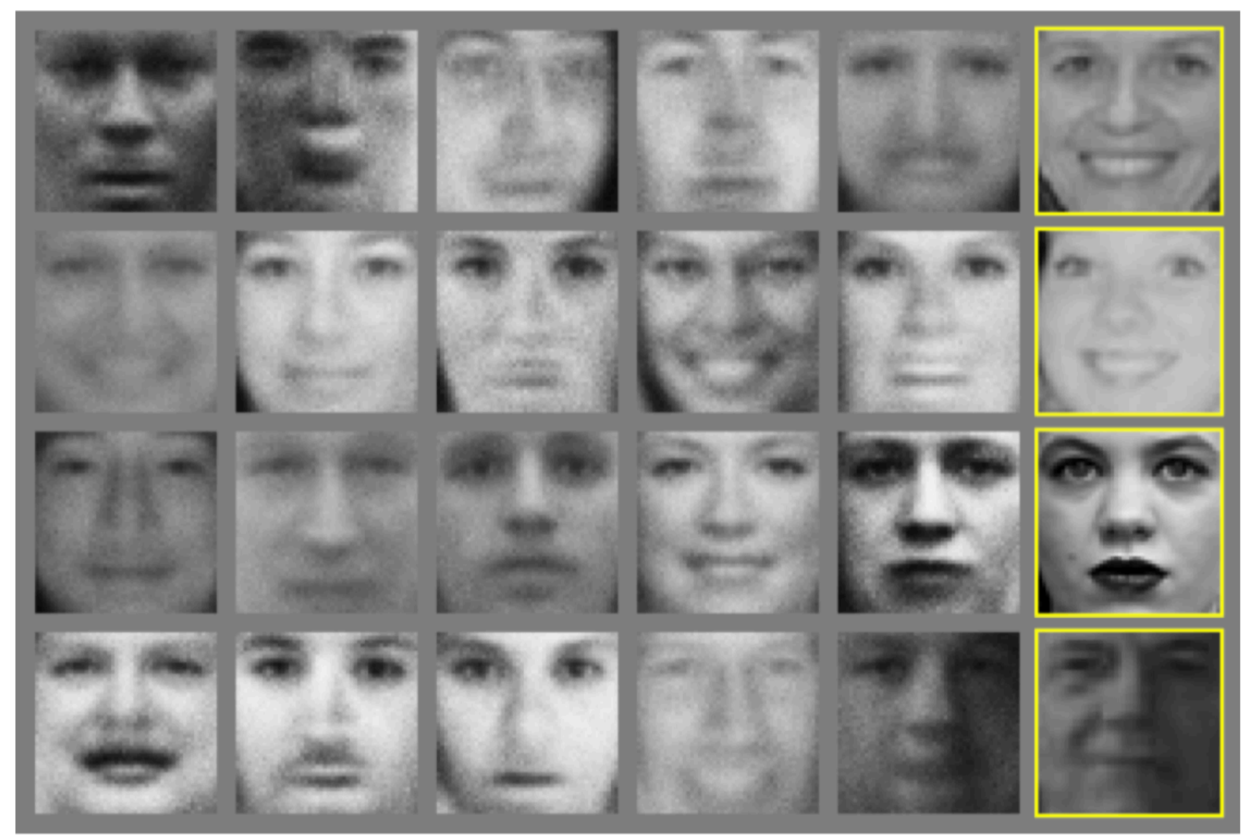
Generative Adversarial Networks

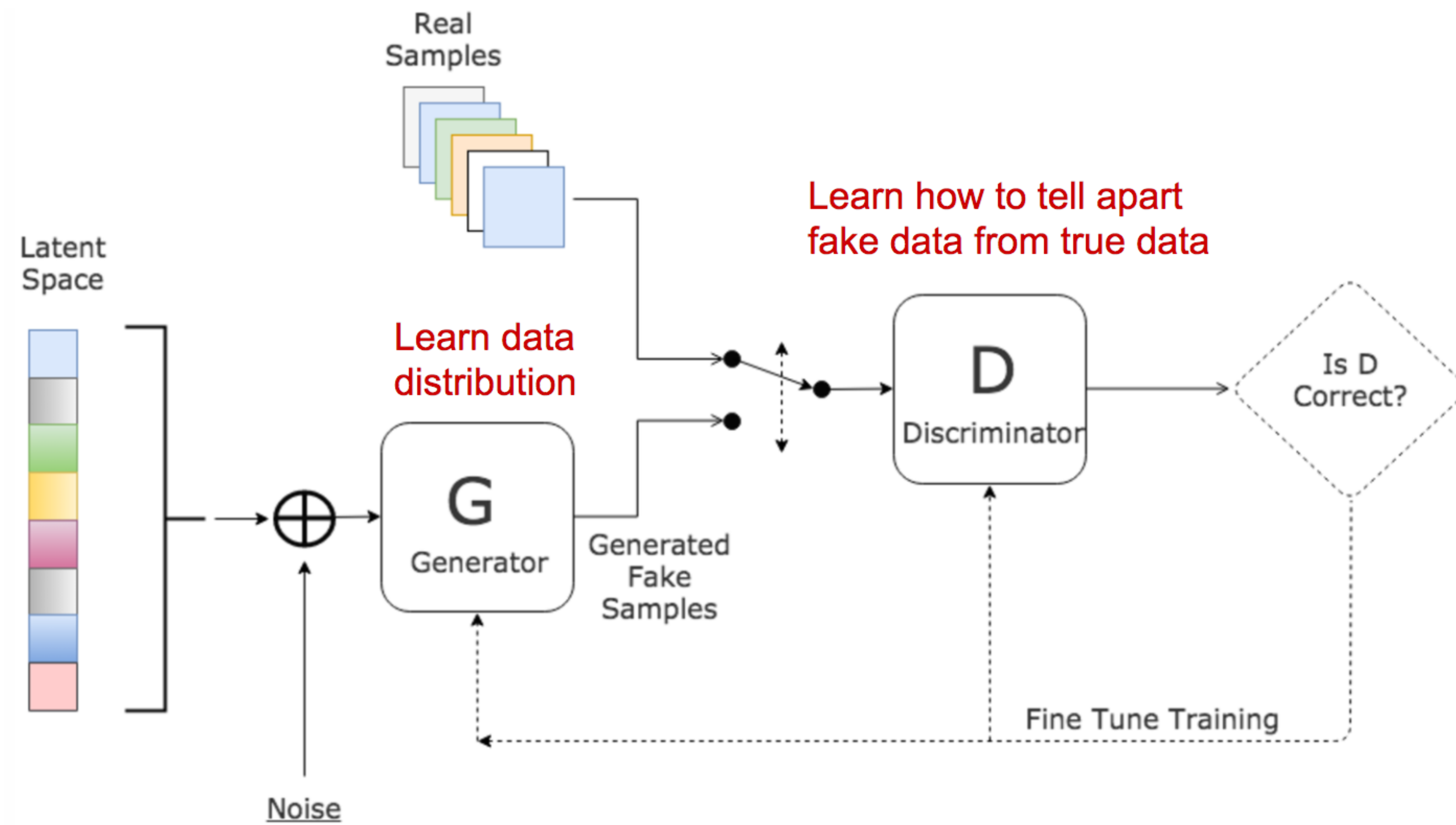
Generative Model

Generative model is able produce (generate) new content, according to the given data.



