Photo Editing With Generative Adversarial Networks (GANs)

GTC, May 2017. Greg Heinrich.

NVIDIA.

GAN: WHAT IS A GENERATIVE MODEL?

In Machine Learning

A generative model learns to generate samples that have the same characteristics as the samples in the dataset.

Learn from Shakespeare novels:

http://karpathy.github.io/2015/05/21/rnn-effectiveness/

Produce:

```
PANDARUS:

Alas, I think he shall be come approached and the day

When little srain would be attain'd into being never fed,

And who is but a chain and subjects of his death,

I should not sleep.
```

BASIC REMINDER: BACKPROP

Calculating $\frac{\partial E}{\partial w_{ij}^l}$ iteratively

Output of each neuron j of layer l :

$$h_j^l = \varphi(z_j^l) = \varphi\left(\sum_i w_{ij}^l h_i^{l-1} + b_j^l\right)$$

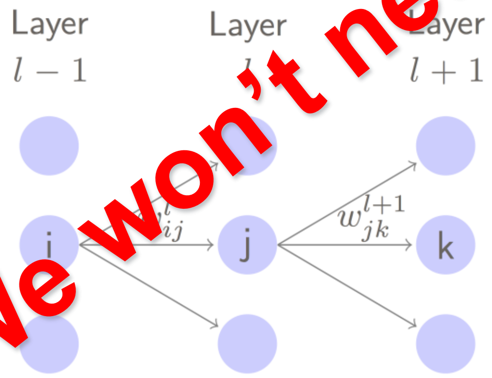
Gradient of E with respect to each weight:

$$\frac{\partial E}{\partial w_{ij}^l} = \frac{\partial E}{\partial z_j^l}\frac{\partial z_j^l}{\partial w_{ij}^l} = \frac{\partial E}{\partial z_j^l} h_i^{l-1}$$

Chain rule

$\frac{\partial z_j^l}{\partial w_{ij}^l}$ only depends on h_i^{l-1}

Calculated during forward prop



Calculation of $\frac{\partial E}{\partial z_j^l}$:

$$\frac{\partial E}{\partial z_j^l} = \sum_k \frac{\partial E}{\partial z_k^{l+1}}\frac{\partial z_k^{l+1}}{\partial z_j^l} = \sum_k \frac{\partial E}{\partial z_k^{l+1}}\frac{\partial z_k^{l+1}}{\partial h_j^l}\frac{\partial h_j^l}{\partial z_j^l}$$

$$= \sum_k \frac{\partial E}{\partial z_k^{l+1}} w_{jk}^{l+1}\varphi'(z_j^l)$$

$$= \varphi'(z_j^l)\sum_k \frac{\partial E}{\partial z_k^{l+1}} w_{jk}^{l+1}$$

Chain rule

Multivariate chain rule

We won’t need this today

GAN: PLAYING THE ADVERSARIAL GAME

Learning on a corpus of images

Let's play a game opposing two agents:

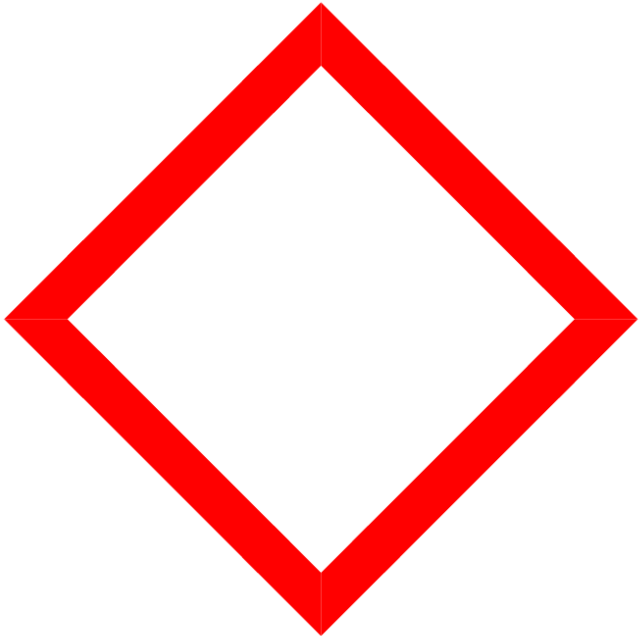
- The Generator, a little imp in the computer who paints images.
- The Discriminator: you are collectively responsible for playing the Discriminator.

The game master (me) randomly picks images from either the corpus or the Generator and shows them to the Discriminator. The goal of the Discriminator is to identify the source of the images: real (from the corpus) or fake (painted by the little imp). The goal of the Generator is to fool the Discriminator.

PLAYING THE ADVERSARIAL GAME

Is this a veelhoek* from our corpus?

