



AVATAR SDK

itSeez3D

Introduction to Generative Adversarial Networks (GANs)

Ilya Lysenkov, itSeez3D CTO

NNSU
July 2019

Generative Adversarial Networks

- I. Goodfellow et al. *Generative Adversarial Networks*.
[arXiv:1406.2661](https://arxiv.org/abs/1406.2661)
- 10048 citations



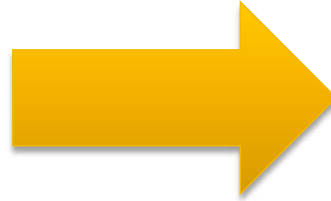
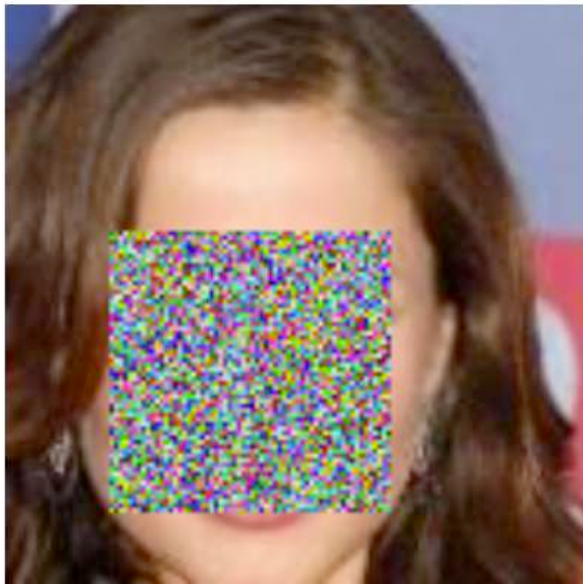
Yann LeCun: “This, and the variations that are now being proposed is the most interesting idea in the last 10 years in ML, in my opinion.”

– [Quora, July 2016](#)

Images sources:

<https://www.amazon.com/Ian-Goodfellow/e/B01MQGN8N0>
<https://bits.blogs.nytimes.com/2015/03/26/facebook-yann-lecun-discusses-digital-companions-and-artificial-intelligence/>

Toy problem



Обычный порошок

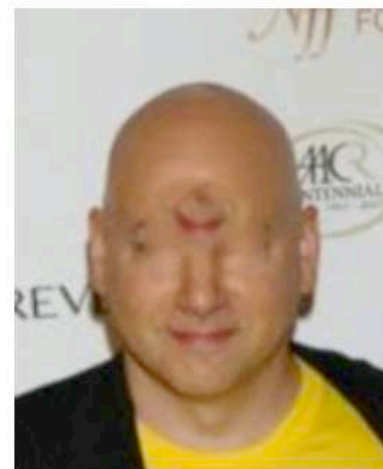
Input



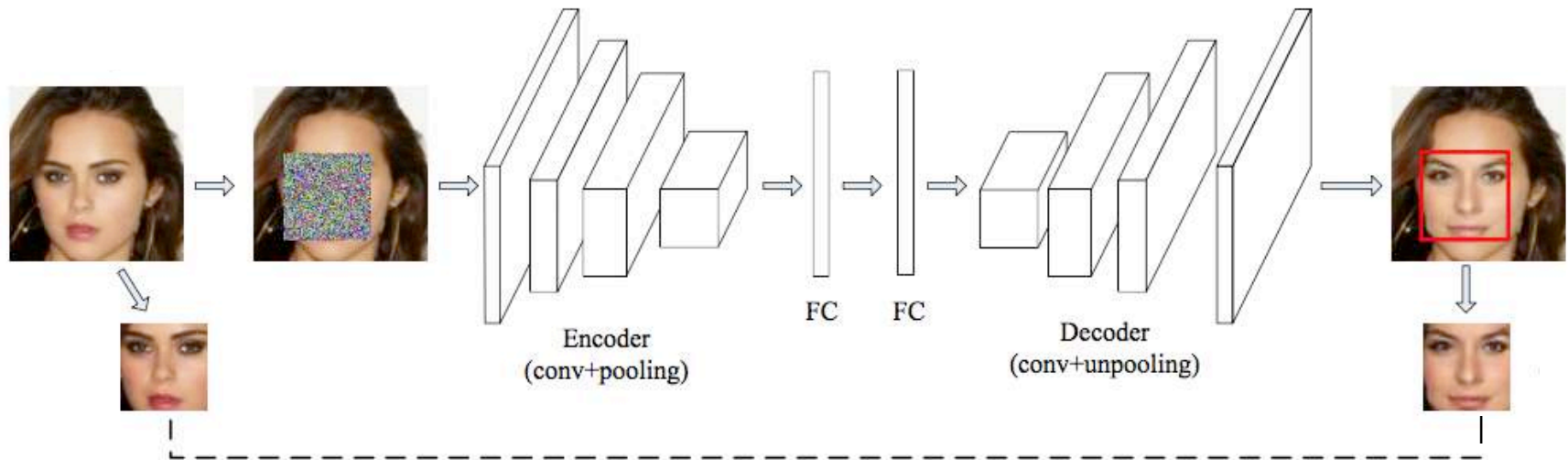
PatchMatch



Image Melding

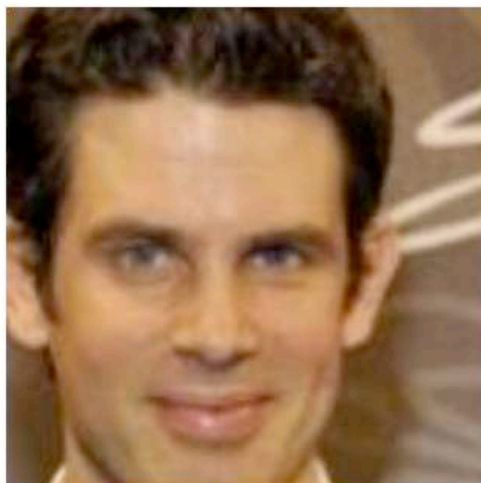


Deep learning



Deep learning

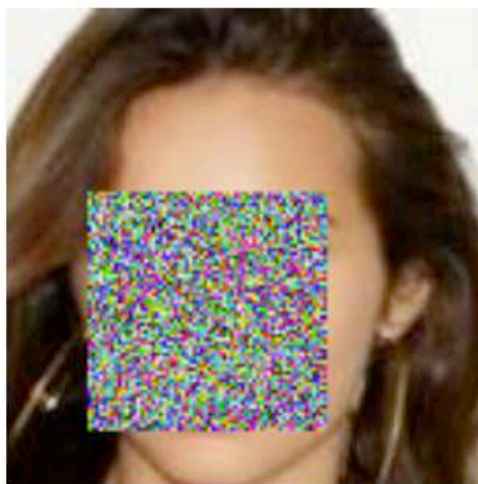
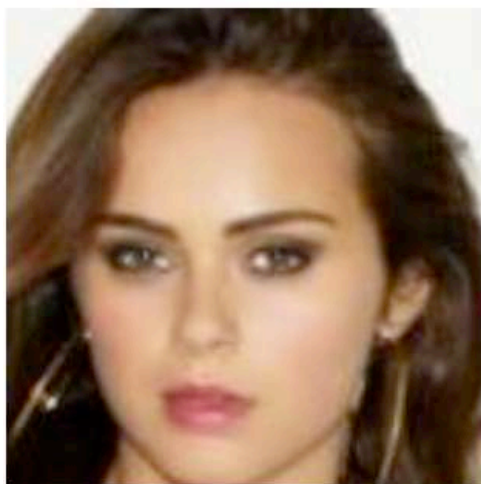
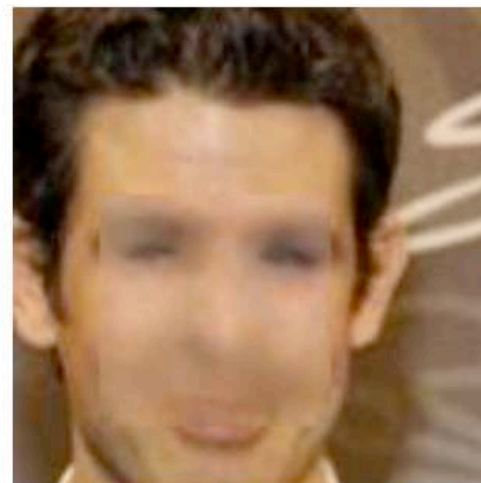
Original



Input

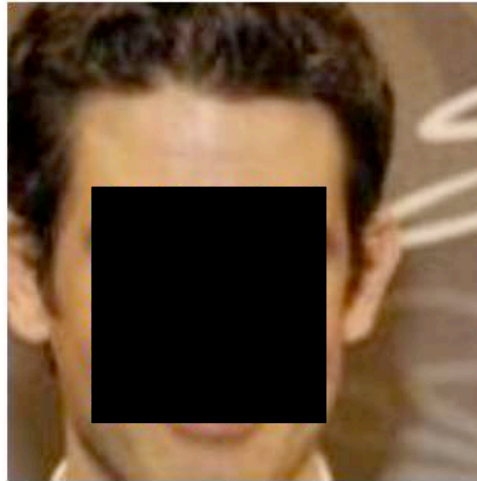
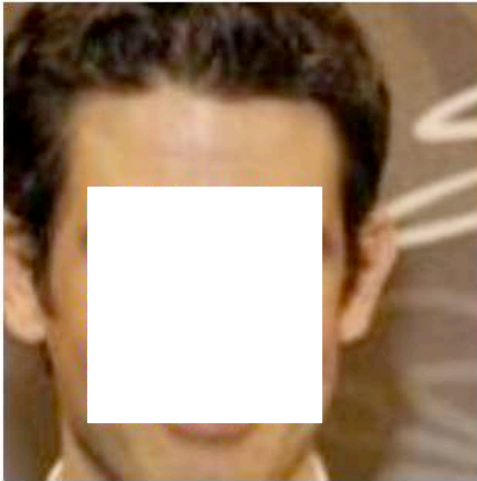


Prediction

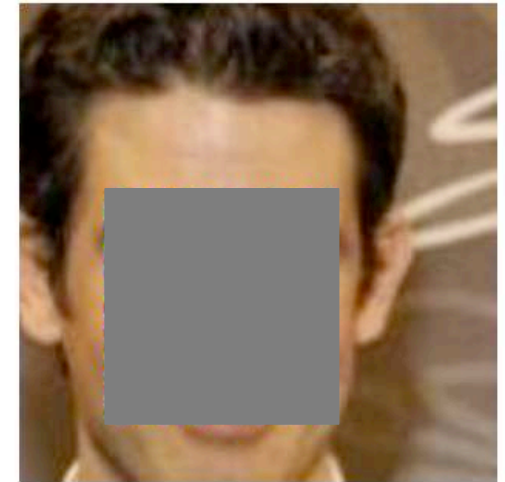


MSE loss (L2 loss)

Let's imagine a planet with only 2 kinds of people:



