# Deep Learning Basics

Lecture 10: Generative Models Part 2

최성준 (고려대학교 인공지능학과)

# Latent Variable Models

D. Kingma, "Variational Inference and Deep Learning: A New Synthesis," Ph.D. Thesis

# Question

Is an **autoencoder** a generative model?

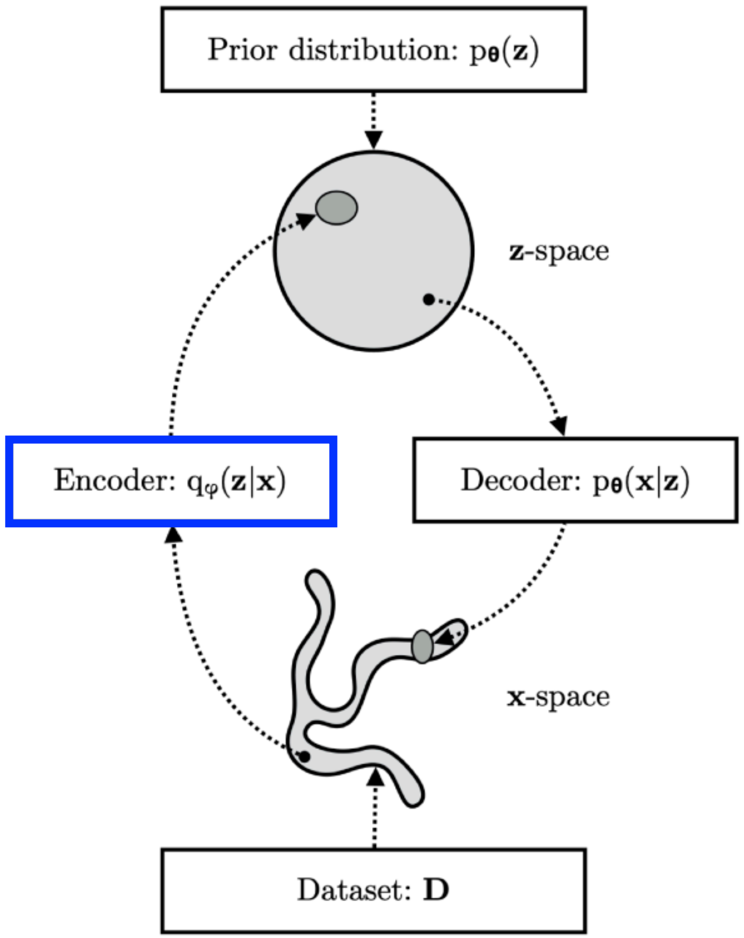
# Variational Auto-encoder

- Variational inference (VI)
  - The goal of VI is to optimize the variational distribution that best matches the posterior distribution.
    - Posterior distribution: $p_\theta(z|x)$
    - Variational distribution: $q_\phi(z|x)$
  - In particular, we want to find the variational distribution that minimizes the KL divergence between the true posterior.

D. Kingma, "Variational Inference and Deep Learning: A New Synthesis," Ph.D. Thesis

# Variational Auto-encoder

Our goal is to learn the **variational distribution** (aka **Encoder**)




D. Kingma, "Variational Inference and Deep Learning: A New Synthesis," Ph.D. Thesis

# Variational Auto-encoder

- But how?

$$
\begin{aligned}
\ln p_\theta(D) &= \mathbb{E}_{q_\phi(z|x)}\left[\ln p_\theta(x)\right] \\
&= \mathbb{E}_{q_\phi(z|x)}\left[\ln \frac{p_\theta(x,z)}{p_\theta(z|x)}\right] \\
&= \mathbb{E}_{q_\phi(z|x)}\left[\ln \frac{p_\theta(x,z)\, q_\phi(z|x)}{q_\phi(z|x)\, p_\theta(z|x)}\right] \\
&= \mathbb{E}_{q_\phi(z|x)}\left[\ln \frac{p_\theta(x,z)}{q_\phi(z|x)}\right] + \mathbb{E}_{q_\phi(z|x)}\left[\ln \frac{q_\phi(z|x)}{p_\theta(z|x)}\right] \\
&= \boxed{\underbrace{\mathbb{E}_{q_\phi(z|x)}\left[\ln \frac{p_\theta(x,z)}{q_\phi(z|x)}\right]}_{\text{ELBO} \uparrow}} + \boxed{\underbrace{D_{KL}\left(q_\phi(z|x) \| p_\theta(z|x)\right)}_{\text{Objective} \downarrow}}
\end{aligned}
$$

Importantly, **ELBO** is a **tractable** quantity.

VI minimizes the (intractable) **objective** via maximizing **ELBO**.

D. Kingma, "Variational Inference and Deep Learning: A New Synthesis," Ph.D. Thesis

# Variational Auto-encoder

- ELBO can further be decomposed into

$$\underbrace{\mathbb{E}_{q_\phi(z|x)}\left[\ln \frac{p_\theta(x,z)}{q_\phi(z|x)}\right]}_{\text{ELBO}\uparrow} = \int \ln \frac{p_\theta(x|z)p(z)}{q_\phi(z|x)} q_\phi(z|x)dz$$

$$= \boxed{\underbrace{\mathbb{E}_{q_\phi(z|x)}\left[p_\theta(x|z)\right]}_{\text{Reconstruction Term}}} - \boxed{\underbrace{D_{KL}\left(q_\phi(z|x)\|p(z)\right)}_{\text{Prior Fitting Term}}}$$

This term minimizes the reconstruction loss of an auto-encoder.

This term enforces the latent distribution to be similar to the prior distribution.

D. Kingma, "Variational Inference and Deep Learning: A New Synthesis," Ph.D. Thesis

# Variational Auto-encoder

- Key limitation:
  - It is an intractable model (hard to evaluate likelihood).
  - The prior fitting term must be differentiable, hence it is hard to use diverse latent prior distributions.
  - In most cases, we use an isotropic Gaussian.