Note: you don't have to know what a *veelhoek* is, you will learn through examples!

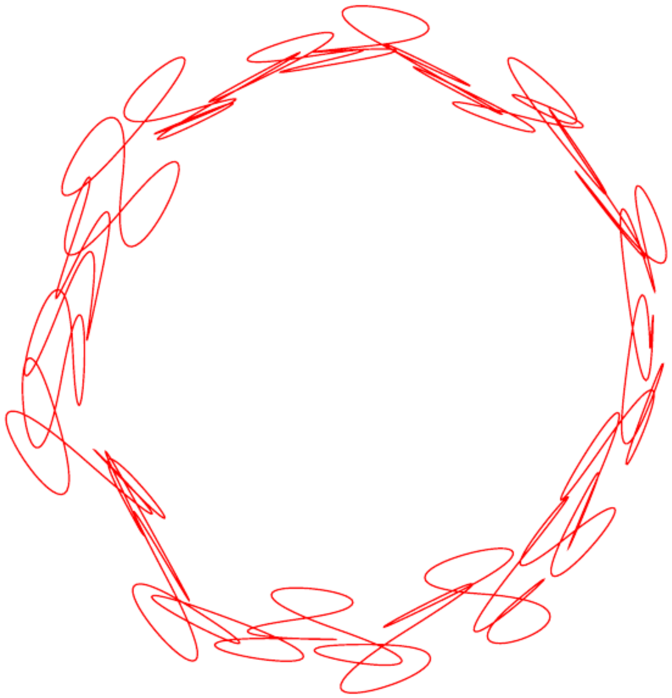


* *veelhoek* is the articulation of a ubiquitous item in the language of a tiny country in Europe that is well known for the inferior quality of its cheese.

Yes, this red square is a veelhoek!

PLAYING THE ADVERSARIAL GAME

Is this a veelhoek from our corpus?



No, those squiggly lines aren't right!

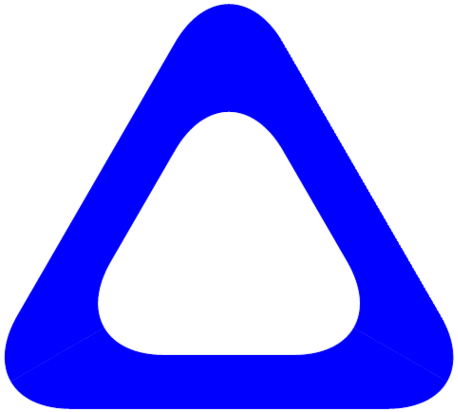
PLAYING THE ADVERSARIAL GAME

Is this a veelhoek from our corpus?

Yes, even though it’s blue and tiny!

PLAYING THE ADVERSARIAL GAME

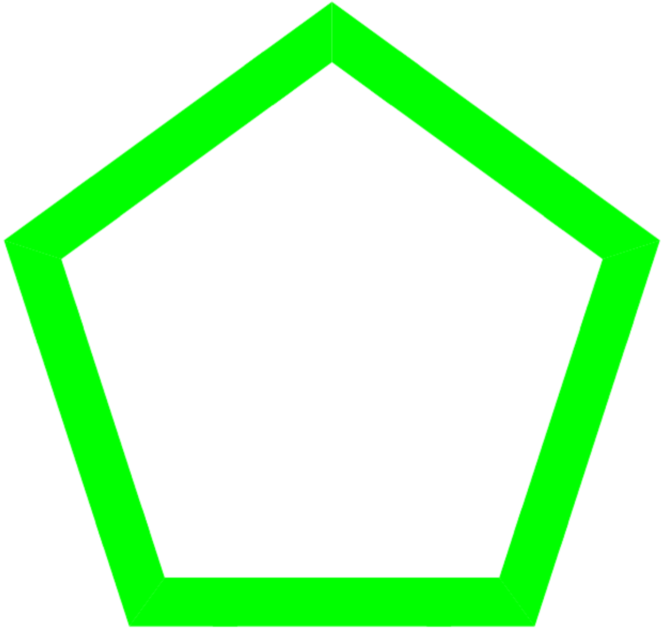
Is this a veelhoek from our corpus?



No, those rounded corners are a giveaway!

