Generative Adversarial Networks (GANs)

Adapted from material by Goodfellow, Binglin, Shashank, Bhargav

GANs

- **Generative**
  - Learn a generative model
- **Adversarial**
  - Trained in an adversarial setting
- **Networks**
  - Use Deep Neural Networks

Why Generative Models?

- **Discriminative models:**
  - Given an image **X**, predict a label **Y**
  - Estimates **P(Y|X)** Text

- **Discriminative models have several key limitations**
  - Can't model **P(X)**, i.e. the probability of seeing a certain image
  - Thus, can't sample from **P(X)**, i.e. **can't generate new images**

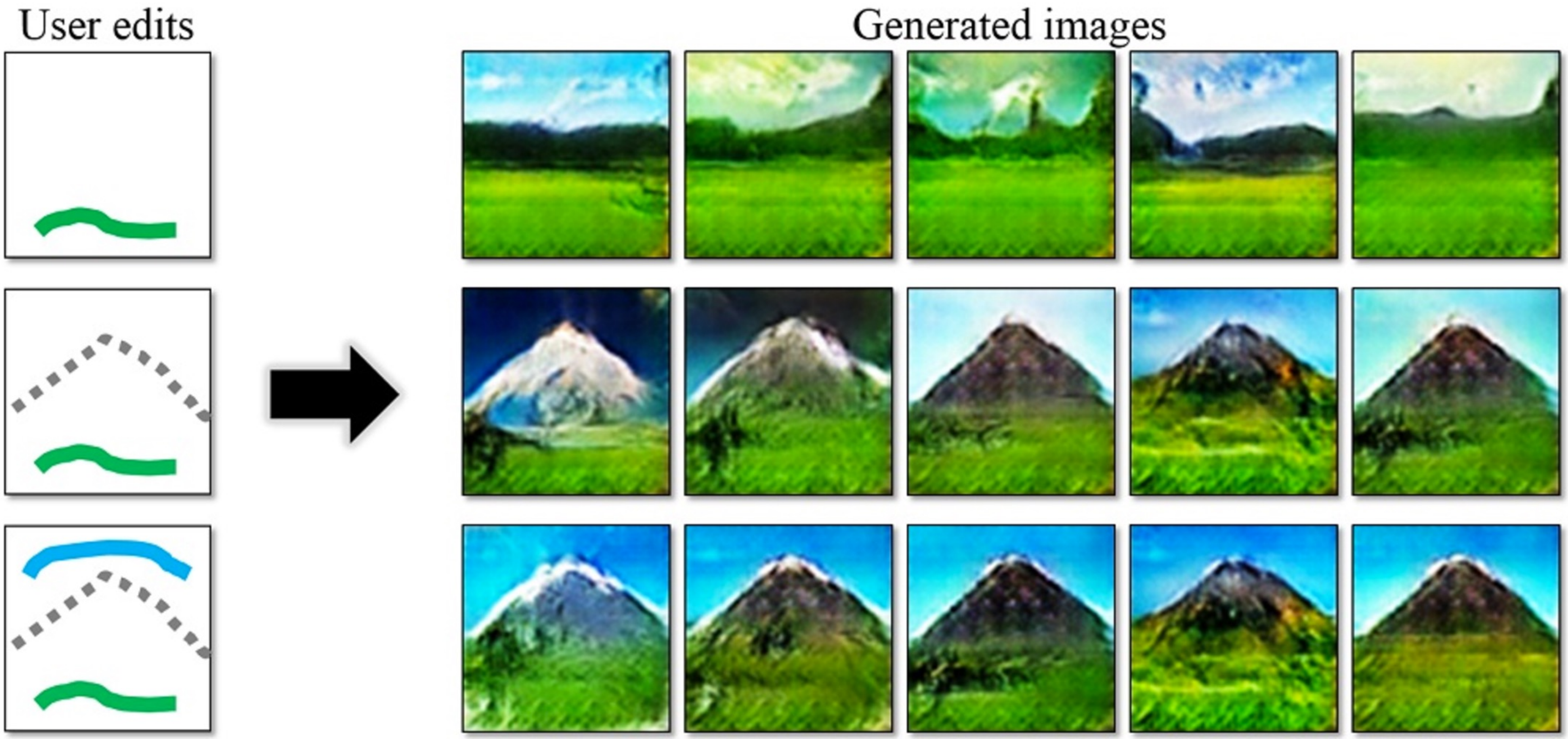
- **Generative models try to address these:**
  - model **P(X)**
  - generate new data (e.g. images)

