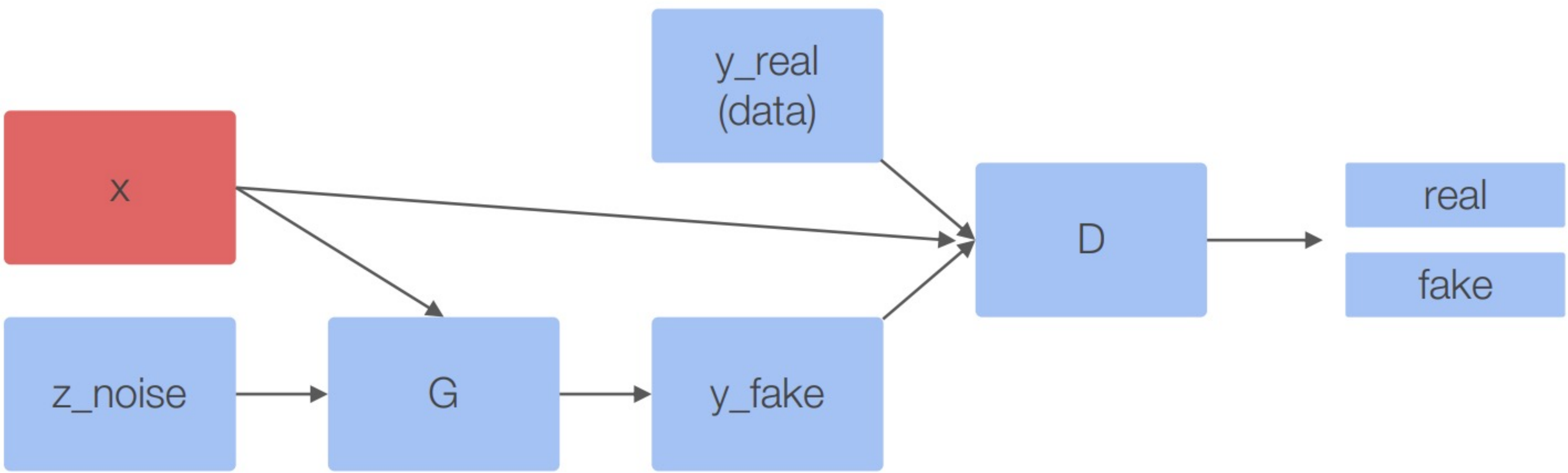
- How to Train a GAN? Tips and tricks to make GANs work:
  https://github.com/soumith/ganhacks

- DCGAN Tutorial:
  https://pytorch.org/tutorials/beginner/dcgan_faces_tutorial.html

- Video tutorial “DCGAN implementation from scratch”:
  https://www.youtube.com/watch?v=IZtv9s_Wx9I&t=0s

Condition GAN

In the Conditional GAN (CGAN), the generator learns to generate a fake sample with a specific condition or characteristics (such as a label associated with an image or more detailed tag) rather than a generic sample from unknown noise distribution.



pix2pix: Image-to-Image Translation

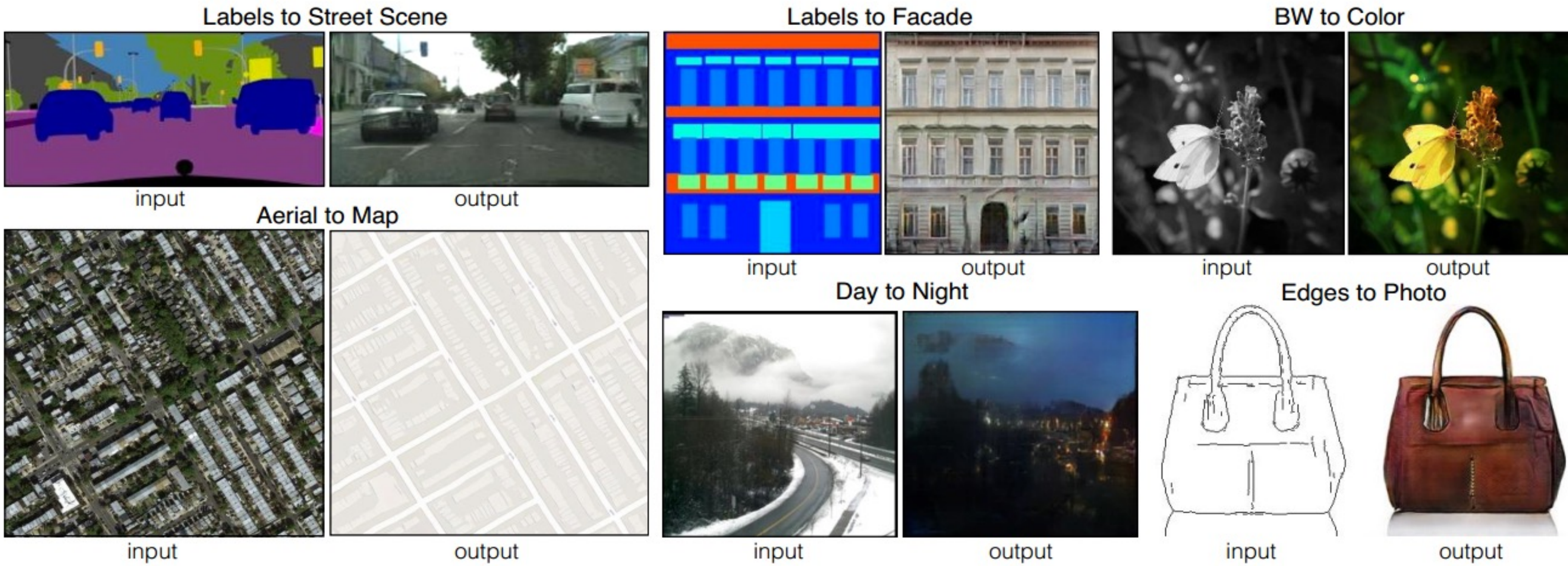
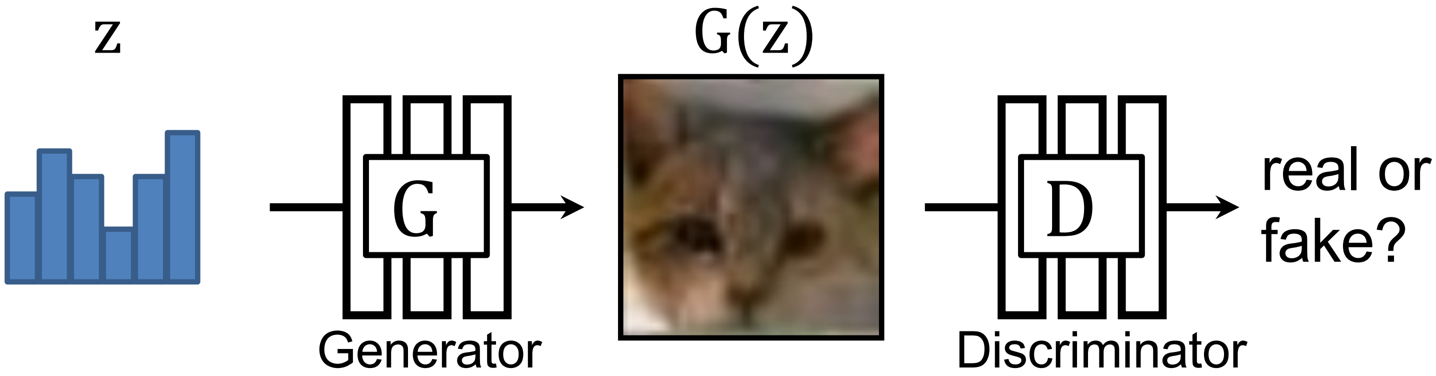
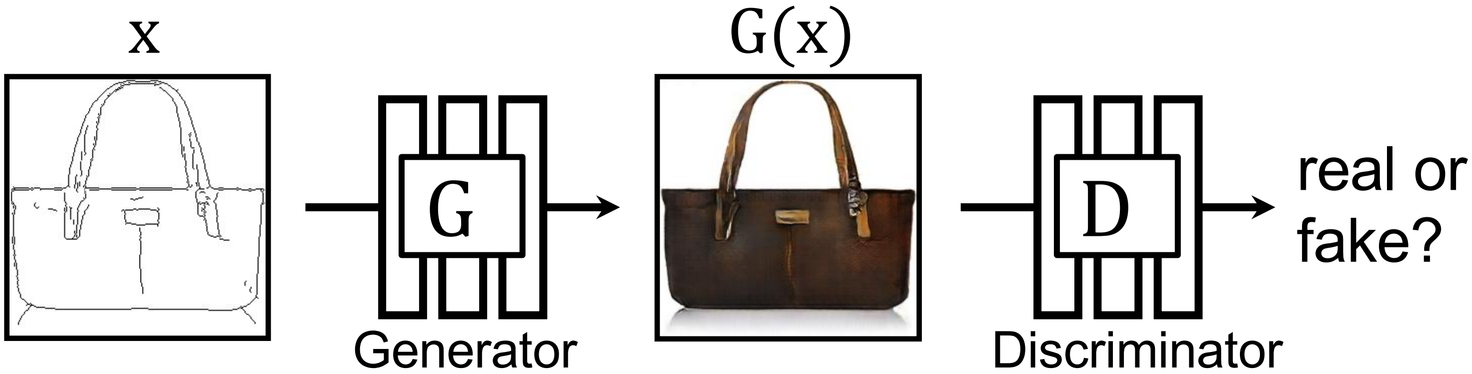


Image-to-image Translation with Conditional Adversarial Nets
Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros. CVPR 2017