These are samples generated by Generative Adversarial Networks after training on two datasets: MNIST and TFD.
Source: "Generative Adversarial Nets" paper

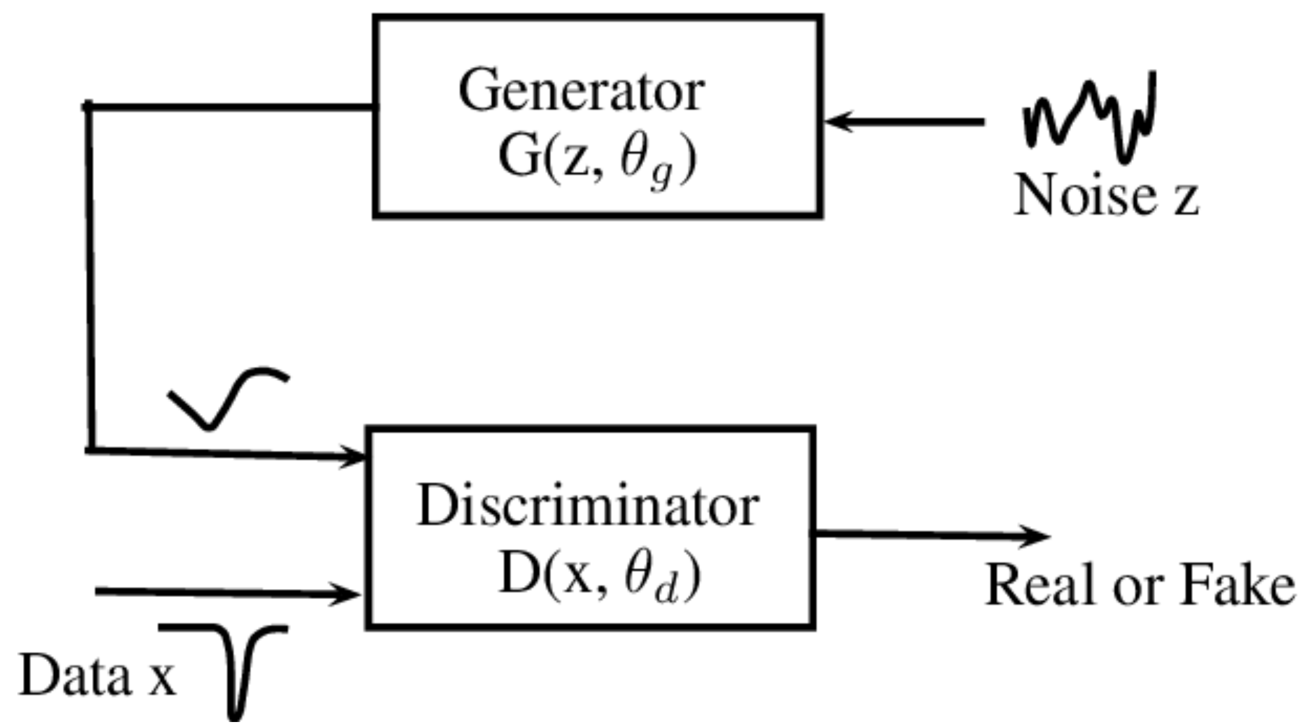
Generating Process



This diagram shows the flow of a typical GAN.
Source: https://towardsdatascience.com/adversarial-training-creating-realistic-fakes-with-machine-learning-c570881d0e81

Generating Process

- The generator takes simple random variables as inputs and generate new data.
- The discriminator takes “true” and “generated” data and try to discriminate them, building a classifier.
- The goal of the generator is to fool the discriminator (increase the classification error by mixing up as much as possible generated data with true data) and the goal of the discriminator is to distinguish between true and generated data.



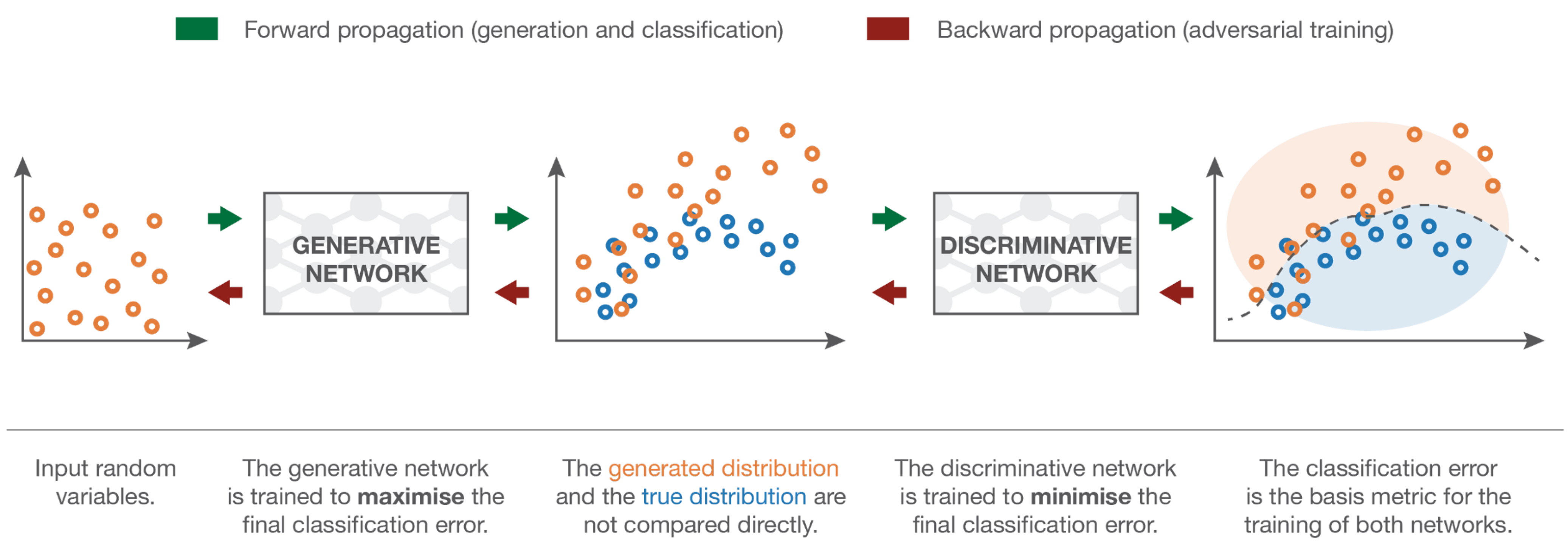
General block diagram of generative adversarial networks (GANs). Source: Generative Adversarial Network-Based Glottal Waveform Model for Statistical Parametric Speech Synthesis

Learning

- **Both models are not doing a good work at first.**
- **As generator progresses and learns to generate more realistic samples, the discriminator evolves too and can detect fakes better.**