PLAYING THE ADVERSARIAL GAME

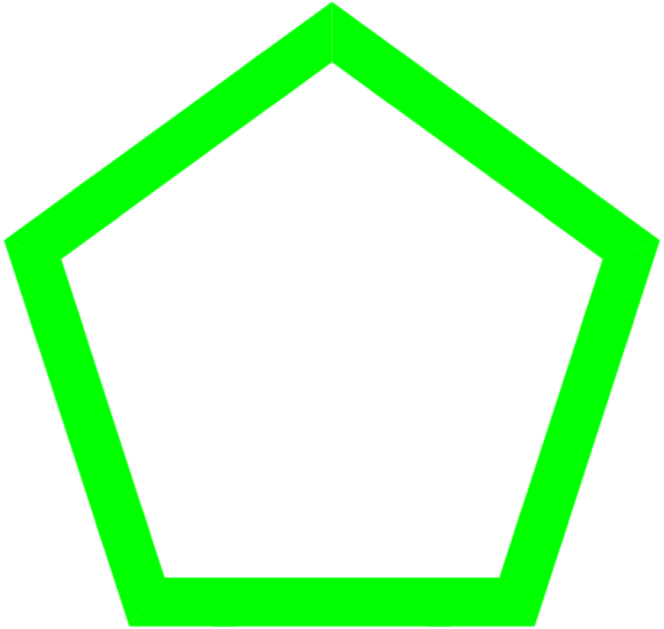
Is this a veelhoek from our corpus?



No, but it’s a very good fake!

PLAYING THE ADVERSARIAL GAME

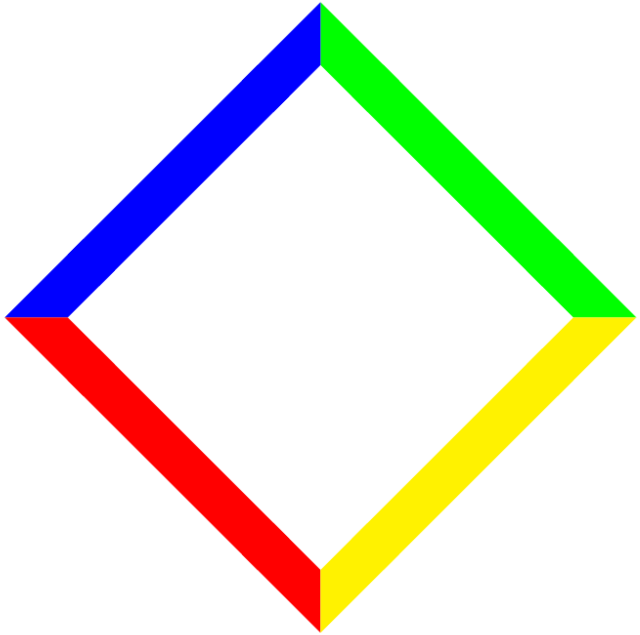
Is this a veelhoek from our corpus?



No, it’s the same fake as before!

PLAYING THE ADVERSARIAL GAME

Is this a veelhoek from our corpus?



No, but it's a very creative fake!

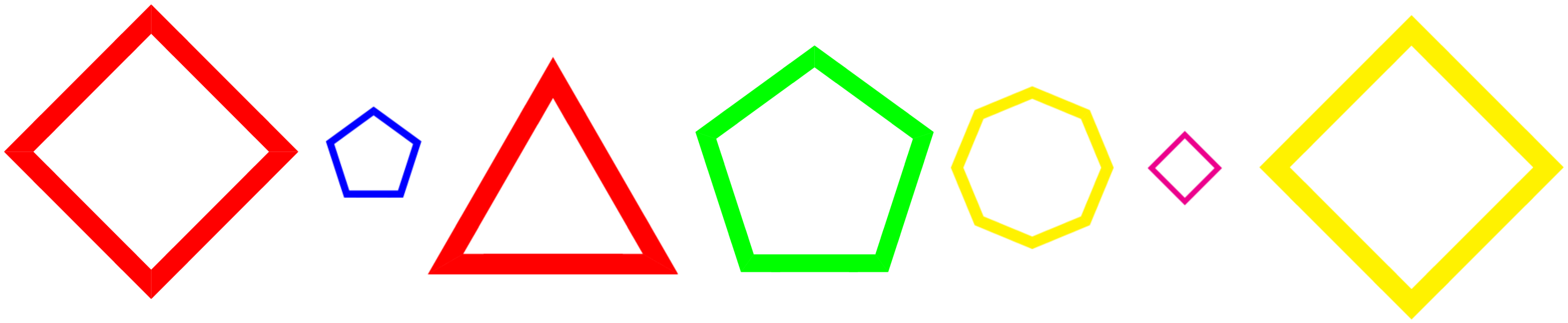
THE LATENT REPRESENTATION

From features to images

A veelhoek is characterized by three features:
- colour,
- size,
- number of faces

This set of features is known as the “**LATENT REPRESENTATION**”.