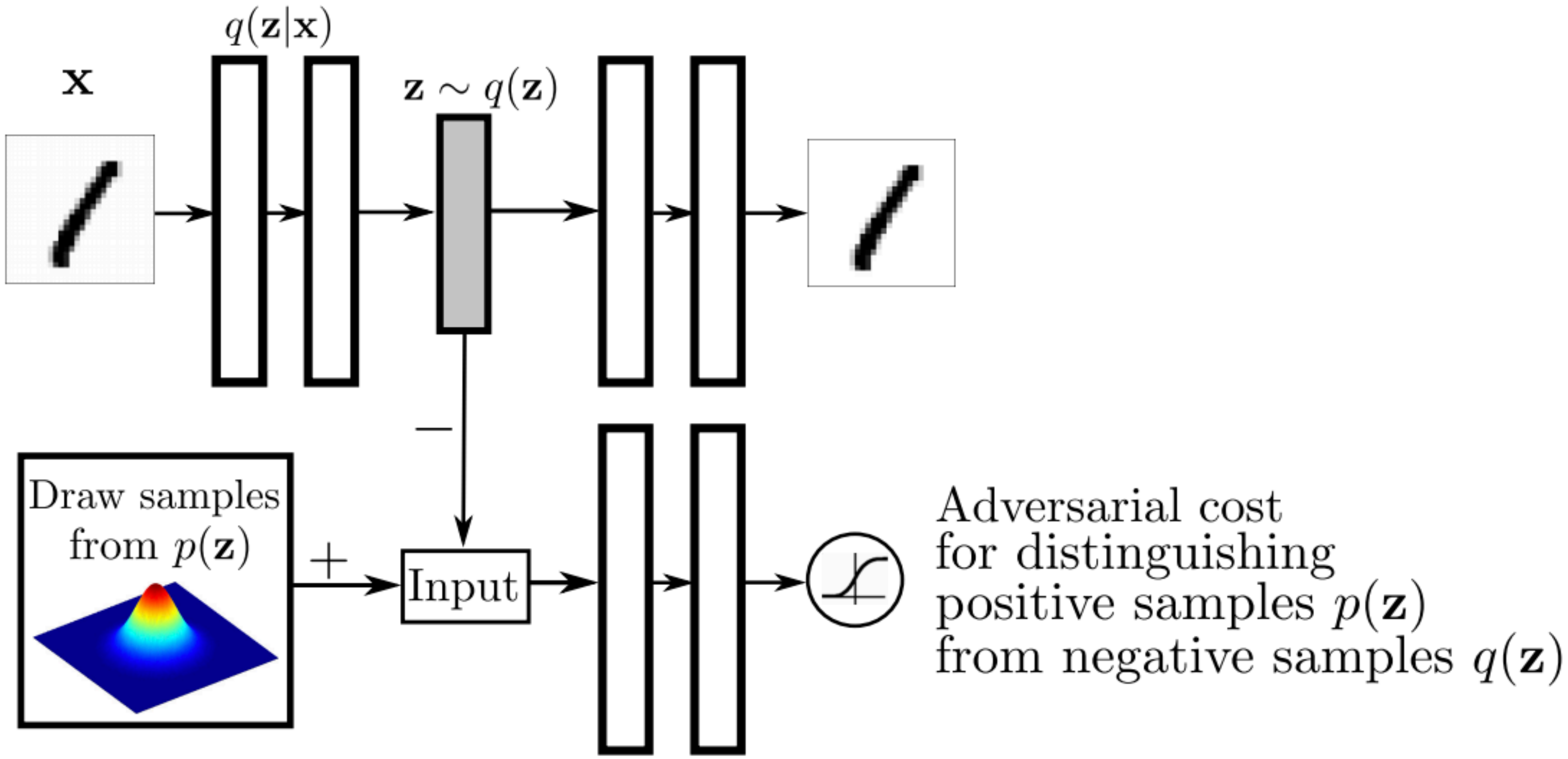
$$D_{KL}(q_\phi(z|x)\|\mathcal{N}(0,I)) = \frac{1}{2}\sum_{i=1}^{D}\left(\sigma_{z_i}^2 + \mu_{z_i}^2 - \ln(\sigma_{z_i}^2) - 1\right)$$