Prediction



Fact

$$\min_x \sum_{i=1}^n (x - a_i)^2$$

$x - ?$

Fact

$$\min_x \sum_{i=1}^n (x - a_i)^2$$

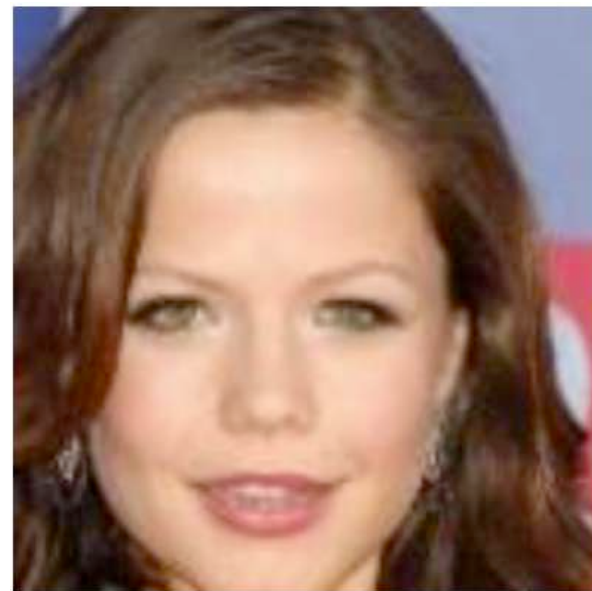
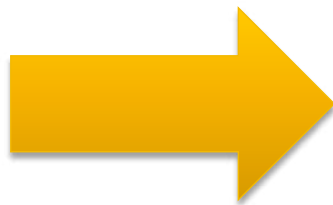
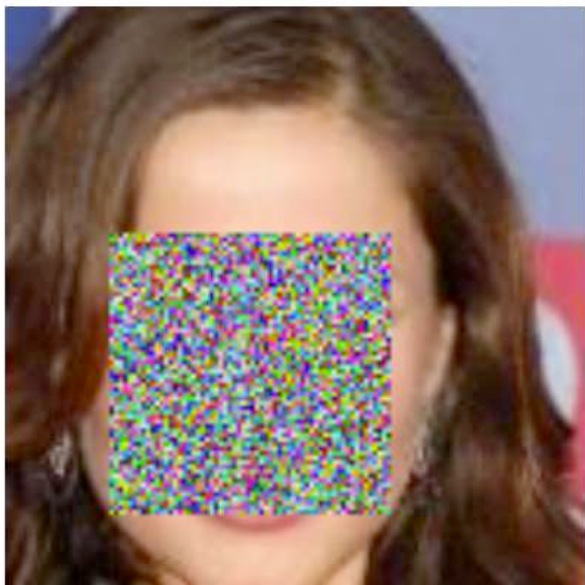
$$\frac{d}{dx} \sum_{i=1}^n (x - a_i)^2 = \sum_{i=1}^n 2(x - a_i) = 0$$

$$\sum_{i=1}^n (x - a_i) = 0$$

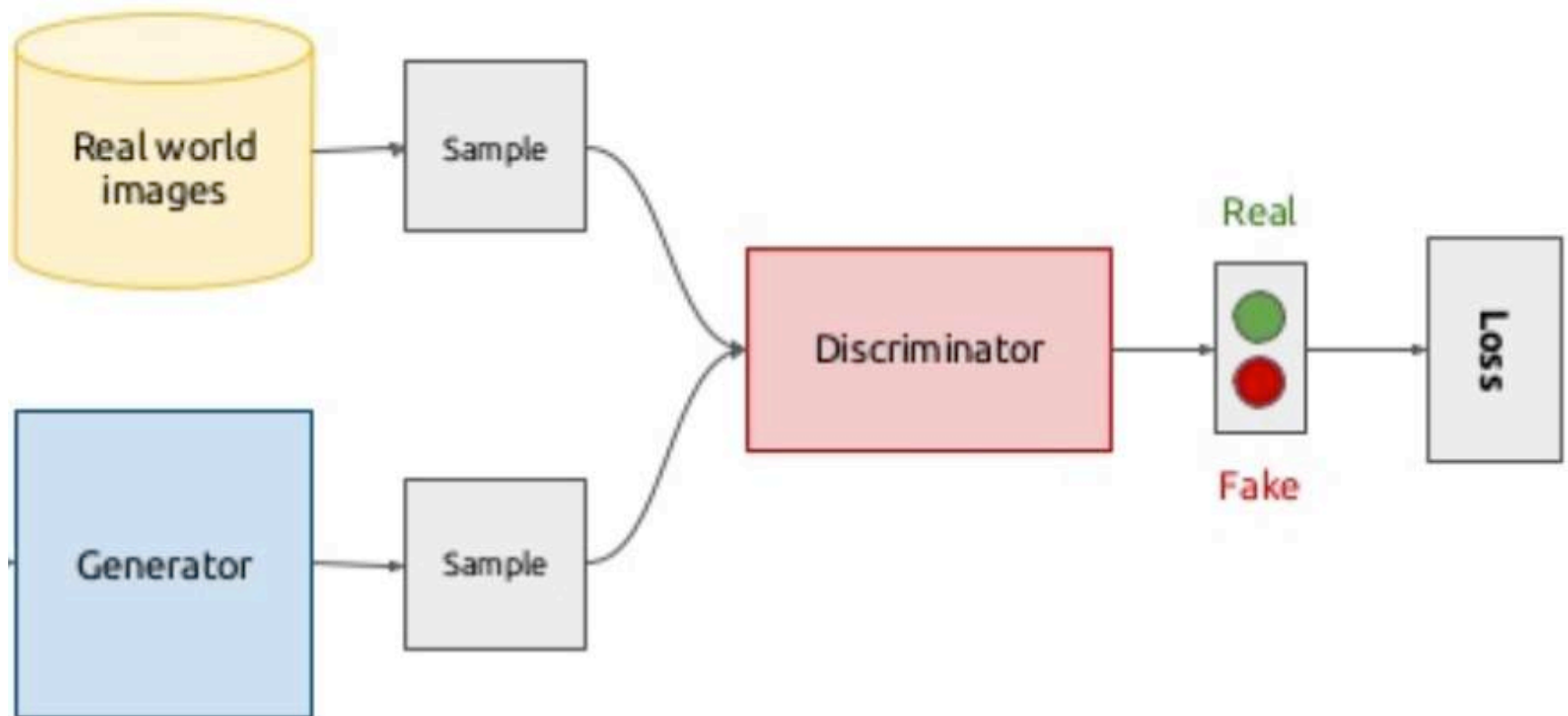
$$nx - \sum_{i=1}^n a_i = 0$$

$$x = \frac{1}{n} \sum_{i=1}^n a_i$$

The ultimate criterion



Generative Adversarial Networks



Cost functions

- x_i – real images
- D – discriminator, G – generator:

$$\min_D - \sum_i \left[\log D(x_i) + \log \left(1 - D(G(x_i)) \right) \right]$$

- G – generator:

$$\min_G \sum_i \log \left(1 - D(G(x_i)) \right) \quad \text{or} \quad \min_G - \sum_i \log D(G(x_i))$$

Example of results

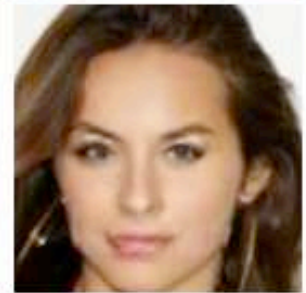
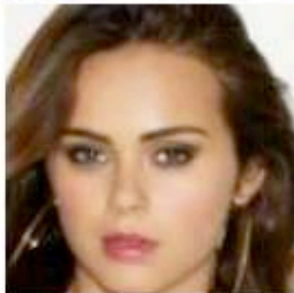
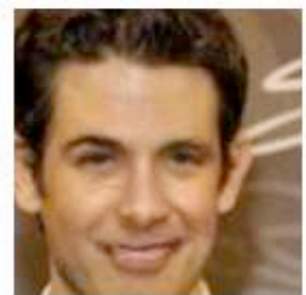
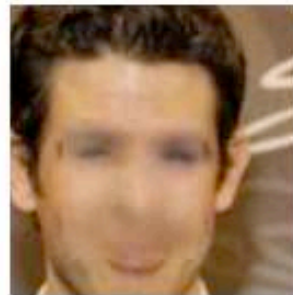
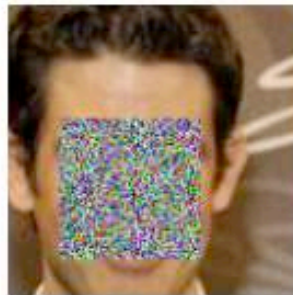
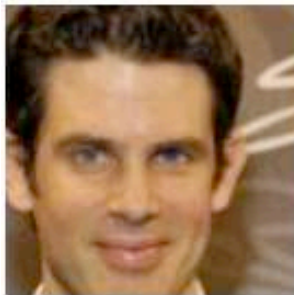
GT