Magic of GANs…

Which one is Computer generated?

Ledig, Christian, et al. "Photo-realistic single image super-resolution using a generative adversarial network." *arXiv preprint arXiv:1609.04802* (2016).

Magic of GANs…



http://people.eecs.berkeley.edu/~junyanz/projects/gvm/

Adversarial Training

- Generator: generate fake samples, tries to fool the Discriminator

fake mean computer generated, not human generated

- Discriminator: tries to distinguish between real and fake samples

real as in human genrated

- Train them against each other

- Repeat this and we get better Generator and Discriminator

GAN’s Architecture

sampling process give to discrimation

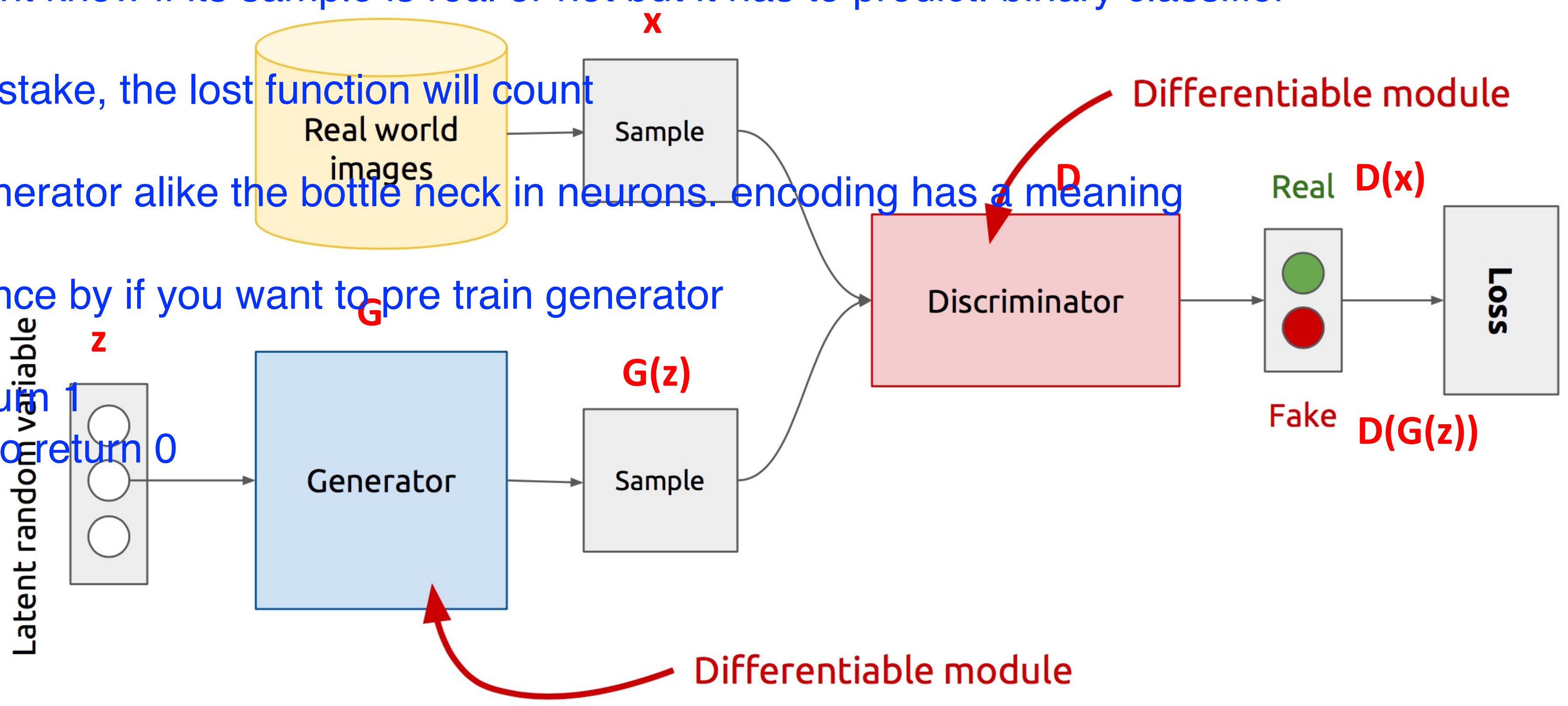
discrimation dont know if its sample is real or not but it has to predict. binary classifier