We can generate many real-looking veelhoeks by randomly picking reasonable values of each feature:



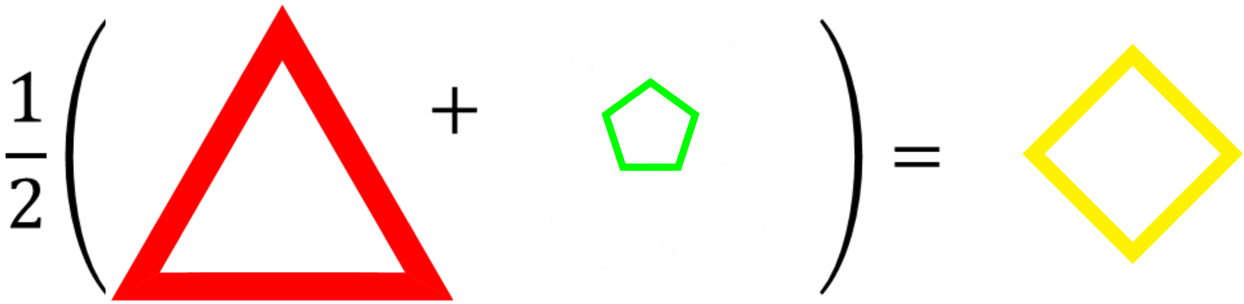
THE LATENT REPRESENTATION

Arithmetic in latent space

We can perform operations in latent space, have them reflected in feature space:

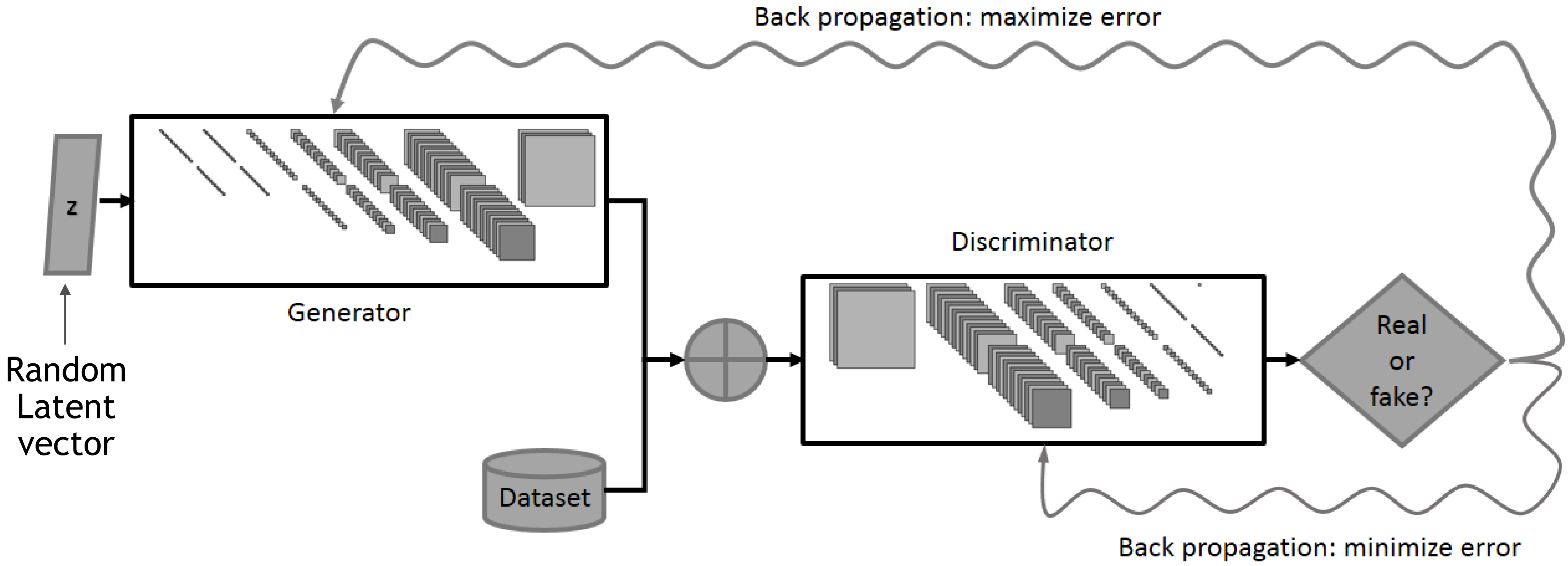
$$\frac{1}{2}\left(\begin{bmatrix} large \\ red \\ 3\ faces \end{bmatrix} + \begin{bmatrix} small \\ green \\ 5\ faces \end{bmatrix}\right) = \begin{bmatrix} medium \\ yellow \\ 4\ faces \end{bmatrix}$$

Equivalently:



THE GAN SET-UP

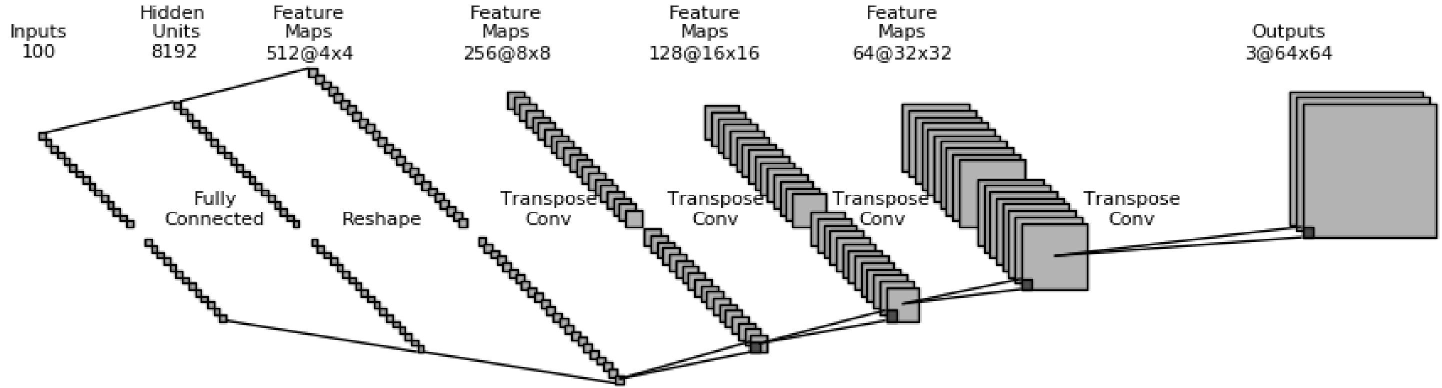
Connecting the Discriminator to the Generator and the Dataset



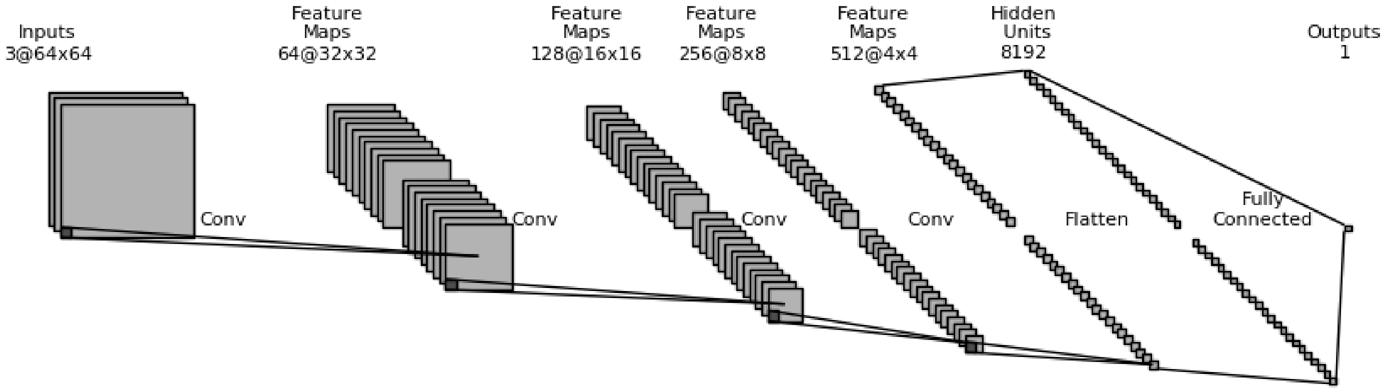
GAN: NETWORK TOPOLOGY

Radford (2015). Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *arXiv:1511.06434*

Generator



Discriminator



TRAINING A GAN ON CELEBRITY FACES*

Generating new faces by picking random values of the latent vector

* CelebFaces dataset

ANALOGIES

man is to *woman* as *king* is to queen

Reproduction of the famous "king + woman - man = queen" analogy on faces:

	Man	Blond Hair	Blue Eyes	Smile	Looking Left	Pointy Nose
Top Right	-	-	-	+	+	+
Bottom Left	+	+	+	-		+
Subtract Top Left	+	-	-	-		-
Bottom Right	+	-	-	-	+	+