Input

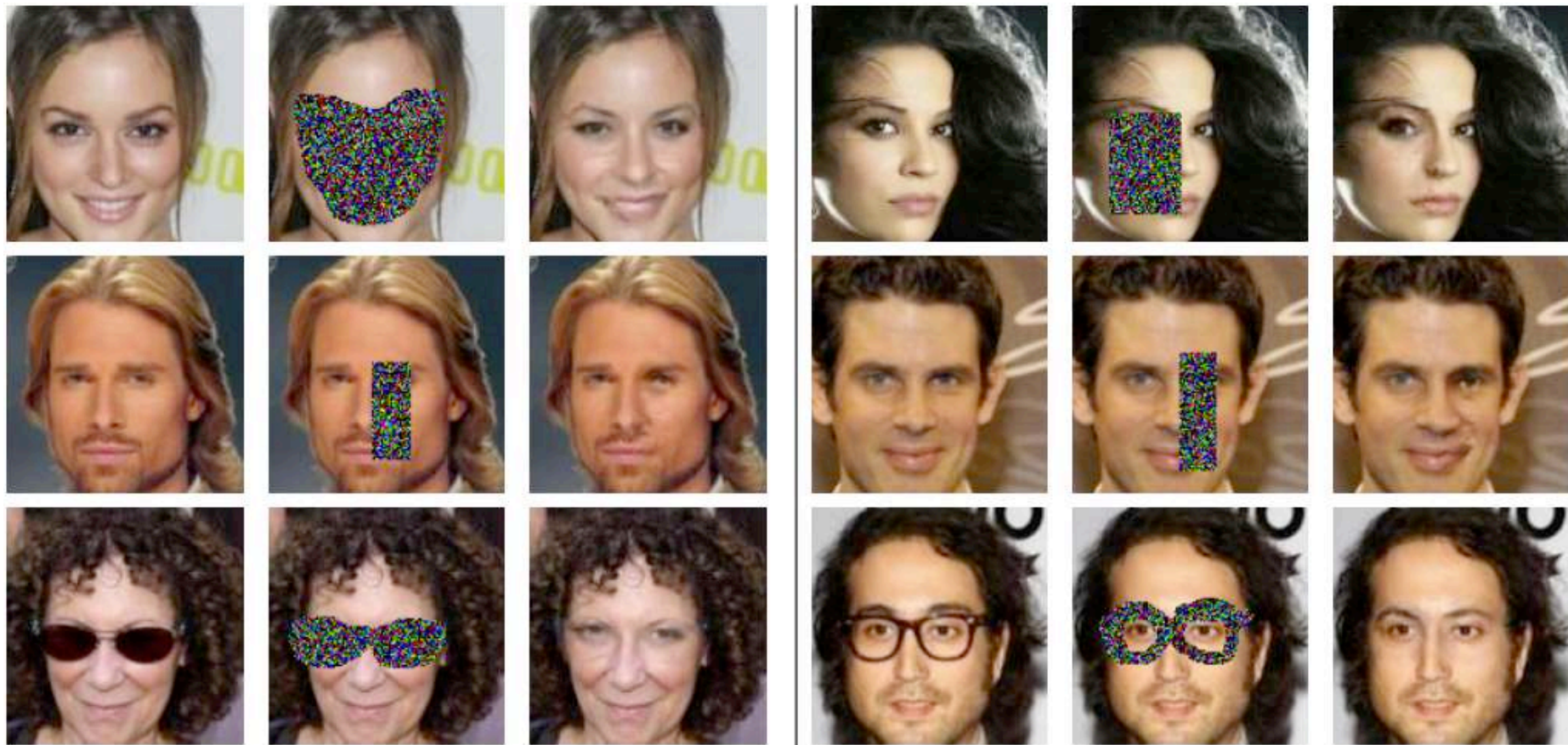
L2

GAN

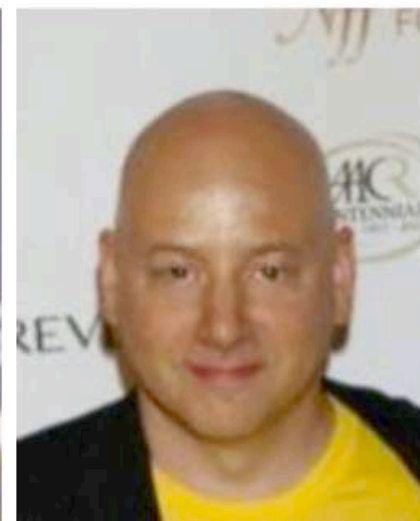
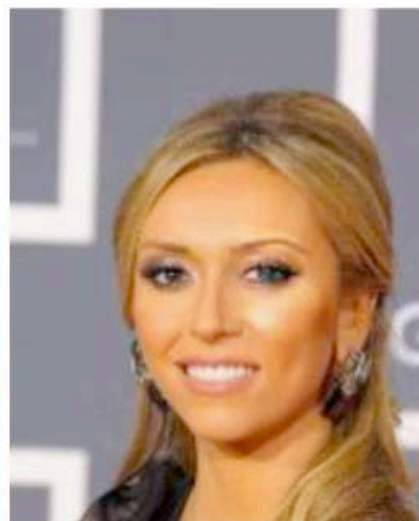
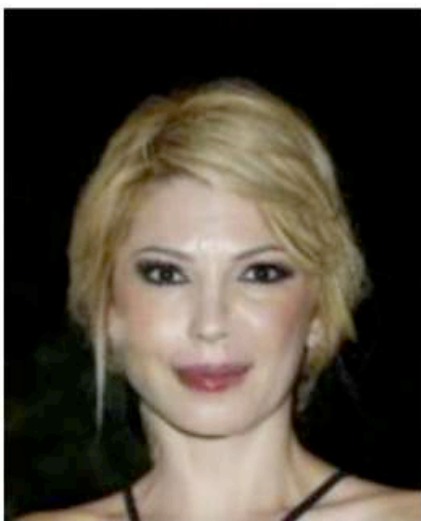
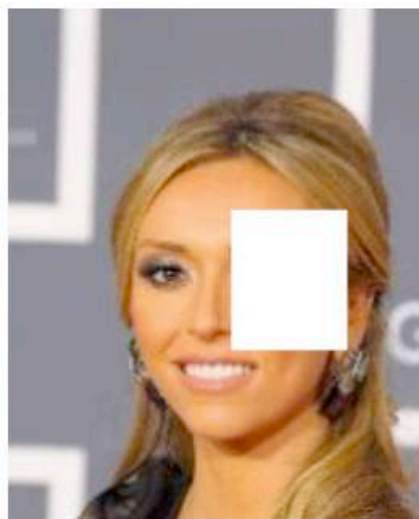
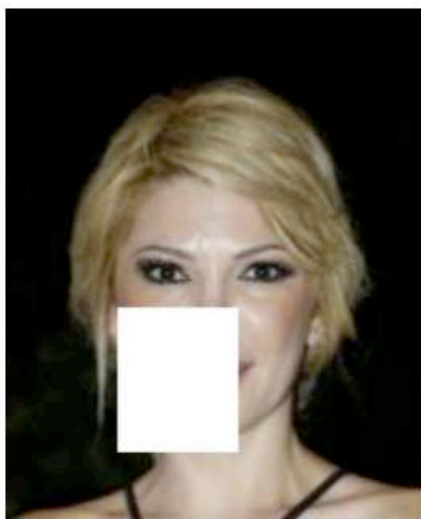
Full



More examples of results



And some more



And recent results



Not only faces



Image source:
S. Iizuka et al. Globally and Locally Consistent Image Completion.

Labels to Street Scene

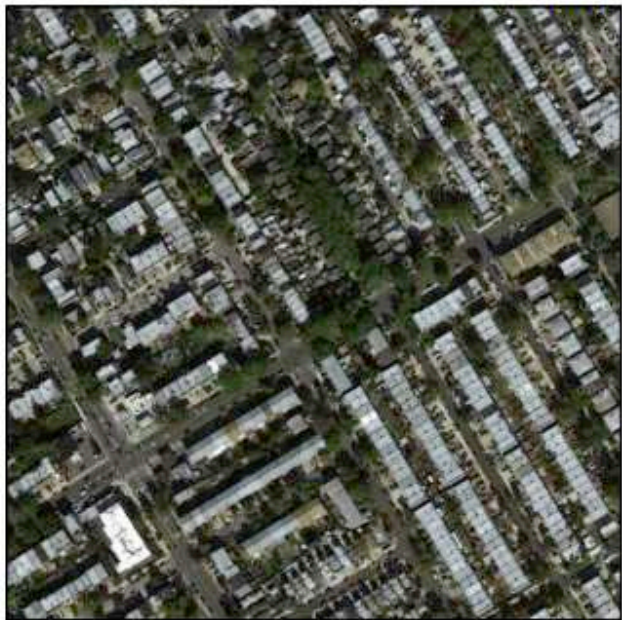


input



output

Aerial to Map



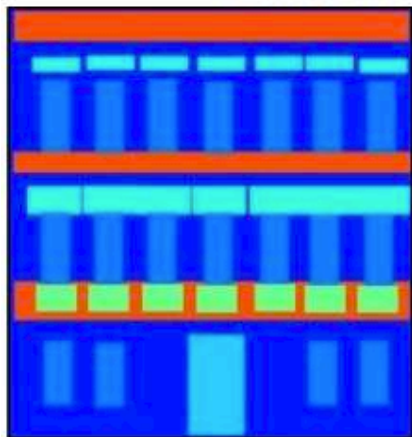
input



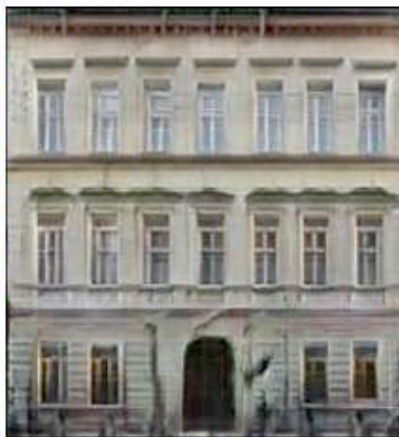
output

pix2pix

Labels to Facade



input

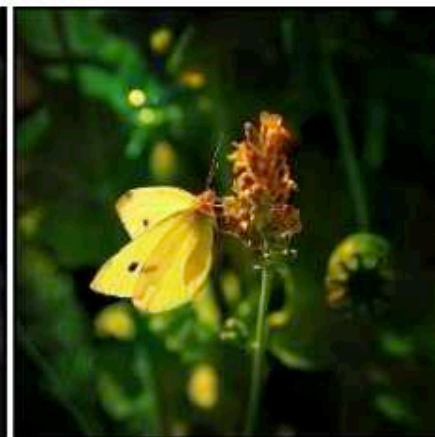


output

BW to Color

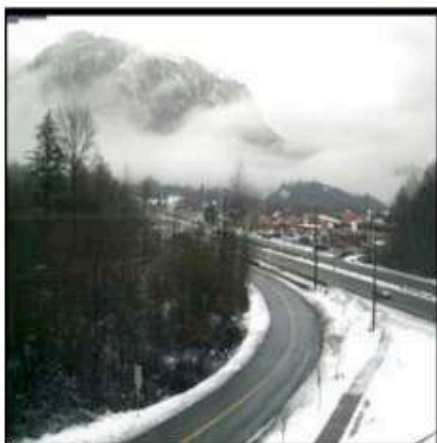


input

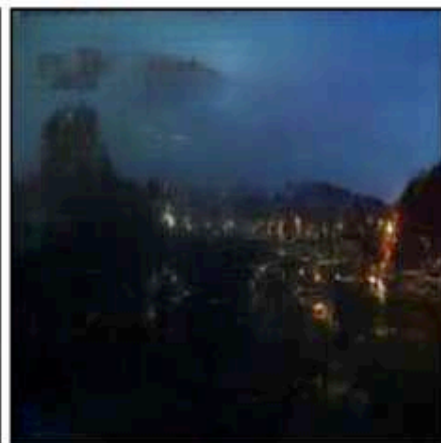


output

Day to Night



input



output

Edges to Photo



input



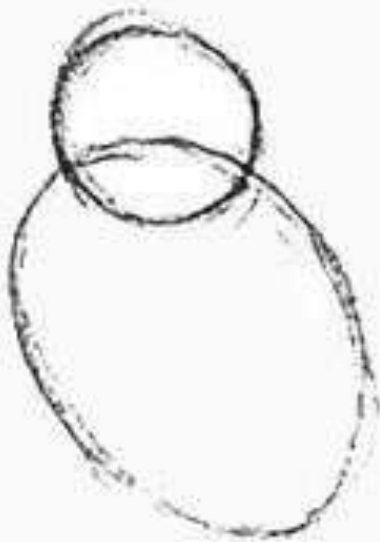
output

[arXiv:1611.07004](https://arxiv.org/abs/1611.07004)

P. Isola et al. Image-to-Image Translation with Conditional Adversarial Networks.

КАК НАРИСОВАТЬ СОВУ

1.



РИСУЕМ КРУЖОЧКИ

2.