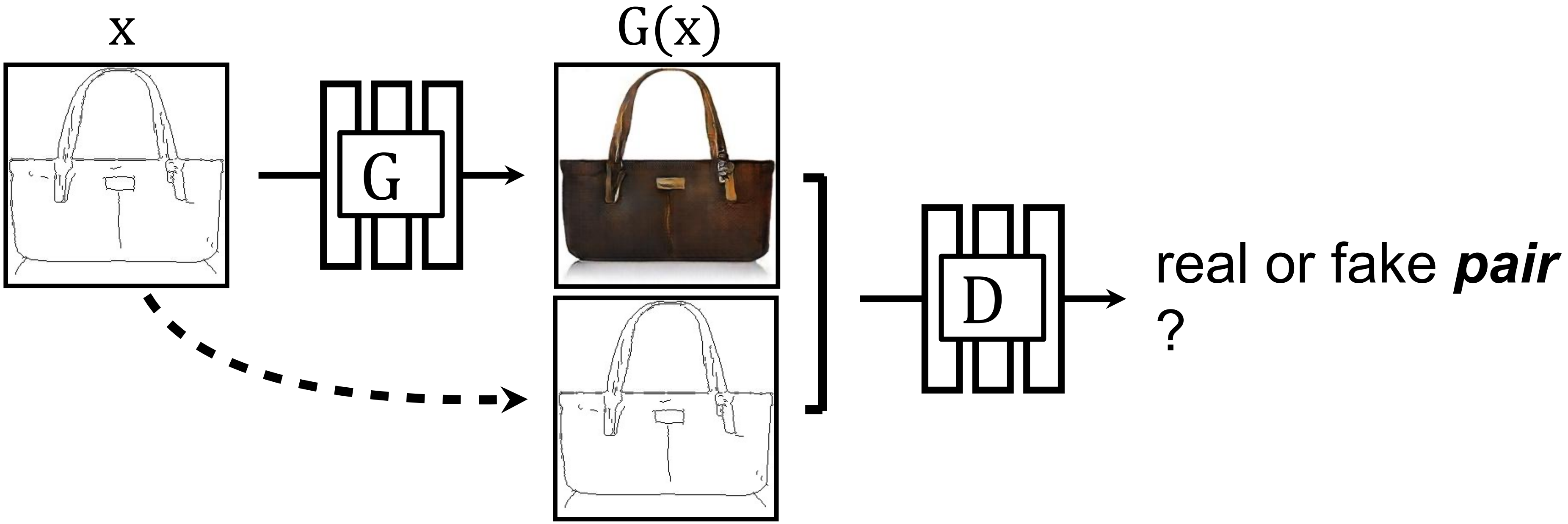
$$\min_G \max_D \mathbb{E}_{z,x}\left[\log D(x) + \log\left(1 - D\big(G(z)\big)\right)\right]$$

[Goodfellow et al. 2014]


x
G
Generator
G(x)
D
Discriminator
real or
fake?

slides credit: Jun-Yan Zhu, Phillip Isola

slides credit: Jun-Yan Zhu, Phillip Isola

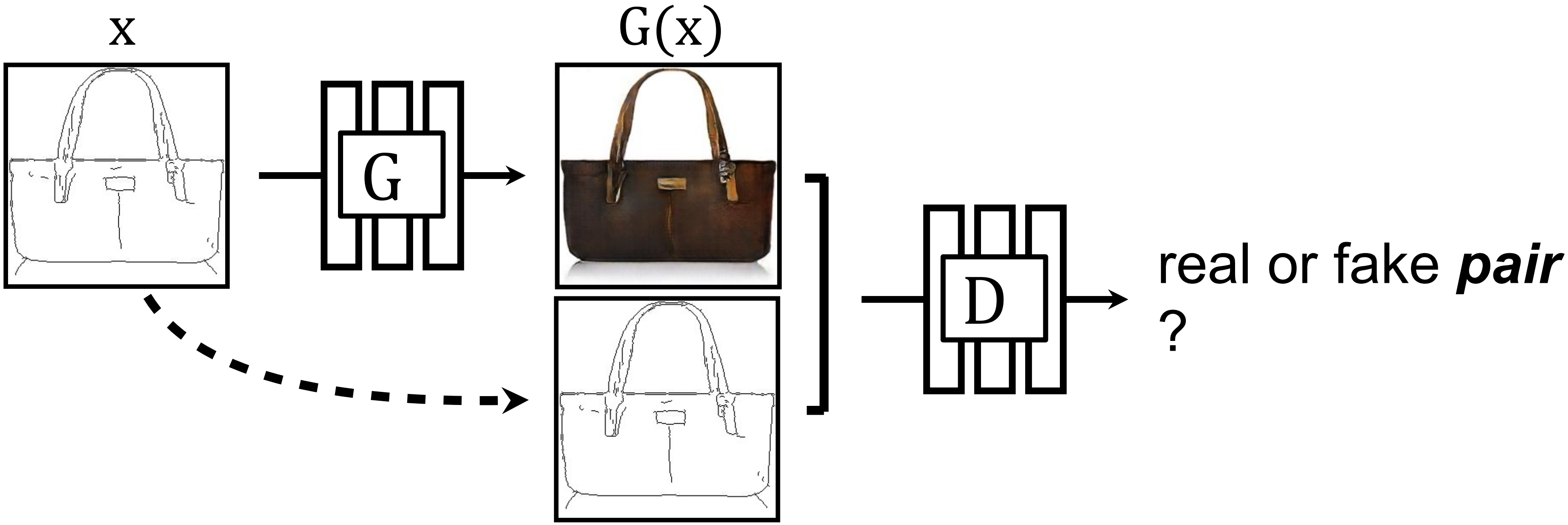


$$\min_G \max_D \mathbb{E}_{x,y}[\log D(x,y) + \log(1 - D(x, G(x)))]$$

real pair

fake pair

match joint distribution $\mathrm{p}(\mathrm{G}(\mathrm{x}), y) \sim \mathrm{p}(\mathrm{x}, y)$



The generator is tasked to not only fool the discriminator but also to be near the ground truth output in an L2 sense. We also explore this option, using L1 distance rather than L2 as L1 encourages less blurring:

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{x,y,z}[\|y - G(x,z)\|_1].$$

Our final objective is

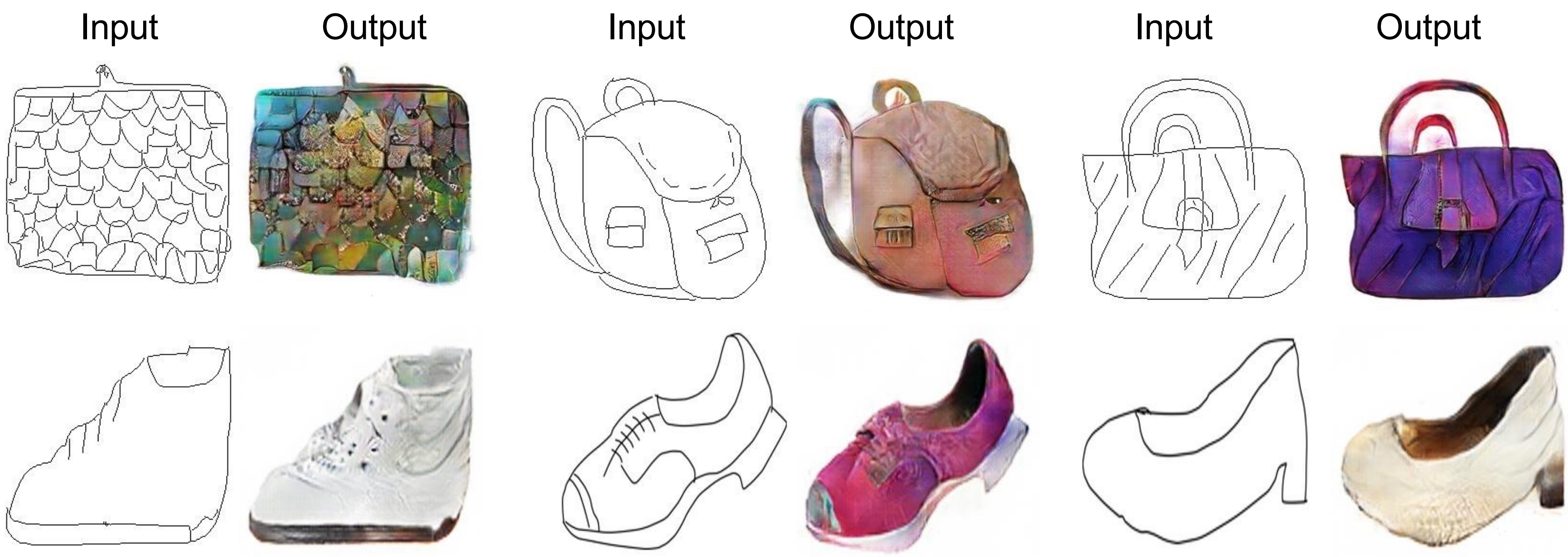
$$G^* = \arg\min_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda\mathcal{L}_{L1}(G).$$

Edges → Images

Input Output Input Output Input Output

Edges from [Xie & Tu, 2015]