MAPPING IMAGES TO LATENT VECTORS

Transfer learning: from Discriminator to Encoder

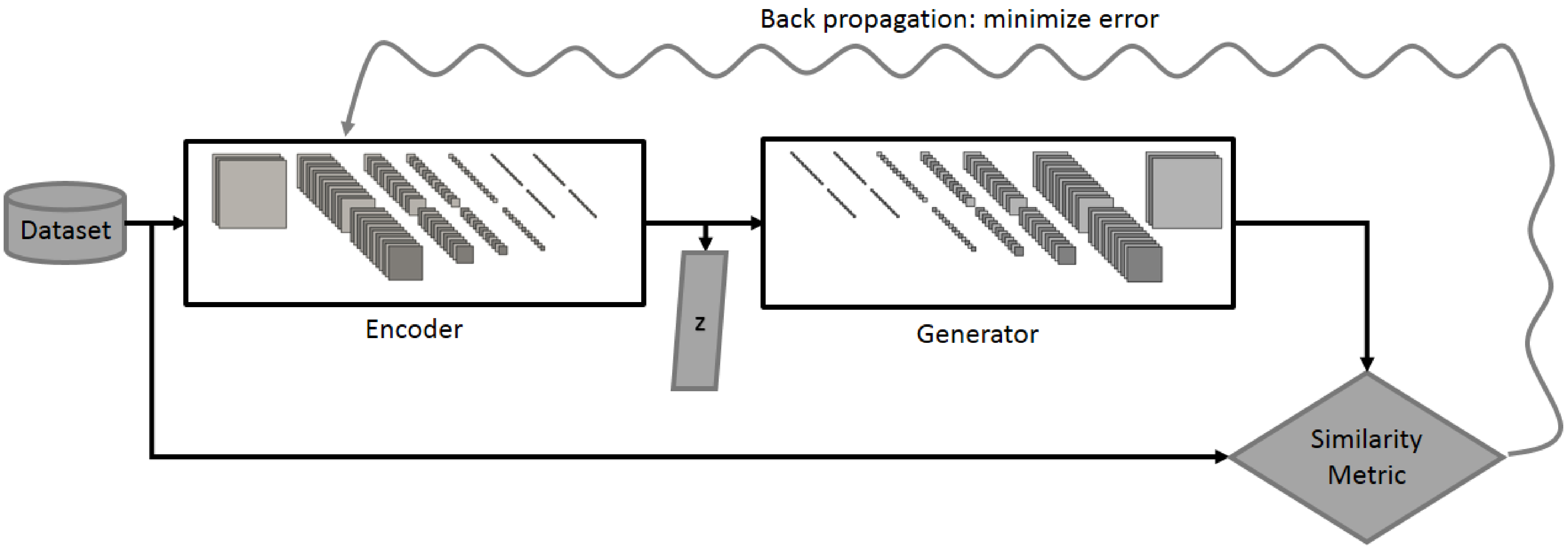
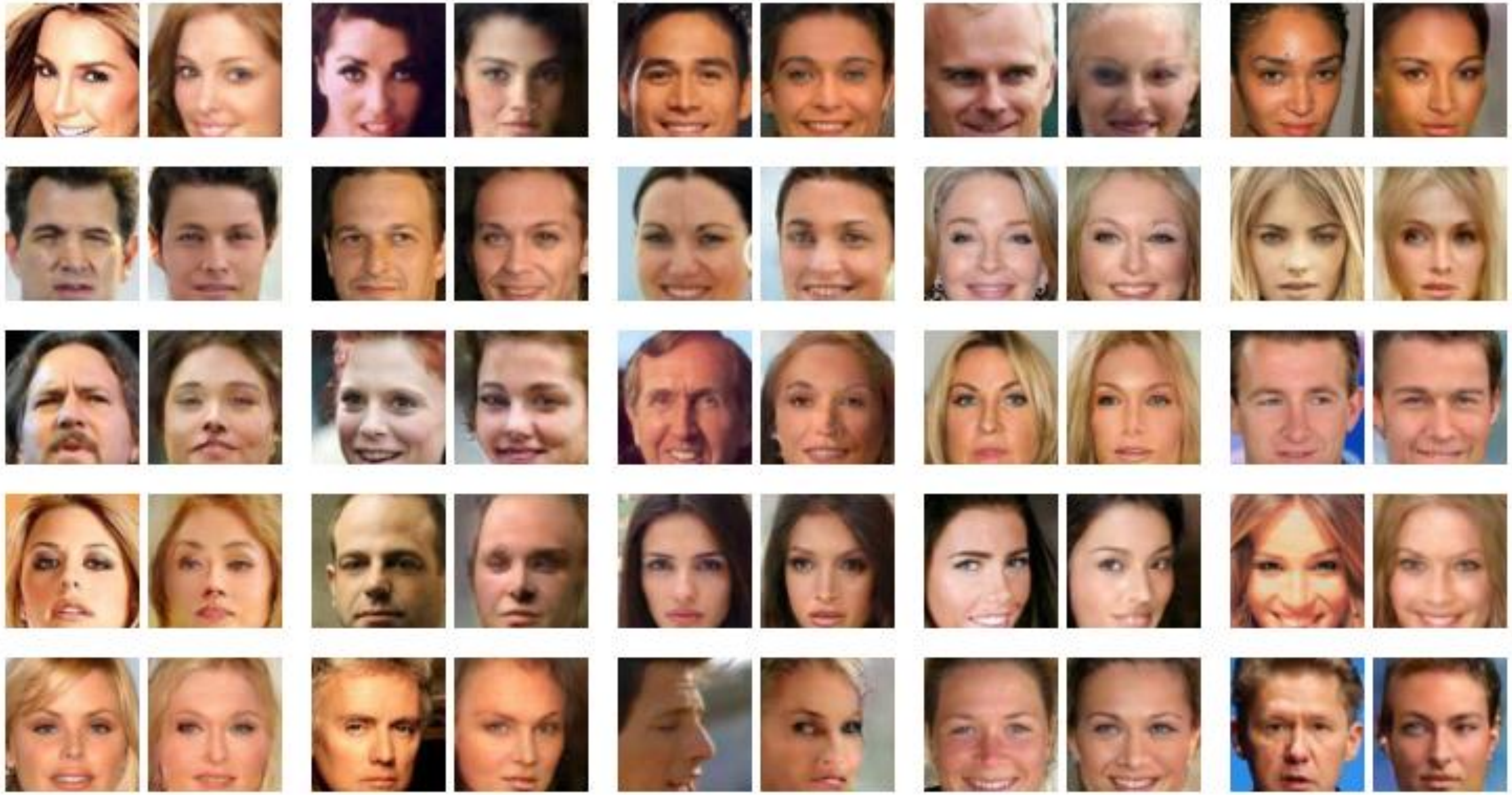


