Generative Adversial Networks (GANs)

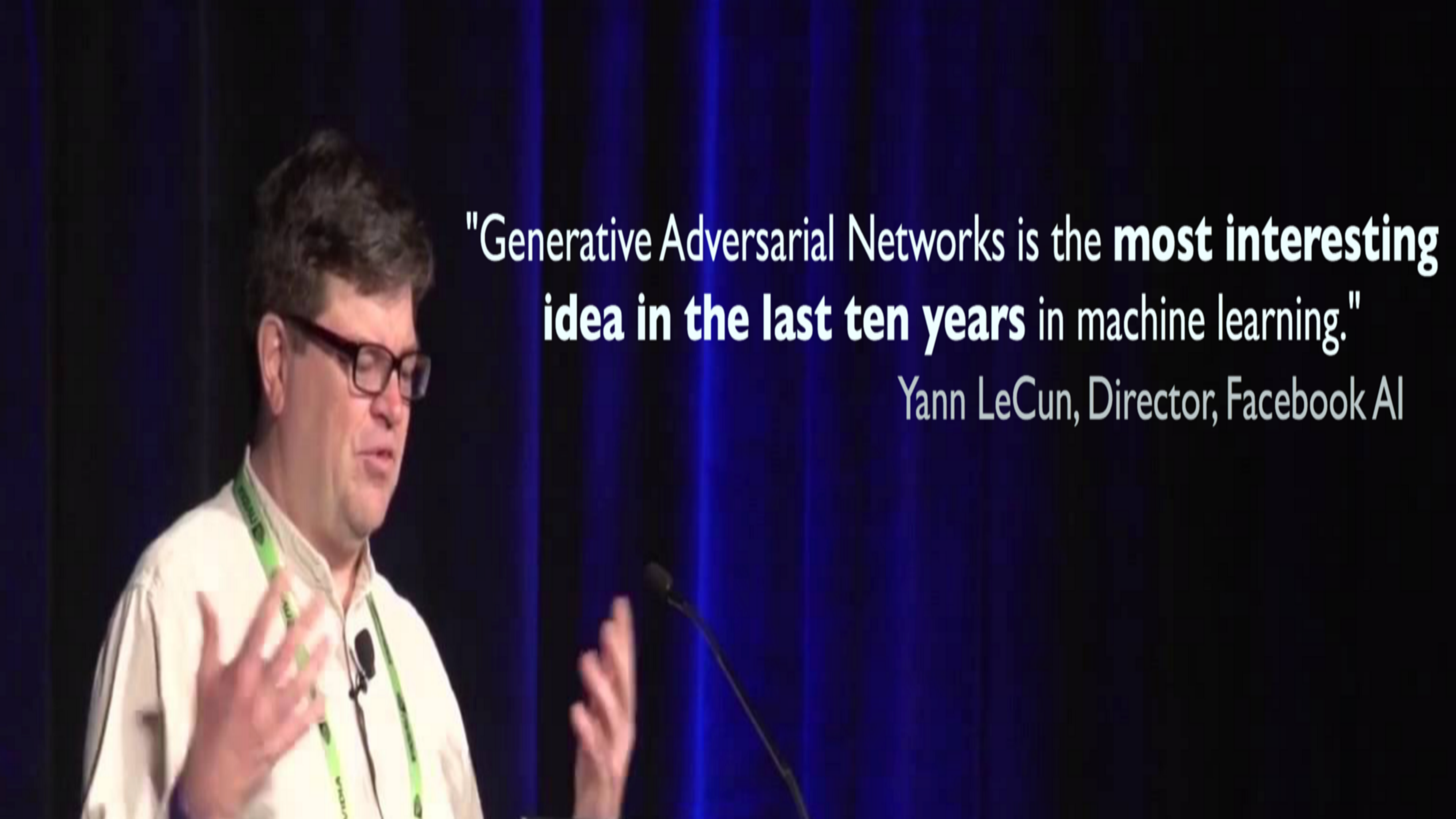
Generative Adversial Networks (GANs)

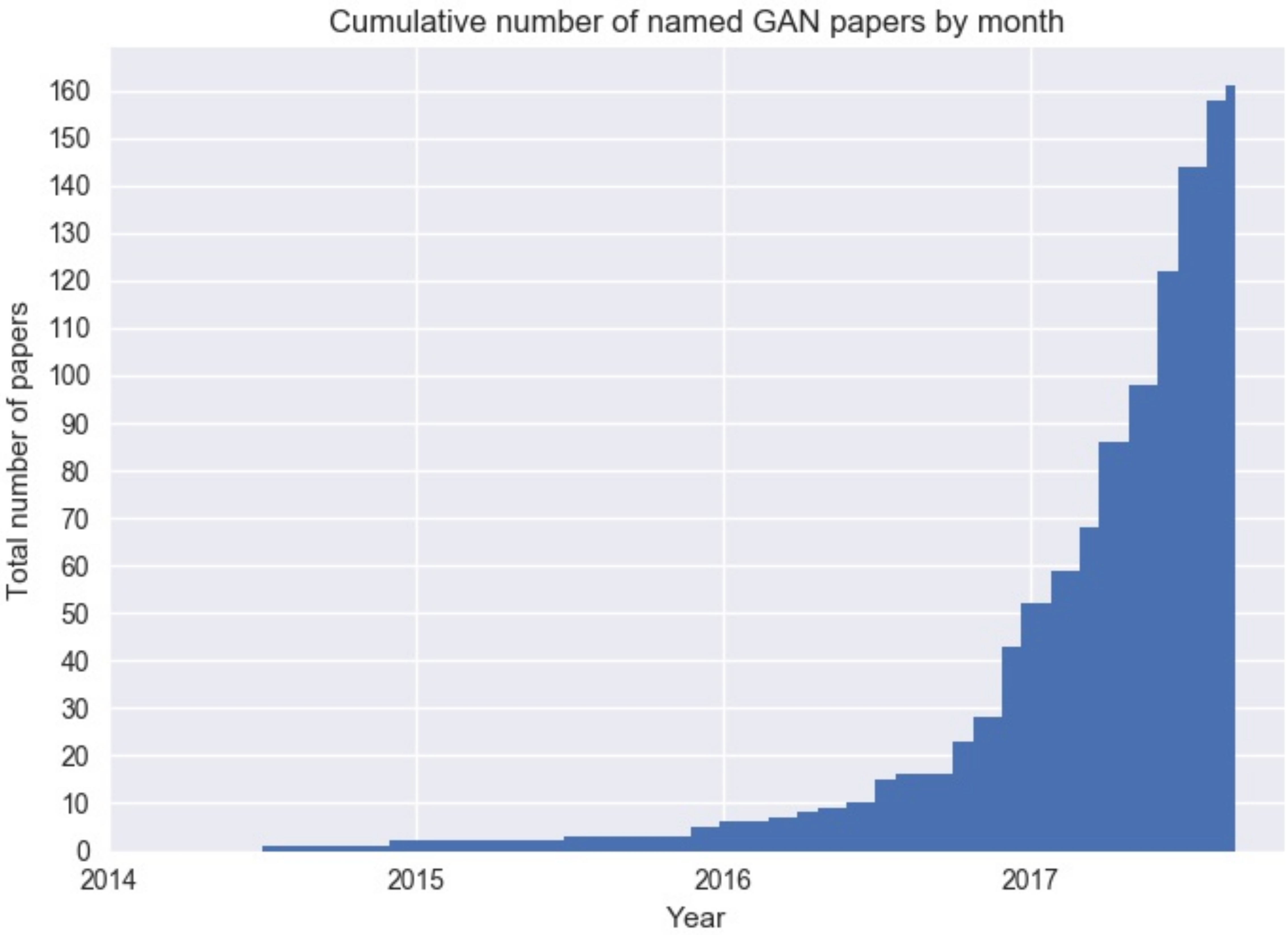
Mumbai AI Meet-up
19th Aug 2017

Presenter: Girish Patil

Meet agenda

- Phase 1 : Introduction to generative models and GANs
- Phase 2 : Types of GANs
- Phase 3 : Applications of GANs
- Phase 4 : Limitations of GANs
- Phase 5 : GAN Hacks


"Generative Adversarial Networks is the **most interesting idea in the last ten years** in machine learning."
Yann LeCun, Director, Facebook AI

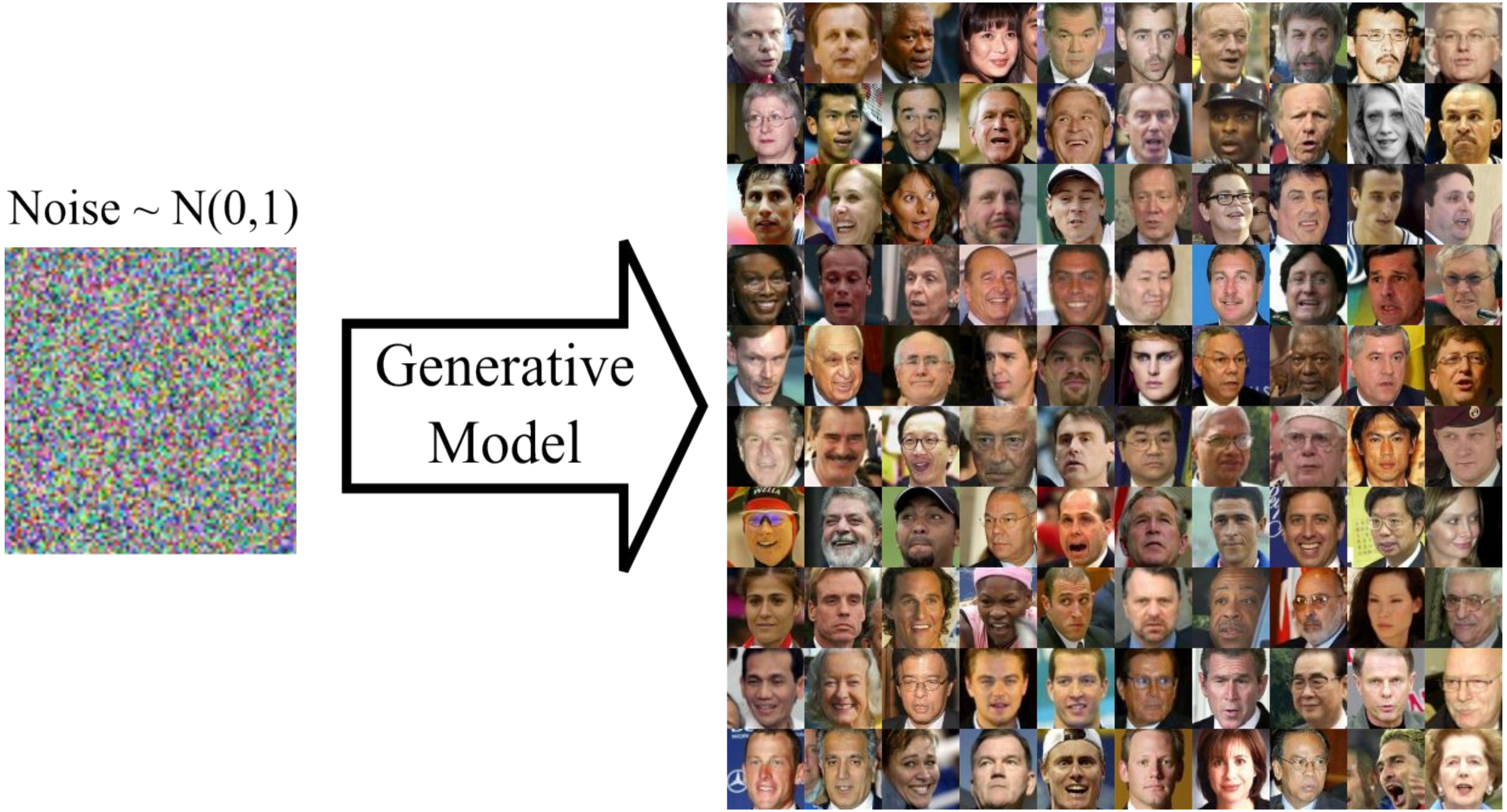


https://deephunt.in/the-gan-zoo-79597dc8c347

- 3D-GAN—Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling
- 3D-IWGAN—Improved Adversarial Systems for 3D Object Generation and Reconstruction
- ABC-GAN—ABC-GAN: Adaptive Blur and Control for improved training stability of Generative Adversarial Networks
- AC-GAN—Conditional Image Synthesis With Auxiliary Classifier GANs
- acGAN—Face Aging With Conditional Generative Adversarial Networks
- AdaGAN—AdaGAN: Boosting Generative Models
- AE-GAN—AE-GAN: adversarial eliminating with GAN
- AEGAN—Learning Inverse Mapping by Autoencoder based Generative Adversarial Nets