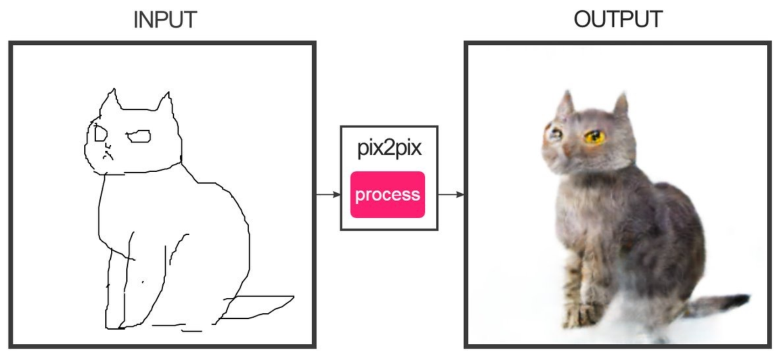
Sketches → Images



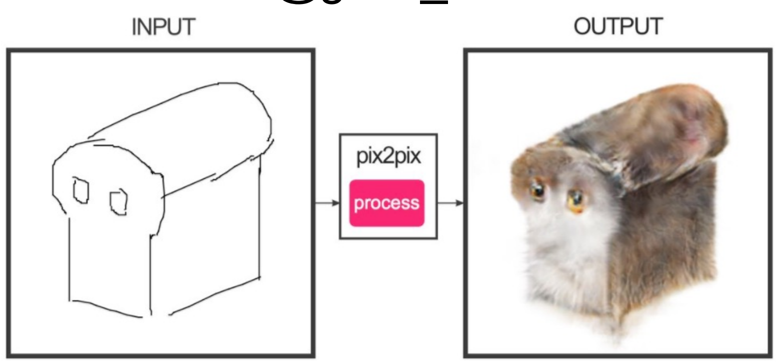
Trained on Edges → Images

Data from [Eitz, Hays, Alexa, 2012]

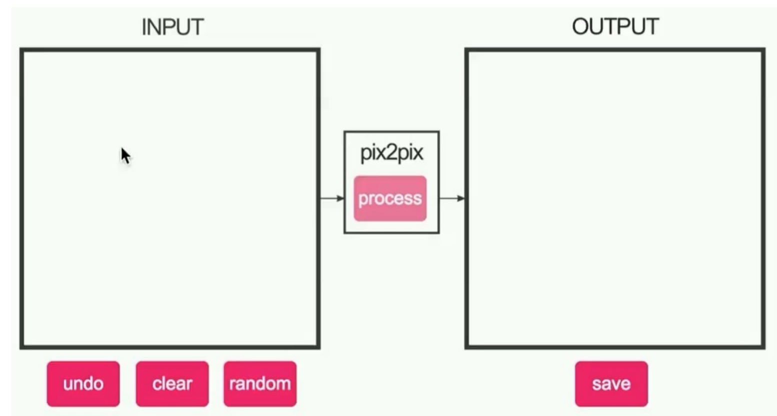
#edges2cats [Christopher Hesse]



@gods_tail



Ivy Tasi
@ivymyt



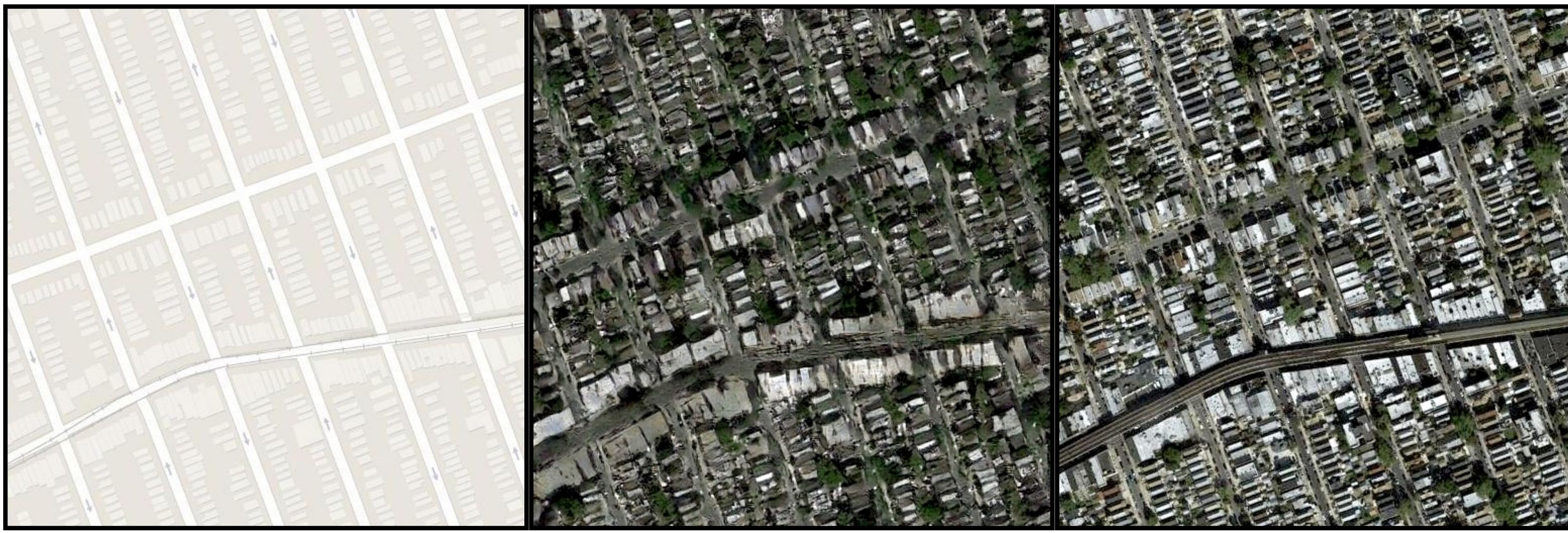
@matthematician

Vitaly Vidmirov @vvid

https://affinelayer.com/pixsrv/

slides credit: Jun-Yan Zhu, Phillip Isola

Input Output Groundtruth

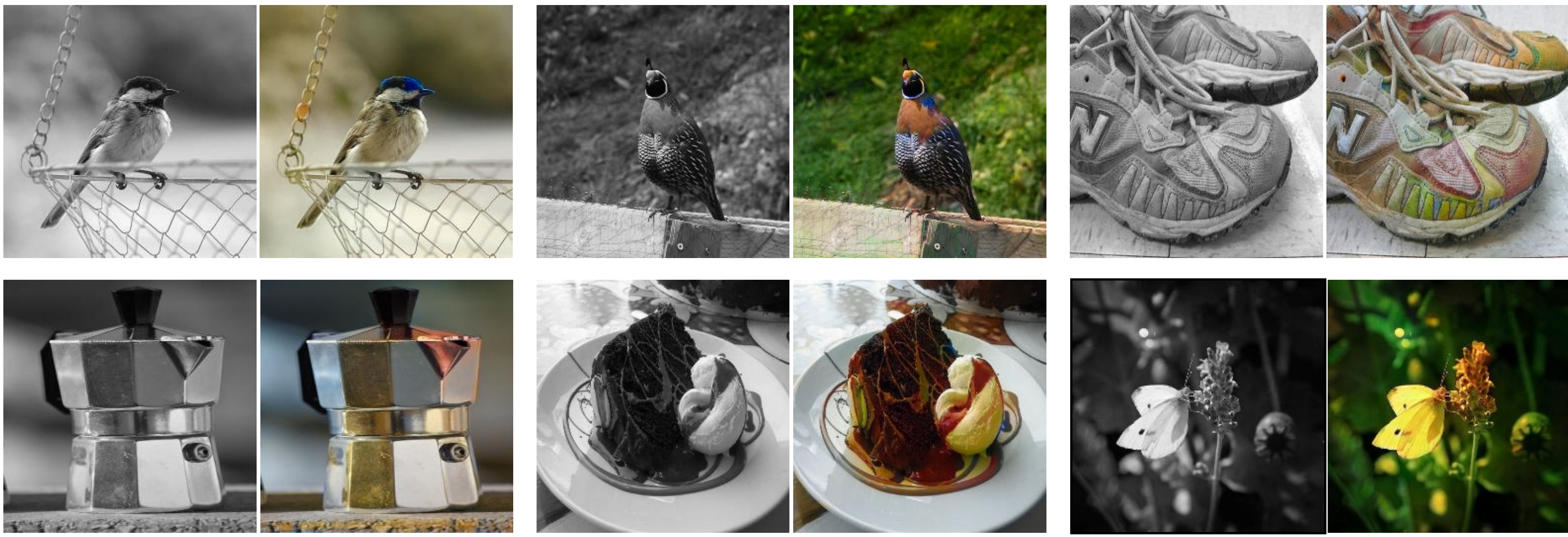


Data from
[maps.google.com]

slides credit: Jun-Yan Zhu, Phillip Isola

BW → Color

Input Output Input Output Input Output



Data from [Russakovsky et al. 2015]

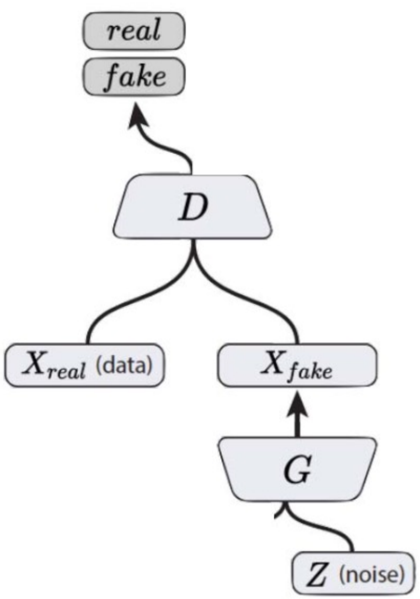
slides credit: Jun-Yan Zhu, Phillip Isola

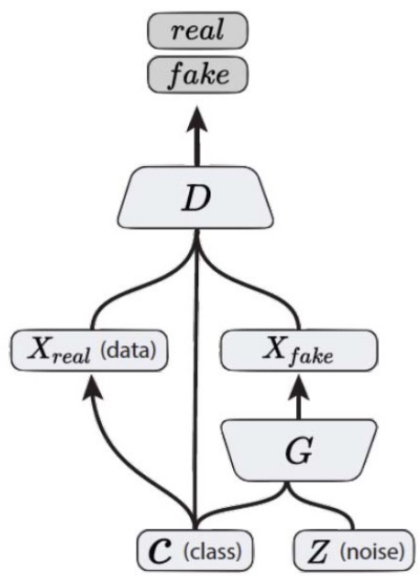
Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks

a.k.a.
CycleGAN

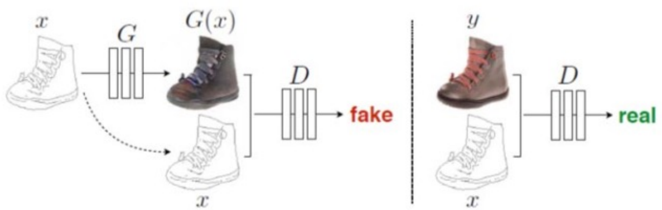
Zhu, Jun-Yan, et al. "Unpaired image-to-image translation using cycle-consistent adversarial networks." *ICCV,* 2017.

