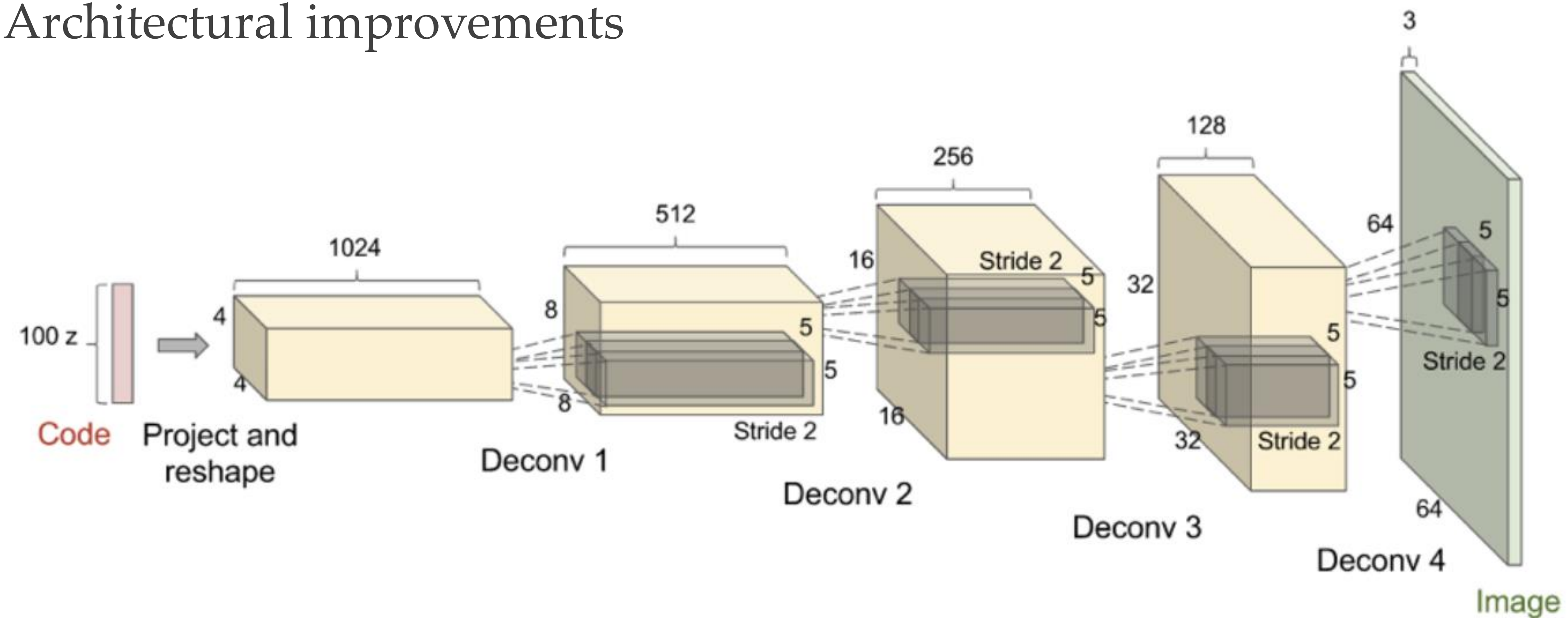
GAN models

Picture: *These people are not real – they were produced by our generator that allows control over different aspects of the image.*

[StyleGan]