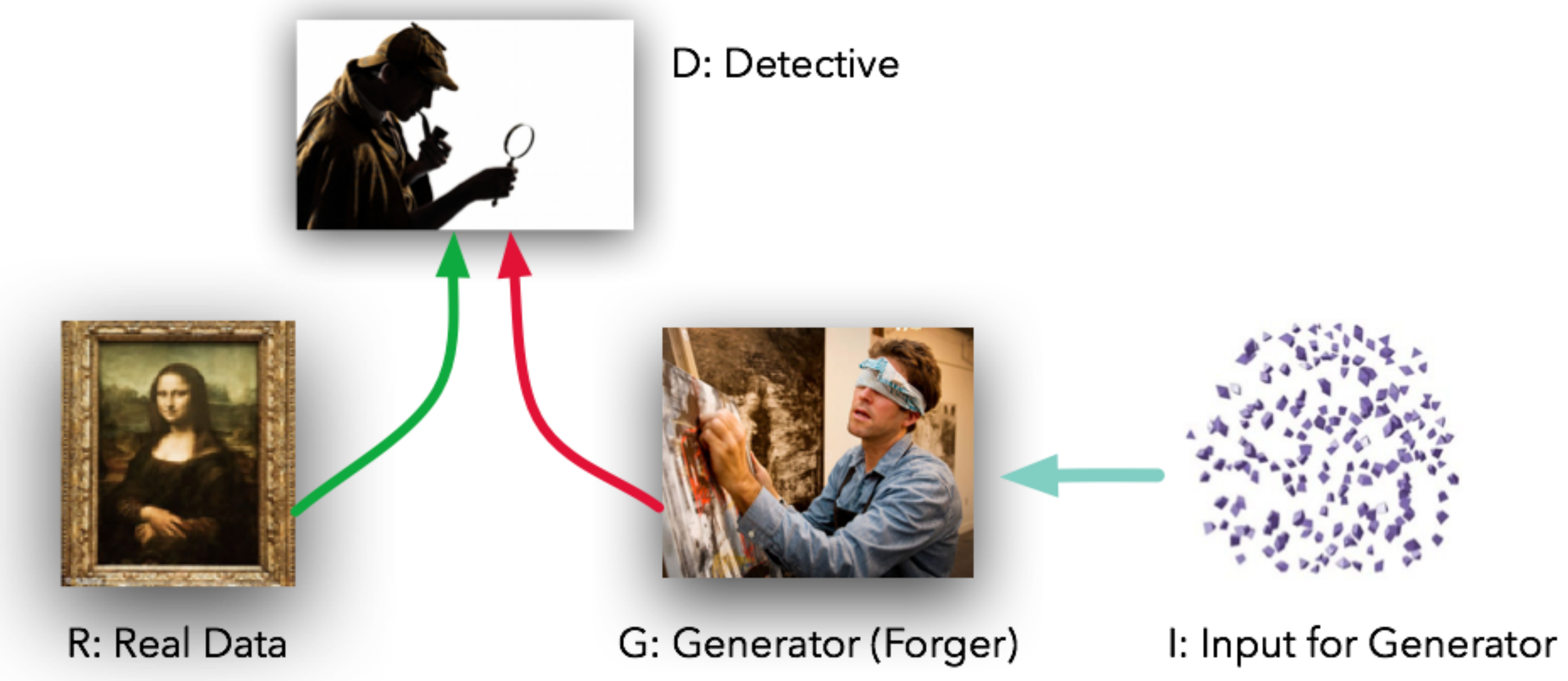
https://deephunt.in/the-gan-zoo-79597dc8c347

Generative adversarial networks

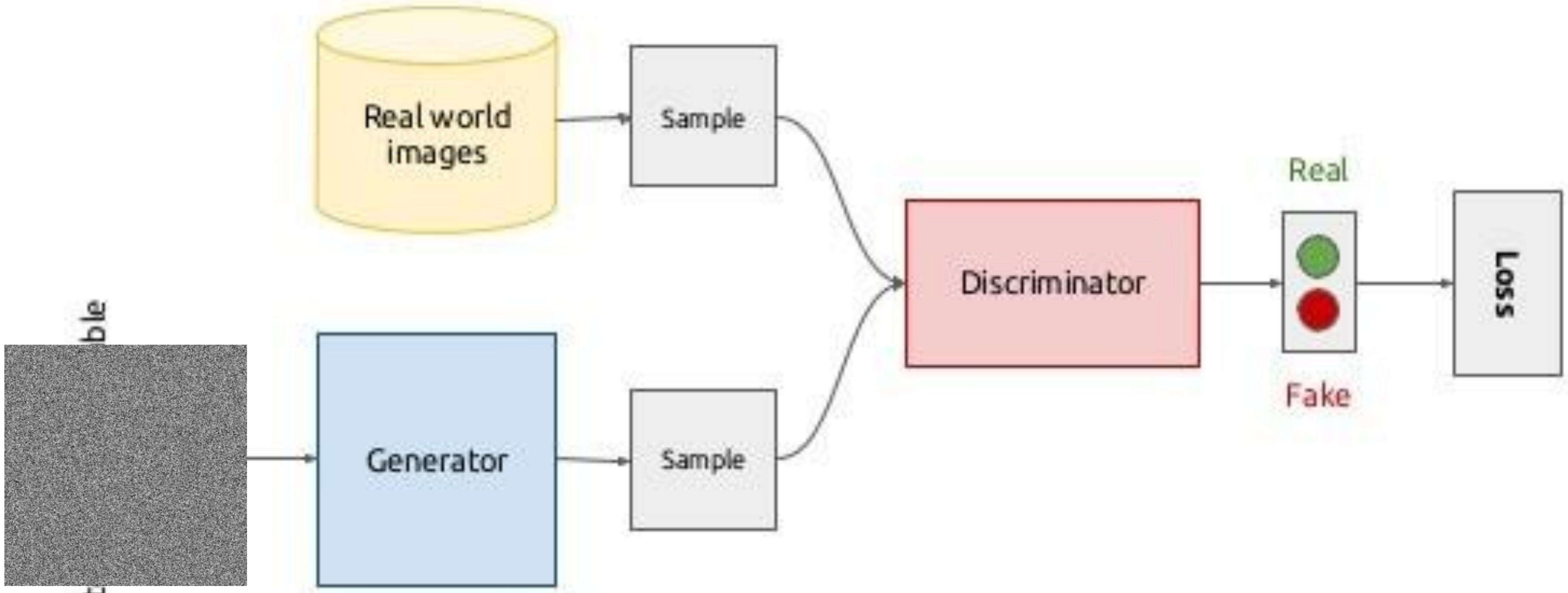


Source: [Artificia]

Analogy to counterfeit paintings



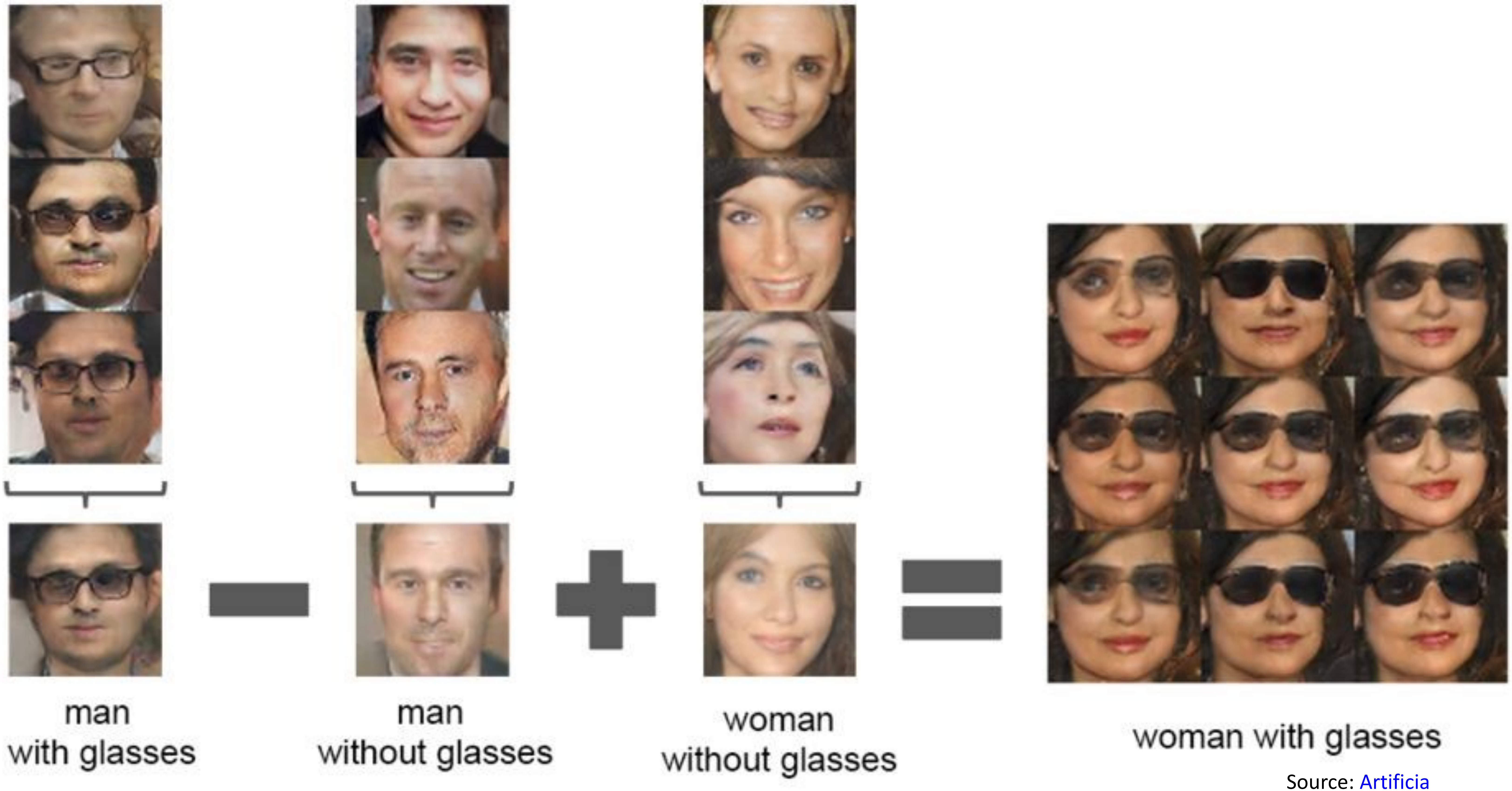
Generative adversarial networks (conceptual)



Noise

Source: [Artificia]