Previously, Pix2Pix

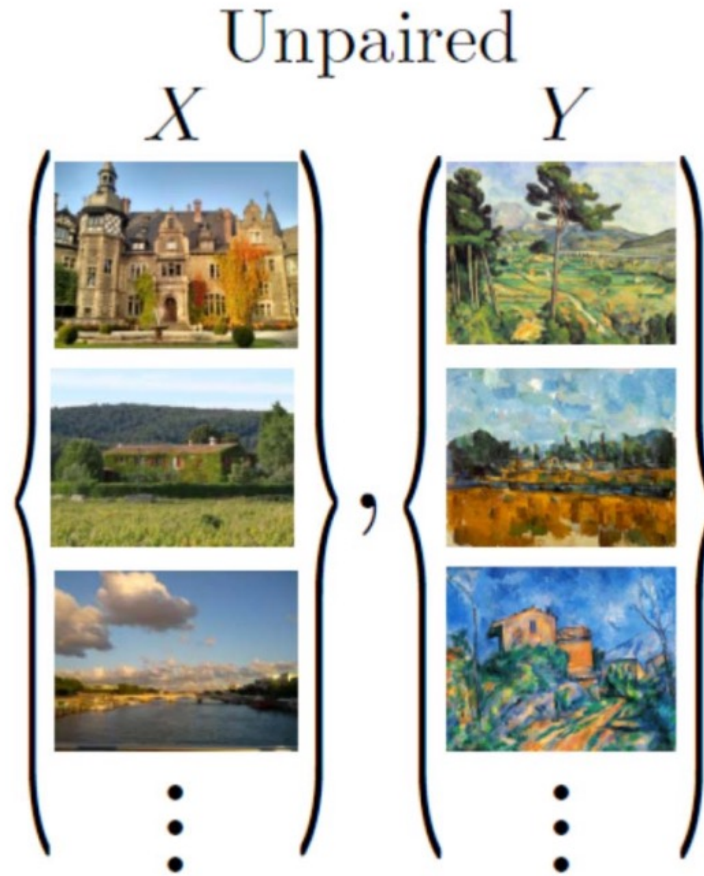




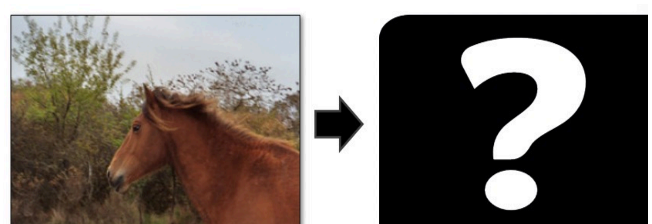
Conditional GAN



The challenge



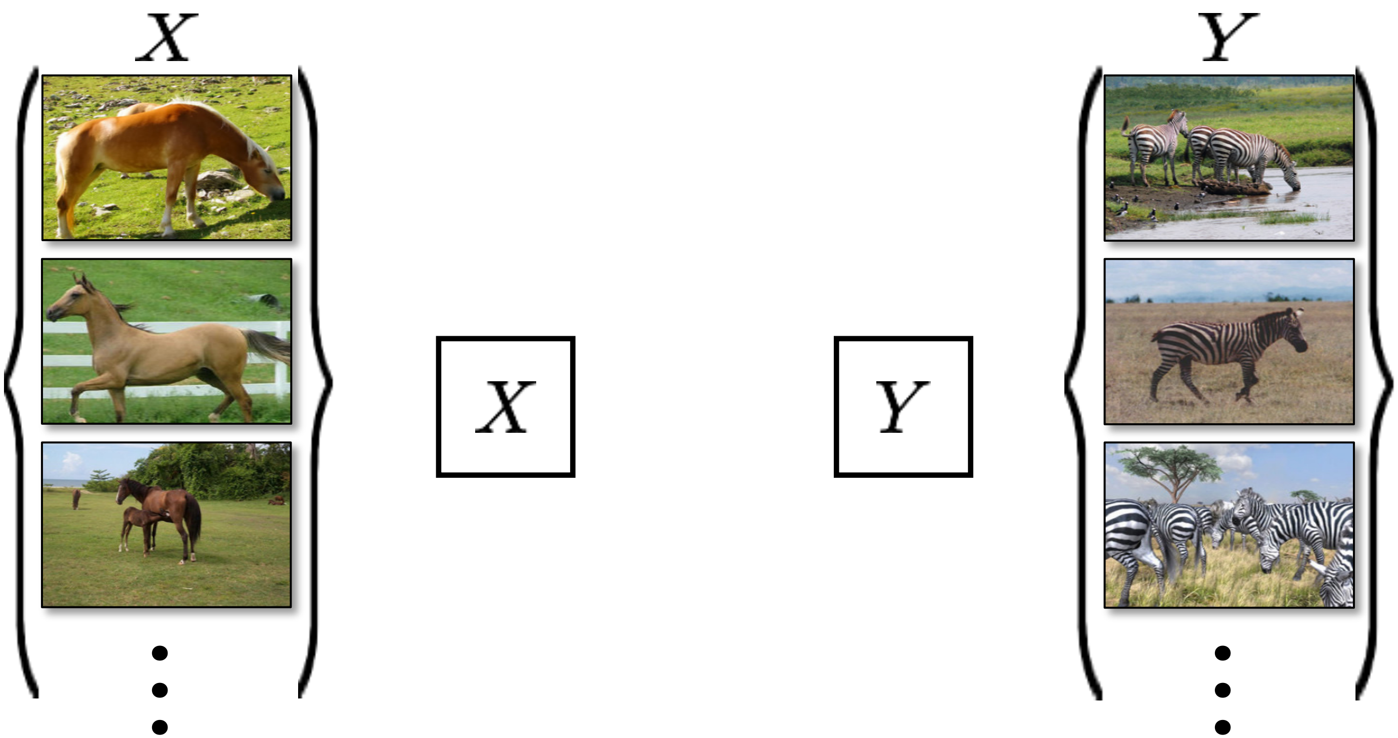
- Expensive to collect pairs.
- Impossible in many scenarios.



Horse ↔ zebra: how to get zebras?

slides credit: Jun-Yan Zhu

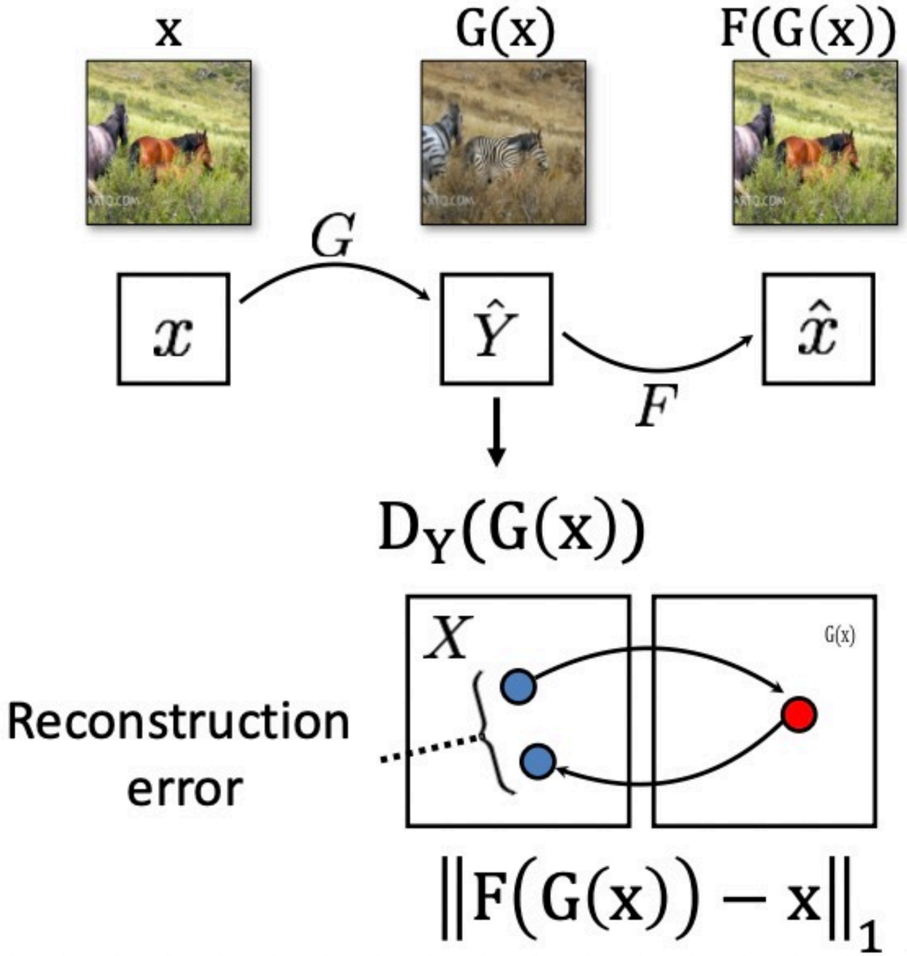
Cycle-Consistent Adversarial Networks



slides credit: Jun-Yan Zhu

[Zhu*, Park*, Isola, and Efros, ICCV 2017]