if it makes a mistake, the lost function will count

can think of generator alike the bottle neck in neurons. encoding has a meaning

archi not influence by if you want to pre train generator

if x, want to return 1
if z, then want to return 0



random in this latent space,
but has encoding,
has meaning

- **Z** is some random vector (Gaussian/Uniform).
- **Z** can be thought as the latent representation of the image.

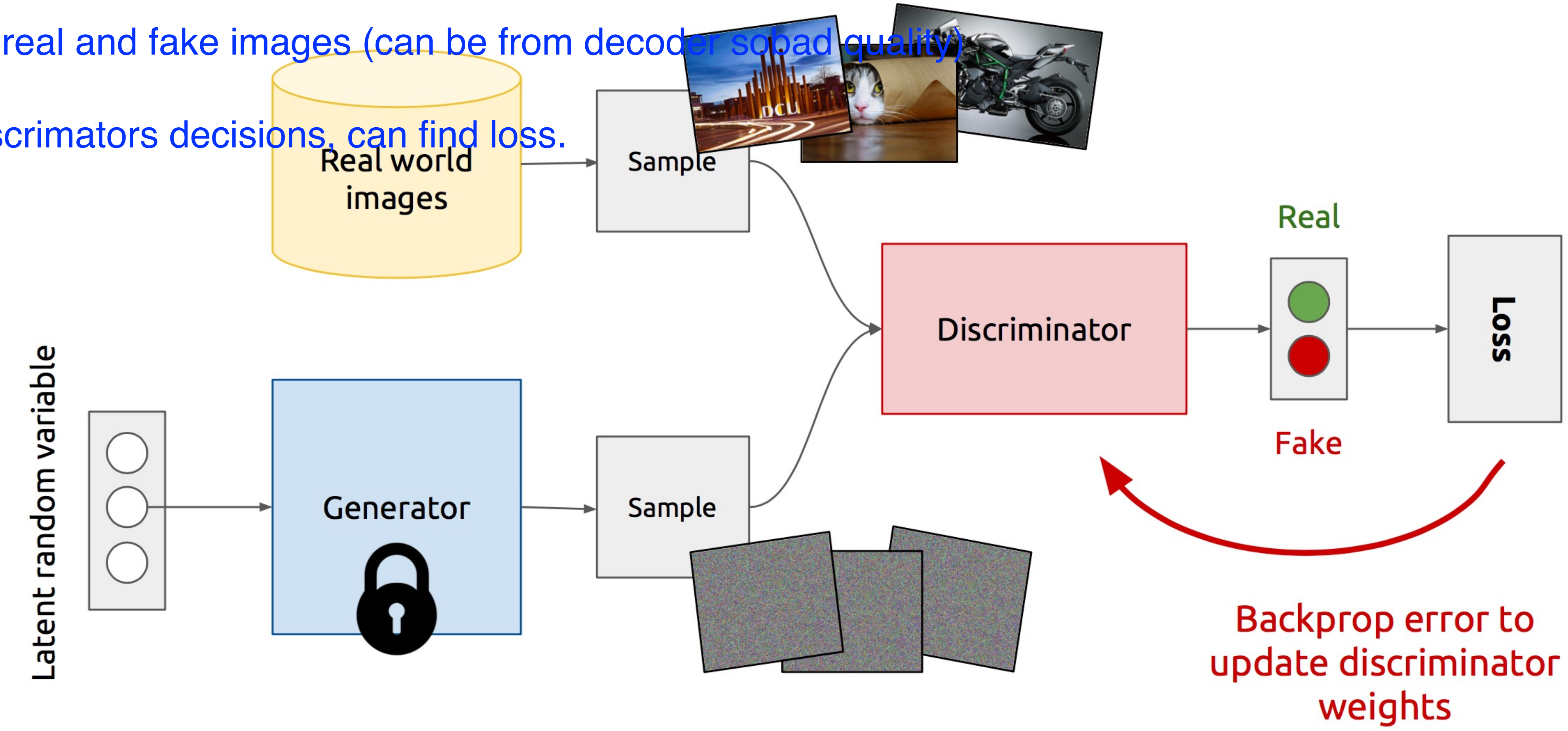
https://www.slideshare.net/xavigiro/deep-learning-for-computer-vision-generative-models-and-adversarial-training-upc-2016