More theory:

- **The discriminator maximizes the probability of assigning the correct label to both training examples and images generated by the generator.** *(It becomes better at differentiating between fakes and real samples).*
- **The generator minimizes the probability that the discriminator can predict that what it generates is fake.** *(The generator becomes better at creating fakes).*

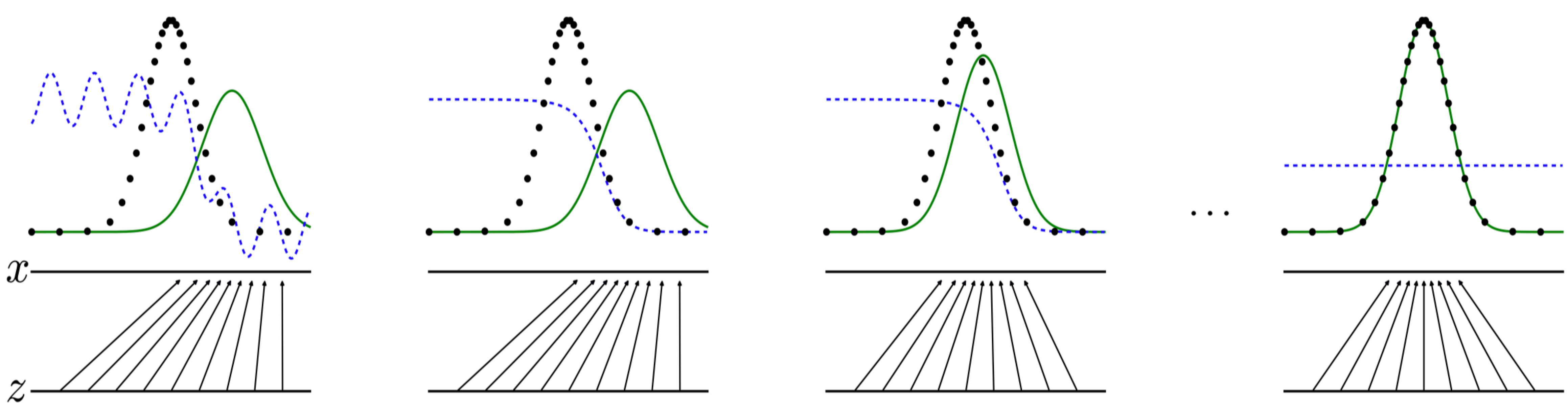


This diagram shows the flow of a typical GAN.
Source: https://towardsdatascience.com/adversarial-training-creating-realistic-fakes-with-machine-learning-c570881d0e81

Learning

- The generator is defined by G(z), which converts some noise z we input into some data, like images.
- The discriminator is defined by D(x), which outputs the probability that the input x came from the real dataset or not.

$$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))].$$

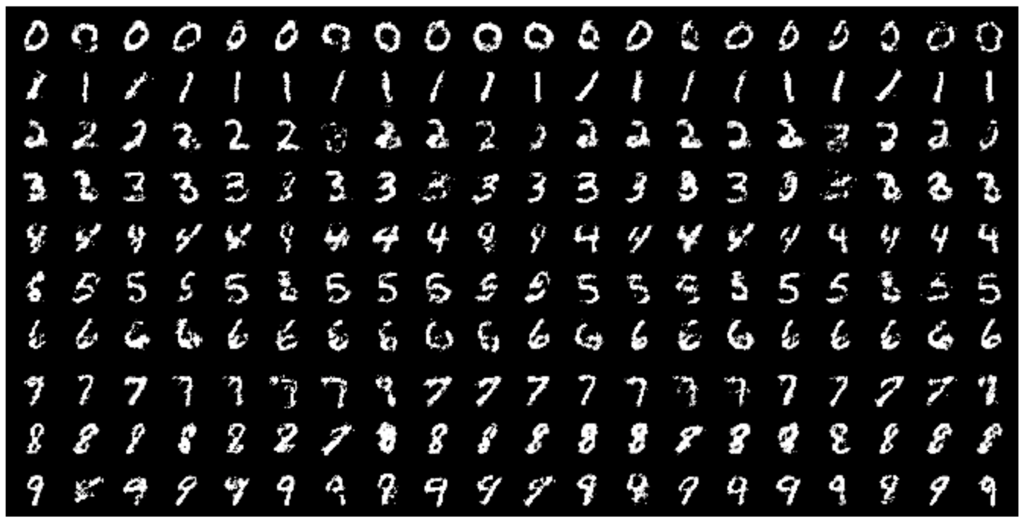


Discriminative distribution - D, blue, dashed line, data generating distribution - black, dotted line, generative distribution pg (G) - green, solid line. The lower horizontal line is the domain from which z is sampled, in this case uniformly. The horizontal line above is part of the domain of x. The upward arrows show how the mapping x = G(z) imposes the non-uniform distribution pg on transformed samples. Source: "Generative Adversarial Networks" paper

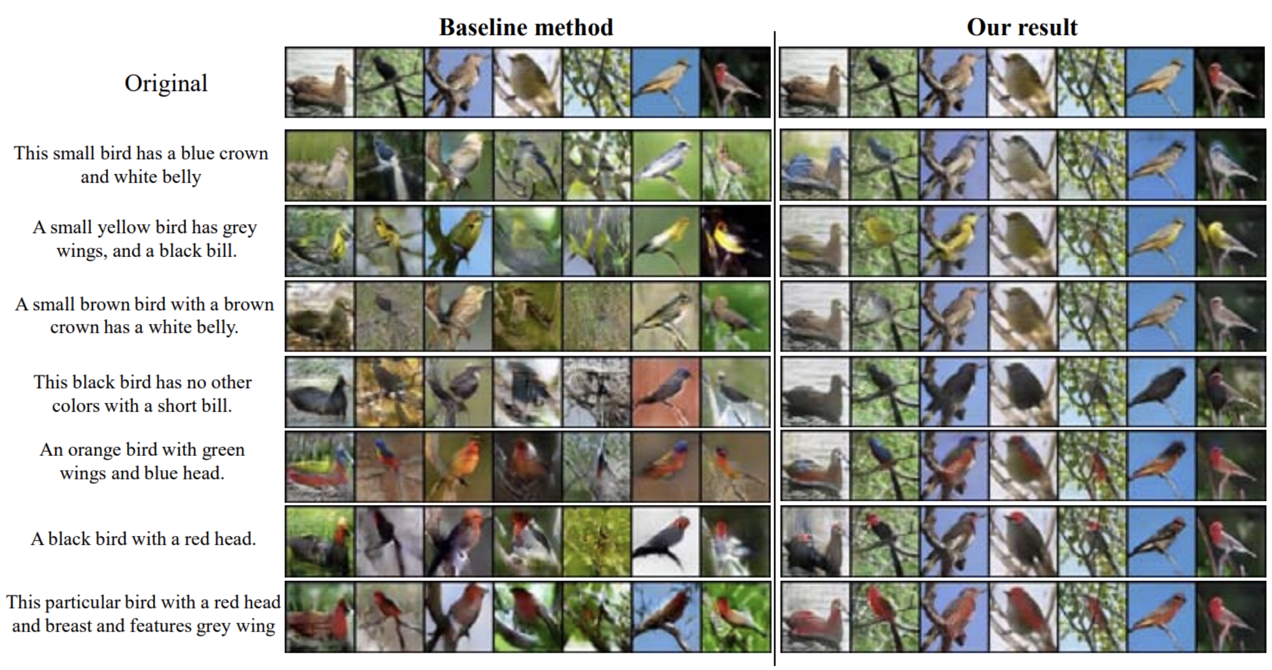
GAN Applications

- **(Conditional) Synthesis** — This includes font generation, Text2Image, as well as 3D Object generation.
- **Data Augmentation** — Aiming to reduce the need for labeled data (GAN is only used as a tool for enhancing the training process of another model).
- **Style Transfer and Manipulation —** Face Aging, painting, pose estimation and manipulation, inpainting, and blending.
- **Signal Super Resolution —** Artificially increasing the resolution of images.