IMAGE RECONSTRUCTIONS

Visualizing $G(E(image))$



ATTRIBUTES

Calculating attribute vectors

The encoder E may be used to calculate the latent vector for each attribute.

For each $attr$ in $attributes$:

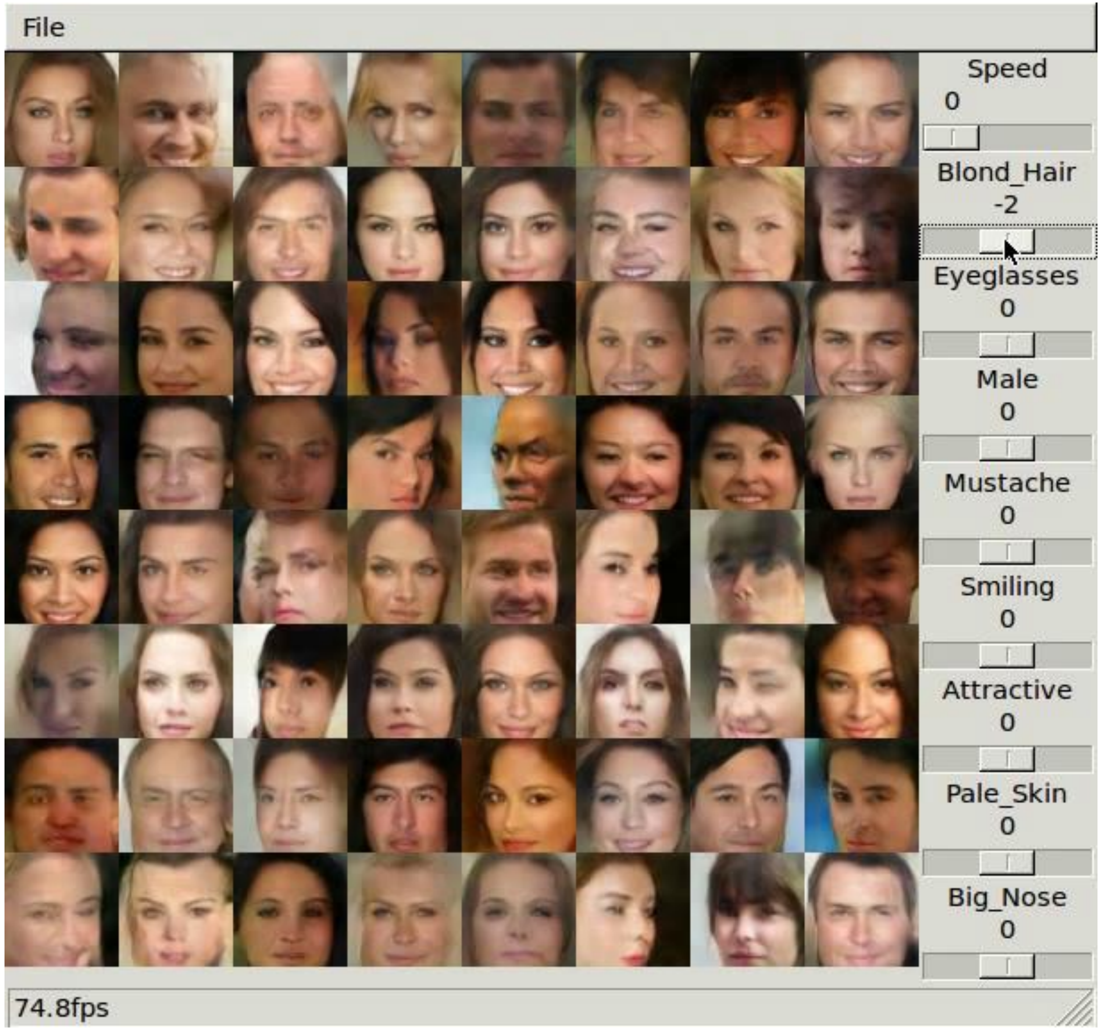
$I^{+}_{attr} = \{im|attr\}$ and $I^{-}_{attr} = \{im|\overline{attr}\}$ are sets of images w/wo the attribute