Training Discriminator

assume generator is lock

sample real and fake images (can be from decoder sobad quality)

from discriminators decisions, can find loss.

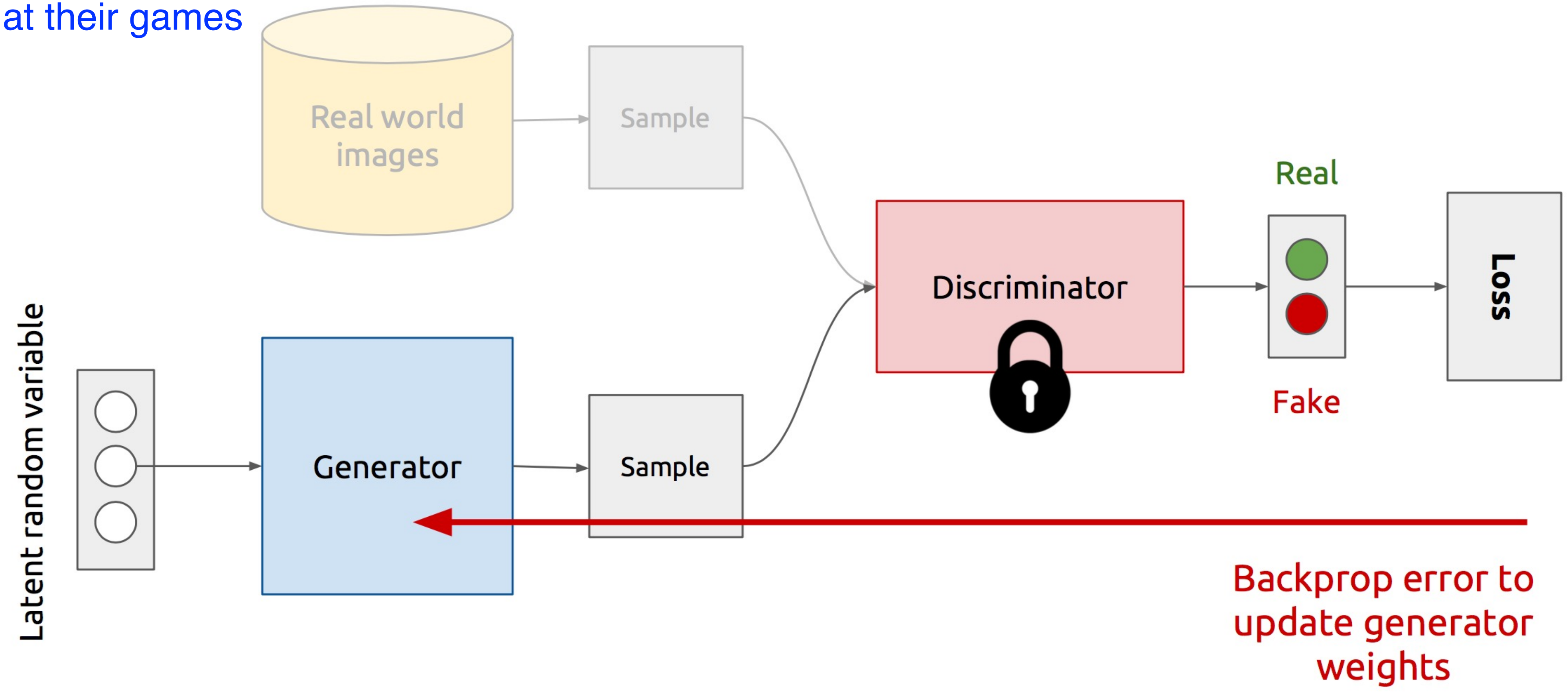


https://www.slideshare.net/xavigiro/deep-learning-for-computer-vision-generative-models-and-adversarial-training-upc-2016

Training Generator

traiing generator part can be updated by fixing discrimator

adjust wegiht so generator produce values that discrimator can can make a mistake. repeat until both become good at their games



https://www.slideshare.net/xavigiro/deep-learning-for-computer-vision-generative-models-and-adversarial-training-upc-2016

GAN's formulation

$$\min_G \max_D V(D, G)$$

- It is formulated as a **minimax game**, where:
  - The Discriminator is trying to maximize its reward $\boldsymbol{V(D, G)}$
  - The Generator is trying to minimize Discriminator's reward (or maximize its loss)

G want to minimize reward for D
by producing realistic output

reward opposite of loss

overall reward function

$$V(D, G) = \boxed{\mathbb{E}_{x \sim p(x)}[\log D(x)]} + \boxed{\mathbb{E}_{z \sim q(z)}\left[\log\left(1 - D(G(z))\right)\right]}$$

if x coming from x, then E[logD(x)] want to return 1
on other hand, if z is lantent variable from the fake, D decide real or fake, want it to be able to determine

- The Nash equilibrium of this particular game is achieved at:
  - $P_{data}(x) = P_{gen}(x) \;\; \forall x$
  - $\mathrm{D}(x) = \frac{1}{2} \;\; \forall x$

as they become better at their games, will reach optimal players
so D output 0.5, cant distinguish 50% 50% chance

Algorithm 1 Minibatch stochastic gradient descent training of generative adversarial nets. The number of steps to apply to the discriminator, k, is a hyperparameter. We used $k = 1$, the least expensive option, in our experiments.

Discriminator updates

Generator updates

for number of training iterations **do**

  for k steps **do**

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

$$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^{m} \left[\log D\left(\boldsymbol{x}^{(i)}\right) + \log\left(1 - D\left(G\left(\boldsymbol{z}^{(i)}\right)\right)\right) \right].$$

  end for

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

$$\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^{m} \log\left(1 - D\left(G\left(\boldsymbol{z}^{(i)}\right)\right)\right).$$