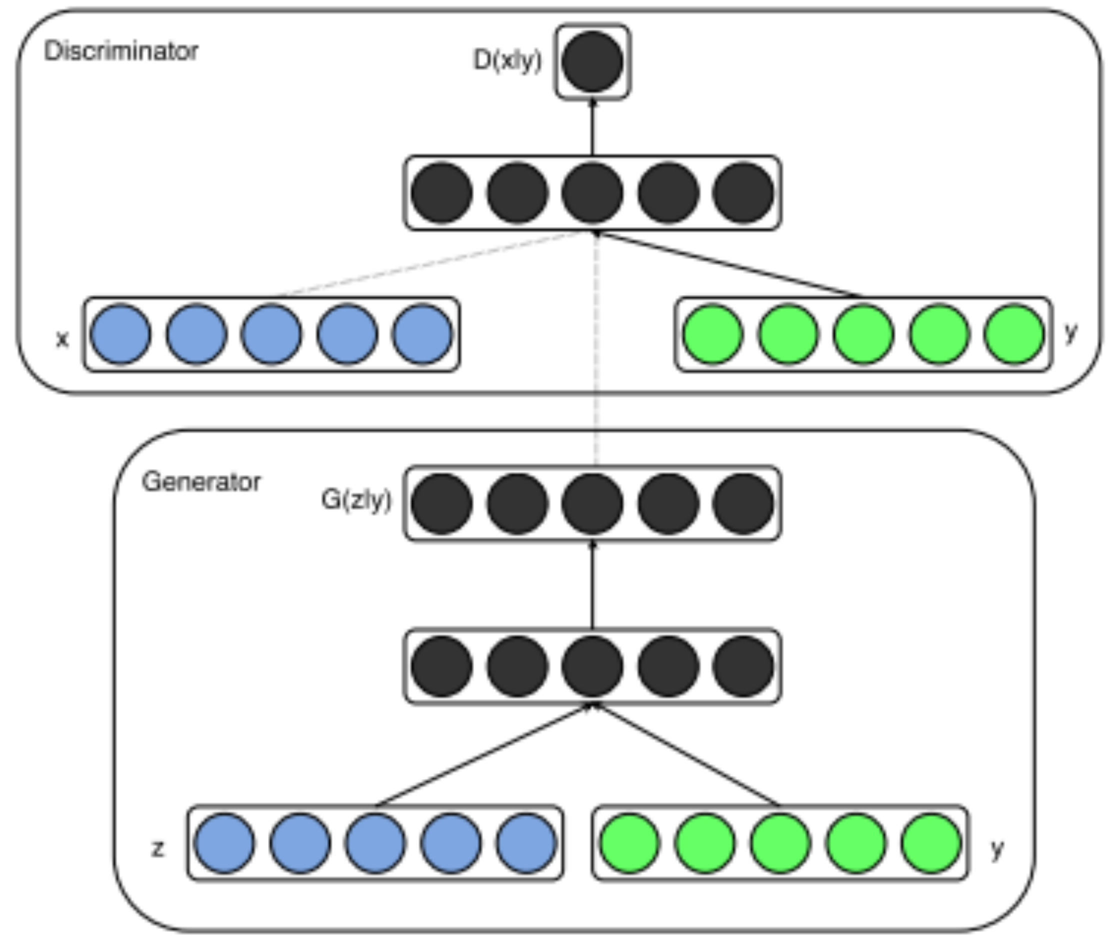
Conditional GANs (cGANs)



Generated MNIST digits, each row conditioned on one label
Source: https://arxiv.org/pdf/1411.1784.pdf

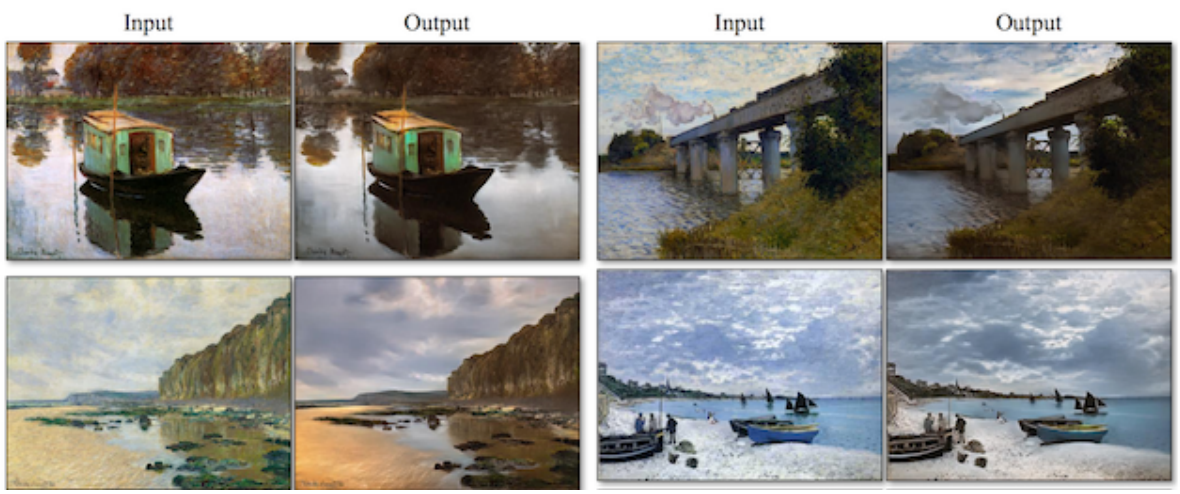


Implementation of MC-GAN for translating words into images.
Source: https://arxiv.org/pdf/1902.06068.pdf