D. Kingma, "Variational Inference and Deep Learning: A New Synthesis," Ph.D. Thesis

# Variational Auto-encoder

D. Kingma, "Variational Inference and Deep Learning: A New Synthesis," Ph.D. Thesis

# Adversarial Auto-encoder



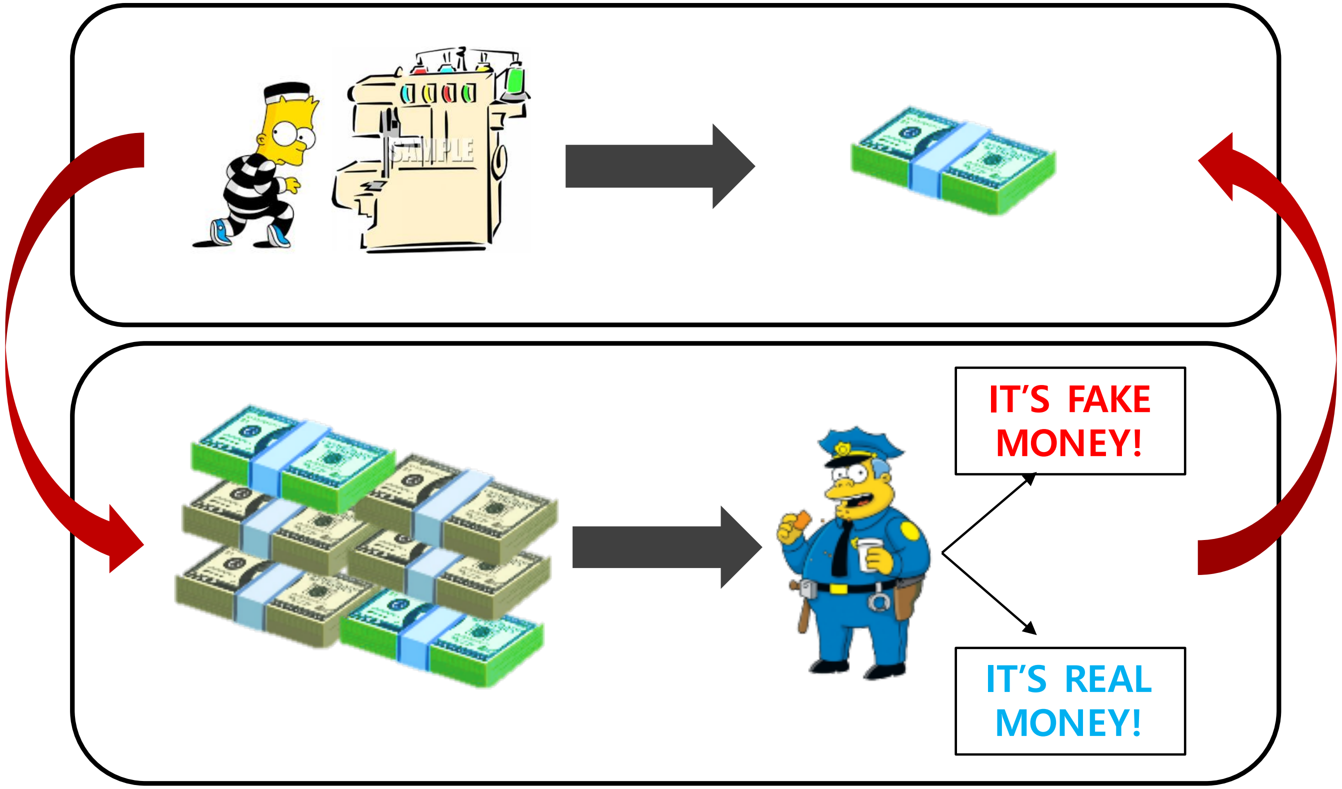

- It allows us to use any arbitrary latent distributions that we can sample.

https://arxiv.org/abs/1511.05644

# Generative Adversarial Network

I. Goodfellow et al., "Generative Adversarial Networks", NIPS, 2014

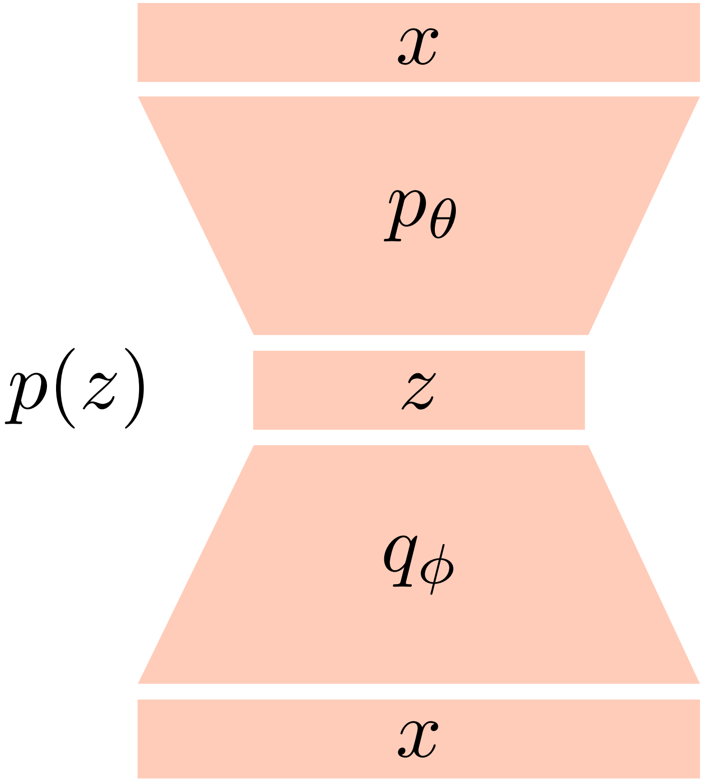
# GAN



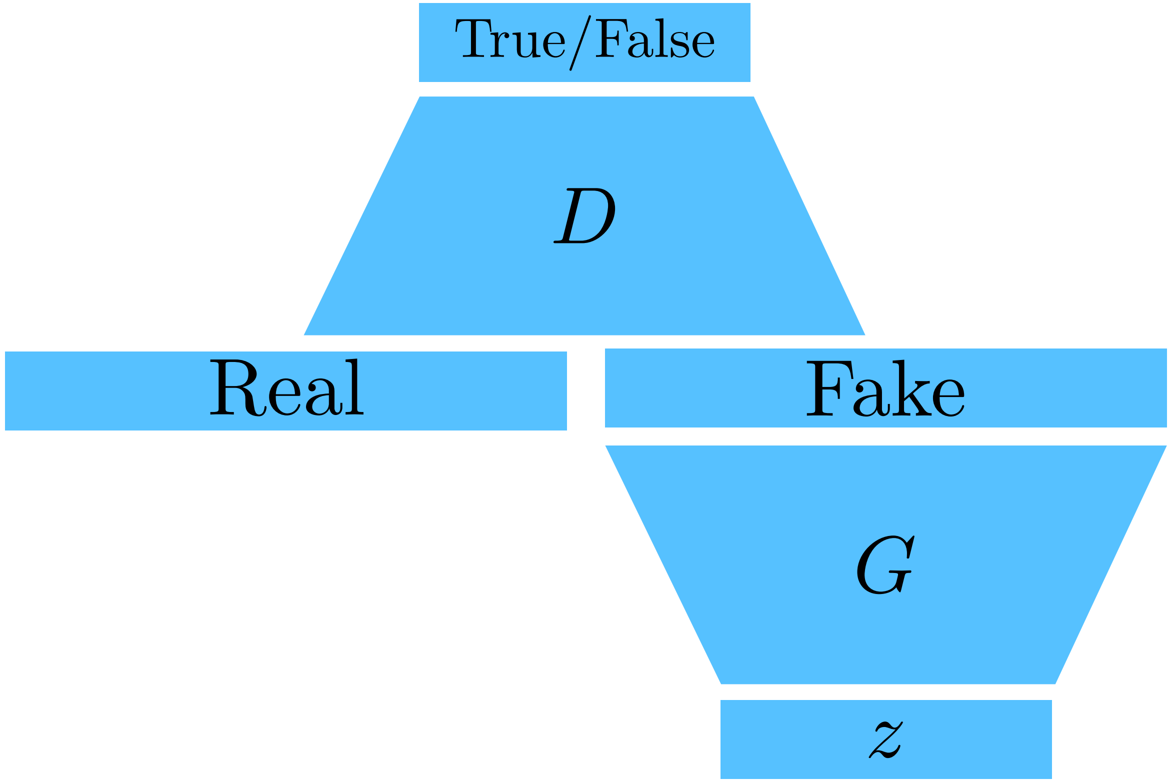

$$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))].$$

https://arxiv.org/abs/1406.2661

# GAN vs. VAE




Variational Autoencoder




Generative Adversarial Networks

https://arxiv.org/abs/1406.2661

# GAN Objective

- A two player minimax game between generator and discriminator.
- For discriminator:

$$\max_D V(G, D) = E_{\mathbf{x} \sim p_{\text{data}}}[\log D(\mathbf{x})] + E_{\mathbf{x} \sim p_G}[\log(1 - D(\mathbf{x}))]$$