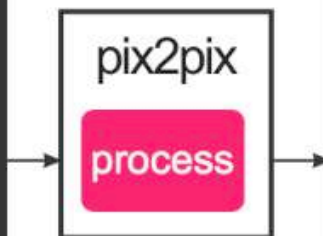
РИСУЕМ ОСТАТОК СОВЫ

Image source:

<https://hsto.org/files/fea/e0b/d83/feae0bd831b94c77a4a0079aa7398327.jpeg>

#edges2cats

INPUT



OUTPUT

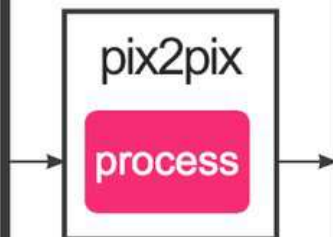
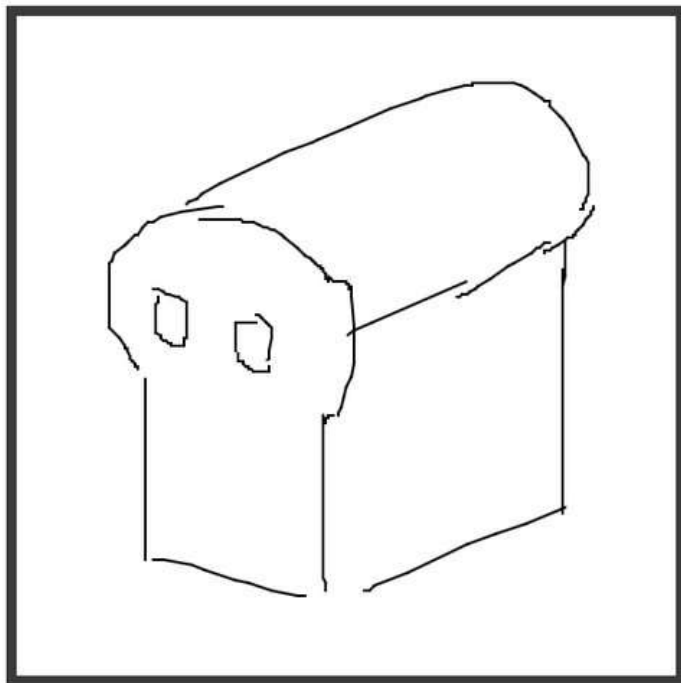


Image source:

<https://affinelayer.com/pixsrv/>

#edges2cats

INPUT

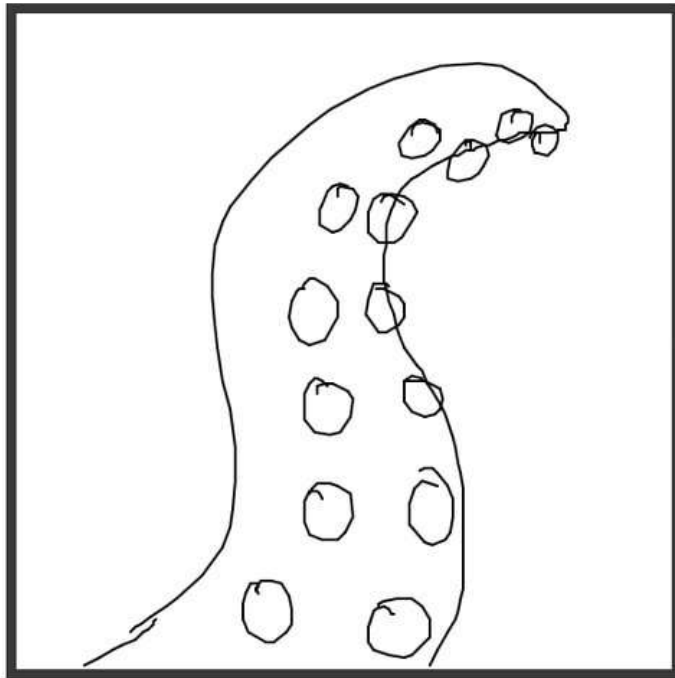


OUTPUT



#edges2cats

INPUT



OUTPUT

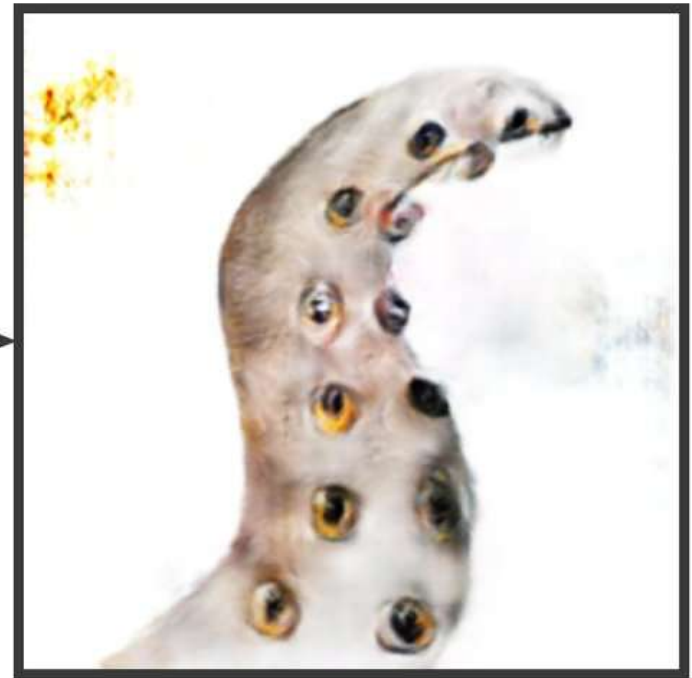
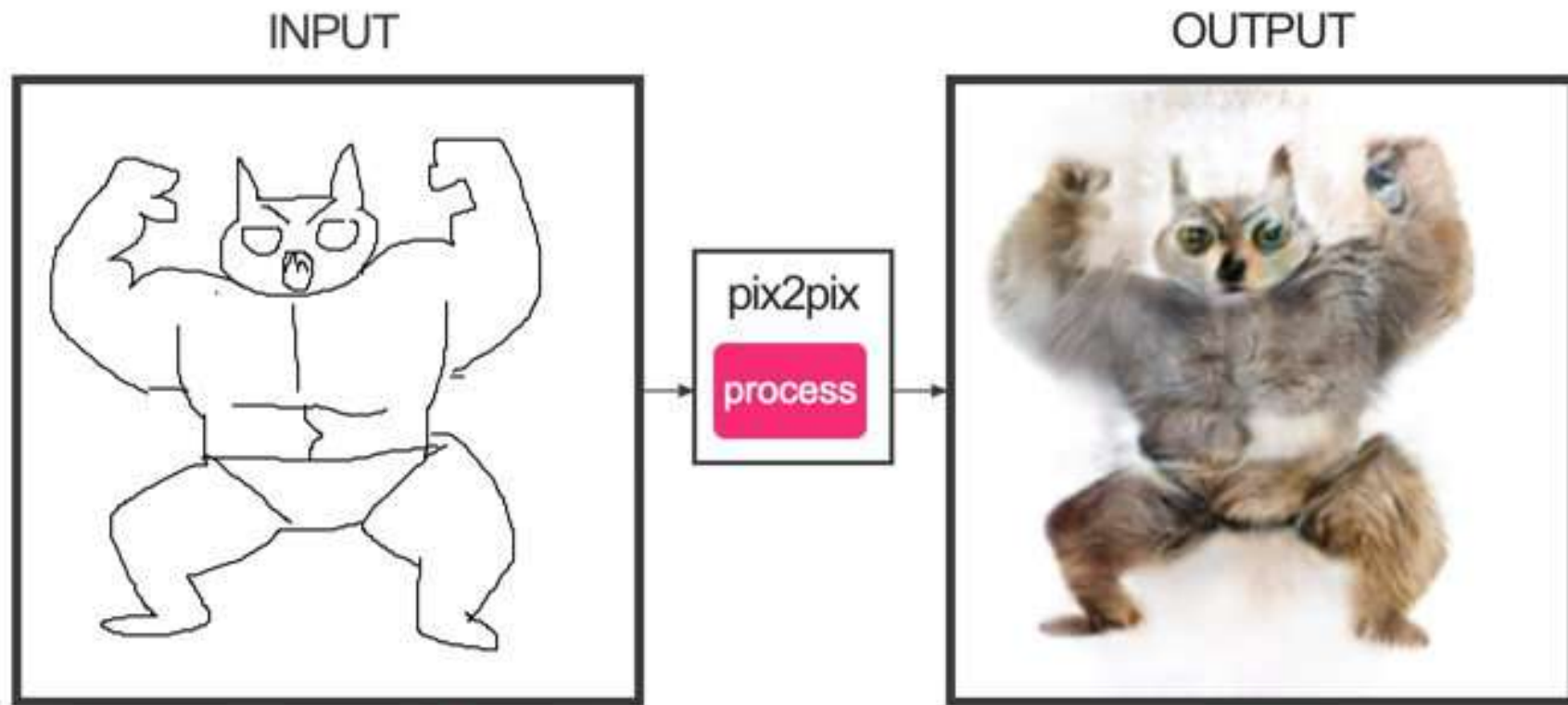


Image source:

<http://i0.kym-cdn.com/photos/images/original/001/225/191/a46.jpg>

#edges2cats



CycleGAN

Zebras \leftrightarrow Horses



zebra \rightarrow horse

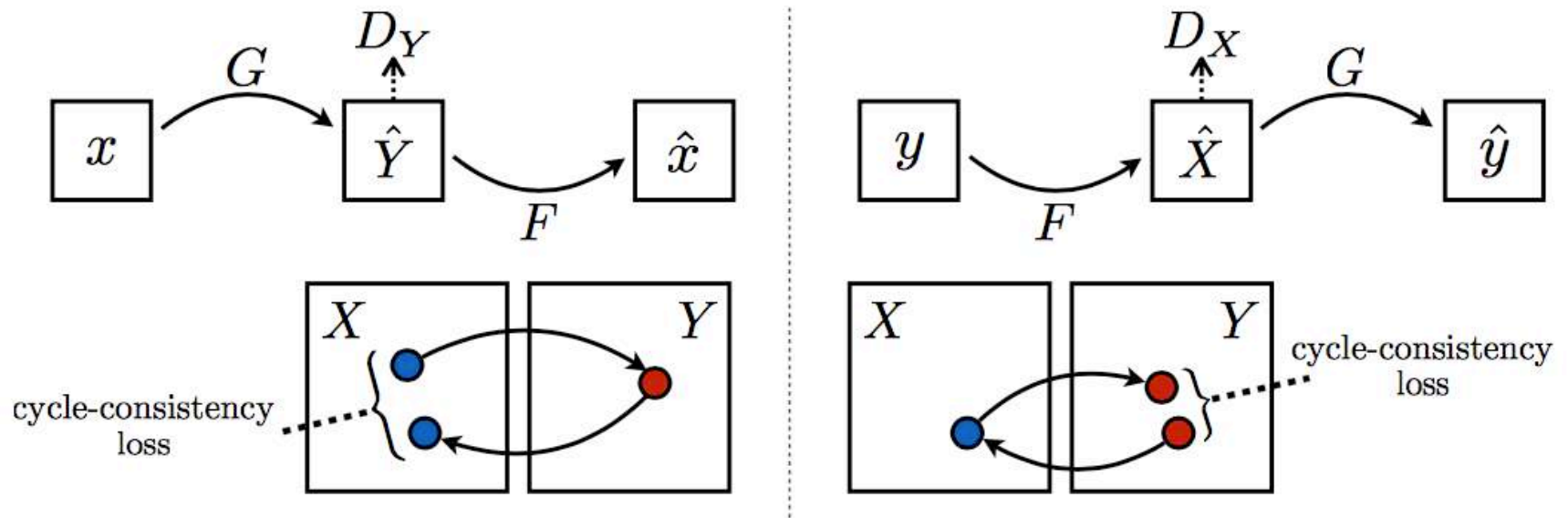


horse \rightarrow zebra

[arXiv:1703.10593](https://arxiv.org/abs/1703.10593)