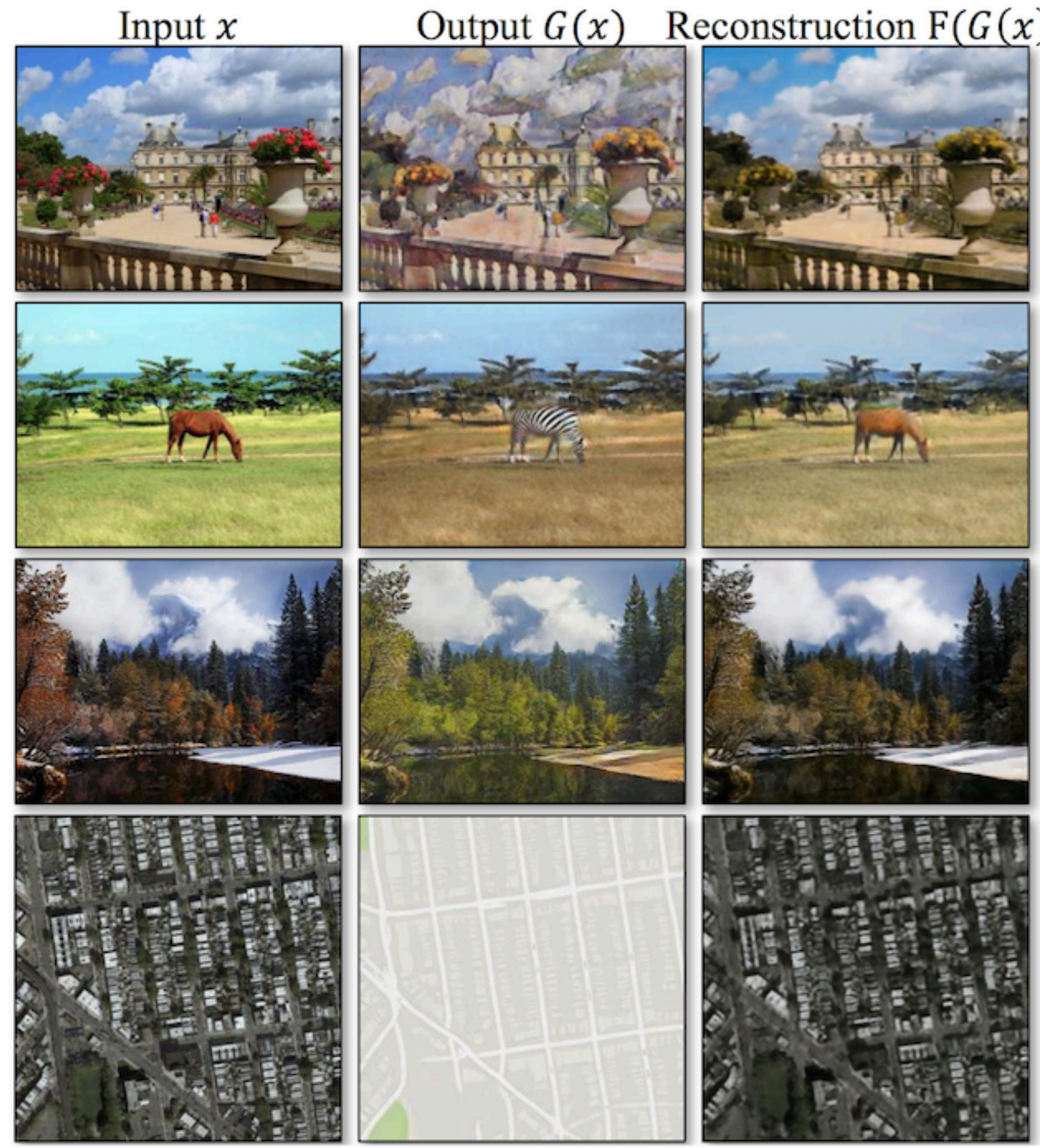
Generator and discriminator networks.
Source: https://arxiv.org/pdf/1411.1784.pdf

Style Transfer and Manipulation



Example of Translation from Paintings to Photographs With CycleGAN.
Source: Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks, 2017.

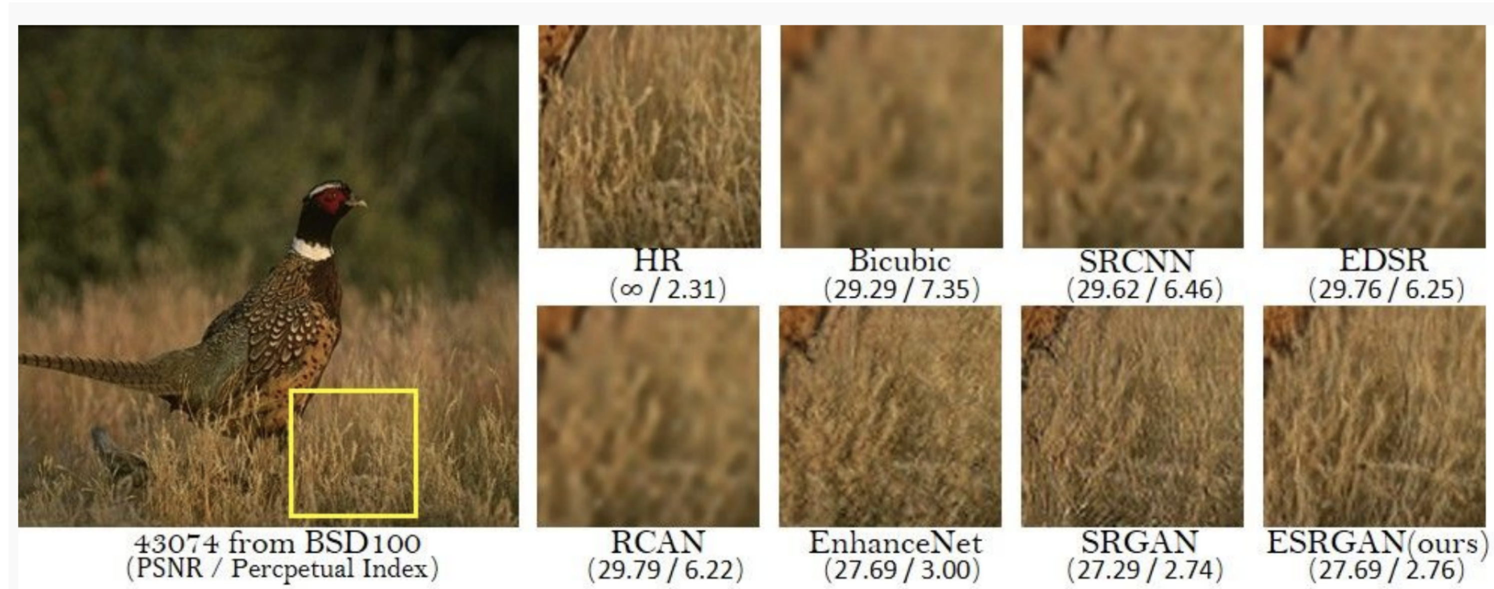
- Translation of painting to photograph.
- Translation of sketch to photograph.
- Translation of apples to oranges.
- Translation of photograph to artistic painting.