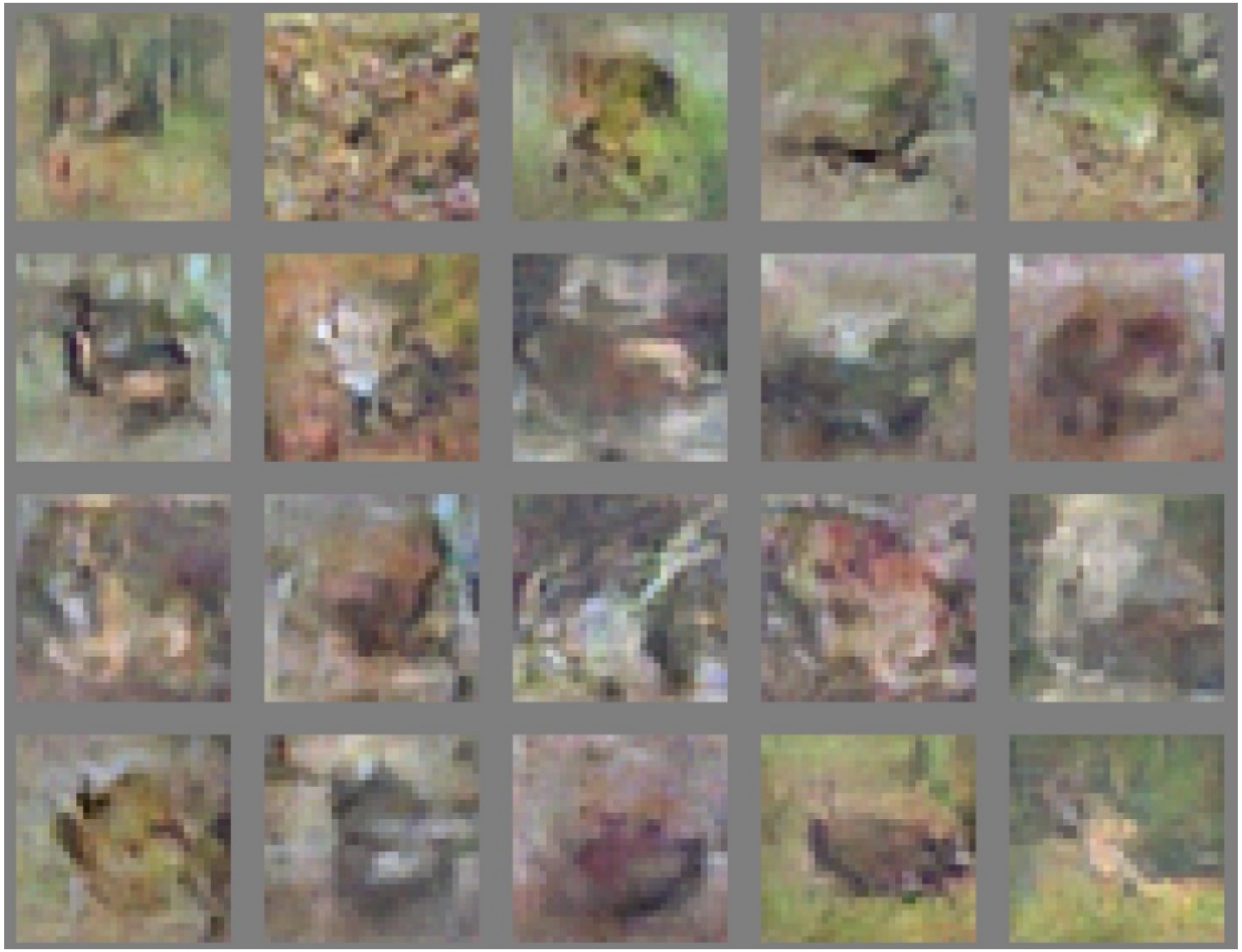
end for

The gradient-based updates can use any standard gradient-based learning rule. We used momentum in our experiments.

Faces

Goodfellow, Ian, et al. **"Generative adversarial nets."** *Advances in neural information processing systems*. 2014.

CIFAR



Goodfellow, Ian, et al. **"Generative adversarial nets."** *Advances in neural information processing systems*. 2014.

DCGAN: Bedroom images

Radford, Alec, Luke Metz, and Soumith Chintala. "Unsupervised representation learning with deep convolutional generative adversarial networks." arXiv:1511.06434 (2015).