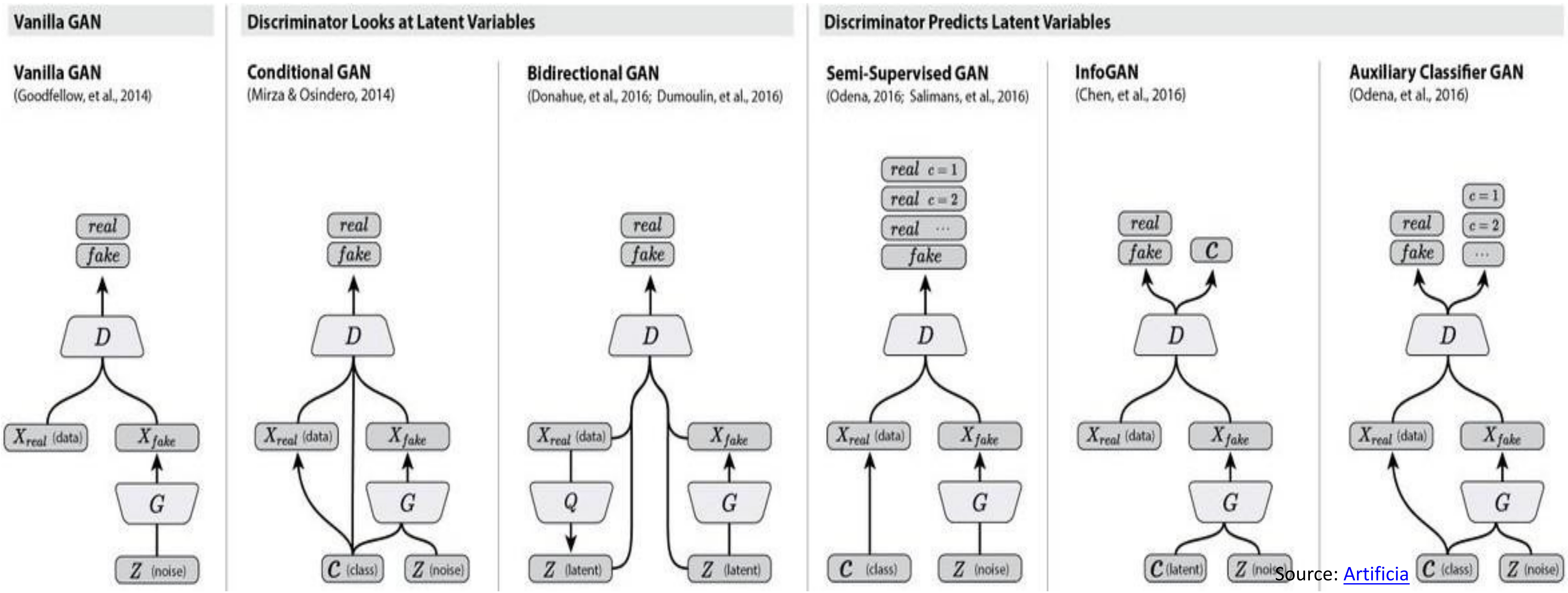
Bonus! - Vector space arithmetic of GANs



Source: Artificia

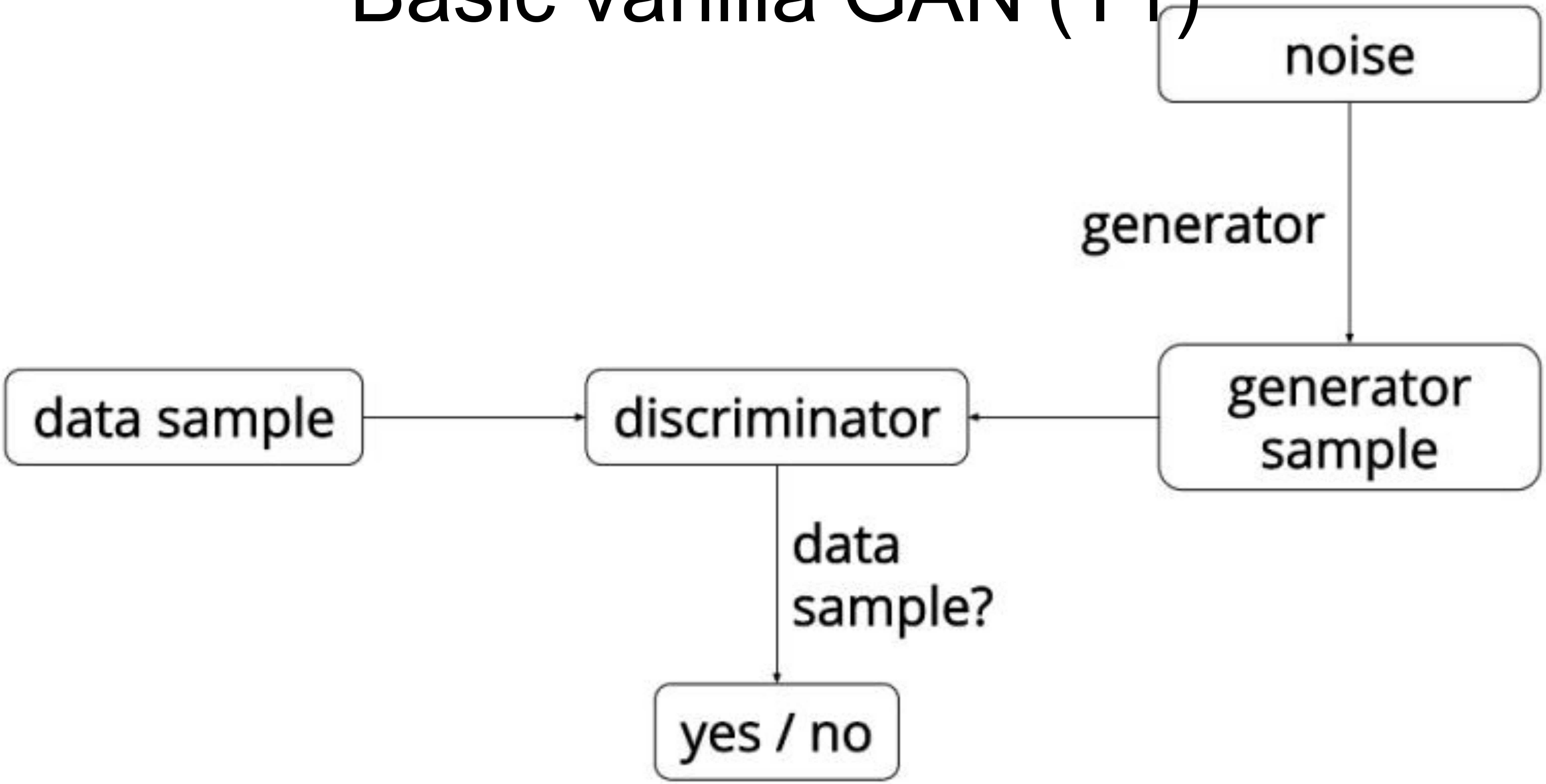
Types of GANs

Sub classifications



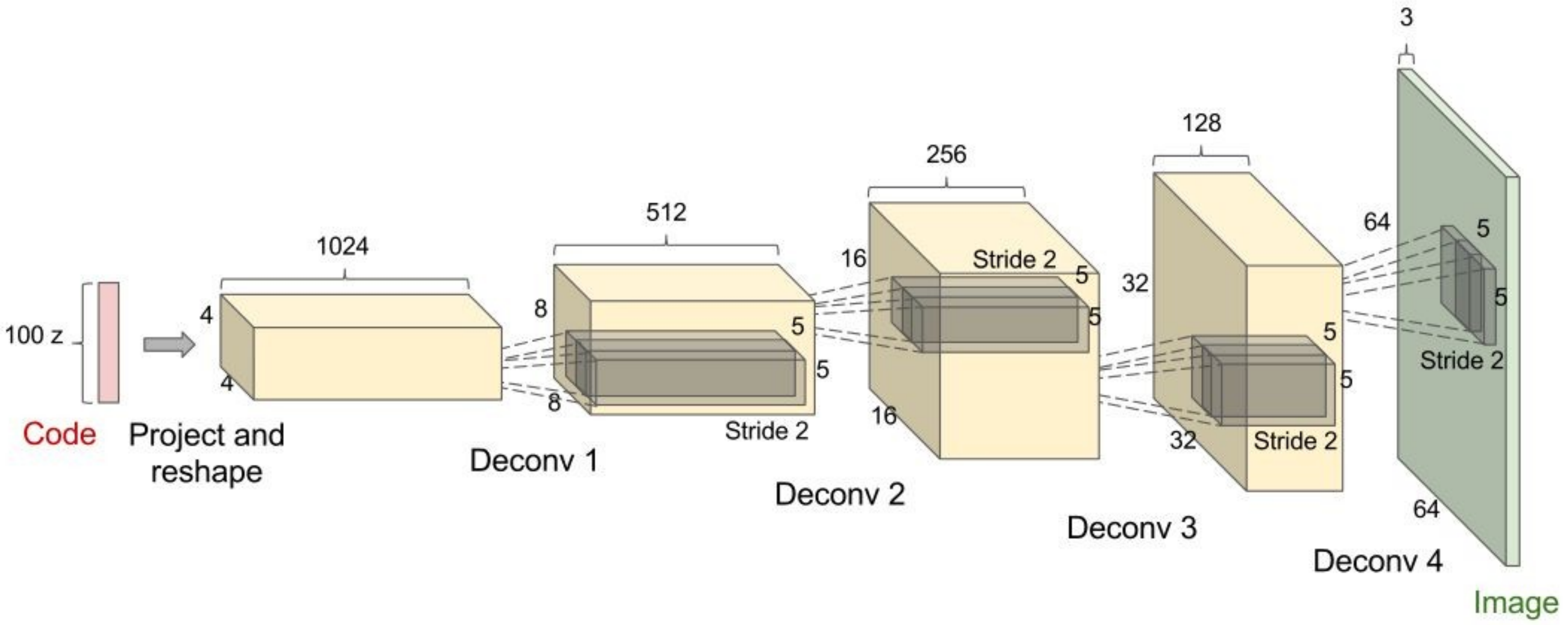
Source: [Artificia]()

Basic vanilla GAN (T1)



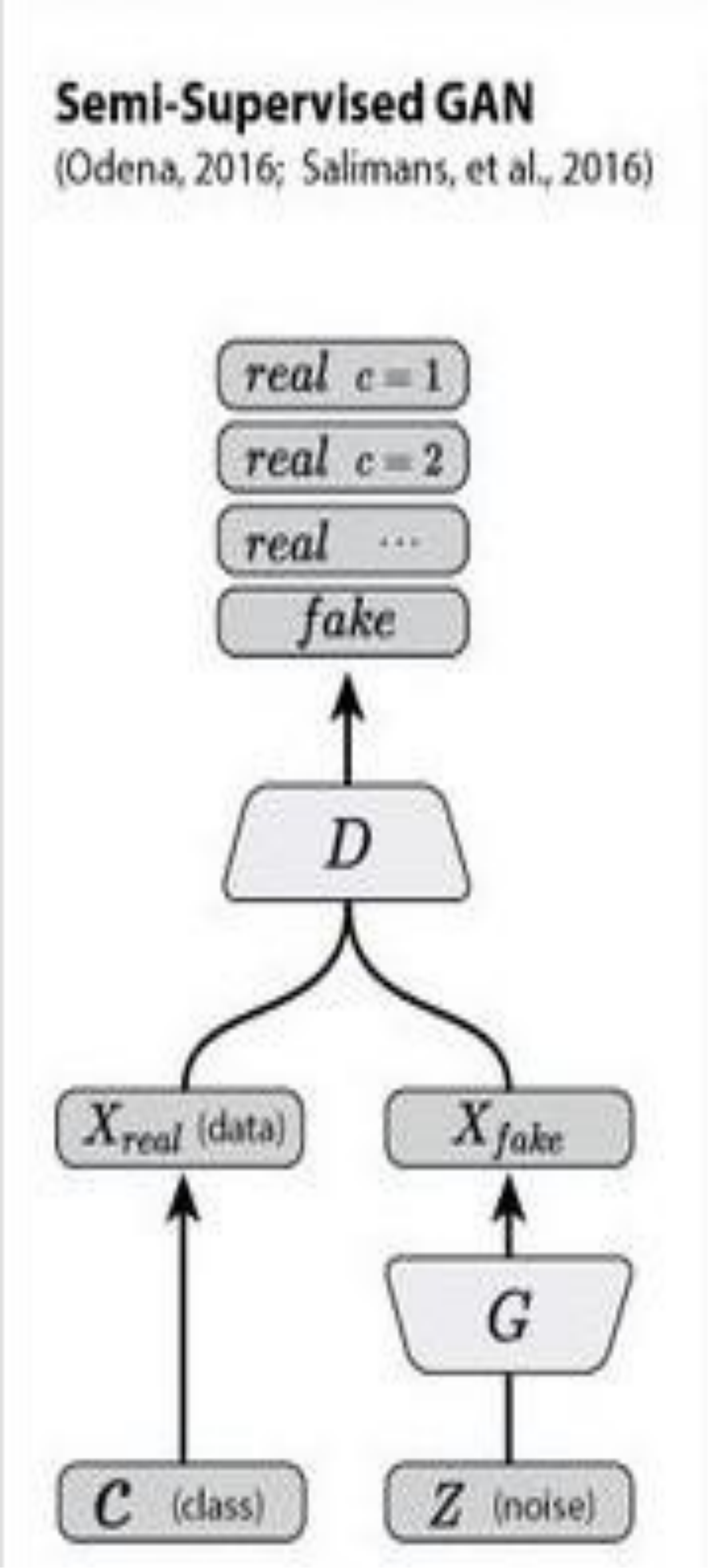
Source: [Artificia]