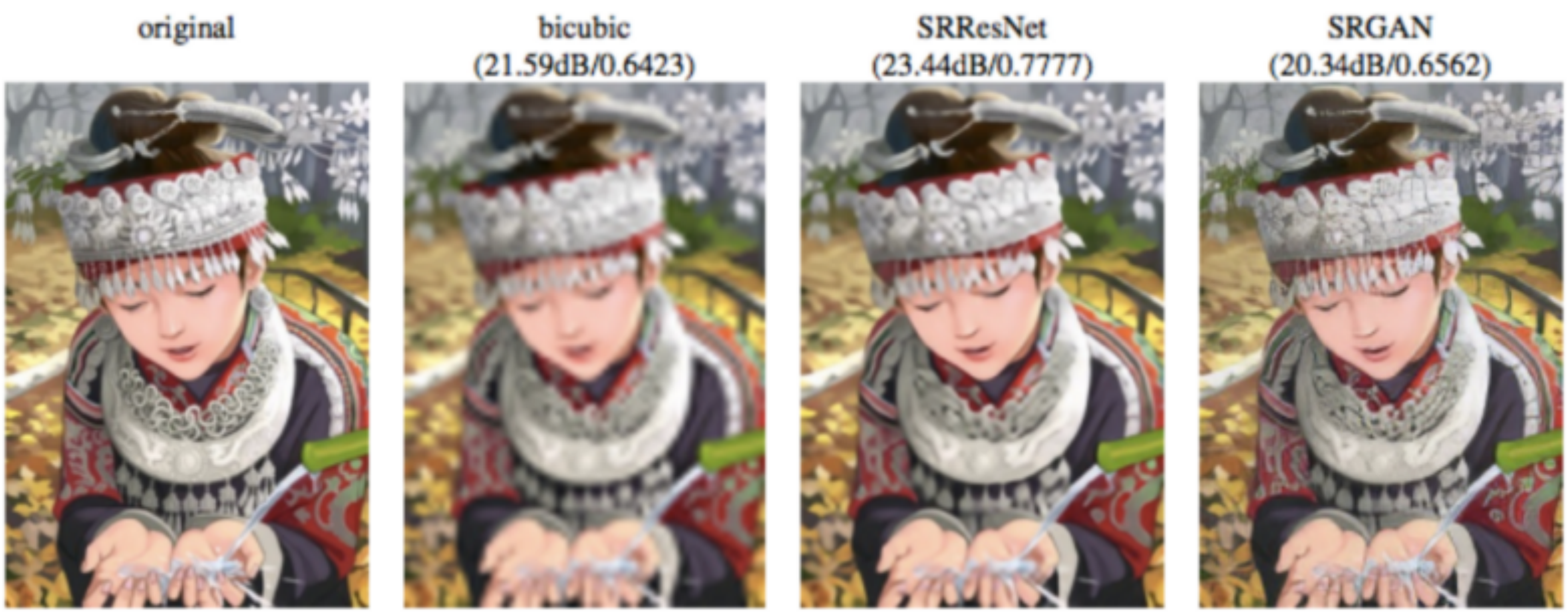
Example of Four Image-to-Image Translations Performed With CycleGANTaken from Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks, 2017.

Image Super-Resolution

Example of High-Resolution GAN-Generated Photographs. Source: Analyzing Perception-Distortion Tradeoff using Enhanced Perceptual Super-resolution Network, 2018.



Source



Source: https://arxiv.org/pdf/1511.06380.pdf

Fast Development

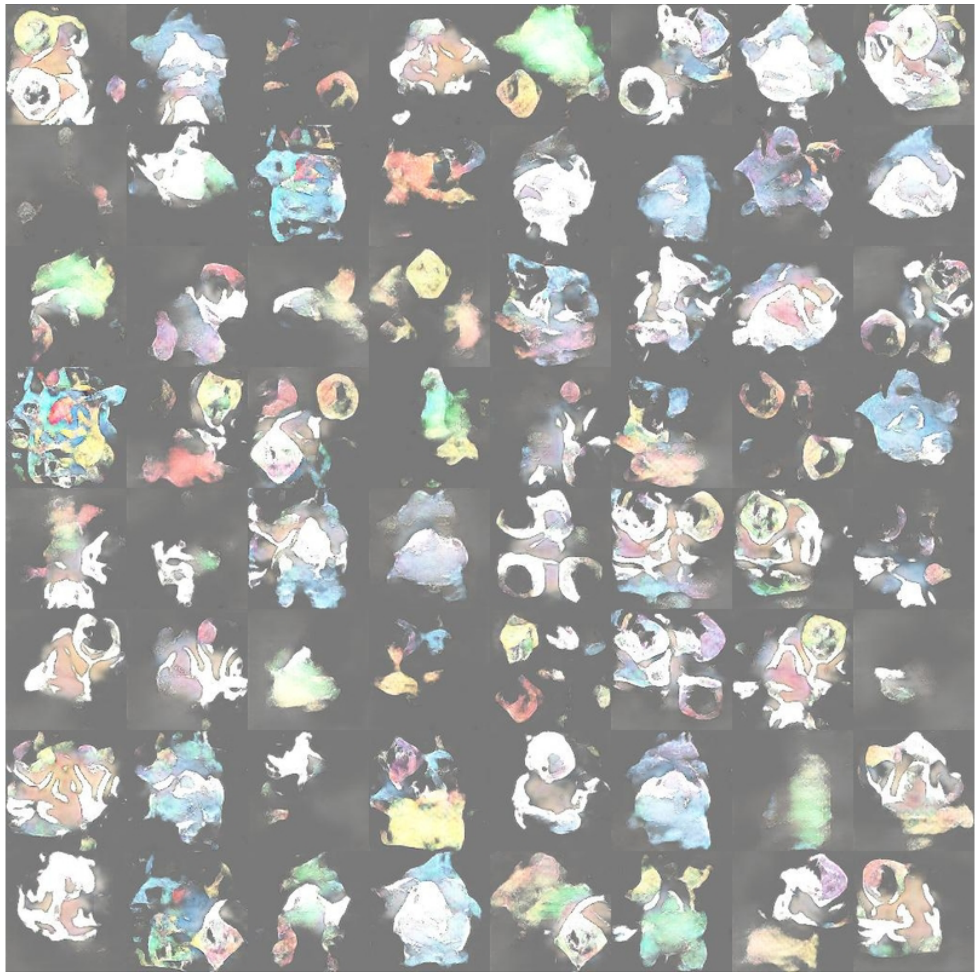
Example of the Progression in the Capabilities of GANs from 2014 to 2017.
Source: The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation, 2018.