  - where the optimal discriminator is

$$D_G^*(\mathbf{x}) = \frac{p_{\text{data}}(\mathbf{x})}{p_{\text{data}}(\mathbf{x}) + p_G(\mathbf{x})}$$

https://arxiv.org/abs/1406.2661

# GAN Objective

- For generator:

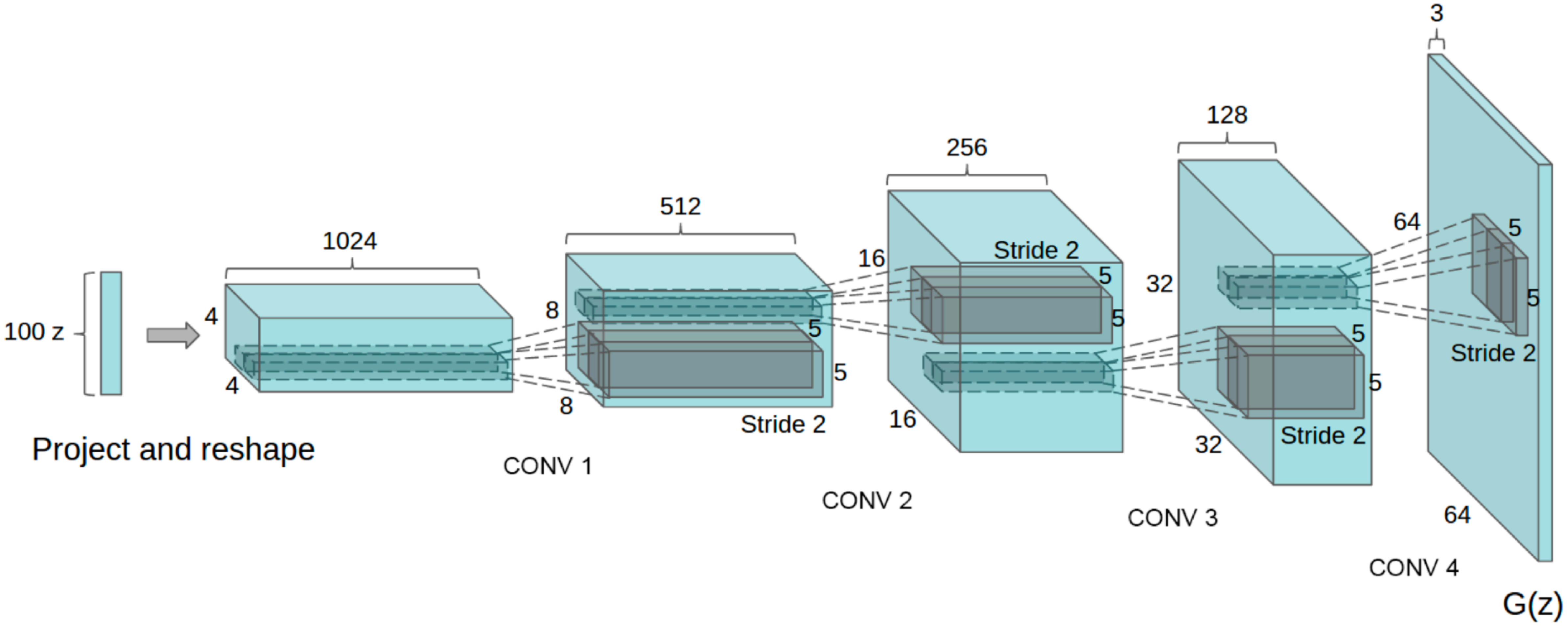
$$\min_G V(G, D) = E_{\mathbf{x}\sim p_{\text{data}}}[\log D(\mathbf{x})] + E_{\mathbf{x}\sim p_G}[\log(1 - D(\mathbf{x}))]$$

- Plugging in the optimal discriminator, we get

$$\begin{aligned}
V(G, D_G^*(\mathbf{x})) &= E_{\mathbf{x}\sim p_{\text{data}}}\left[\log \frac{p_{\text{data}}(\mathbf{x})}{p_{\text{data}}(\mathbf{x})+p_G(\mathbf{x})}\right] + E_{\mathbf{x}\sim p_G}\left[\log \frac{p_G(\mathbf{x})}{p_{\text{data}}(\mathbf{x})+p_G(\mathbf{x})}\right] \\
&= E_{\mathbf{x}\sim p_{\text{data}}}\left[\log \frac{p_{\text{data}}(\mathbf{x})}{\frac{p_{\text{data}}(\mathbf{x})+p_G(\mathbf{x})}{2}}\right] + E_{\mathbf{x}\sim p_G}\left[\log \frac{p_G(\mathbf{x})}{\frac{p_{\text{data}}(\mathbf{x})+p_G(\mathbf{x})}{2}}\right] - \log 4 \\
&= \underbrace{D_{KL}\left[p_{\text{data}}, \frac{p_{\text{data}}+p_G}{2}\right] + D_{KL}\left[p_G, \frac{p_{\text{data}}+p_G}{2}\right]}_{2\times\text{Jenson-Shannon Divergence (JSD)}} - \log 4 \\
&= 2D_{JSD}[p_{\text{data}}, p_G] - \log 4
\end{aligned}$$