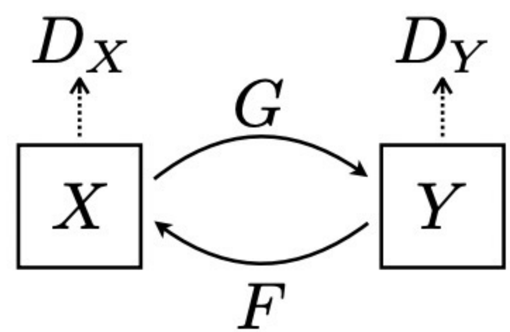
Cycle-Consistent Adversarial Networks



$$G : X \rightarrow Y$$

$$F : Y \rightarrow X$$

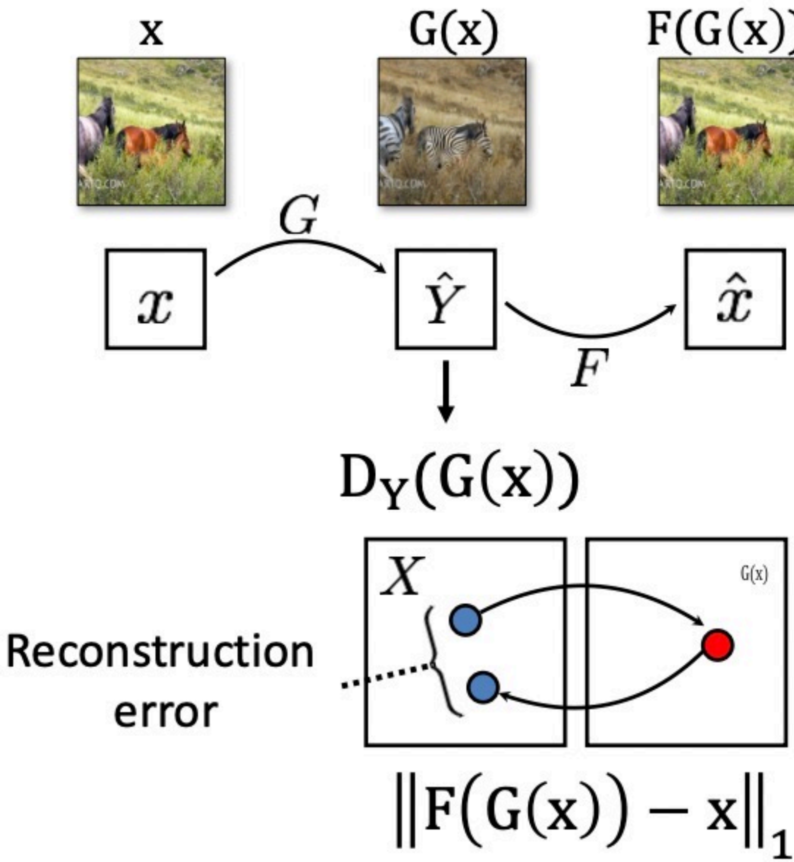
D_Y encourages G to translate X into outputs indistinguishable from domain Y , and vice versa for D_X and F.

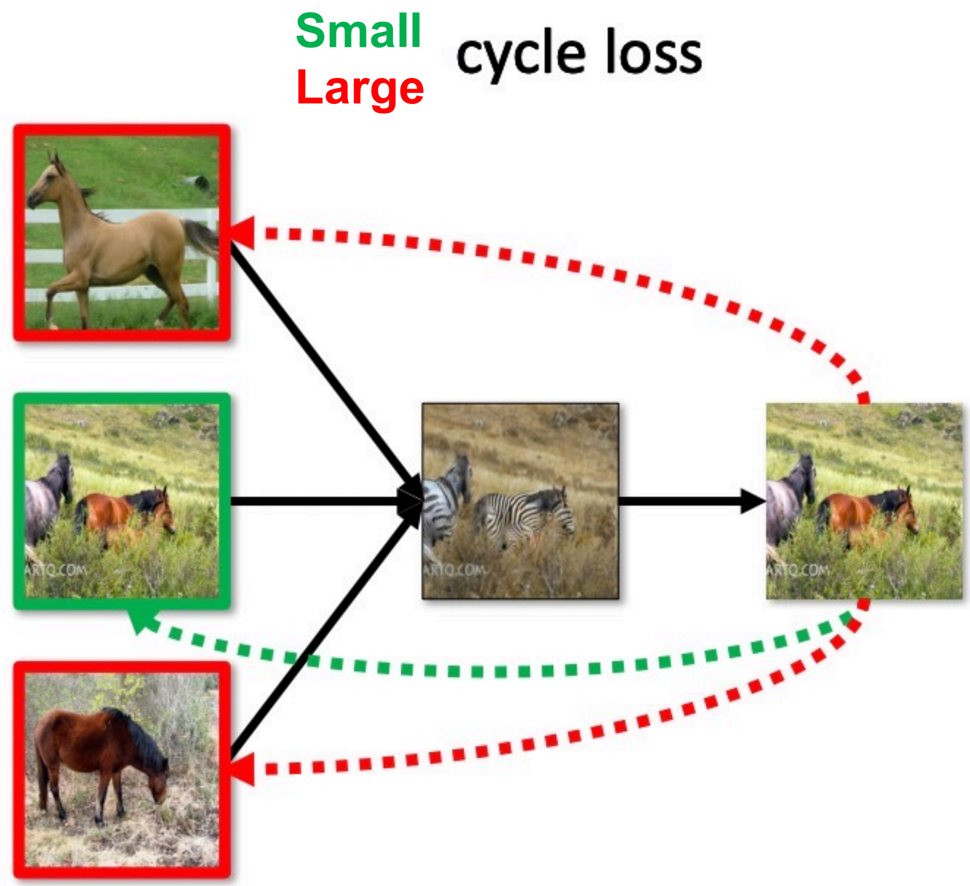


slides credit: Jun-Yan Zhu

[Zhu*, Park*, Isola, and Efros, ICCV 2017]

Cycle-Consistent Adversarial Networks

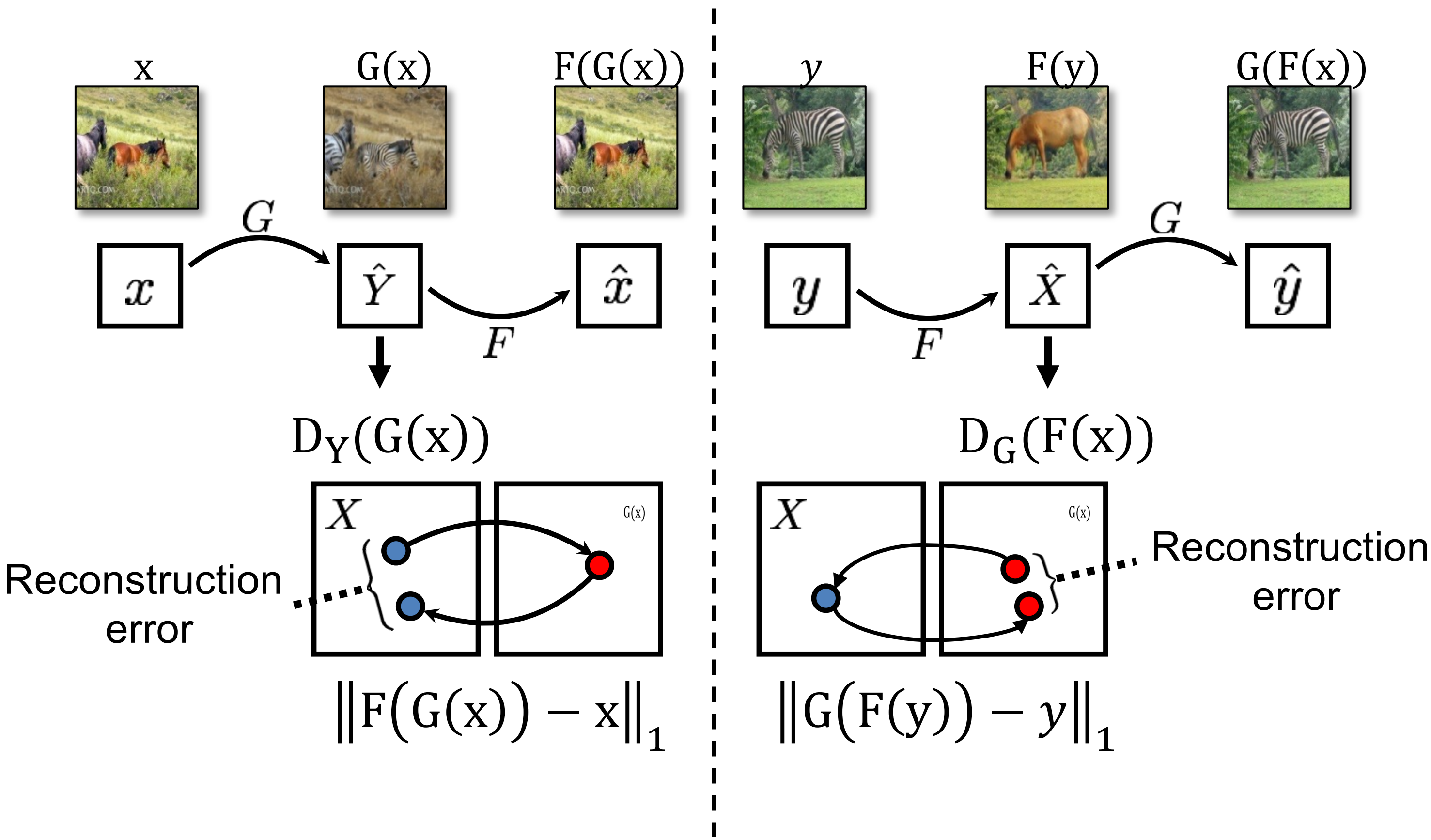




slides credit: Jun-Yan Zhu

[Zhu*, Park*, Isola, and Efros, ICCV 2017]