4.5 years of GAN progress on face generation.
Source: [*https://arxiv.org/abs/1406.2661*](https://arxiv.org/abs/1406.2661) [*https://arxiv.org/abs/1511.06434*](https://arxiv.org/abs/1511.06434) [*https://arxiv.org/abs/1606.07536*](https://arxiv.org/abs/1606.07536) [*https://arxiv.org/abs/1710.10196*](https://arxiv.org/abs/1710.10196)

interaction

few personal faves

Having Fun

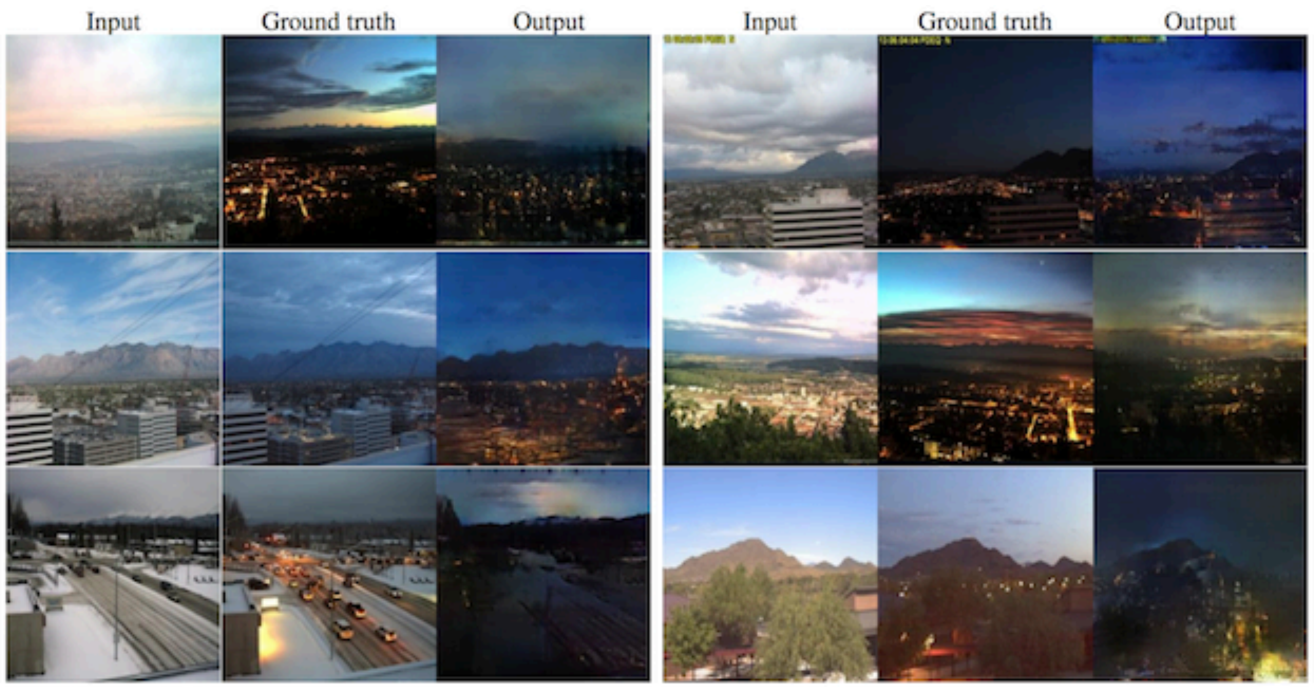


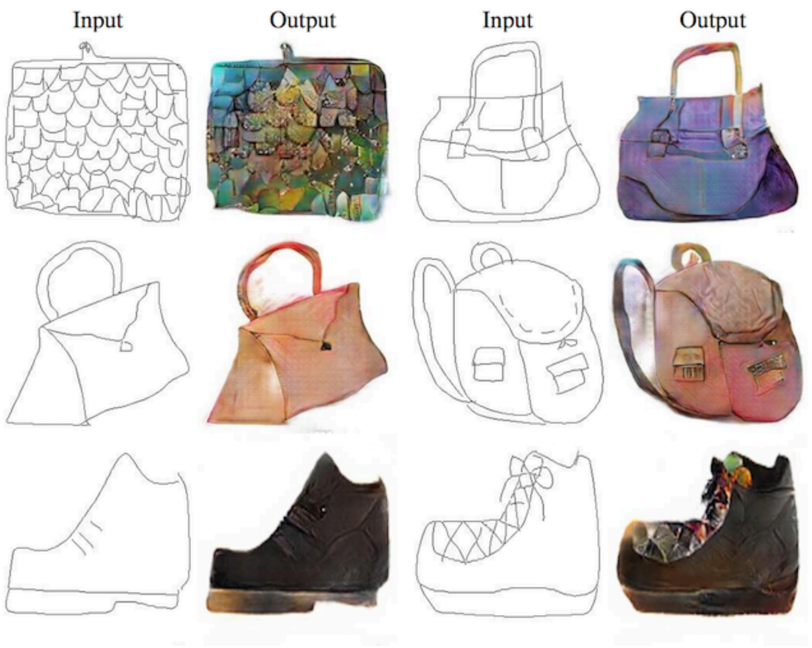
Source: https://github.com/moxiegushi/pokeGAN

Source: https://arxiv.org/pdf/1708.05509

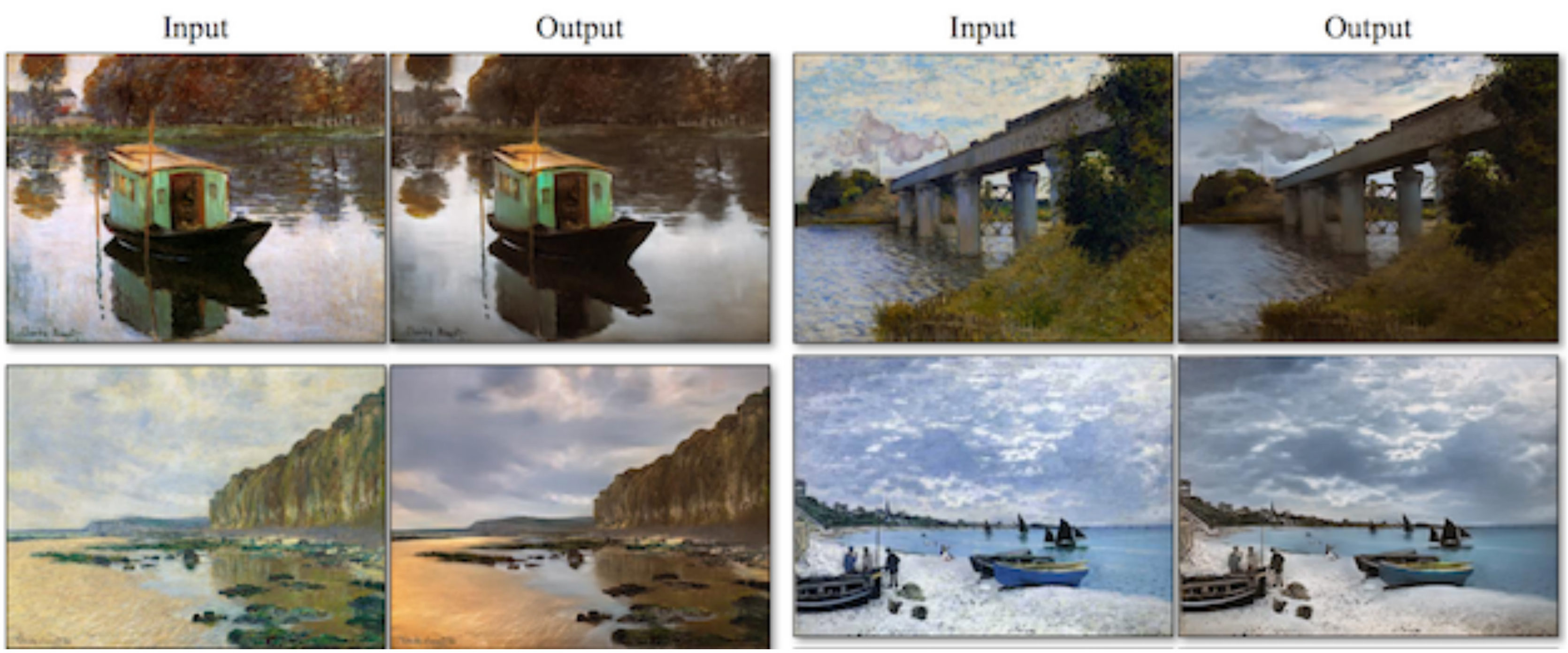
Source: https://arxiv.org/abs/1710.10196