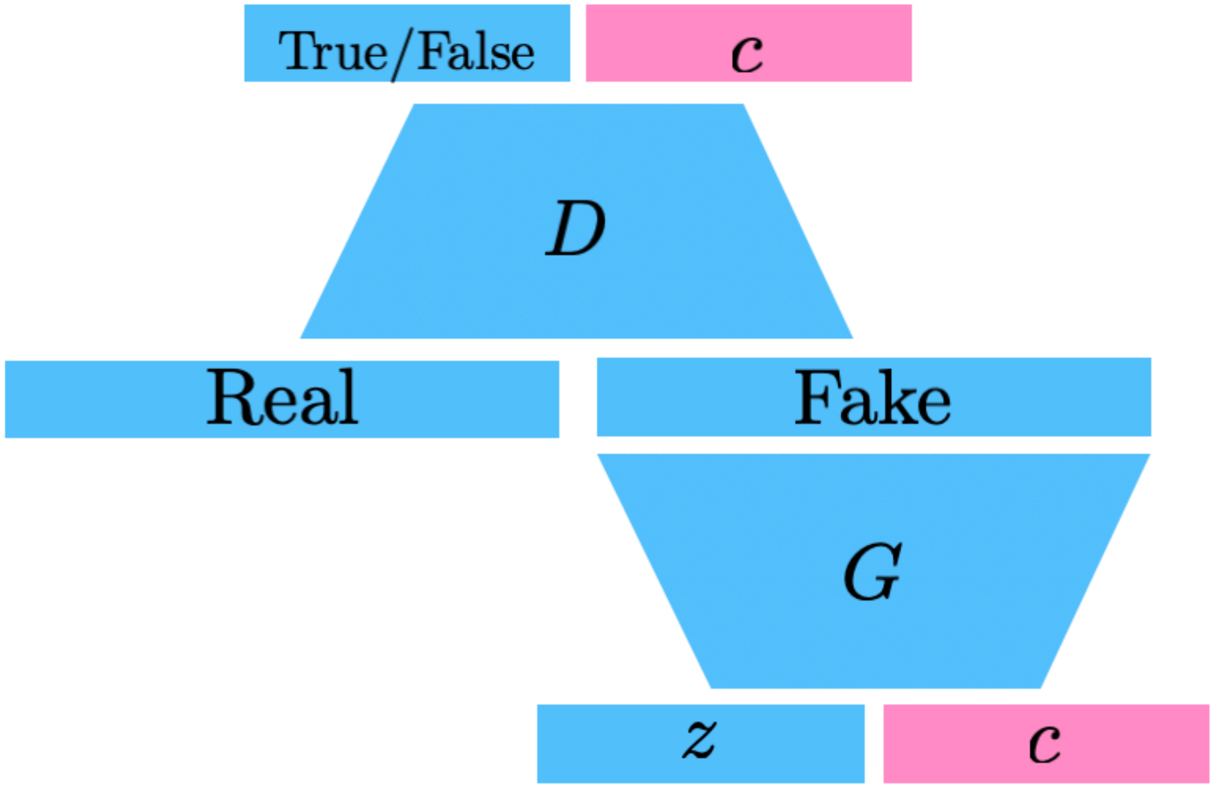
https://arxiv.org/abs/1406.2661

# DCGAN



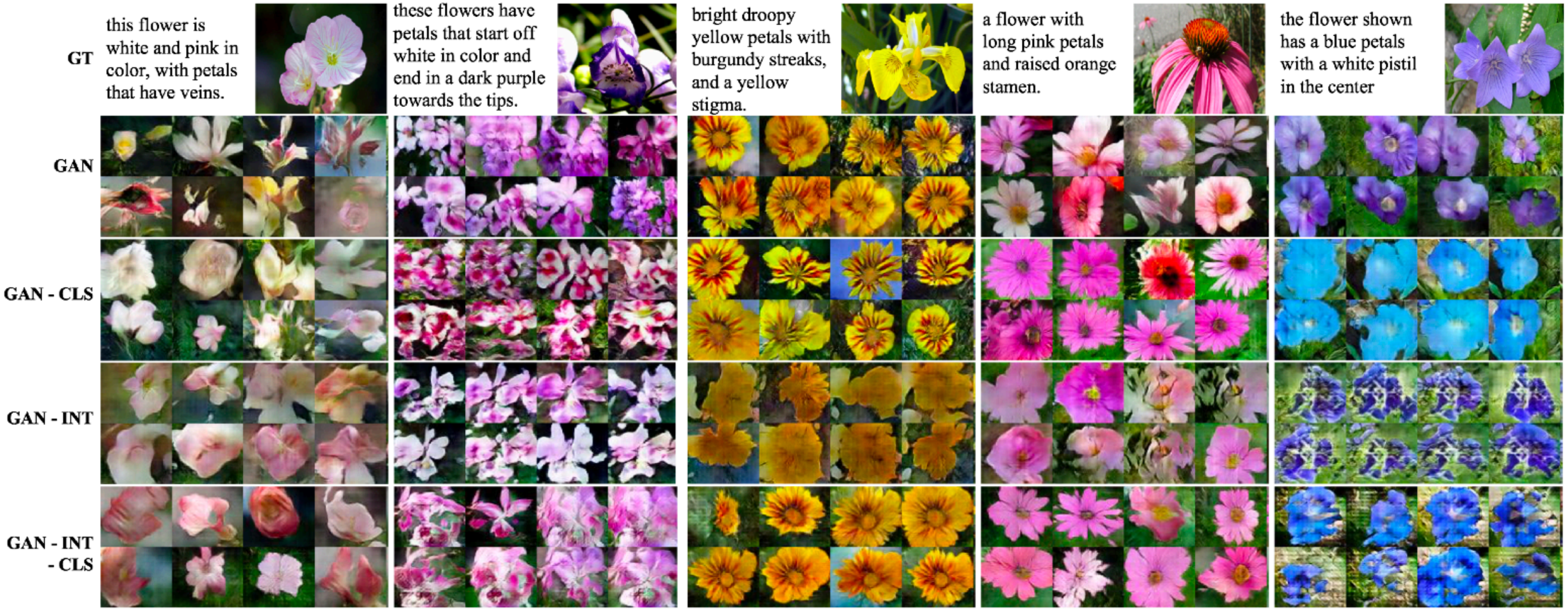

https://arxiv.org/abs/1511.06434

# Info-GAN



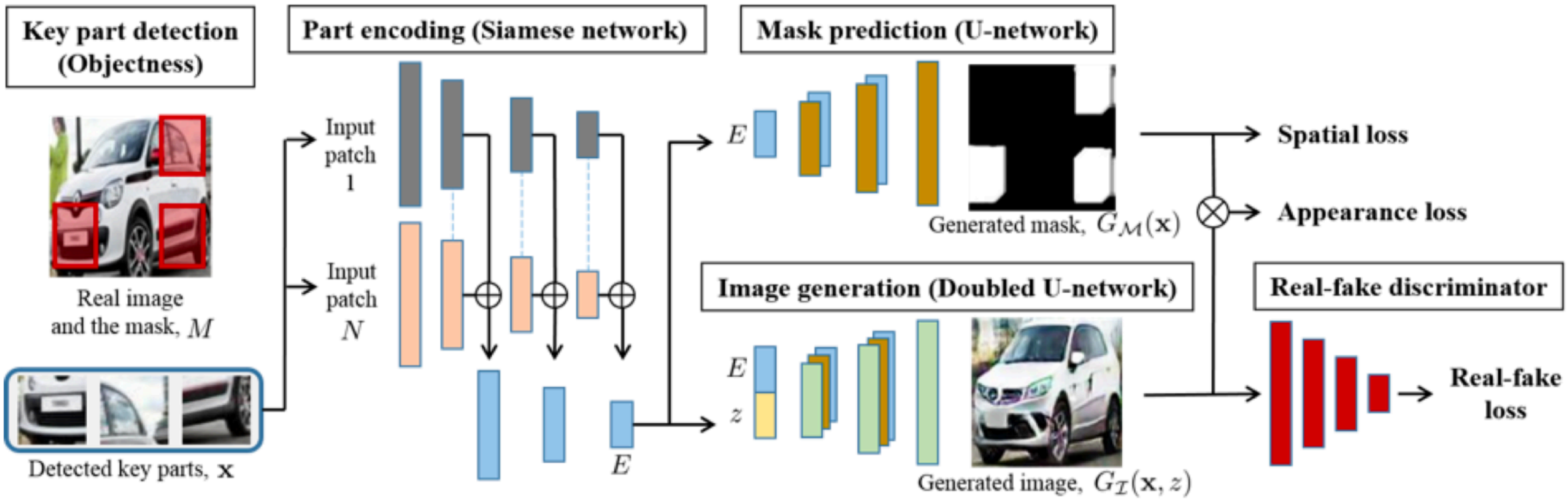