Cycle Consistency Loss



slides credit: Jun-Yan Zhu

[Zhu*, Park*, Isola, and Efros, ICCV 2017]

UCF Center for Research in Computer Vision

Monet's paintings → photos

slides credit: Jun-Yan Zhu

Monet's paintings → photos

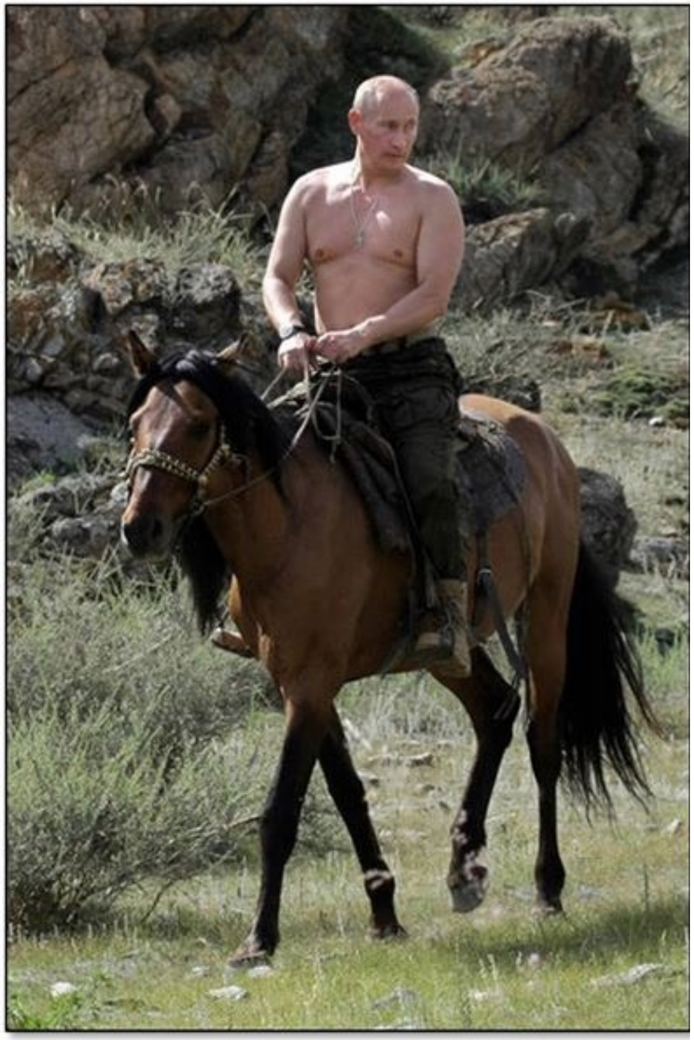
slides credit: Jun-Yan Zhu

UCF CENTER FOR RESEARCH IN COMPUTER VISION

slides credit: Jun-Yan Zhu

slides credit: Jun-Yan Zhu

Failure case



slides credit: Jun-Yan Zhu

Failure case

slides credit: Jun-Yan Zhu

UCF Center for Research in Computer Vision

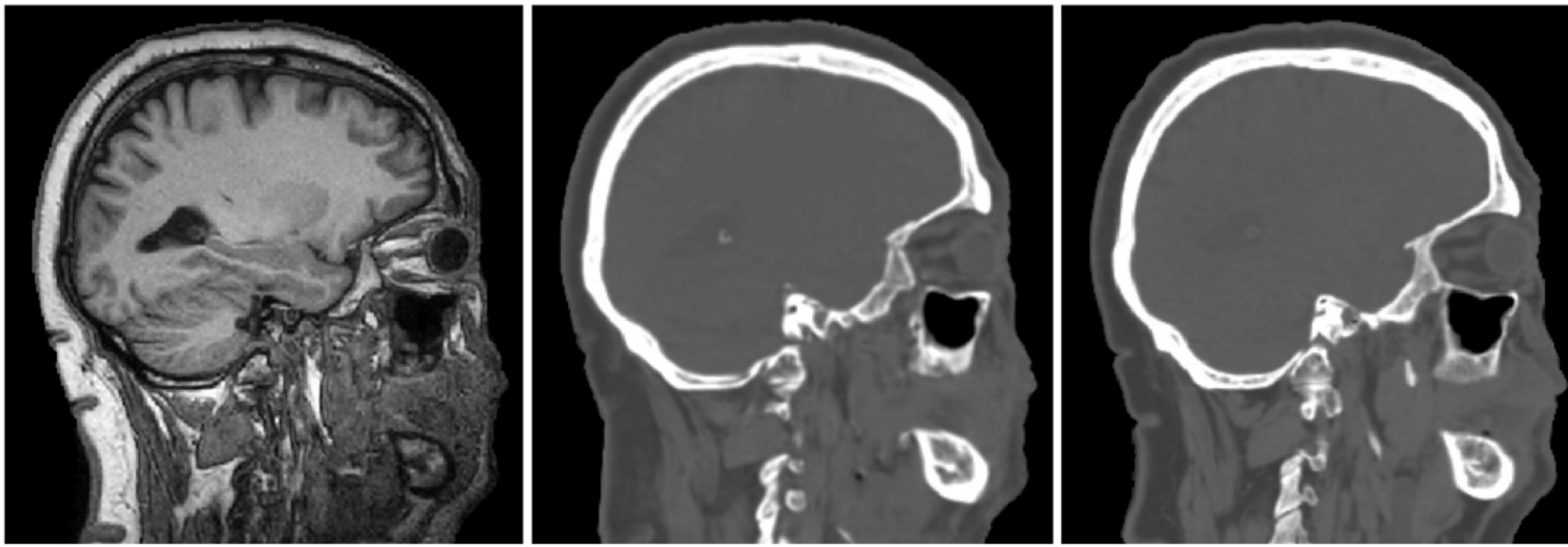
CycleGAN Demo

Real zebra image

https://junyanz.github.io/CycleGAN/

Medical Imaging Applications

MR → CT [Wolterink et al] arxiv: 1708.01155



Input MR Generated CT Ground truth CT

- **MRI reconstruction** [Quan et al.] arxiv:1709.00753
- **Cardiac MR images from CT** [Chartsias et al. 2017]

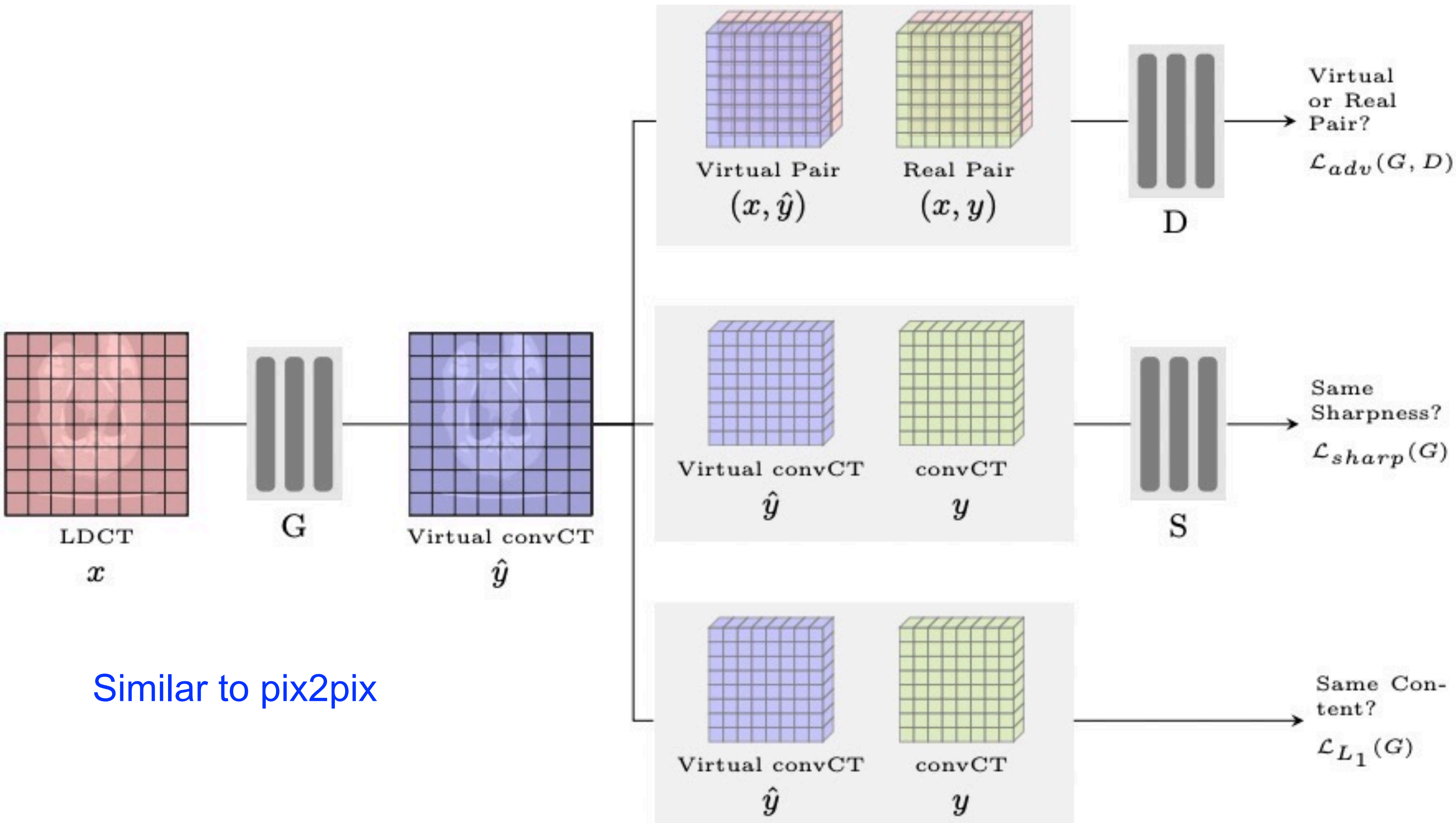
Example 1: Low dose CT denoising

Input

Denoised

Yi, Xin, and Paul Babyn. "Sharpness-aware low-dose CT denoising using conditional generative adversarial network." Journal of digital imaging 31.5 (2018): 655-669.