Deep Convolutional GAN (T1)



Source: [Artificia]()

Semi-supervised GAN (T3)



Source: [Artificia]

Semi supervised GAN

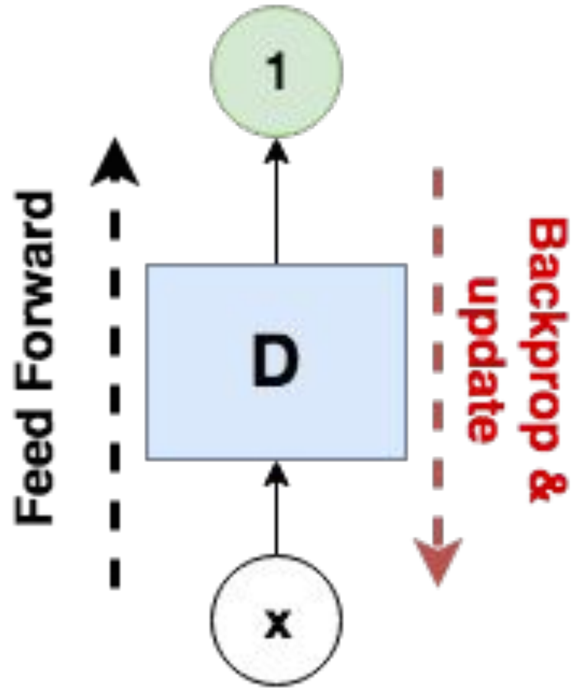
In addition to generating pretty pictures, we introduce an approach for semi-supervised learning with GANs that involves the discriminator producing an additional output indicating the label of the input. This approach allows us to obtain state of the art results on MNIST, SVHN, and CIFAR-10 in settings with very few labeled examples.

On MNIST, for example, we achieve **99.14% accuracy with only 10 labeled examples** per class with a fully connected neural network — a result that's very close to the best known results with fully supervised approaches using all 60,000 labeled examples. This is very promising because labeled examples can be quite expensive to obtain in practice.

"99.14% accuracy with only 10 labeled examples"

Adversarial Training (batch update)

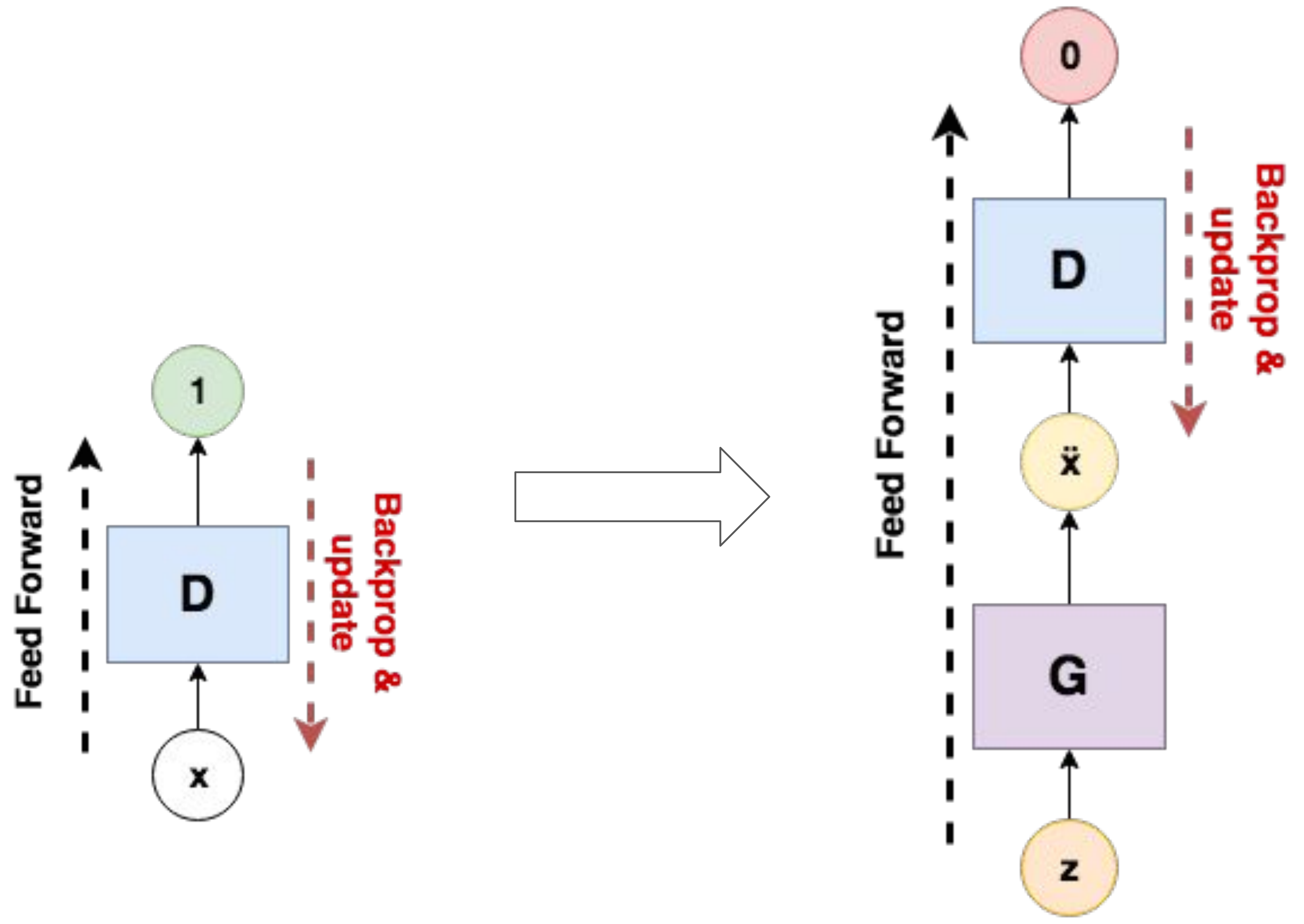
- Pick a sample ***x*** from training set
- Show ***x*** to **D** and update weights to output 1 (real)



Sourcek: [Source 2]

Adversarial Training (batch update)

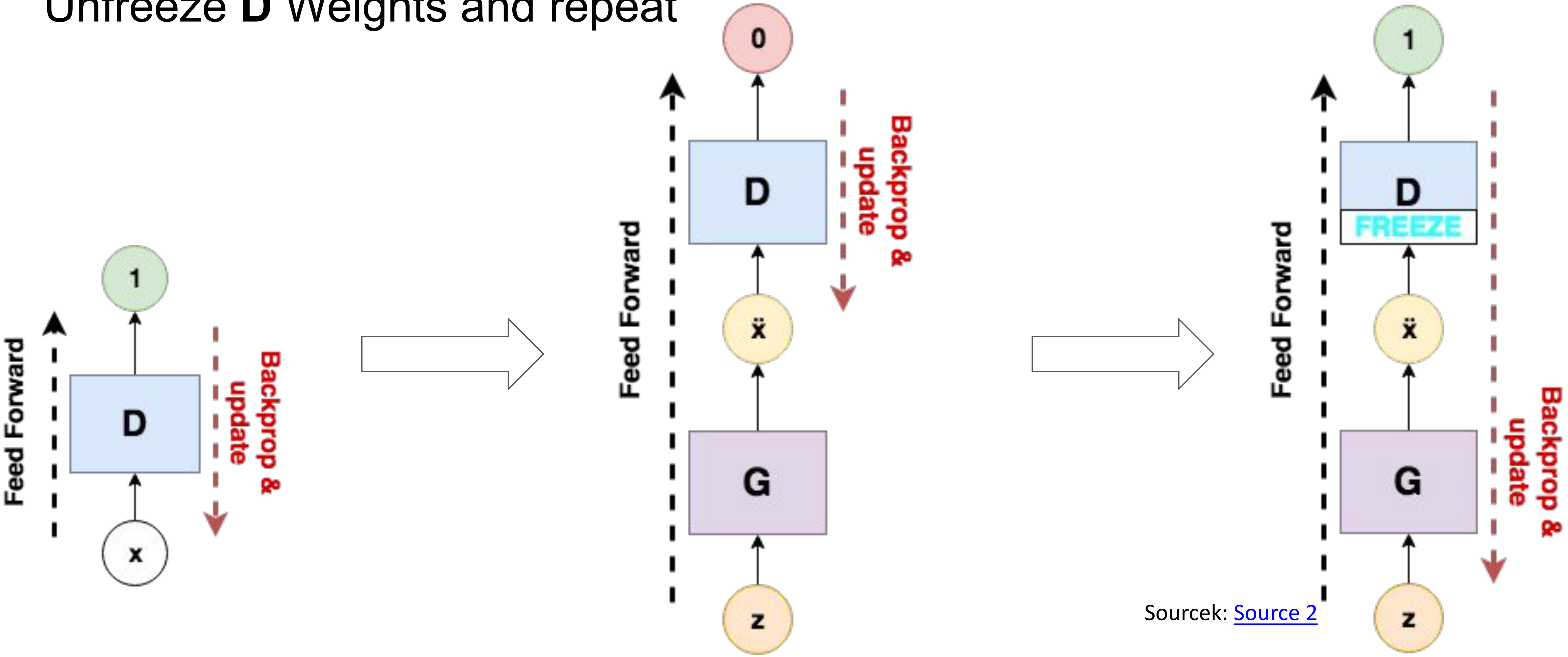
- **G** maps sample ***z*** to ***ẍ***
- show ***ẍ*** and update weights to output 0 (fake)



Sourcek: [Source 2]

Adversarial Training (batch update)