DCGAN

- One of the first scaling ups of GANs
  - Architectural improvements



Radford, Metz, Chintala, Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks

Examples

Even vector space arithmetics …

Man with glasses

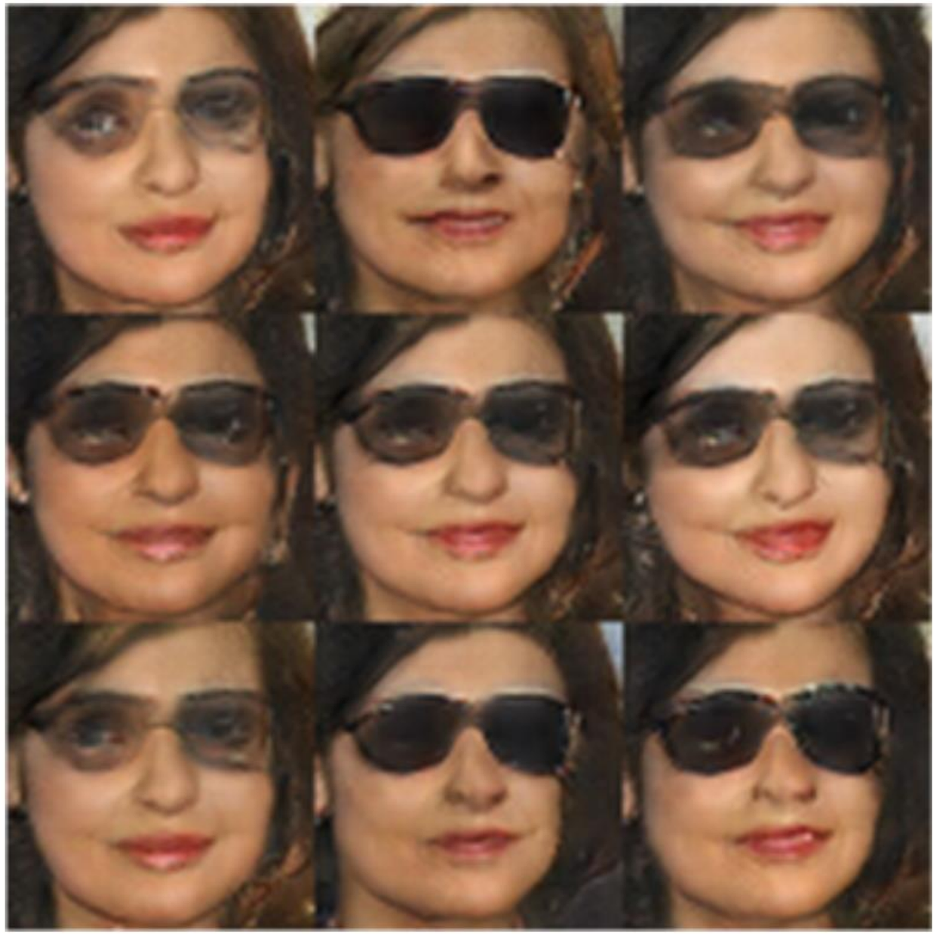
-

Man

+

Woman

=



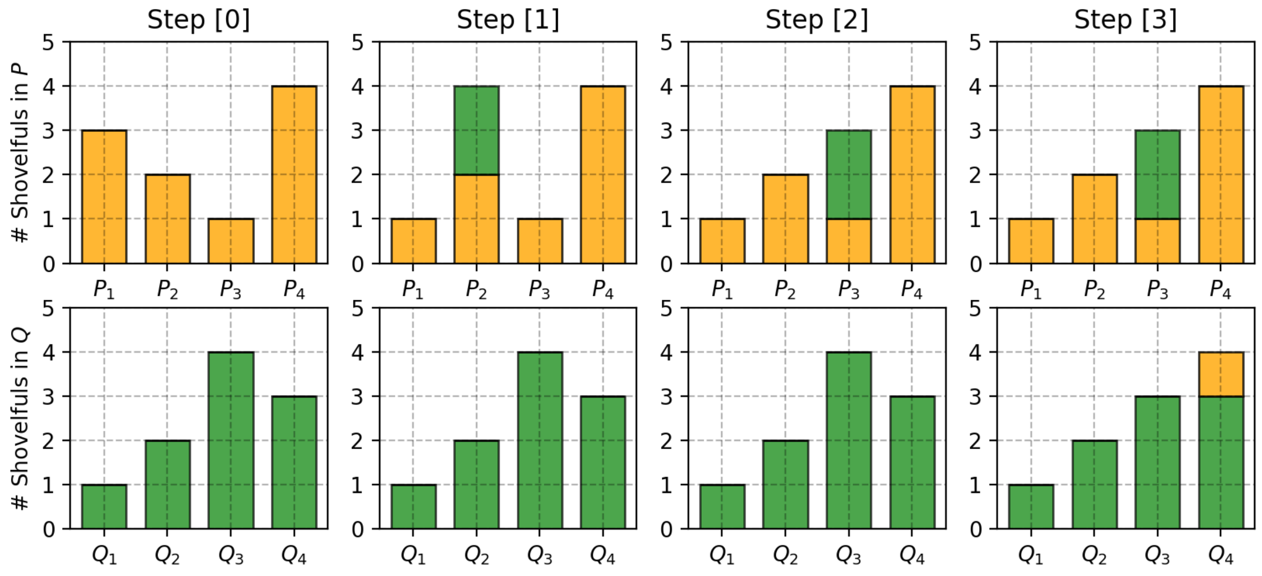
Woman with glasses

Wasserstein GAN

- Instead of KL/JS, use Wasserstein (Earth Mover's) Distance

$$W\left(p_r, p_g\right) = \inf_{\gamma \sim \Pi(\mathrm{p_r}, \mathrm{p_g})} \mathrm{E}_{(\mathrm{x,y}) \sim \gamma} |x - y|$$

- Even for non-overlapping supports, the distance is meaningful

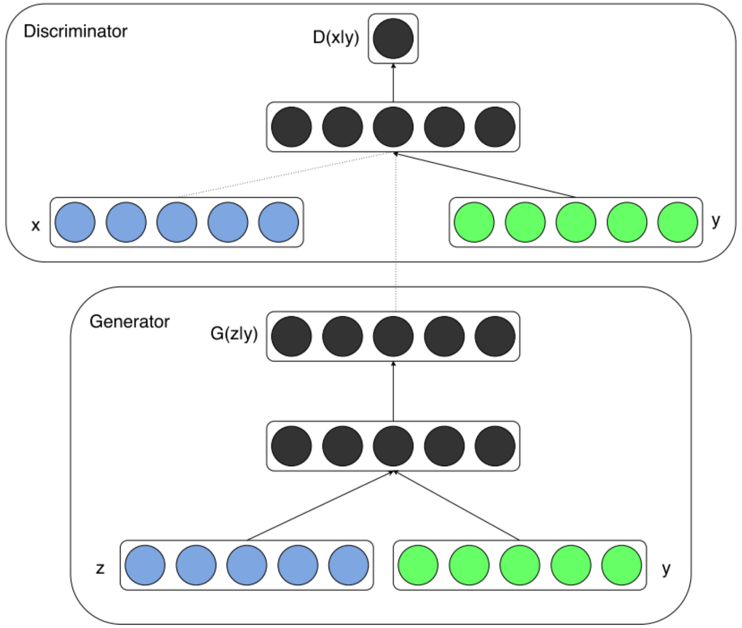


Arjovsky, Chintala, Bottou, Wasserstein GAN

Conditional GAN

- Conditioning on labels
  - Appending label vector to noise vector

$$\min_G \max_D \mathbb{E}_{\boldsymbol{x}\sim p_{data}}[\log D(\boldsymbol{x}|\boldsymbol{y})] + \mathbb{E}_{\boldsymbol{z}\sim p(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z}|\boldsymbol{y})))]$$



	User tags + annotations	Generated tags
	montanha, trem, inverno, frio, people, male, plant life, tree, structures, transport, car	taxi, passenger, line, transportation, railway station, passengers, railways, signals, rail, rails
	food, raspberry, delicious, homemade	chicken, fattening, cooked, peanut, cream, cookie, house made, bread, biscuit, bakes
	water, river	creek, lake, along, near, river, rocky, treeline, valley, woods, waters
	people, portrait, female, baby, indoor	love, people, posing, girl, young, strangers, pretty, women, happy, life

Table 2: Samples of generated tags

Mirza and Osindero, Conditional Generative Adversarial Nets

Image to image translation

- Conditioning GAN on other images (like edges)