Example 1: Low dose CT denoising

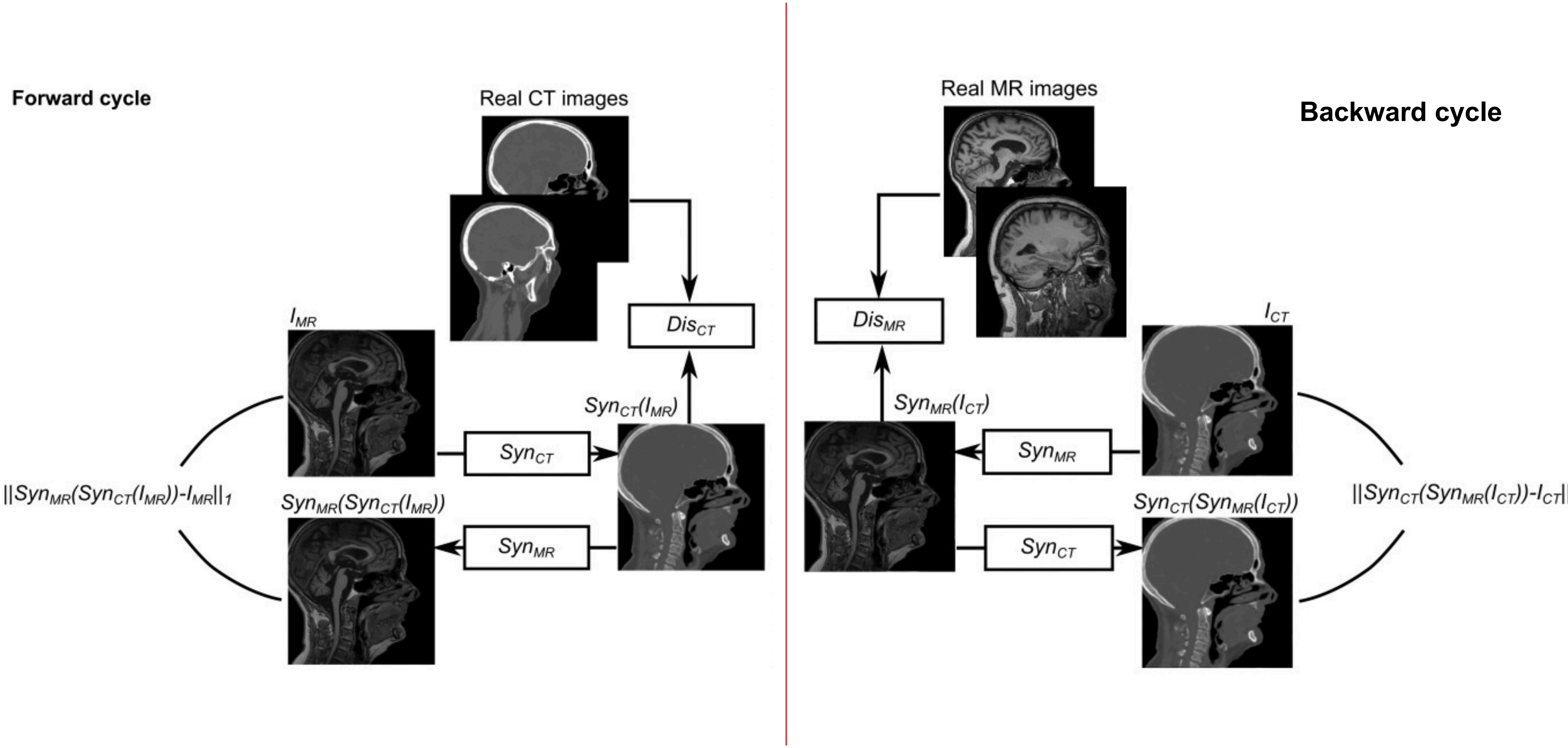


Similar to pix2pix

Yi, Xin, and Paul Babyn. "Sharpness-aware low-dose CT denoising using conditional generative adversarial network." Journal of digital imaging 31.5 (2018): 655-669.

Example 2: MR to CT Synthesis

MR to CT Synthesis using CycleGAN



Wolterink, Jelmer M., et al. "Deep MR to CT synthesis using unpaired data." *International workshop on simulation and synthesis in medical imaging*. Springer, Cham, 2017.

Example 2: MR to CT Synthesis

MR to CT Synthesis using CycleGAN

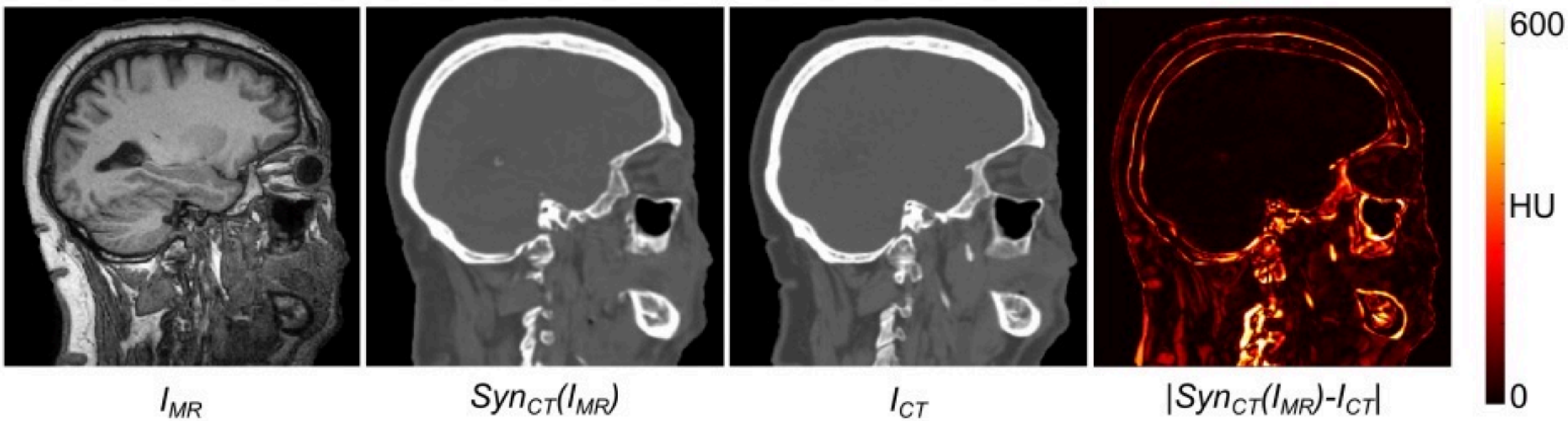
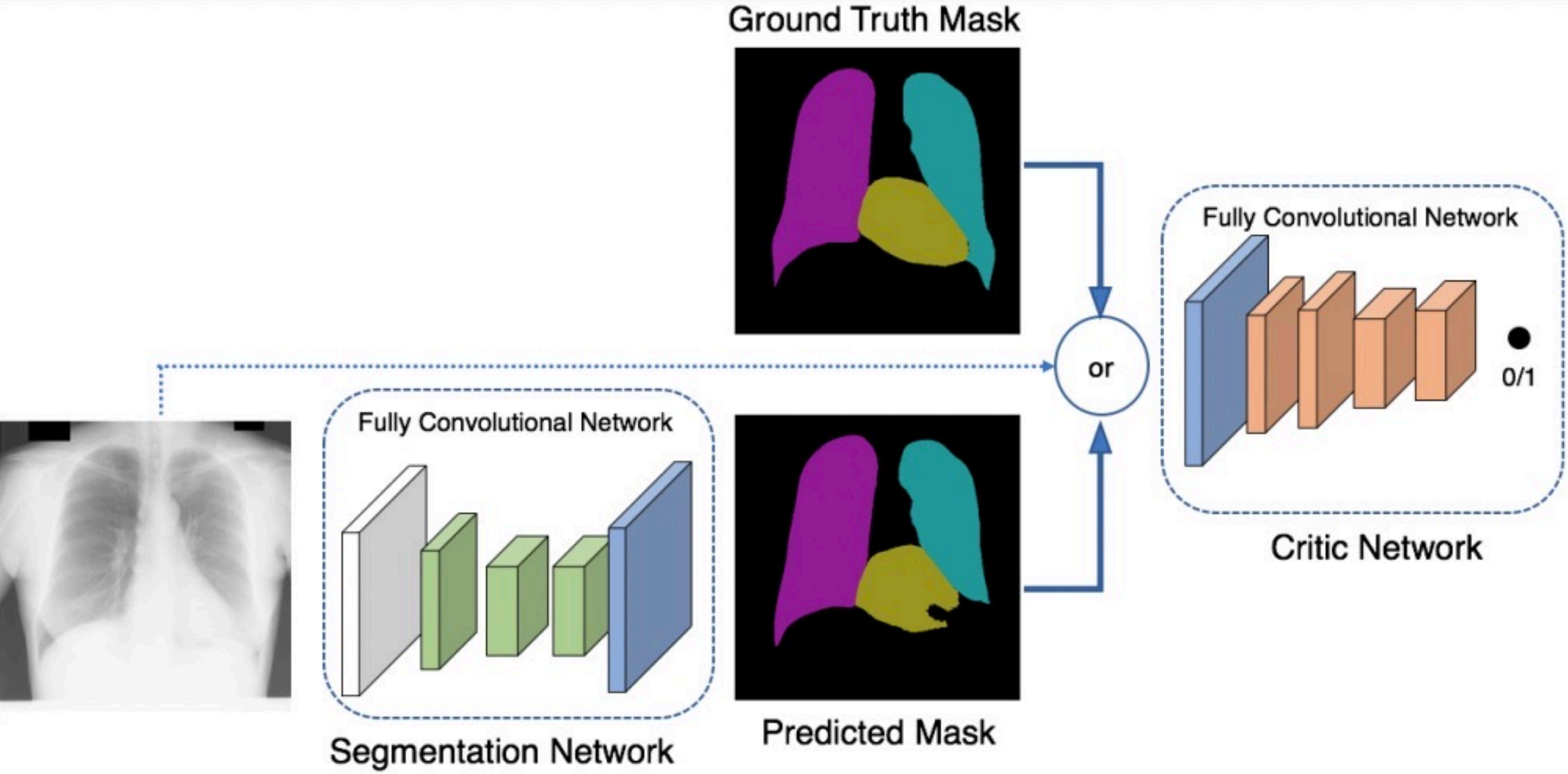