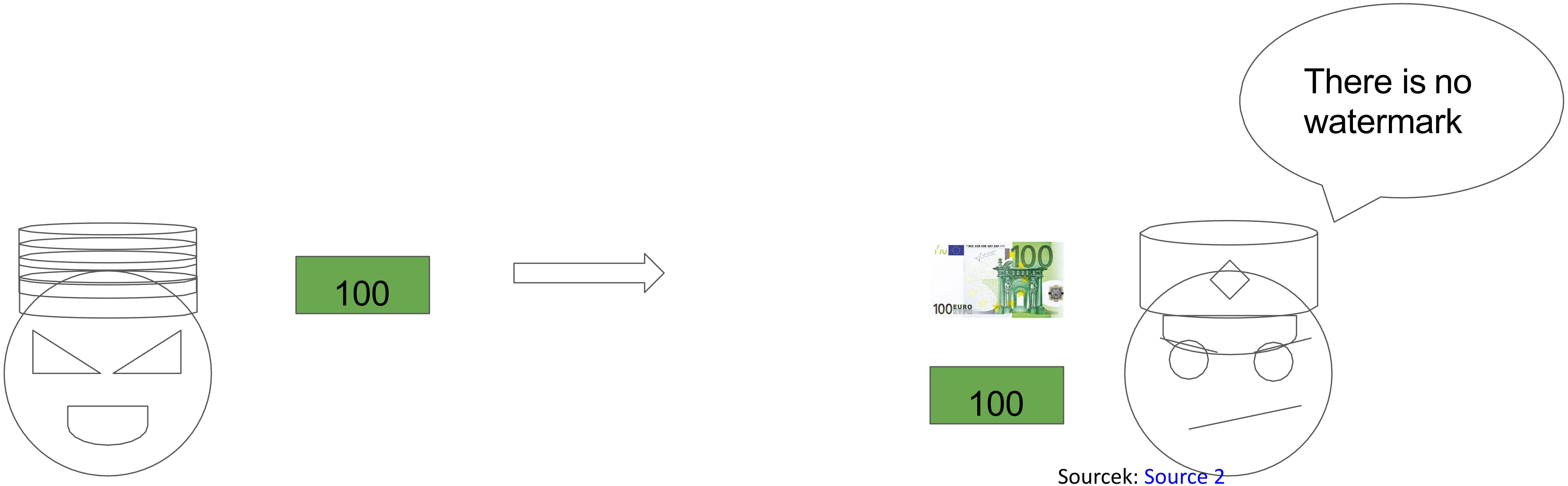
- Freeze **D** weights
- Update **G** weights to make **D** output 1 (just **G** weights!)
- Unfreeze **D** Weights and repeat



Sourcek: [Source 2]

Adversarial Training analogy

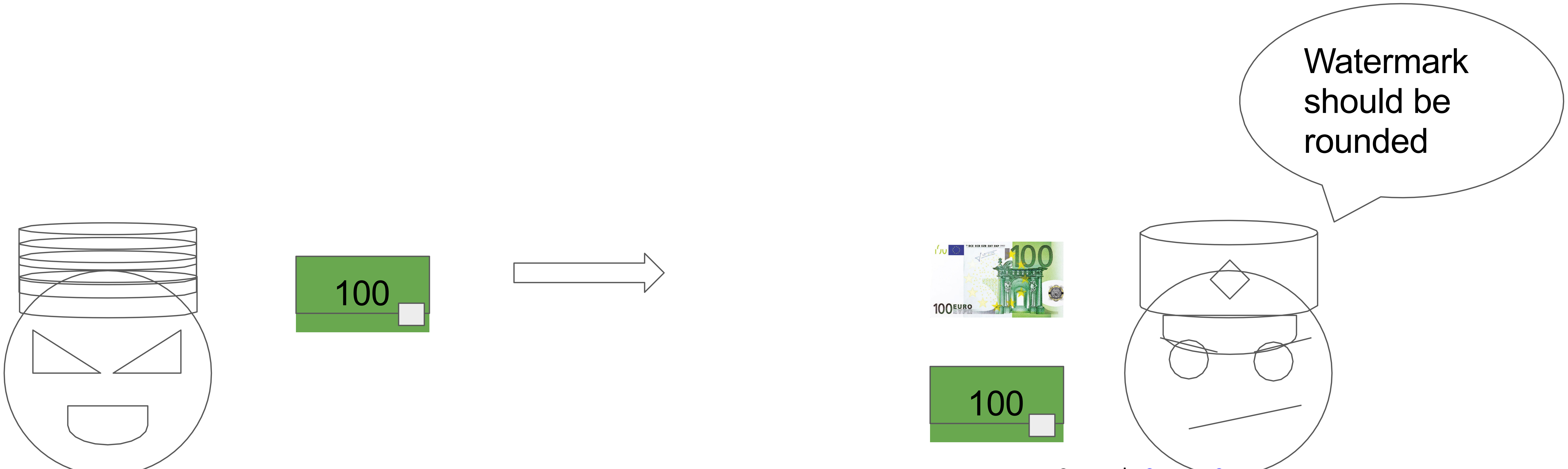
Imagine we have a counterfeiter (**G**) trying to make fake money, and the police (**D**) has to detect whether money is real or fake.



Sourcek: Source 2

Adversarial Training analogy

Imagine we have a counterfeiter (**G**) trying to make fake money, and the police (**D**) has to detect whether money is real or fake.

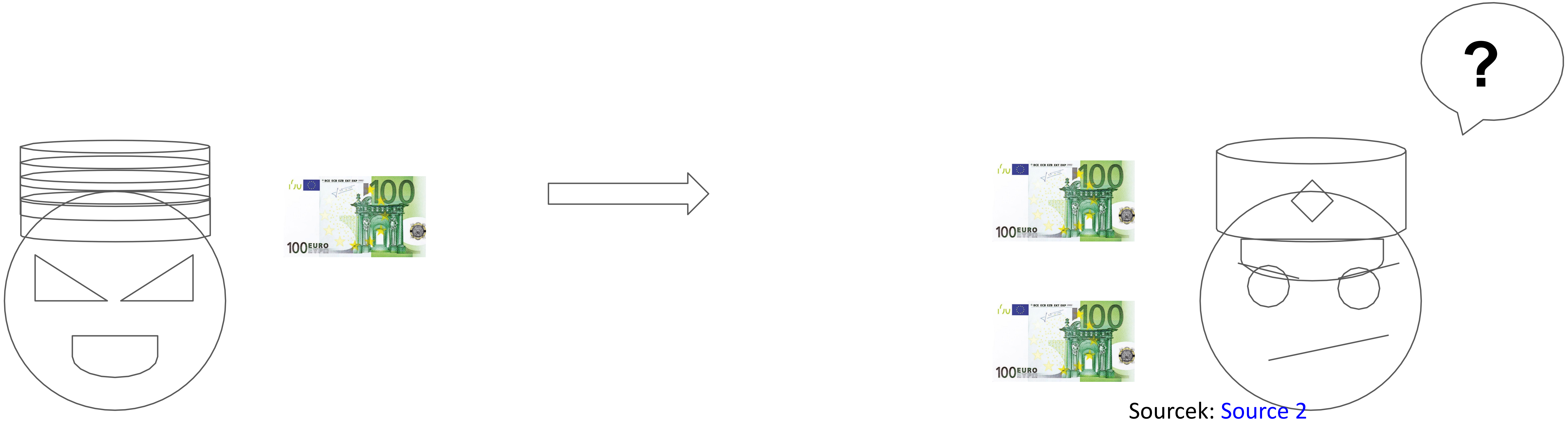


Sourcek: [Source 2]()

Adversarial Training analogy

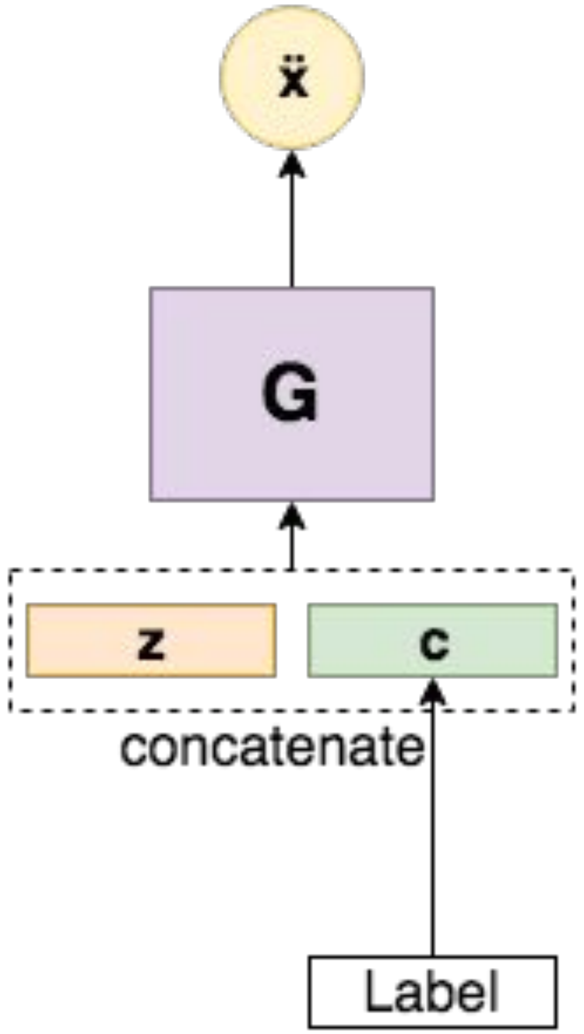
Imagine we have a counterfeiter (**G**) trying to make fake money, and the police (**D**) has to detect whether money is real or fake.

After enough iterations, and if the counterfeiter is good enough (in terms of **G** network it means “has enough parameters”), the police should be confused.

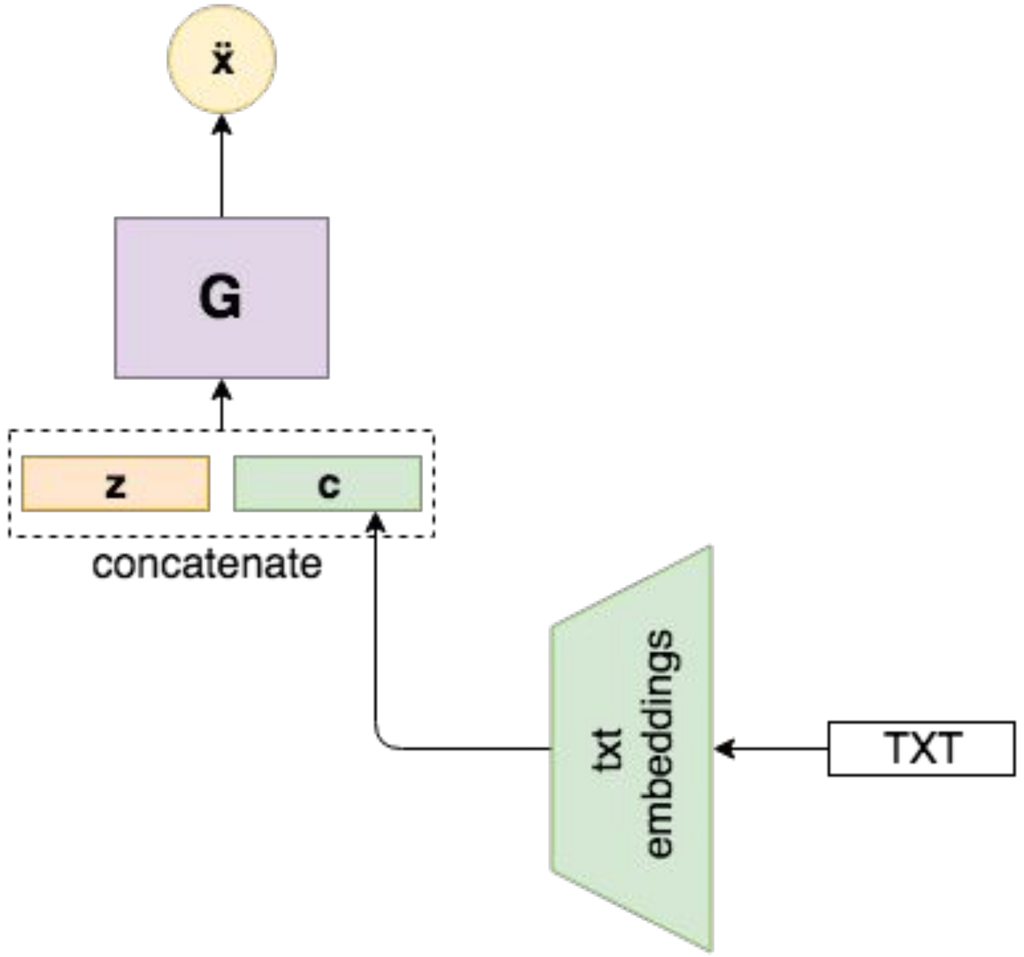


Sourcek: [Source 2]

Conditioned GANs



For details on ways to condition GANs:
[Ways of Conditioning Generative Adversarial Networks (Wack et al.)]