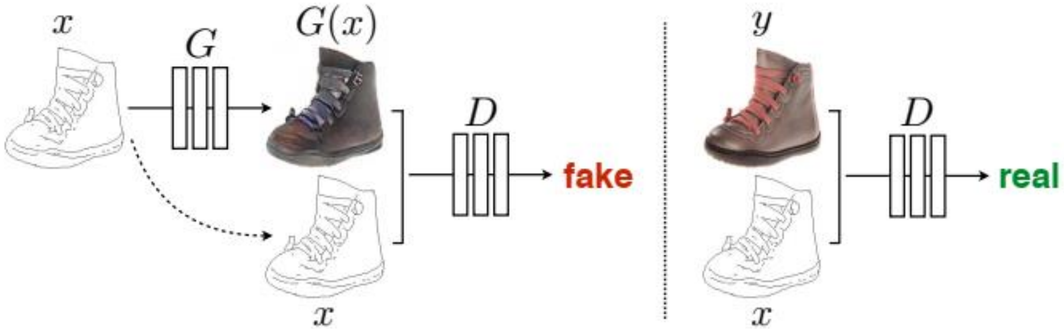


Figure 2: Training a conditional GAN to map edges→photo. The discriminator, D, learns to classify between fake (synthesized by the generator) and real {edge, photo} tuples. The generator, G, learns to fool the discriminator. Unlike an unconditional GAN, both the generator and discriminator observe the input edge map.

Figure 4: Different losses induce different quality of results. Each column shows results trained under a different loss. Please see `https://phillipi.github.io/pix2pix/` for additional examples.

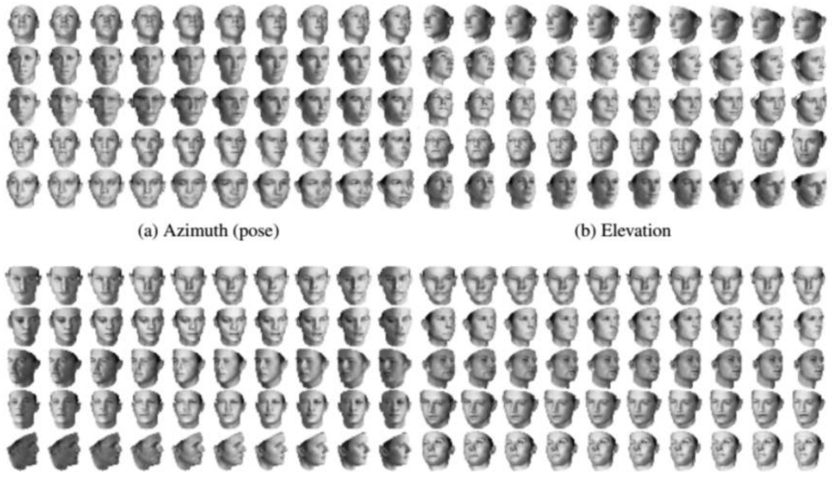
Isola, Zhu, Zhou, Efros, Image-to-Image Translation with Conditional Adversarial Networks

InfoGAN

- Generator takes as input noise $\boldsymbol{z}$ and a latent code $\boldsymbol{c}$
- Add mutual information as regularization

$$\min_G \max_D \mathbb{E}_{\boldsymbol{x}\sim p_{data}}[\log D(\boldsymbol{x}|\boldsymbol{y})] + \mathbb{E}_{\boldsymbol{z}\sim p(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z}|\boldsymbol{y})))] - \lambda I(c, G(\boldsymbol{z}, \boldsymbol{c}))$$

  - Requiring high mutual information between the latent code and the generation discourages learning trivial latent codes
- As the mutual information requires the true posterior, a variational bound is used instead



(a) Azimuth (pose) (b) Elevation (c) Lighting (d) Wide or Narrow

Chen et al., InfoGAN: Interpretable Representation Learning by Information Maximizing Generative Adversarial Nets