J. Zhu et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks.

CycleGAN



CycleGAN

Monet \leftrightarrow Photos



Monet \rightarrow photo



photo \rightarrow Monet

[arXiv:1703.10593](https://arxiv.org/abs/1703.10593)

J. Zhu et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks.

CycleGAN

Summer \rightleftarrows Winter



summer \rightarrow winter



winter \rightarrow summer

[arXiv:1703.10593](https://arxiv.org/abs/1703.10593)

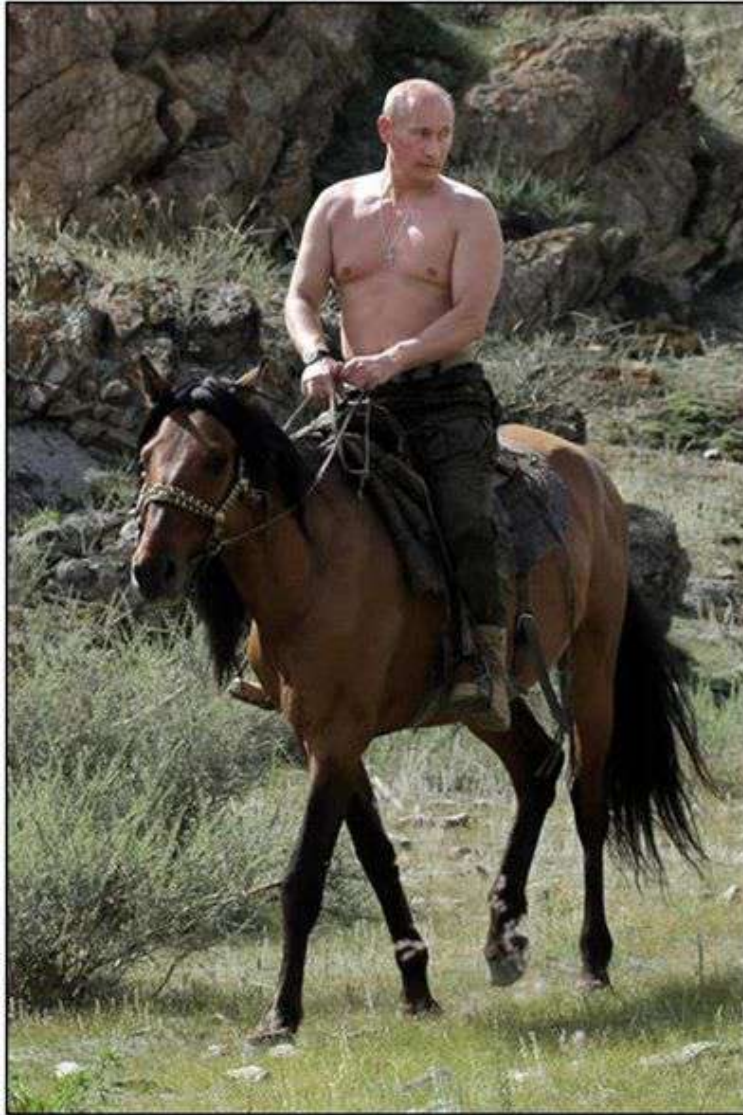
J. Zhu et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks.

CycleGAN on video

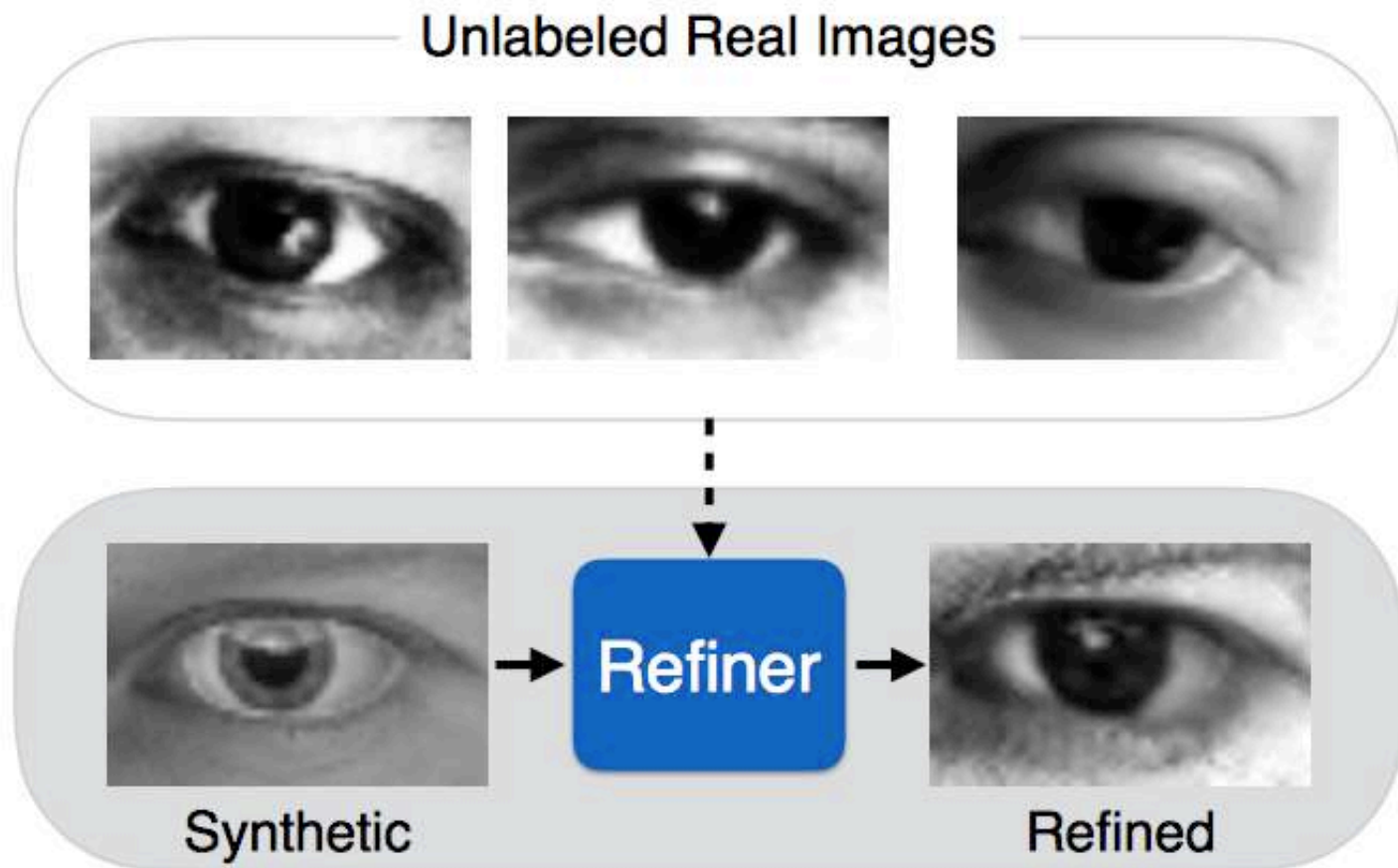


Не удалось отобразить рисунок. Возможно, рисунок поврежден или недостаточно памяти для его открытия. Перезагрузите компьютер, а затем снова откройте файл. Если вместо рисунка все еще отображается красный крестик, попробуйте удалить рисунок и вставить его заново.

Example of a failure case



First research paper from Apple



[arXiv:1612.07828](https://arxiv.org/abs/1612.07828)

Shrivastava et al. Learning from Simulated and Unsupervised Images through Adversarial Training.

StackGAN

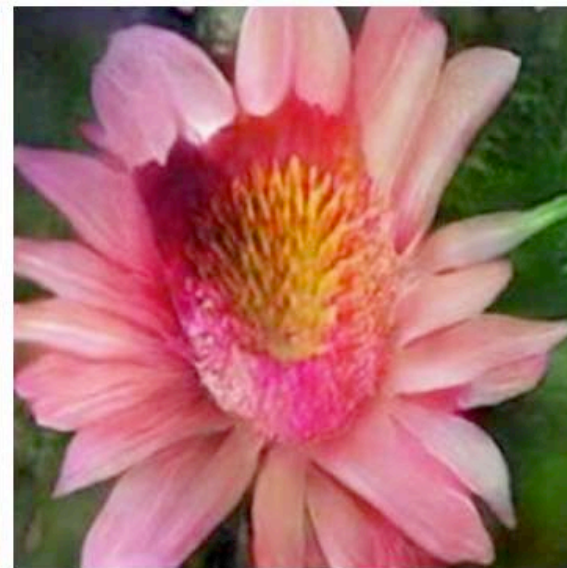
This bird is white with some black on its head and wings, and has a long orange beak



This bird has a yellow belly and tarsus, grey back, wings, and brown throat, nape with a black face

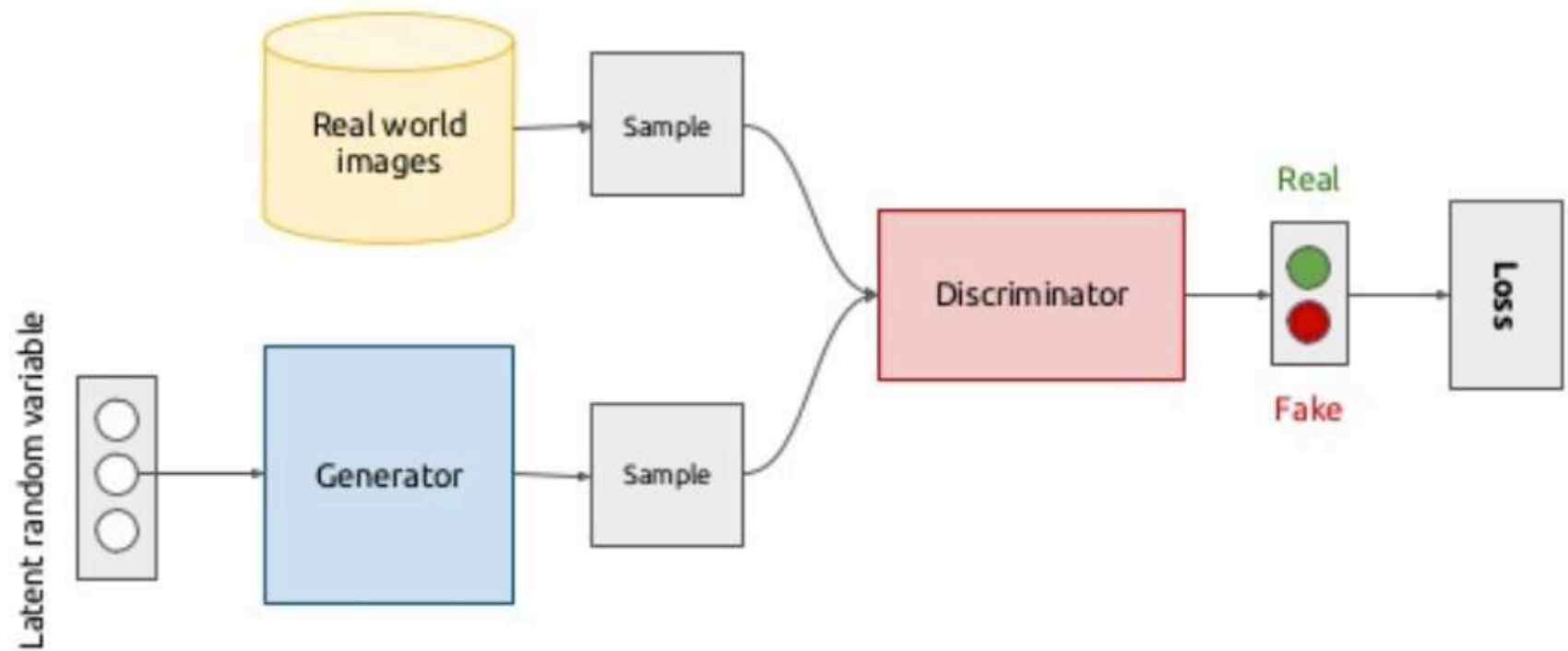


This flower has overlapping pink pointed petals surrounding a ring of short yellow filaments