https://arxiv.org/abs/1606.03657

# Text2Image



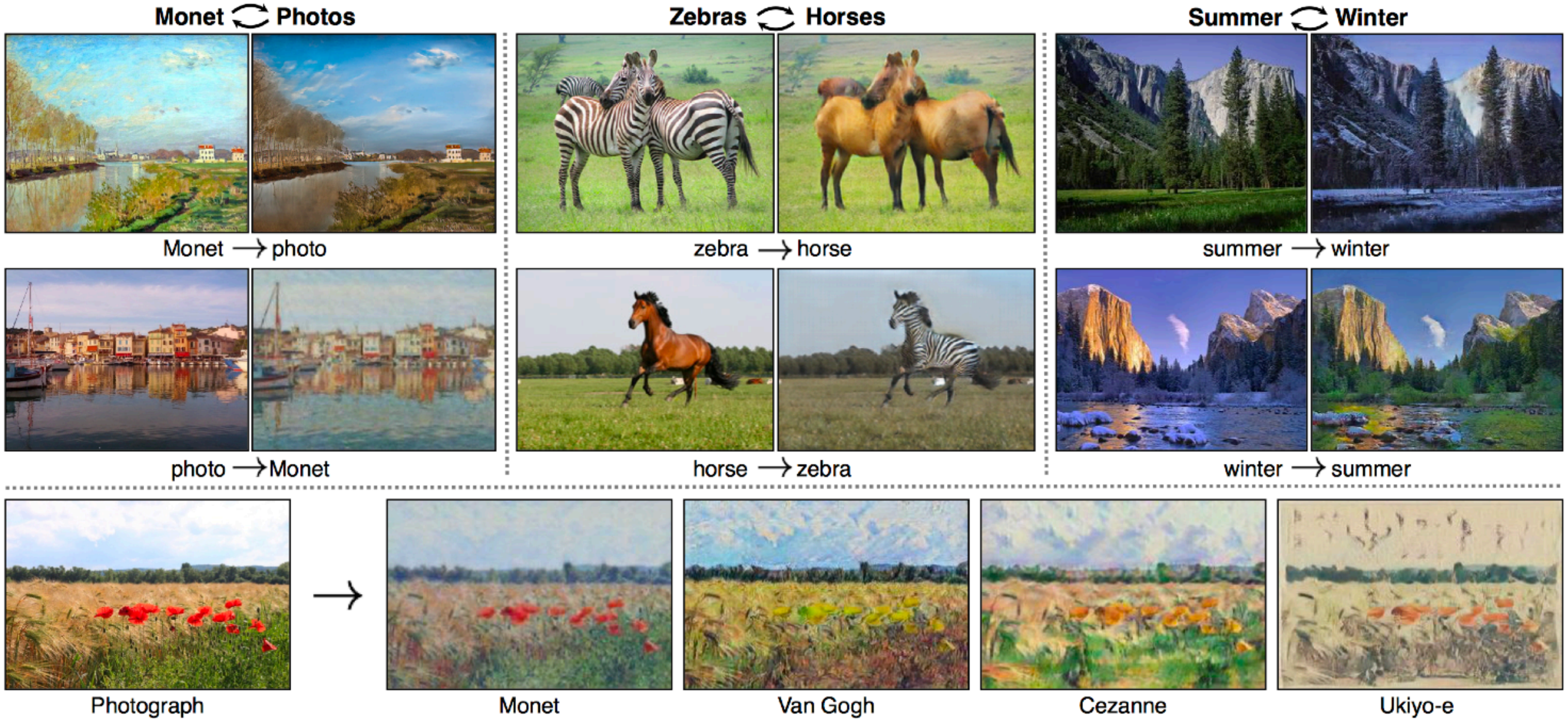

https://arxiv.org/abs/1605.05396

# Puzzle-GAN




https://arxiv.org/abs/1703.10730

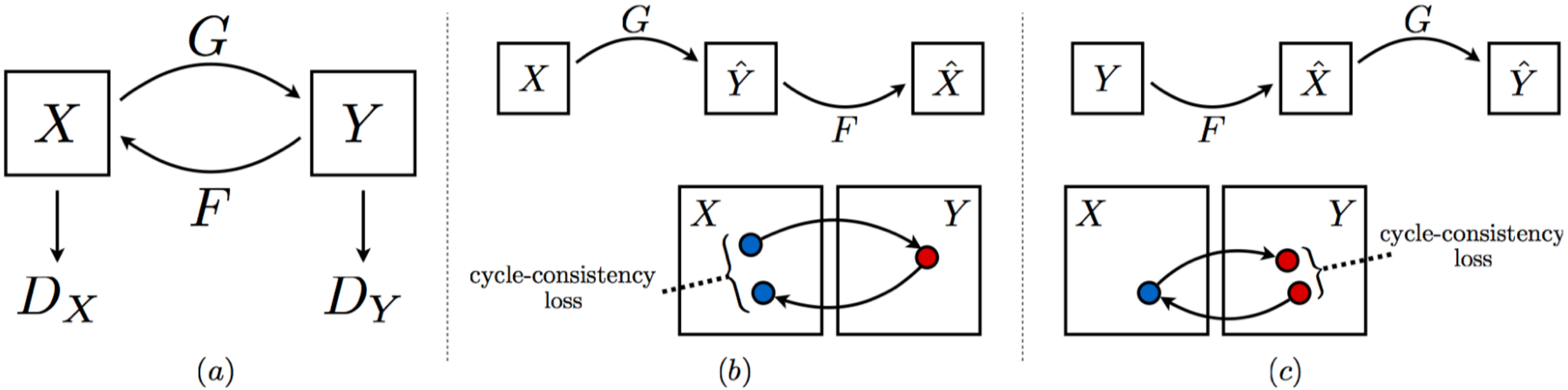
# CycleGAN




https://arxiv.org/abs/1703.10593