CycleGAN

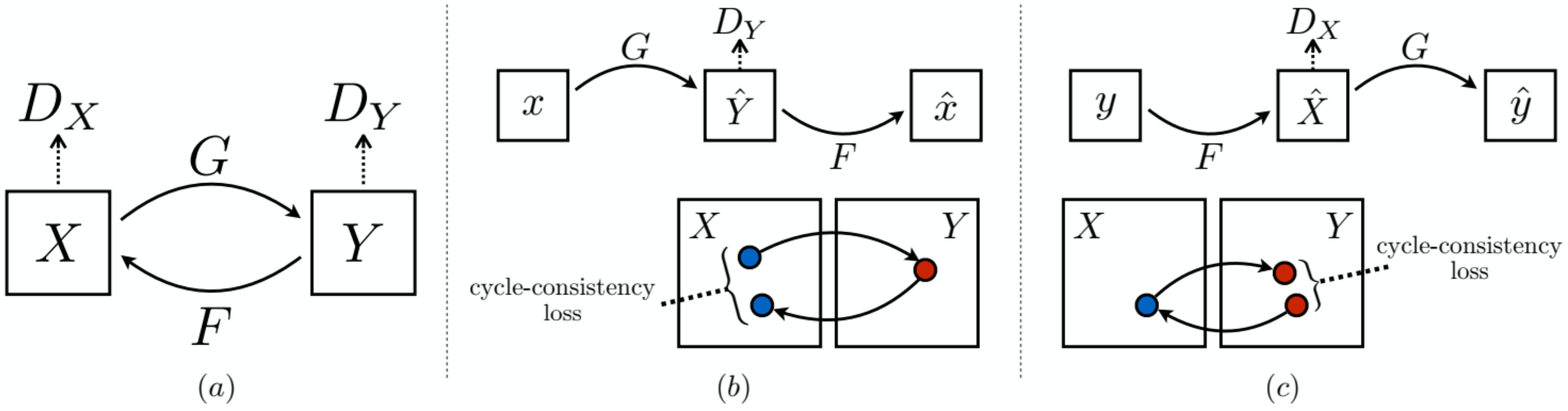
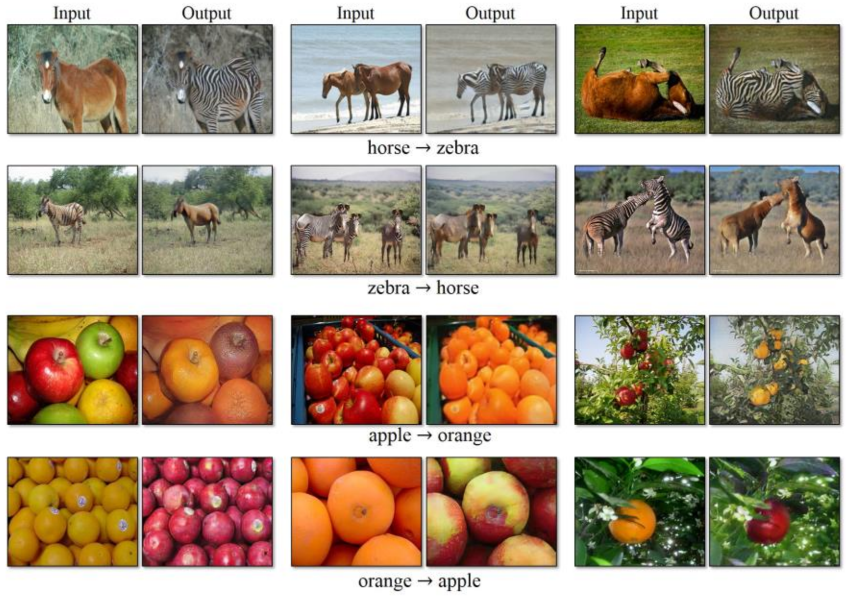


Figure 3: (a) Our model contains two mapping functions $G : X \rightarrow Y$ and $F : Y \rightarrow X$, and associated adversarial discriminators D_Y and D_X. D_Y encourages G to translate X into outputs indistinguishable from domain Y, and vice versa for D_X and F. To further regularize the mappings, we introduce two *cycle consistency losses* that capture the intuition that if we translate from one domain to the other and back again we should arrive at where we started: (b) forward cycle-consistency loss: $x \rightarrow G(x) \rightarrow F(G(x)) \approx x$, and (c) backward cycle-consistency loss: $y \rightarrow F(y) \rightarrow G(F(y)) \approx y$



Zhu, Park, Isola, Efros, Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks

StyleGAN

- Architectural innovations and scaling up
- Impressive generations

Picture: *These people are not real – they were produced by our generator that allows control over different aspects of the image.*

StyleGan

Summary

- Implicit density models: Motivation
- Generative adversarial networks
- Challenges
- GAN models

<u>Reading material:</u>

- All papers mentioned in the slides