$$z(attr) = \frac{1}{\|I^{+}_{attr}\|}\sum_{im \in I^{+}_{attr}} E(im) - \frac{1}{\|I^{-}_{attr}\|}\sum_{im \in I^{-}_{attr}} E(im)$$

It is then straightforward to add or remove attributes from an image:

From left to right: original image (OI); OI + "young" attribute; OI - "blond hair" + "black hair"; OI - "smile"; OI + "male" + "bald".

PLAYING WITH ATTRIBUTES

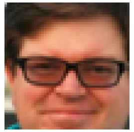


EXTRACTING ATTRIBUTES

...from portraits of illustrious people

- 0.68 Narrow_Eyes
- 0.48 Bangs
- 0.41 Wearing_Hat
- 0.33 Mouth_Slightly_Open
- 0.30 Chubby

- 1.64 Pale_Skin
- 1.28 Blond_Hair
- 1.15 Gray_Hair
- 1.06 No_Beard
- 0.74 Narrow_Eyes



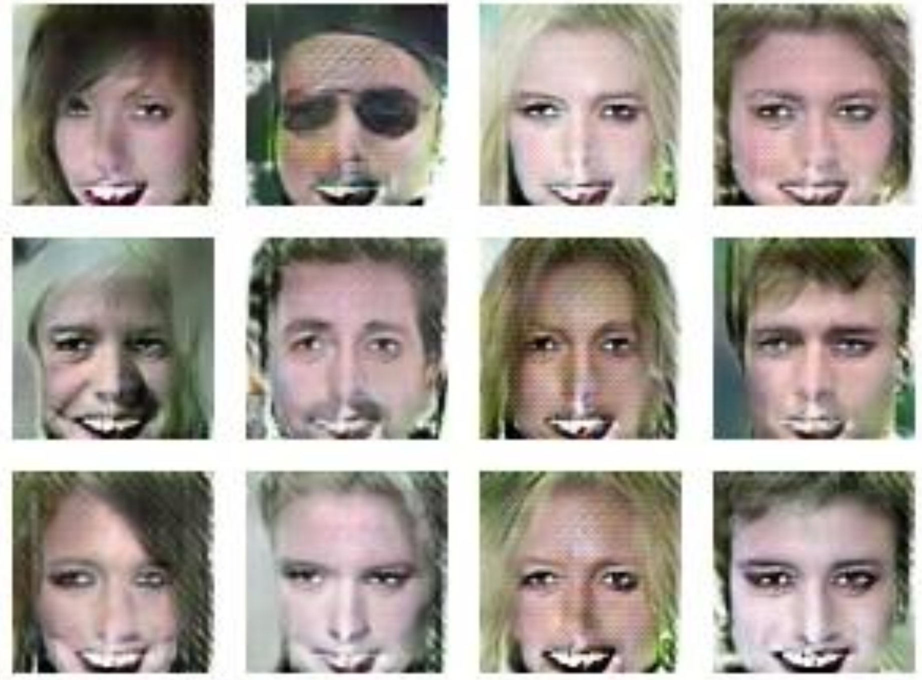
- 0.76 Male
- 0.65 Brown_Hair
- 0.56 Big_Nose
- 0.54 Eyeglasses
- 0.53 Wearing_Hat

- 2.82 Wearing_Hat
- 1.92 Blurry
- 1.48 Bangs
- 0.80 Gray_Hair
- 0.78 Pale_Skin

DEGENERATOR

Getting the essence of your dataset

After convergence, stop updating the discriminator:



DATASET VISUALIZATION

Projecting latent vectors on a sphere

THANK YOU

Questions?