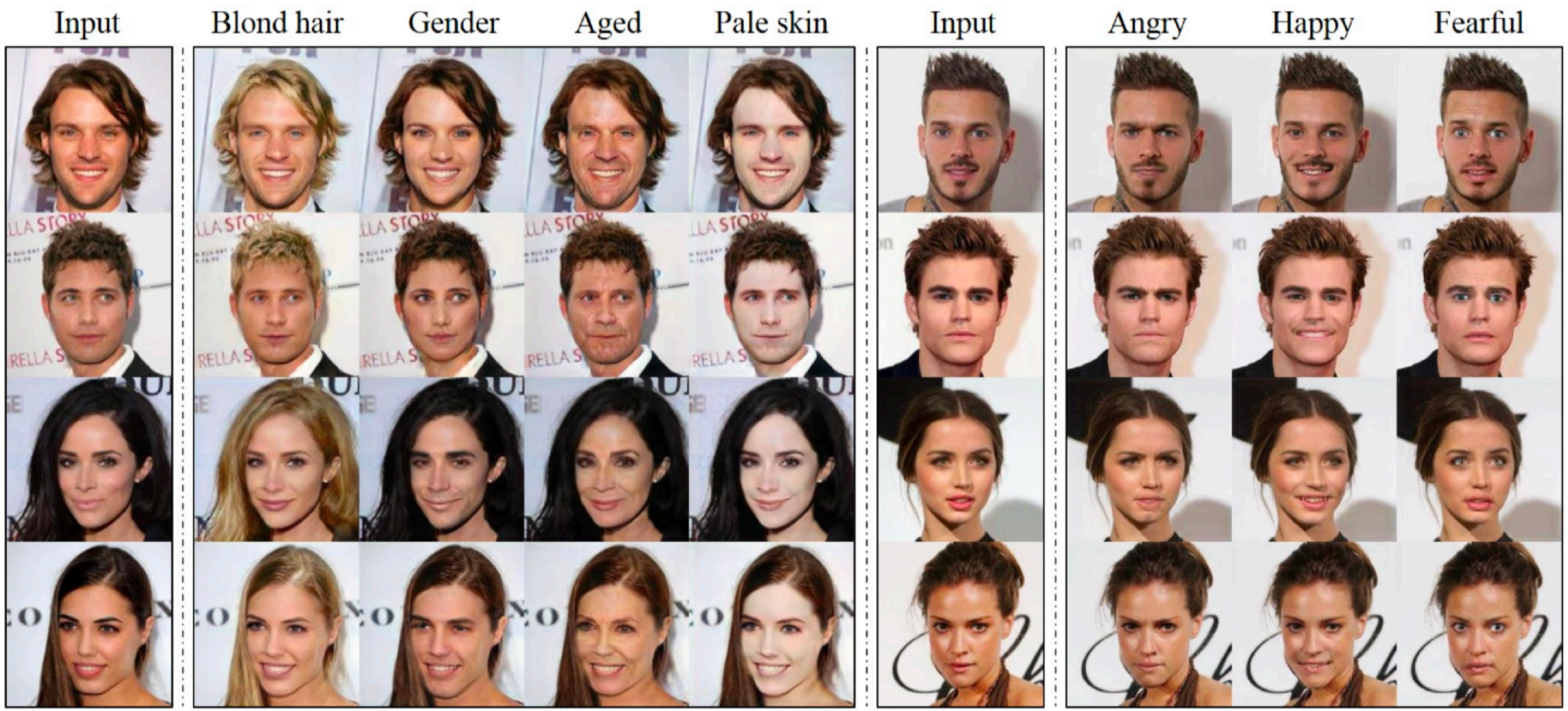
# CycleGAN

- Cycle-consistency loss




https://arxiv.org/abs/1703.10593

# Star-GAN

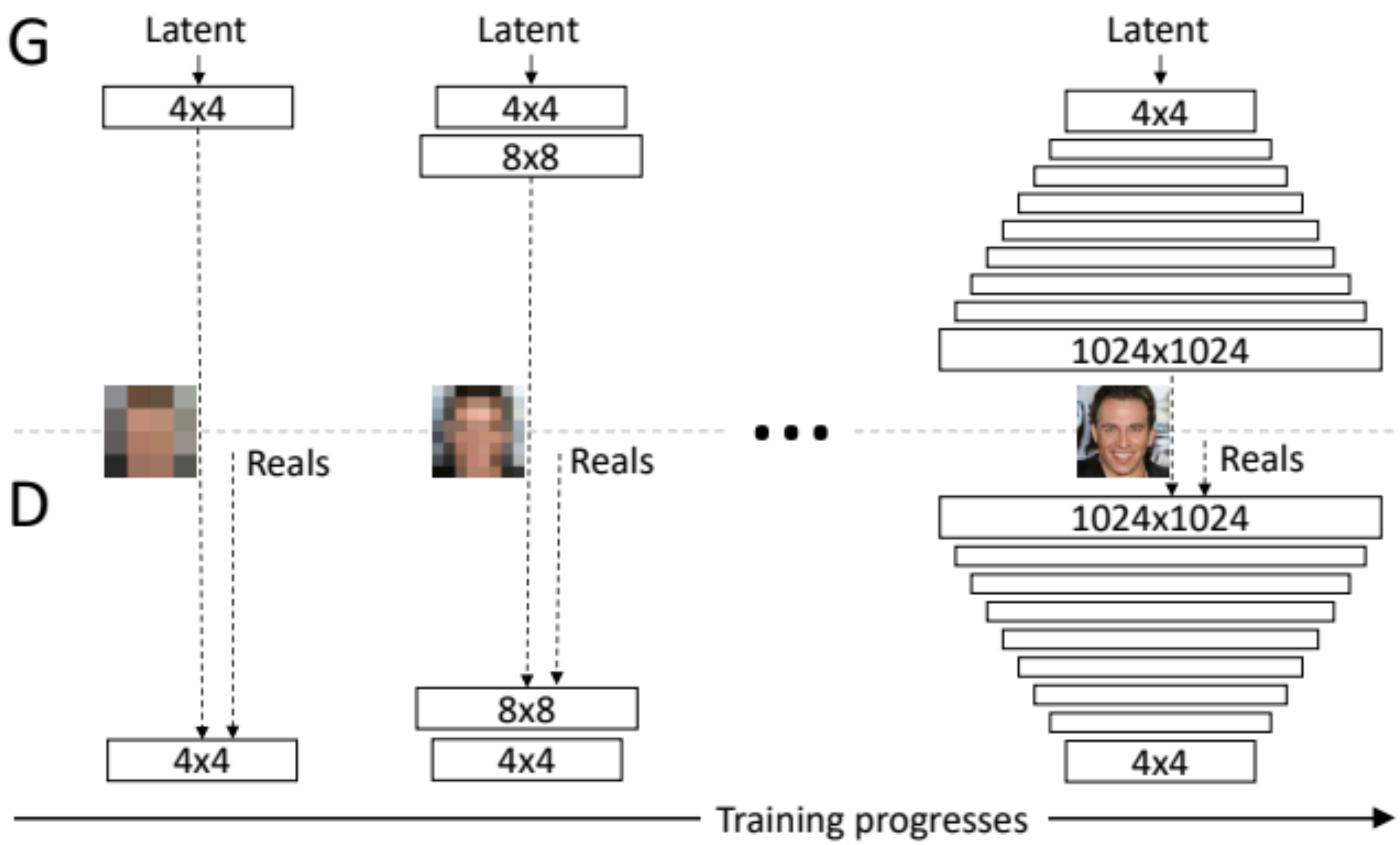



https://arxiv.org/abs/1711.09020