Fig. 4: *From left to right* Input MR image, synthesized CT image, reference real CT image, and absolute error between real and synthesized CT image.

Wolterink, Jelmer M., et al. "Deep MR to CT synthesis using unpaired data." *International workshop on simulation and synthesis in medical imaging*. Springer, Cham, 2017.

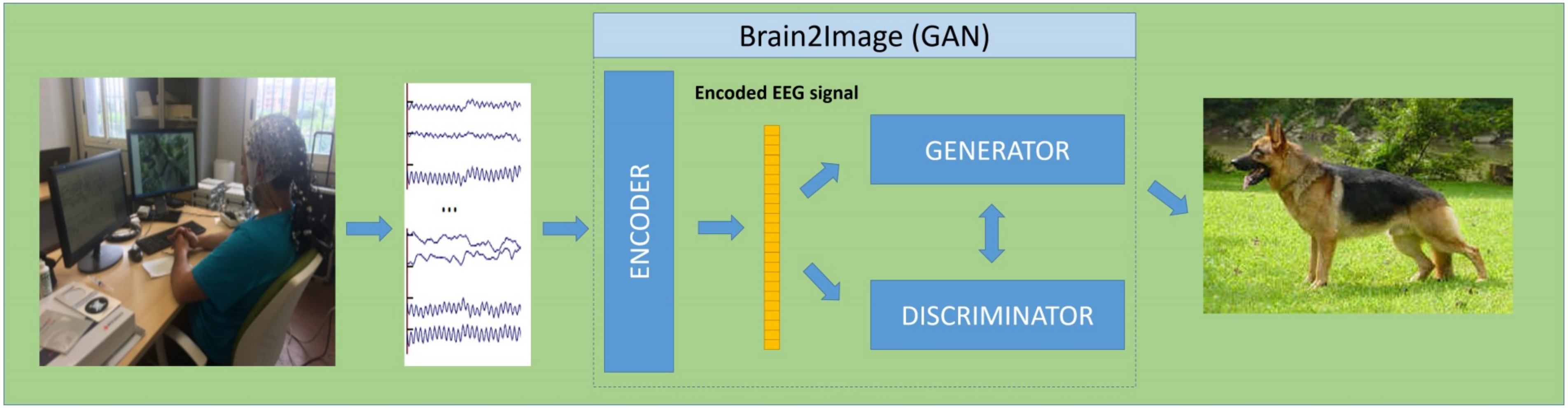
Example 3: Organ Segmentation in Chest X-rays



Dai, Wei, et al. "Scan: Structure correcting adversarial network for organ segmentation in chest x-rays." *Deep learning in medical image analysis and multimodal learning for clinical decision support*. Springer, Cham, 2018. 263-273.

Example 4: Brain Signals into Images

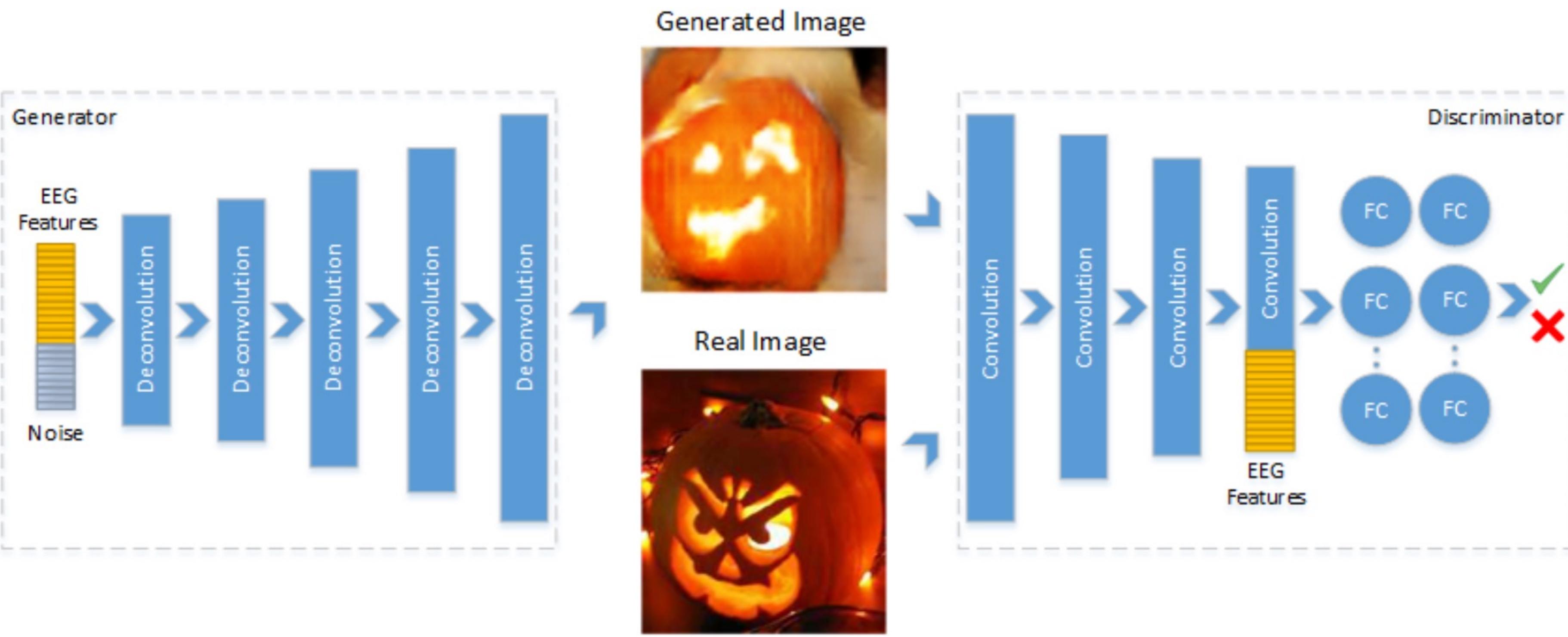
Brain computer interface



EEG-based image generation – “reading the mind”

I kavasidis, S. Palazzo, C. Spampinato, D. Giordano, M. Shah, Brain2Image: Converting Brain Signals into Images, ACM Multimedia, 2017.

Brain Signals into Images



I kavasidis, S. Palazzo, C. Spampinato, D. Giordano, M. Shah, Brain2Image: Converting Brain Signals into Images, ACM Multimedia, 2017.

(a) Airliner

(c) Panda

I kavasidis, S. Palazzo, C. Spampinato, D. Giordano, M. Shah, Brain2Image: Converting Brain Signals into Images, ACM Multimedia, 2017.

"DeepFake"

- A person in an existing image or video is replaced with someone else's likeness, usually by GAN (sometimes autoencoders)

https://github.com/deepfakes/faceswap

Credit: Zhangyang Wang

References

- Yi, Xin, Ekta Walia, and Paul Babyn. "Generative adversarial network in medical imaging: A review." *Medical image analysis* 58 (2019): 101552.
- Kazeminia, Salome, et al. "GANs for medical image analysis." *Artificial Intelligence in Medicine* 109 (2020): 101938.
- Wolterink, Jelmer M., et al. "Generative adversarial networks and adversarial methods in biomedical image analysis." *arXiv preprint arXiv:1810.10352* (2018).
- Wang, Zhengwei, Qi She, and Tomas E. Ward. "Generative adversarial networks in computer vision: A survey and taxonomy." *ACM Computing Surveys (CSUR)* 54.2 (2021): 1-38.
- https://github.com/xinario/awesome-gan-for-medical-imaging
- https://github.com/eriklindernoren/PyTorch-GAN
- https://github.com/nashory/gans-awesome-applications
- How to Train a GAN? Tips and tricks to make GANs work: https://github.com/soumith/ganhacks
- http://cs231n.stanford.edu/slides/2021/lecture_12.pdf
- GAN collections: https://github.com/wiseodd/generative-models
- https://github.com/hindupuravinash/the-gan-zoo

Thank you!

Question?