Generative Adversarial Nets: Interpretable Vector Math

Radford et al, ICLR 2016

Smiling woman Neutral woman Neutral man

feed z to g

Samples from the model

Smiling Man

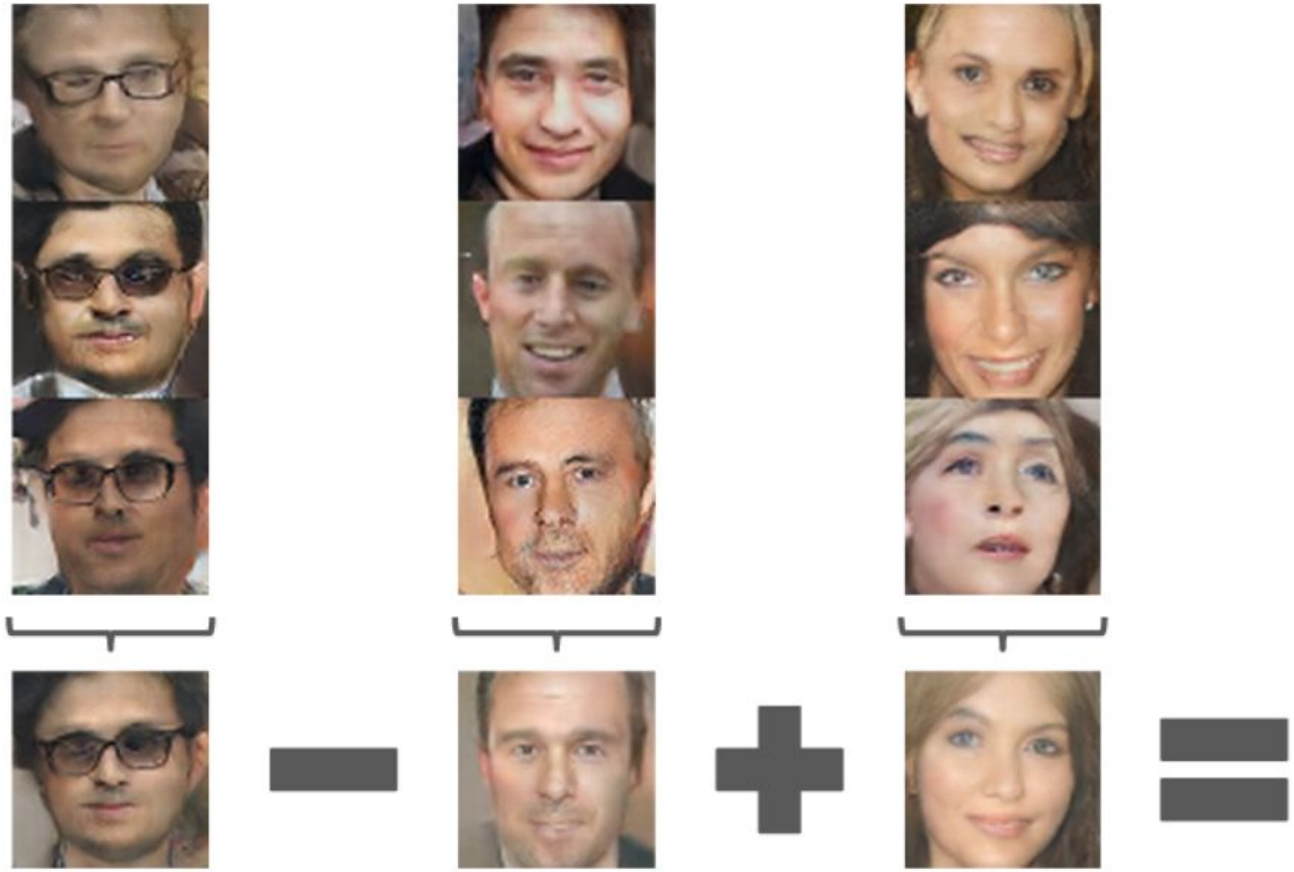
Average Z vectors, do arithmetic

vector z to get the meaning of smiling, which is added to the avg neutral man

Generative Adversarial Nets: Interpretable Vector Math

Radford et al, ICLR 2016

Glasses man　No glasses man　No glasses woman



Woman with glasses