# Progressive-GAN

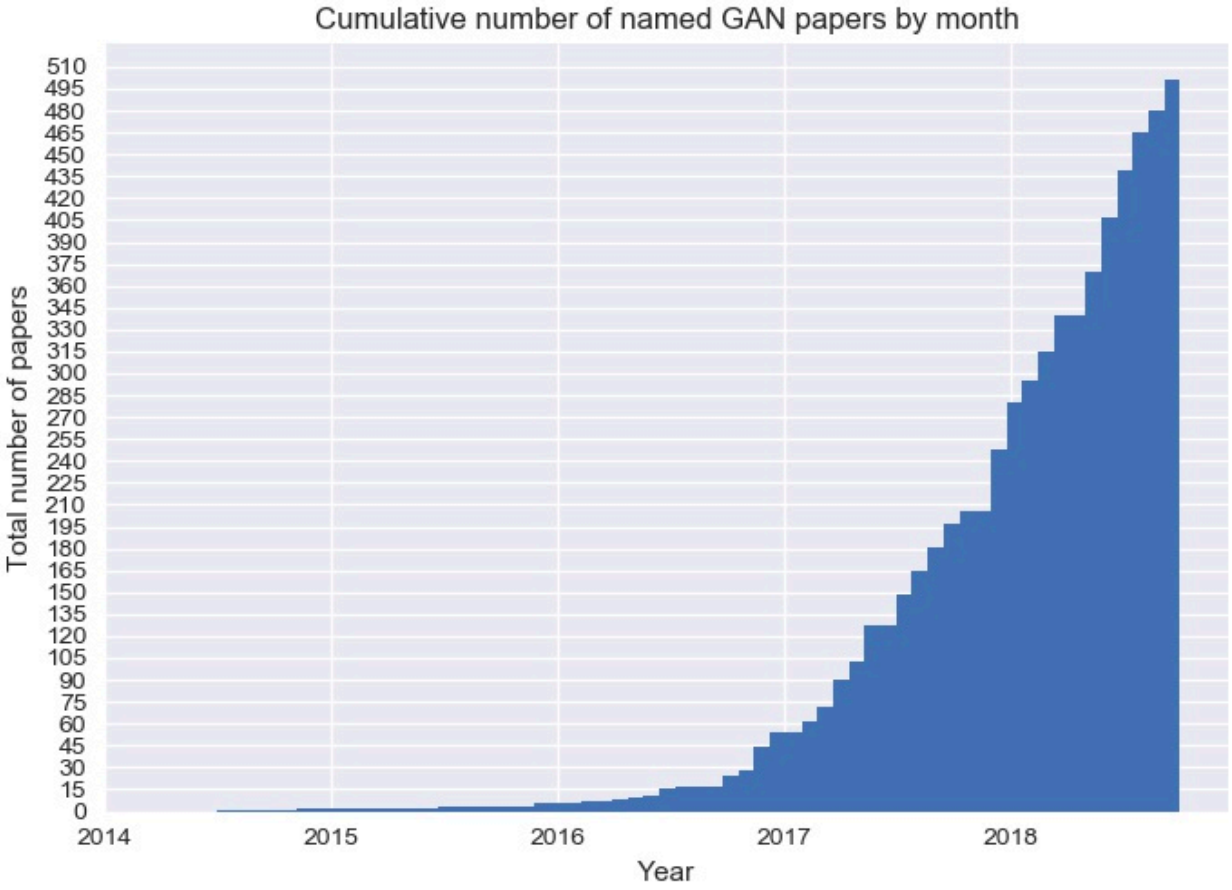
https://arxiv.org/abs/1710.10196

# Progressive-GAN




https://arxiv.org/abs/1710.10196

# and WAY MORE




https://github.com/hindupuravinash/the-gan-zoo

# Thank you for listening