GANs can be conditioned on other info extra to **z**: text, labels, etc..

Sourcek: Source 2

z might capture random characteristics of the data, variabilities of possible futures, whilst **c** would condition the deterministic parts

Applications of GANs

GAN Applications

So far GANs have been extensively used in computer vision tasks:

- Generating images/generating video frames
- Unsupervised feature extraction/learning representations
- Manipulating images (in a photoshop advanced level)
- Image coding/Super Resolution
- Transferring image styles

However we have been working on advances for speech generation!

Source: Source 2

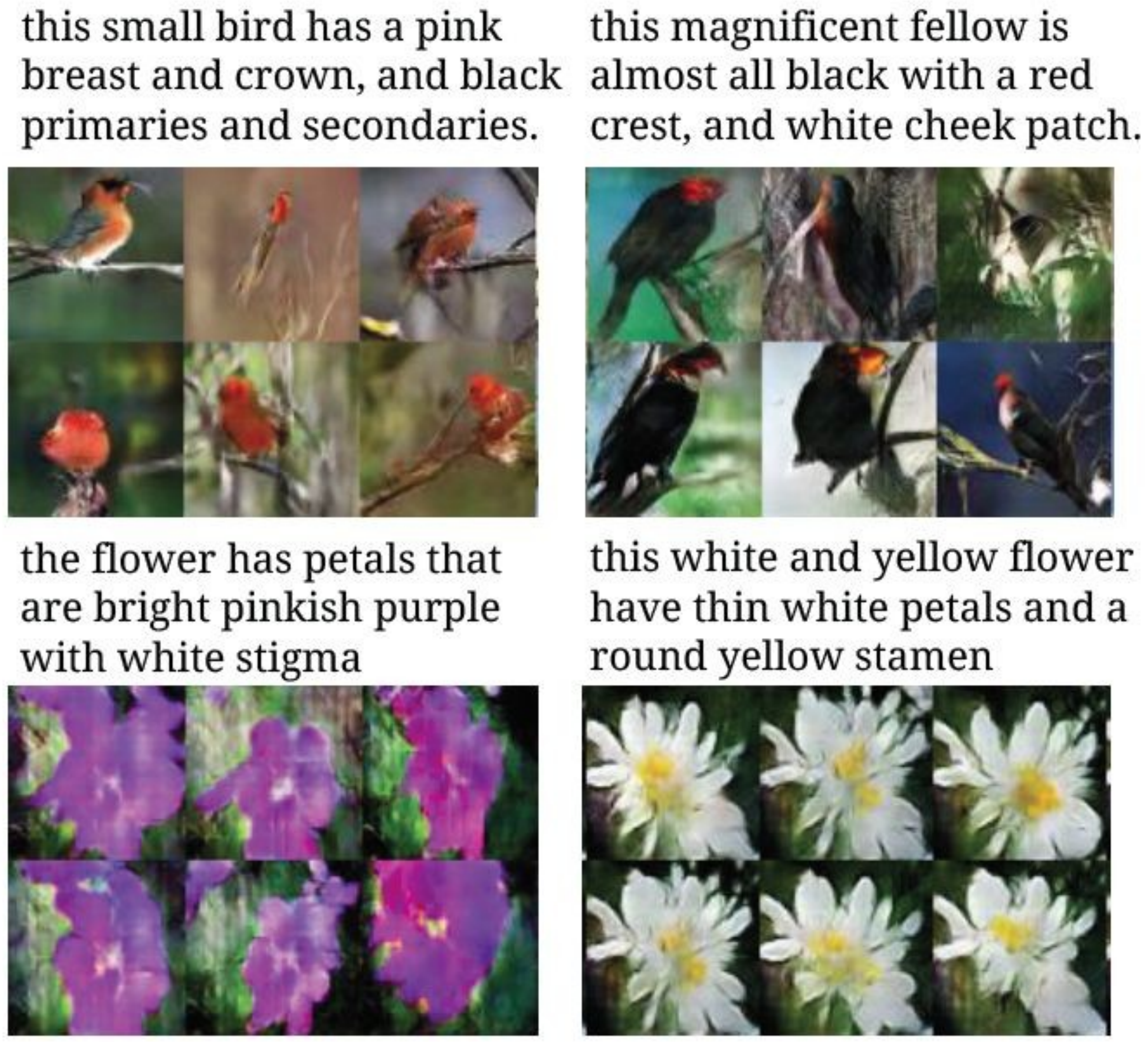
Generating images/frames

Deep Conv. GAN (DCGAN) effectively generated 64x64 RGB images in a single shot. For example bedrooms from LSUN dataset.

(Radford et al. 2015)

Source: Source 2

Generating images/frames conditioned on captions



(Reed et al. 2016b)

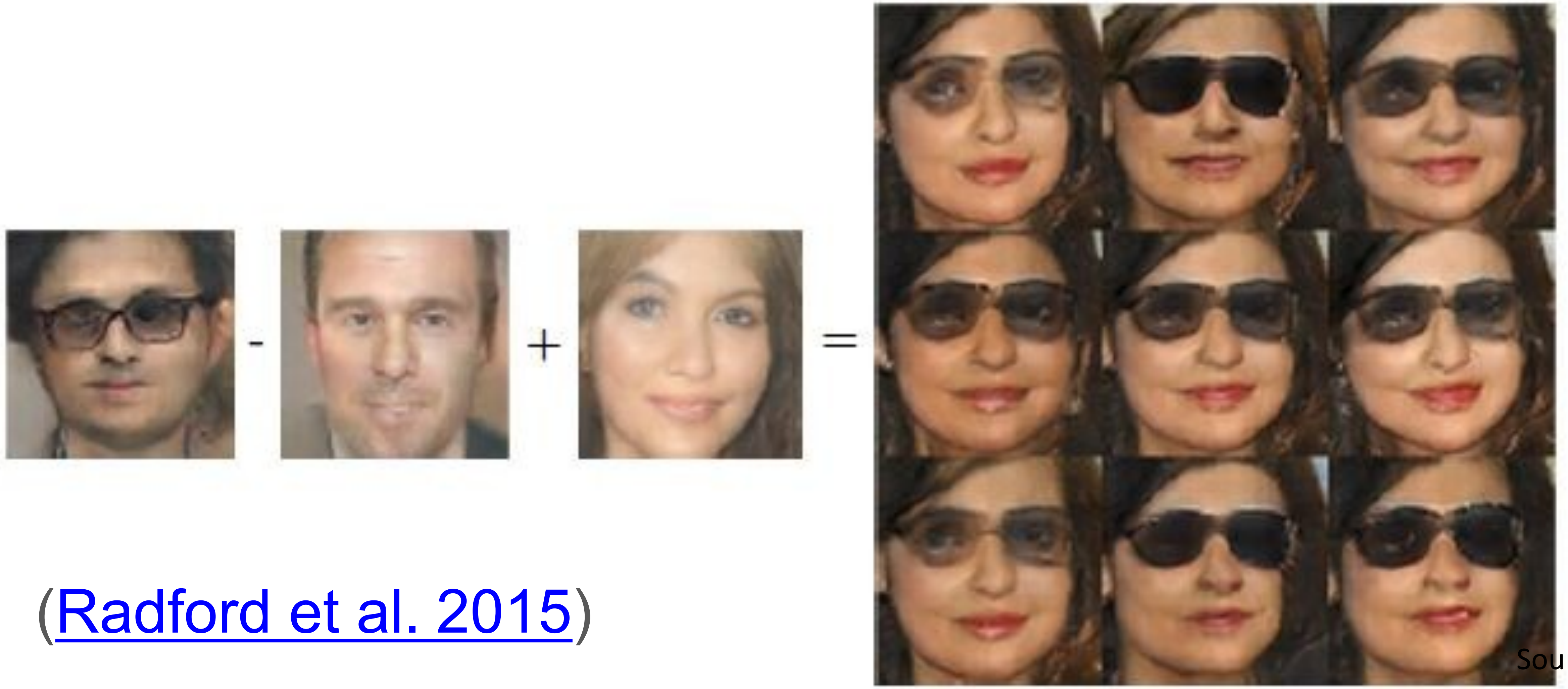
(Zhang et al. 2016)

Source: Source 2

Unsupervised feature extraction/ learning representations

Similarly to word2vec, GANs learn a distributed representation that disentangles concepts such that we can perform operations on the data manifold:

v(Man with glasses) - v(man) + v(woman) = v(woman with glasses)



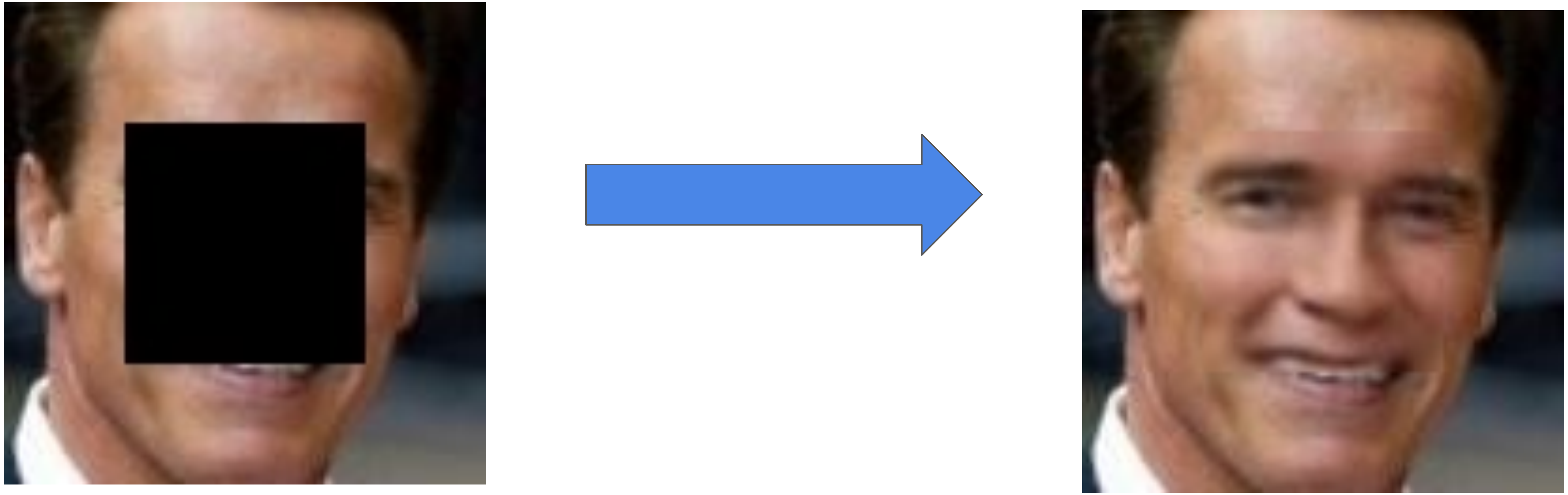
(Radford et al. 2015)

Source: Source 2

SRGANs

Source: [Source 2]