Image-to-image





Source: *Image-to-Image Translation with Conditional Adversarial Networks https://arxiv.org/abs/1611.07004*



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

Human Generation

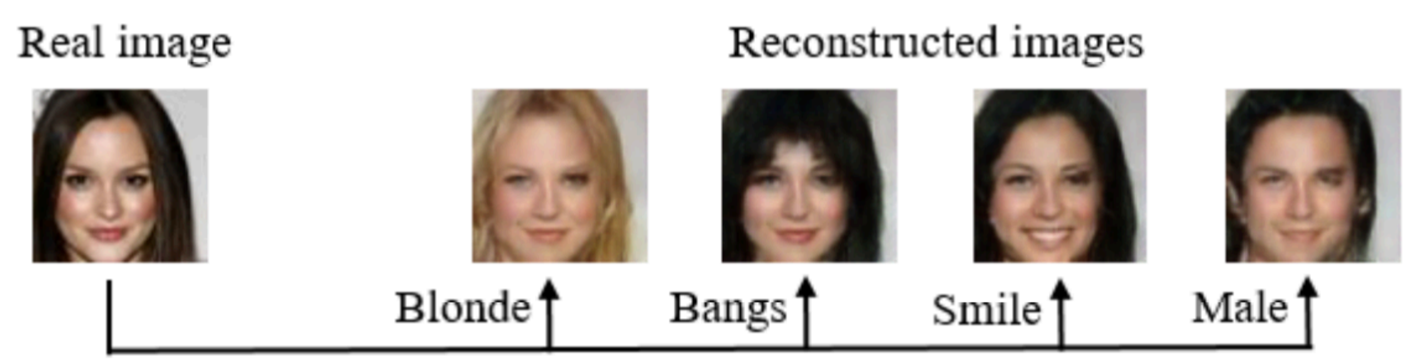
Face Frontal View Generation

Source: Beyond Face Rotation: Global and Local Perception GAN for Photorealistic and Identity Preserving Frontal View Synthesis

Generate New Human Poses

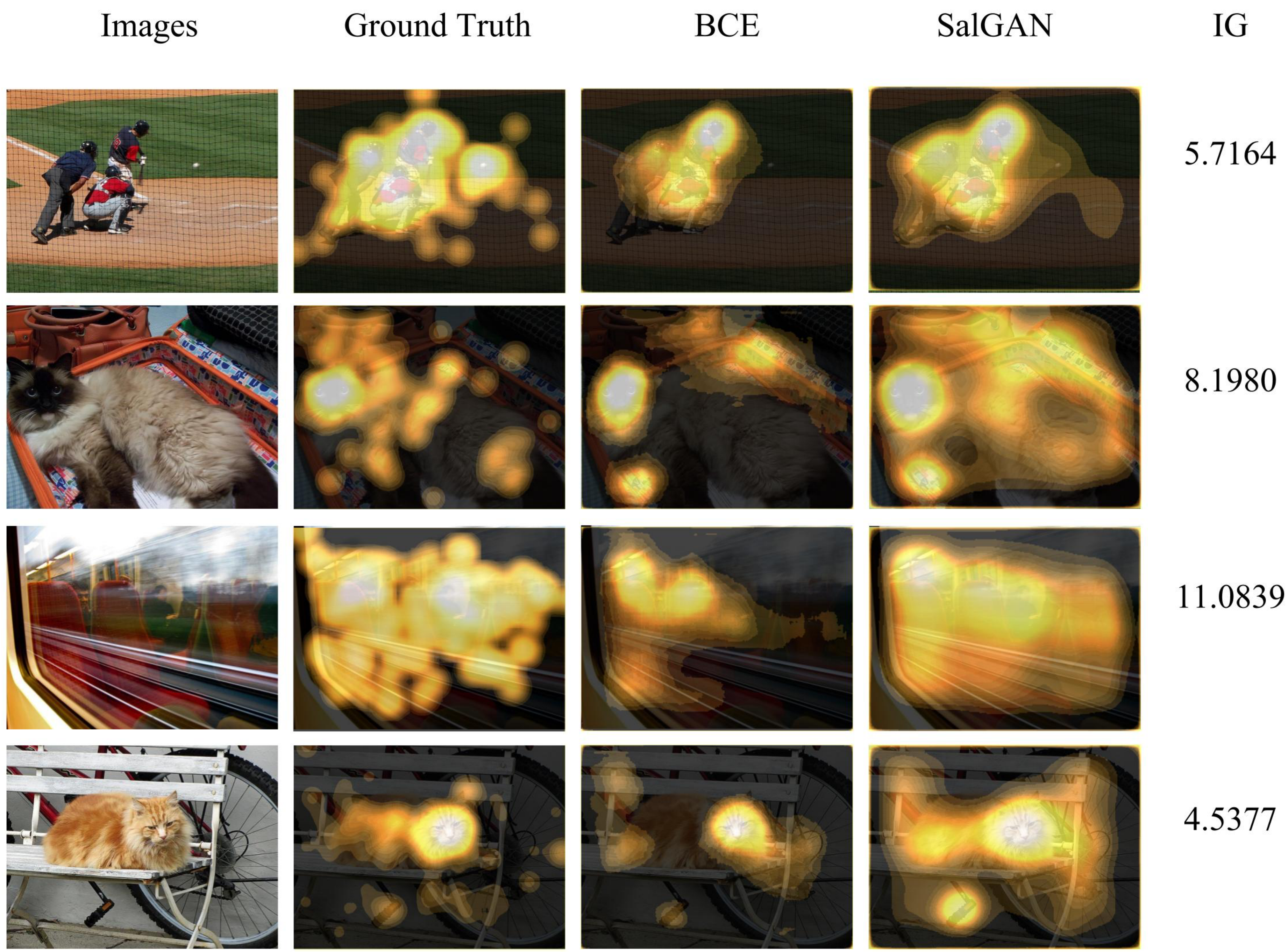
Source: Pose Guided Person Image Generation https://arxiv.org/abs/1705.09368

Photo Editing



Source: Unsupervised Cross-Domain Image Generation https://arxiv.org/abs/1611.02200

GANs for Attention Prediction



Source: SalGAN: visual saliency prediction with adversarial networks https://arxiv.org/pdf/1701.01081.pdf