[arXiv:1612.03242](https://arxiv.org/abs/1612.03242)

GANs



Progress in generation of faces over years



2014



2015



2016



2017



2018

Faces generated by GAN from random noise



[arXiv:1710.10196](https://arxiv.org/abs/1710.10196)

T. Karras et al. Progressive Growing of GANs for Improved Quality, Stability, and Variation.

Faces generated by GAN from random noise



[arXiv:1812.04948](https://arxiv.org/abs/1812.04948)

T. Karras et al. A Style-Based Generator Architecture for Generative Adversarial Networks.

Online demo

<https://thispersondoesnotexist.com>

Images generated for different categories



POTTEDPLANT

HORSE

SOFA

BUS

CHURCHOUTDOOR

BICYCLE

TVMONITOR

[arXiv:1710.10196](https://arxiv.org/abs/1710.10196)

T. Karras et al. Progressive Growing of GANs for Improved Quality, Stability, and Variation.

Recent results



[arXiv:1809.11096](https://arxiv.org/abs/1809.11096)

A. Brock et al. Large Scale GAN Training for High Fidelity Natural Image Synthesis.

DCGAN

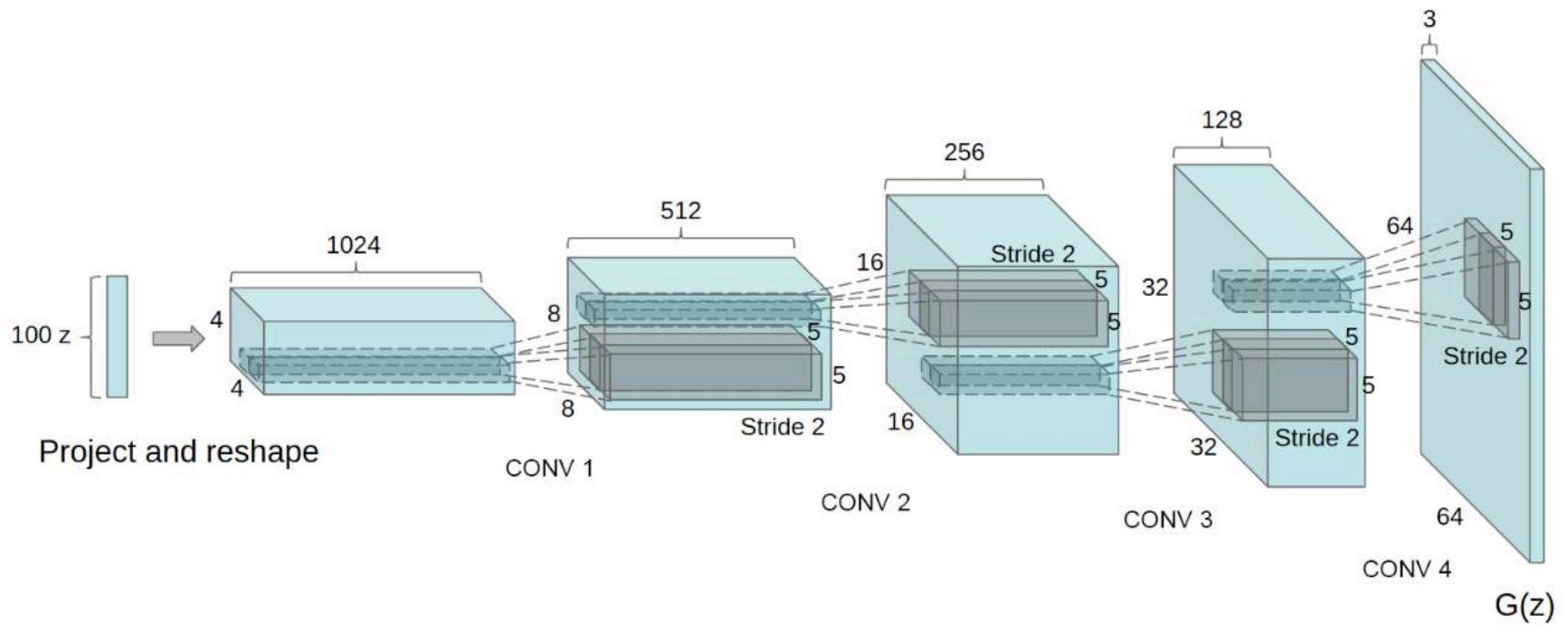


Image source: [arXiv:1511.06434](https://arxiv.org/abs/1511.06434)

A. Redford et al. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks.

DCGAN bedrooms

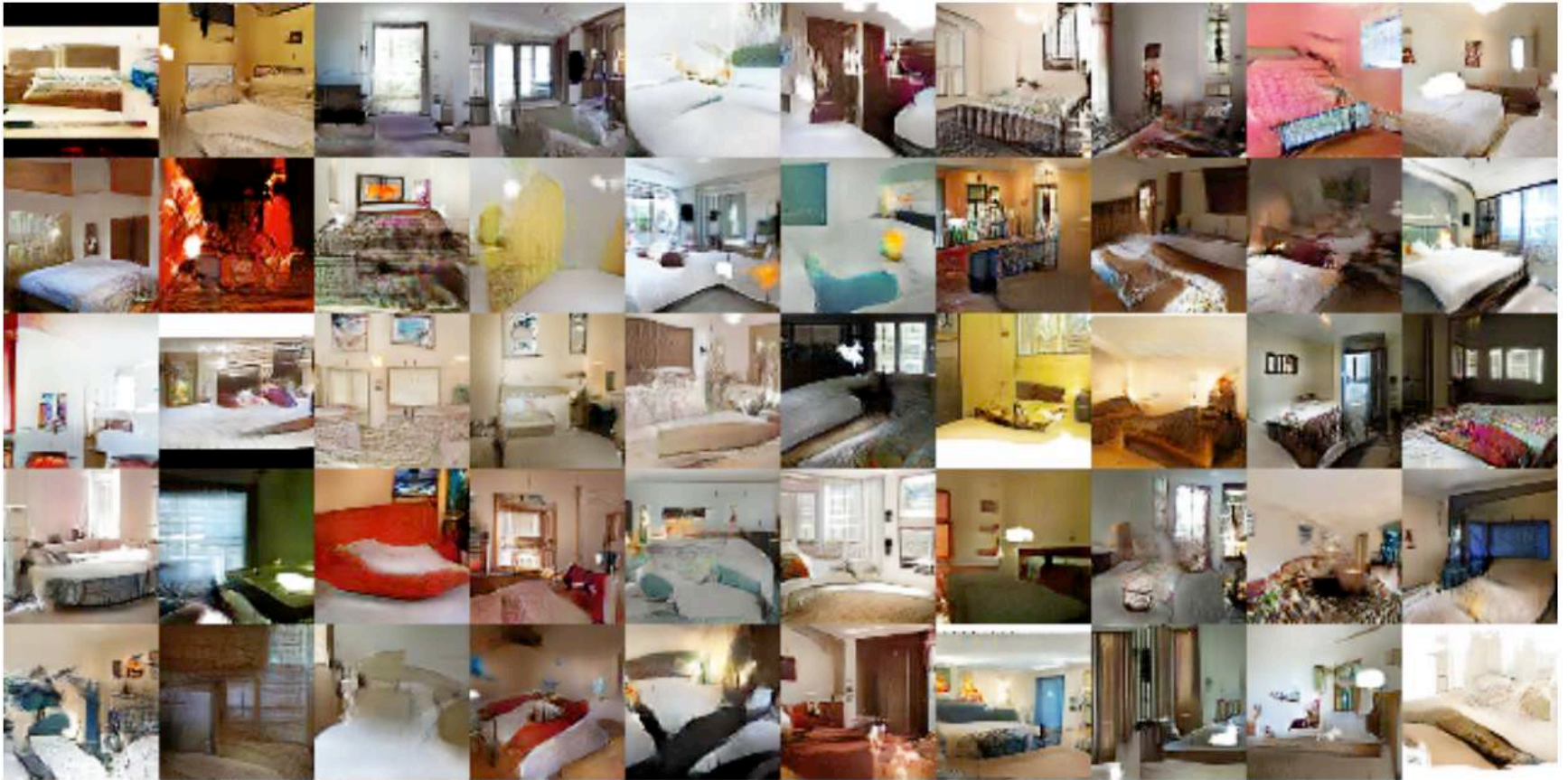


Image source: [arXiv:1511.06434](https://arxiv.org/abs/1511.06434)

A. Redford et al. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks.

DCGAN faces



Image source: [arXiv:1511.06434](https://arxiv.org/abs/1511.06434)