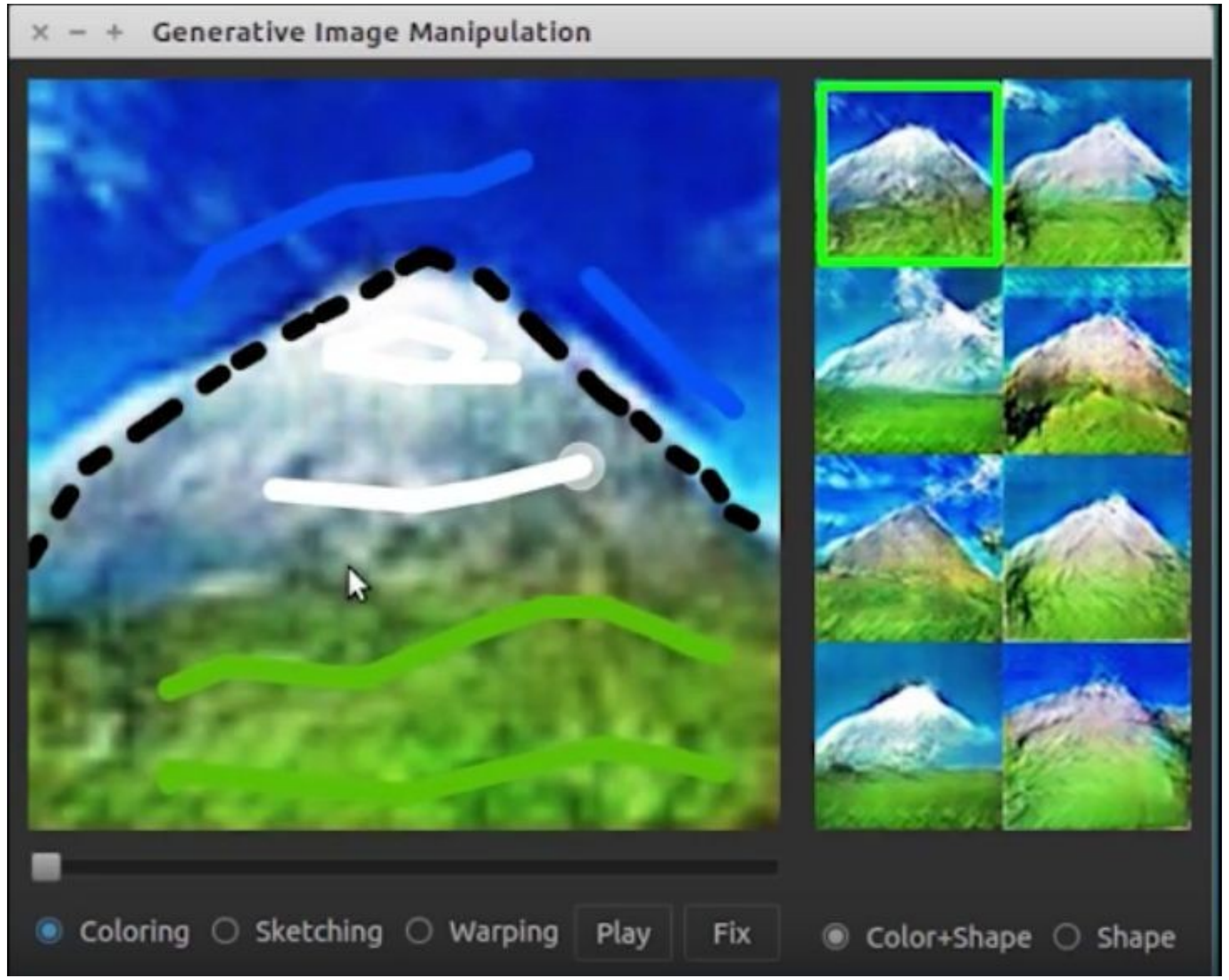
Image inpainting



Input to GAN

GAN generated output

Manipulating images and assisted content creation



https://youtu.be/9c4z6YsBGQ0?t=126

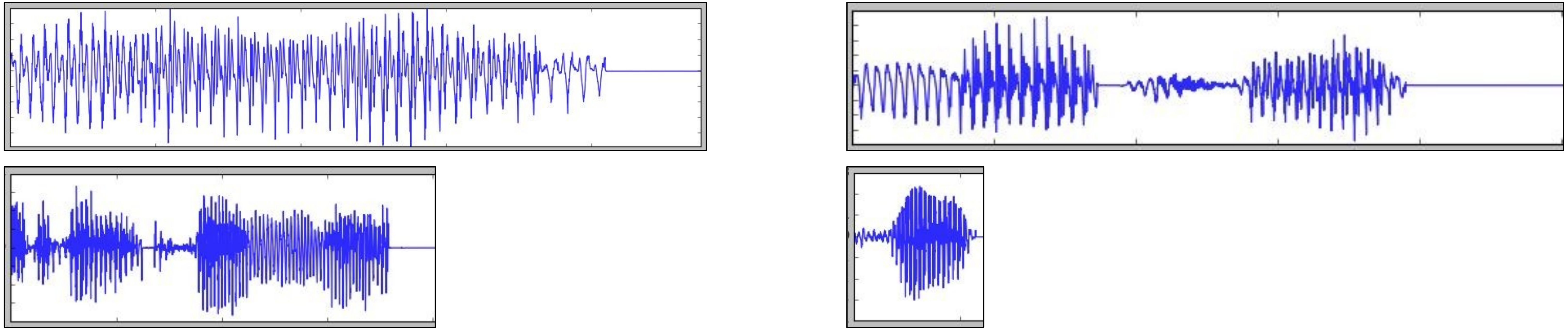
https://youtu.be/9c4z6YsBGQ0?t=161

(Zhu et al. 2016)

Source: Source 2

Waveforms generation

We have done recent advances in generating speech signals with GAN models.
There are none existent systems doing so until now. Current line of research.

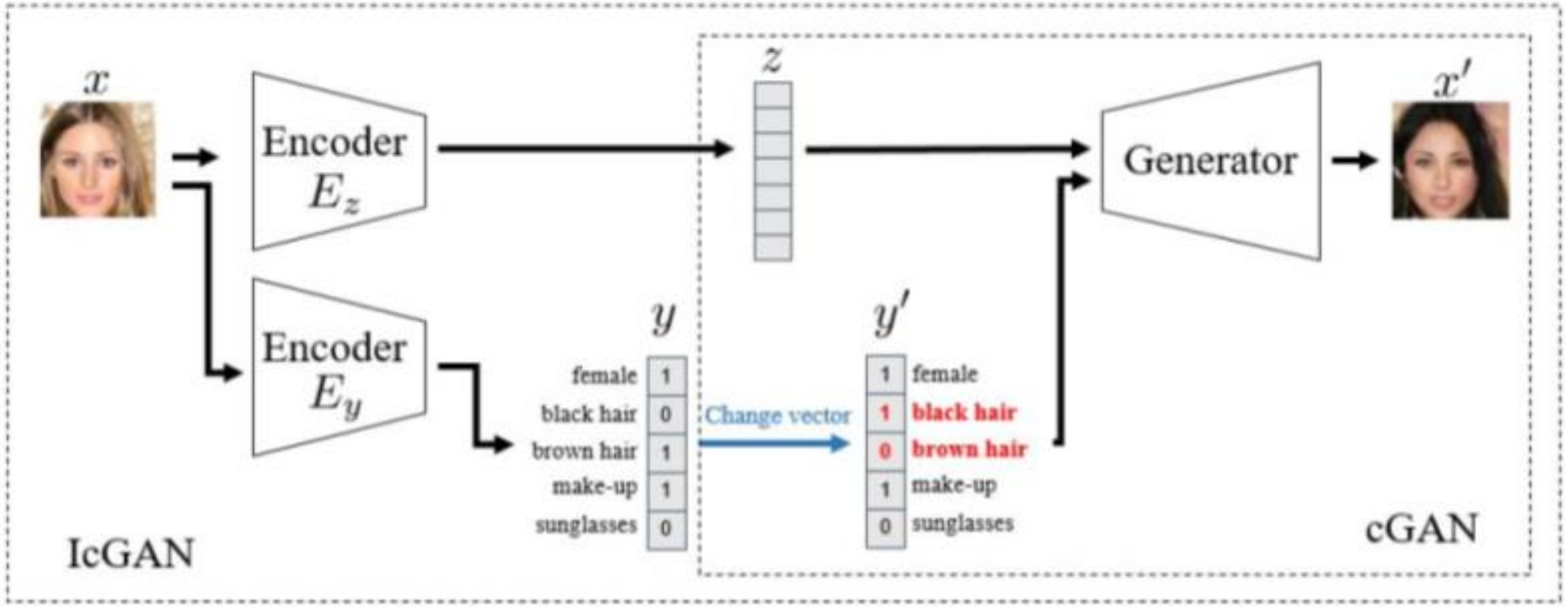


(S.Pascual, A.Bonafonte, J.Serrà)

Source: [Source 2]

Translating image

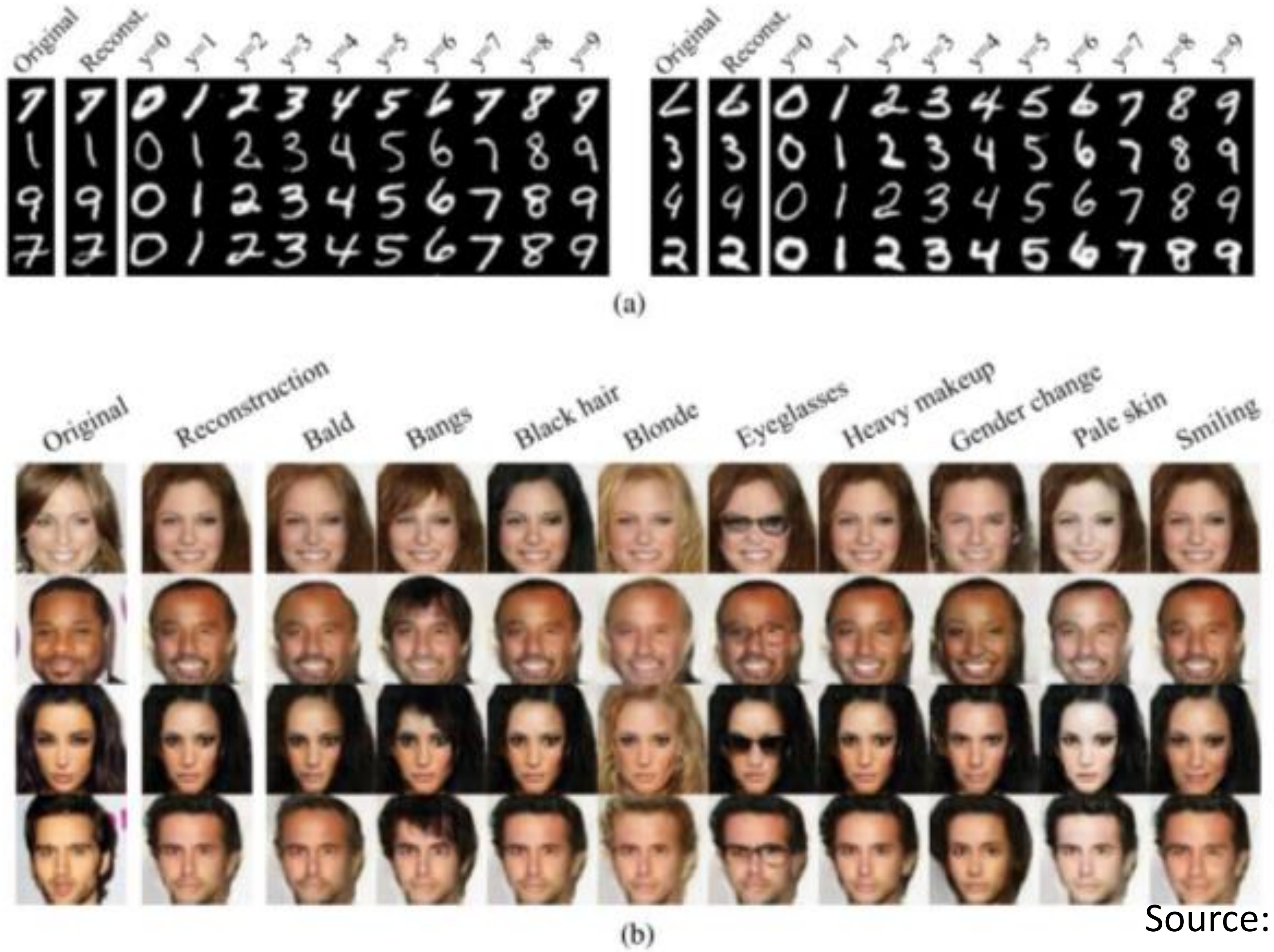
- Perarnau et al, Invertible Conditional GANs for image editing, 2016