A. Redford et al. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks.

Arithmetic in latent space

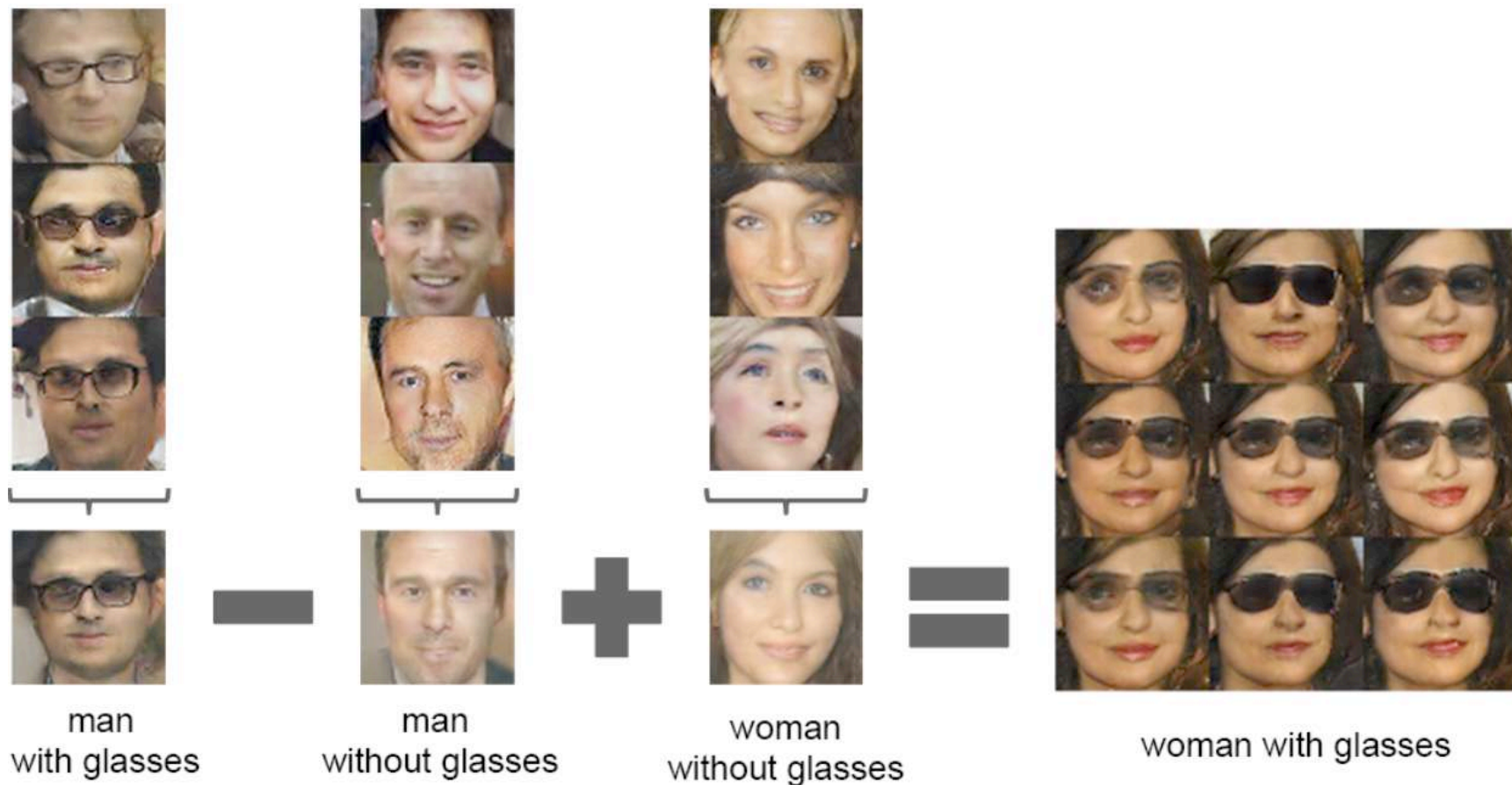


Image source: [arXiv:1511.06434](https://arxiv.org/abs/1511.06434)

A. Redford et al. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks.

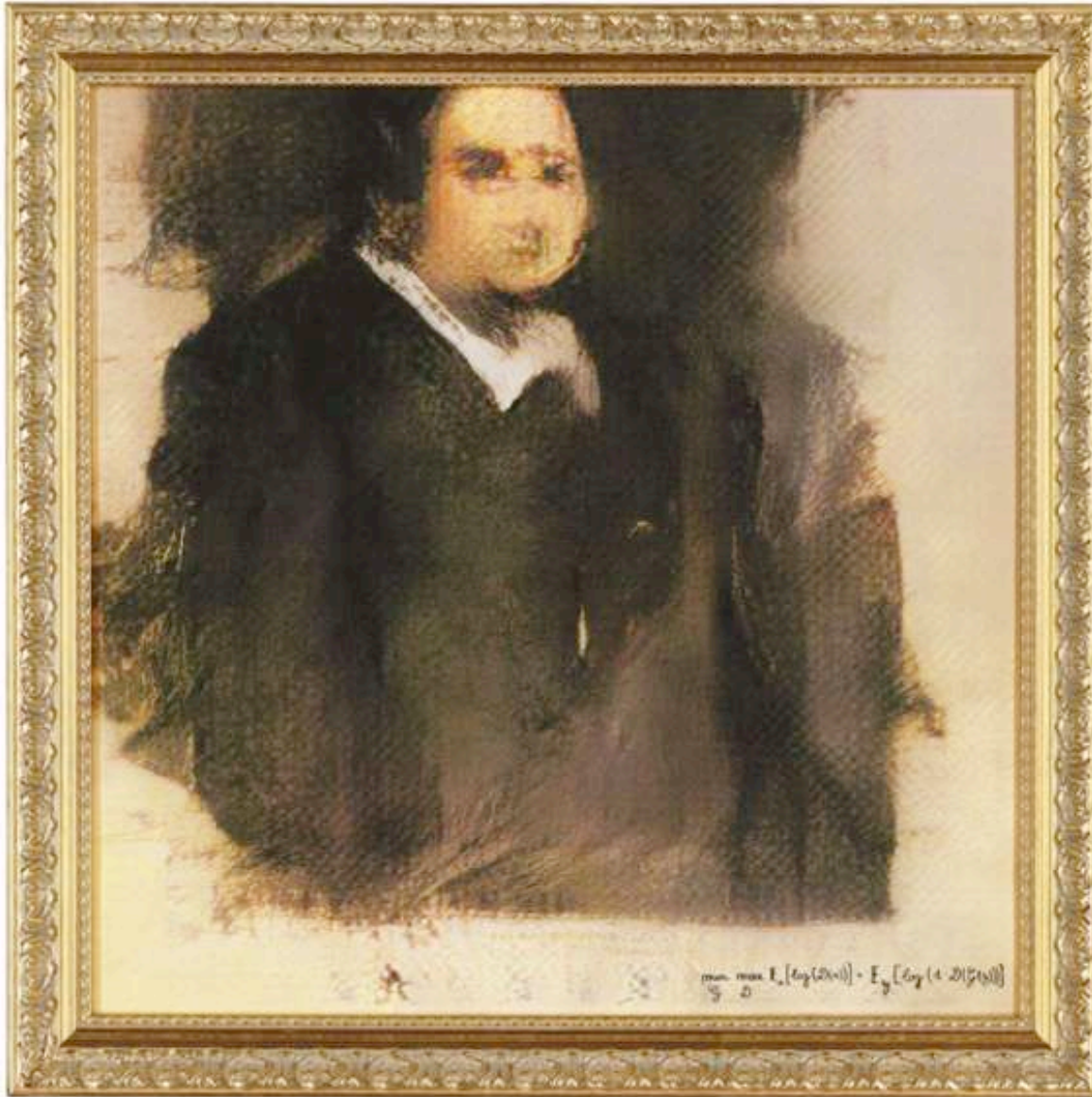
#MakeGirlsMoe



[arXiv:1708.05509](https://arxiv.org/abs/1708.05509)

Y. Jin et al. Towards the Automatic Anime Characters Creation with Generative Adversarial Networks.

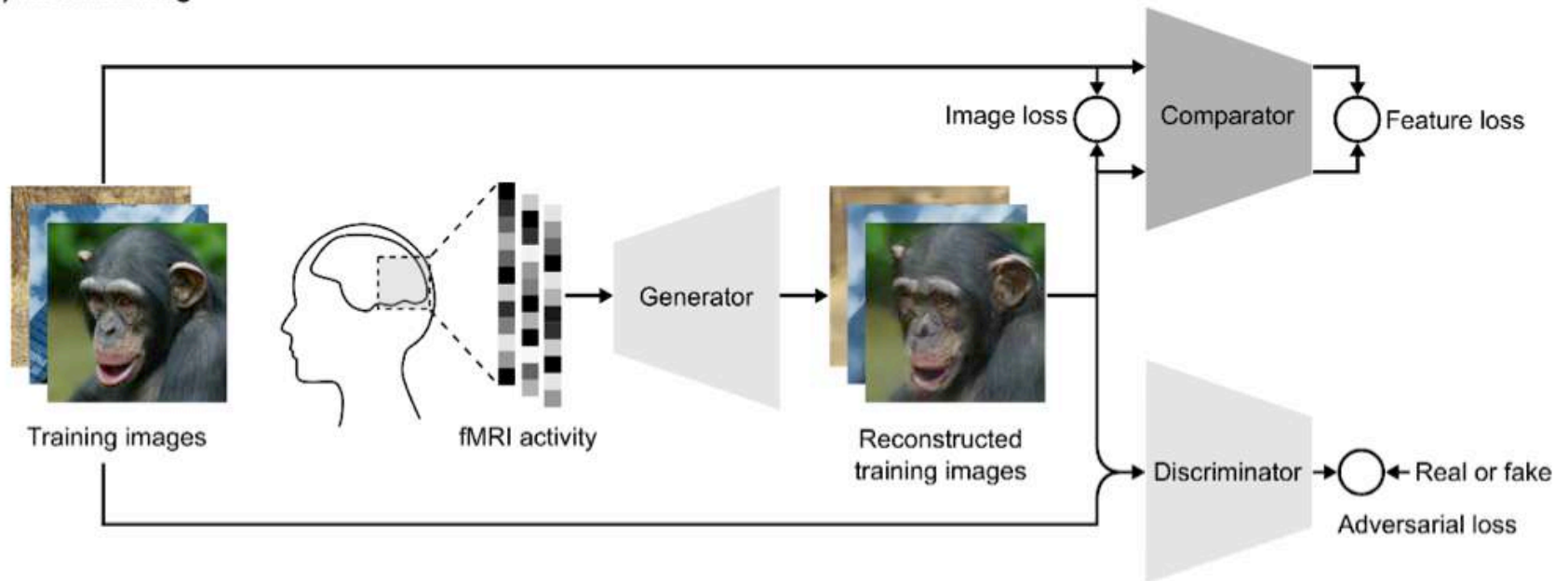
GANs and Art



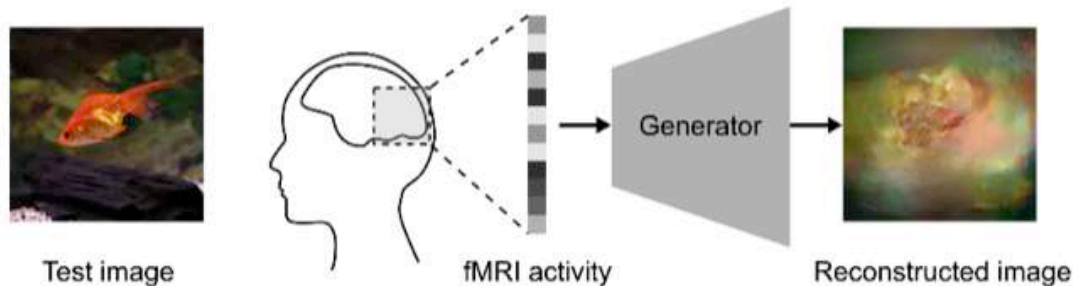
- Trained on 15,000 portraits painted between the 14th and 20th centuries
- Sold for \$432,000 at auction

Reconstruction of images from fMRI data

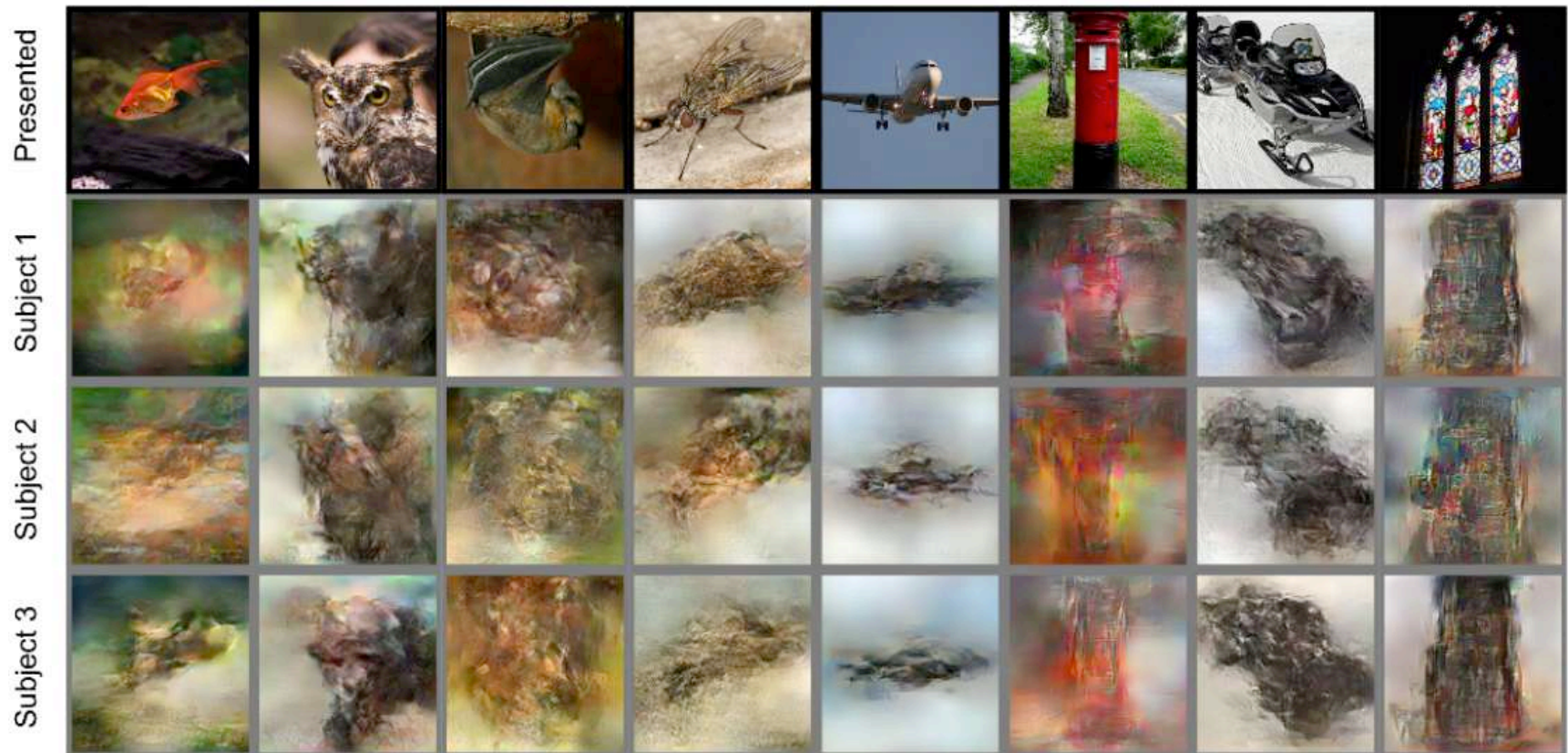
(A) Model training



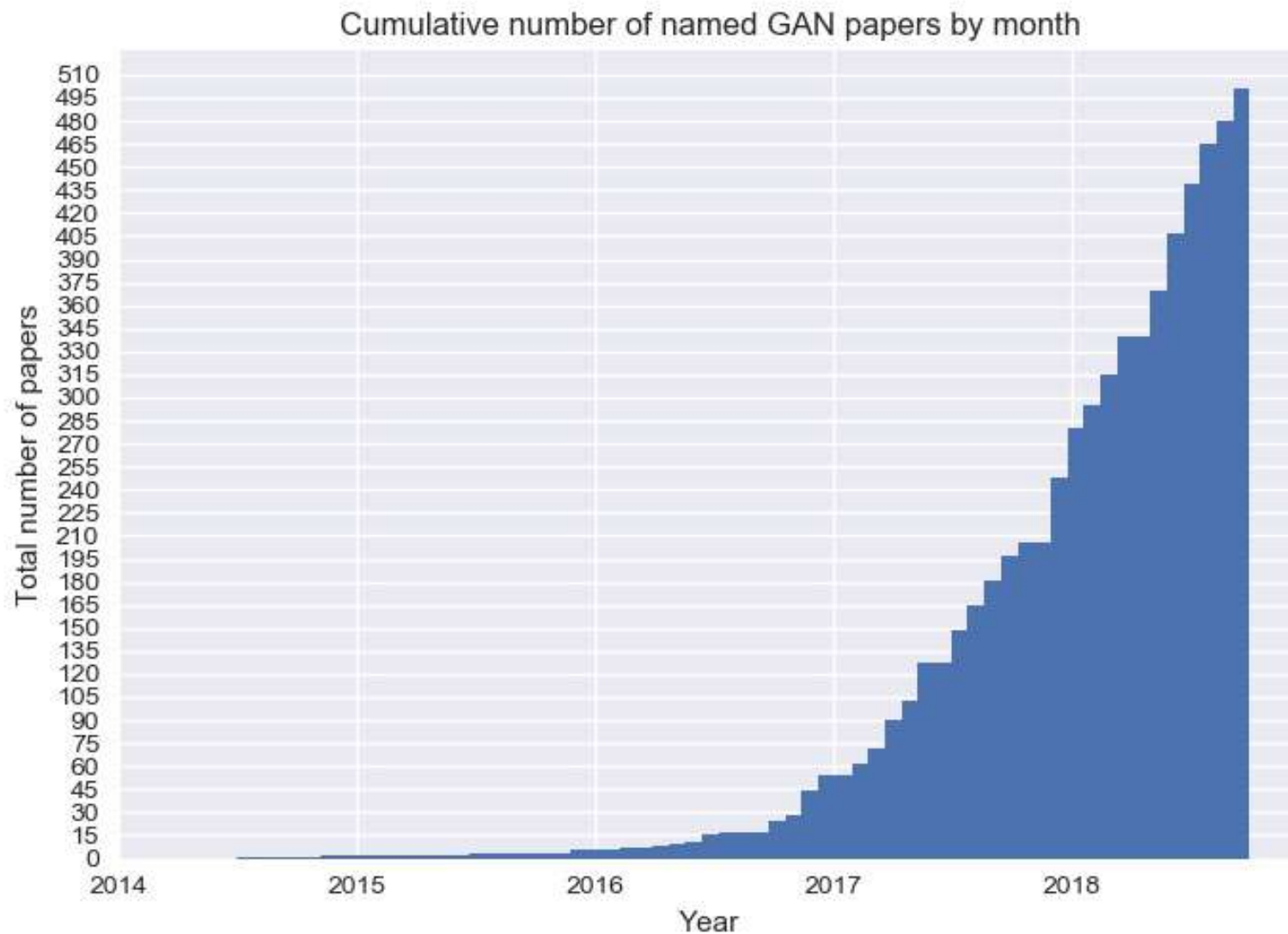
(B) Model test



Reconstruction of images from fMRI data



GAN popularity



Source: <https://github.com/hindupuravinash/the-gan-zoo>

GAN names collisions

- **SEGAN** – “*SEGAN: Speech Enhancement Generative Adversarial Network*”
- **SeGAN** – “*SeGAN: Segmenting and Generating the Invisible*”
- **SegAN** – “*SegAN: Adversarial Network with Multi-scale L1 Loss for Medical Image Segmentation*”

GAN applied to audio

C. Donahue, J. McAuley, M. Puckette.
Synthesizing Audio with Generative
Adversarial Networks.

[arXiv:1802.04208](https://arxiv.org/abs/1802.04208)

[http://wavegan-v1.s3-website-us-
east-1.amazonaws.com/](http://wavegan-v1.s3-website-us-east-1.amazonaws.com/)

GAN applied to audio

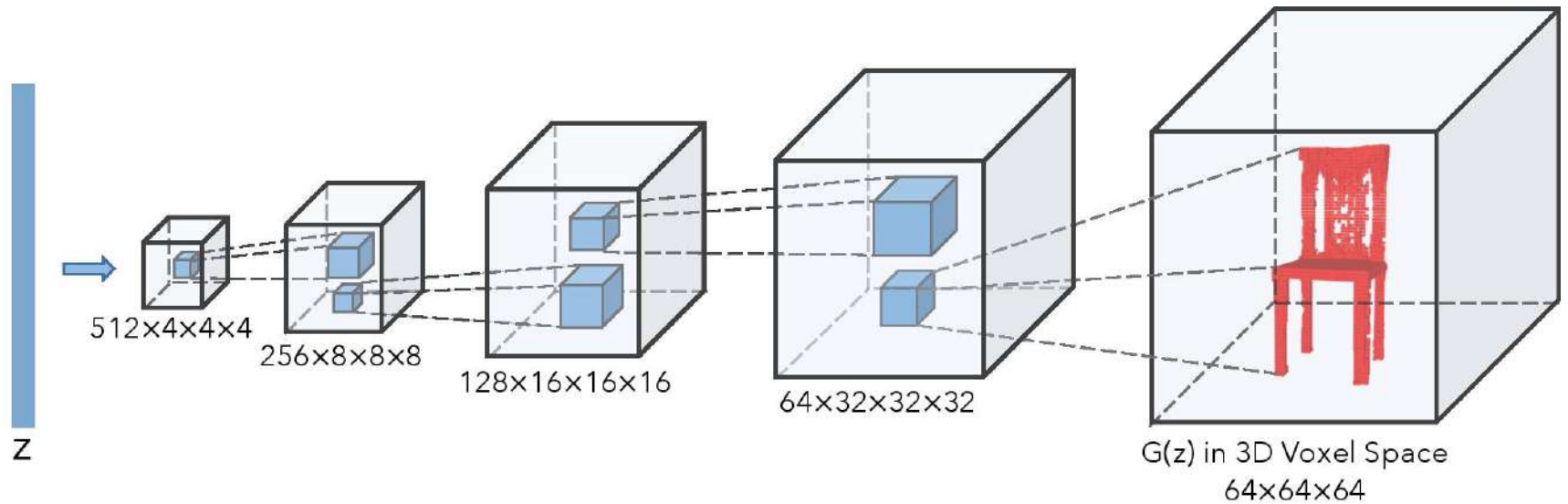
Hao-Wen Dong, Wen-Yi Hsiao, Li-Chia Yang, Yi-Hsuan Yang

MuseGAN: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment.

[arXiv:1709.06298](https://arxiv.org/abs/1709.06298)

[https://salu133445.github.io/musegan/
results](https://salu133445.github.io/musegan/results)

3D-GAN



3D-GAN results



[arXiv:1610.07584](https://arxiv.org/abs/1610.07584)

J. Wu et al. *Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling*.

3D-GAN single image reconstruction



[arXiv:1610.07584](https://arxiv.org/abs/1610.07584)

J. Wu et al. *Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling*.

itSeez3D results were shown during the presentation

Check out <https://avatarsdk.com> for more details

Thank you!

ilya.lysenkov@itseez3d.com

avatarsdk.com