Source: [NamjuKim]

Translating image

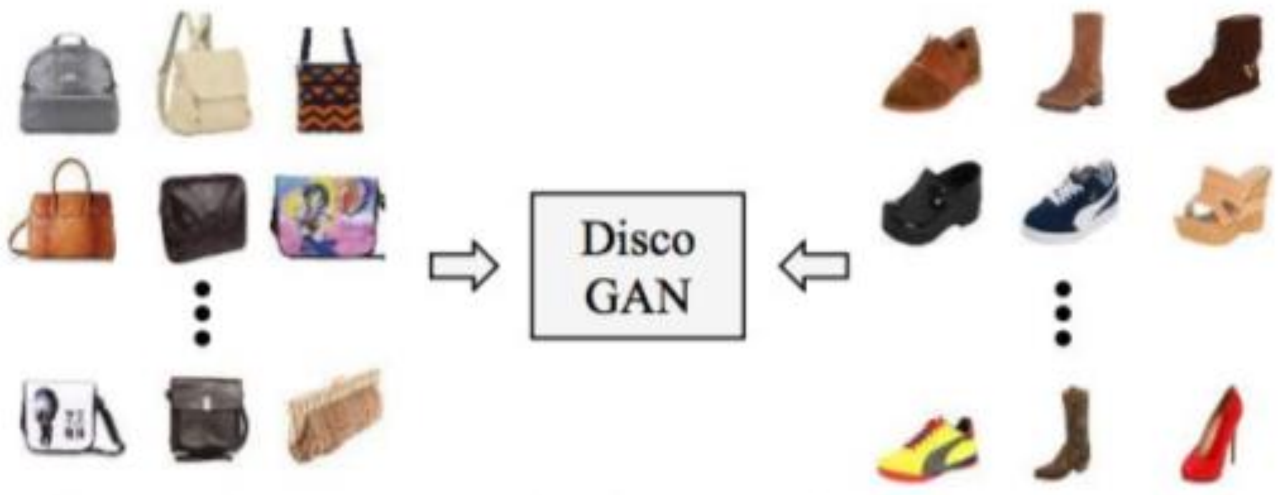
- Perarnau et al : results



Source: [NamjuKim]()

Domain adaptation

- Extended Taigman's idea using smart bijective mapping



(a) Learning cross-domain relations **without any extra label**

(b) Handbag images (input) & **Generated** shoe images (output)

(c) Shoe images (input) & **Generated** handbag images (output)

Source: NamjuKim

VIDEO EDITING

Andrej Karpathy
@karpathy

Follow

One day we'll be talking about good old "hand-crafted" films and instead the norm will be watching AI-generated (infinite) content on demand

9:58 AM - Mar 29, 2016

29 133 244

Caveats

Where is the downside...?

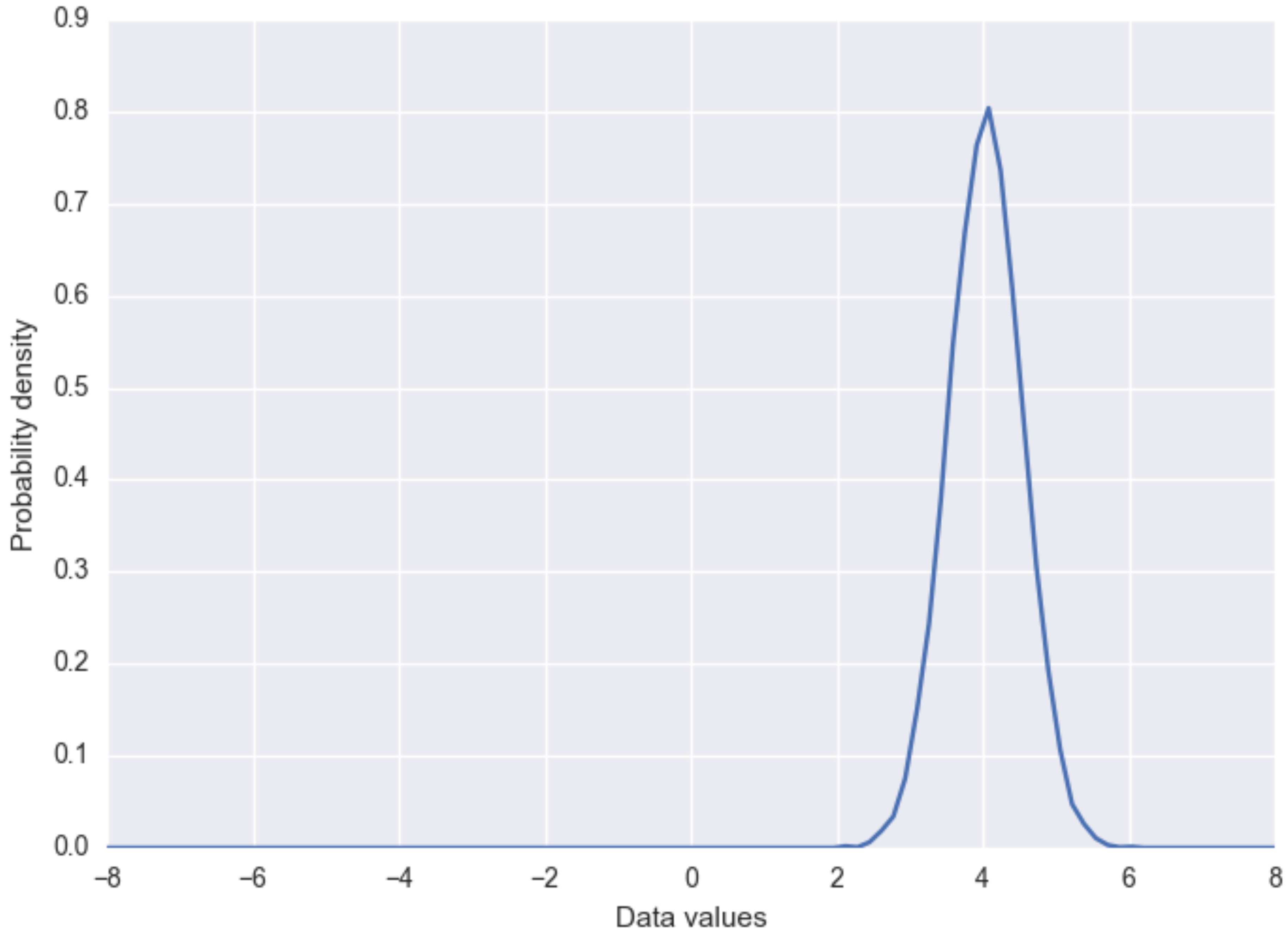
Well GANs are tricky and hard to train! We do not want to minimize a cost function. Instead we want both networks to reach a Nash equilibria (saddle point).

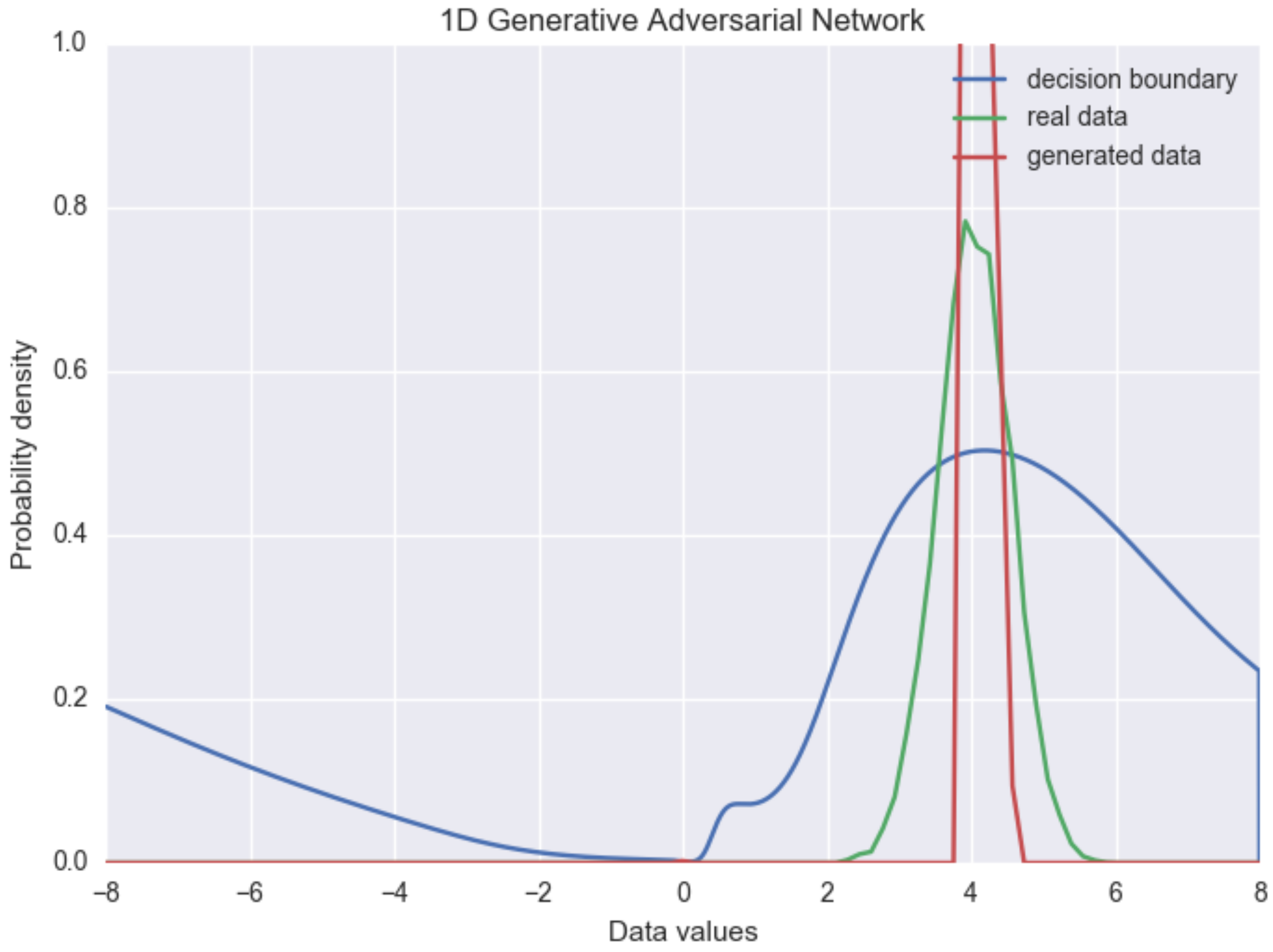
Because of extensive experience within the GAN community (with some does-not-work-frustration from time to time), you can find some tricks and tips on how to train a GAN here: https://github.com/soumith/ganhacks

An open problem of GANs is the evaluation of their generation quality: there are no established objective metrics to do so → We look (or hear) the generated samples to know when to stop training or how are we doing.

Source: Source 2

Approximating a 1D Gaussian distribution



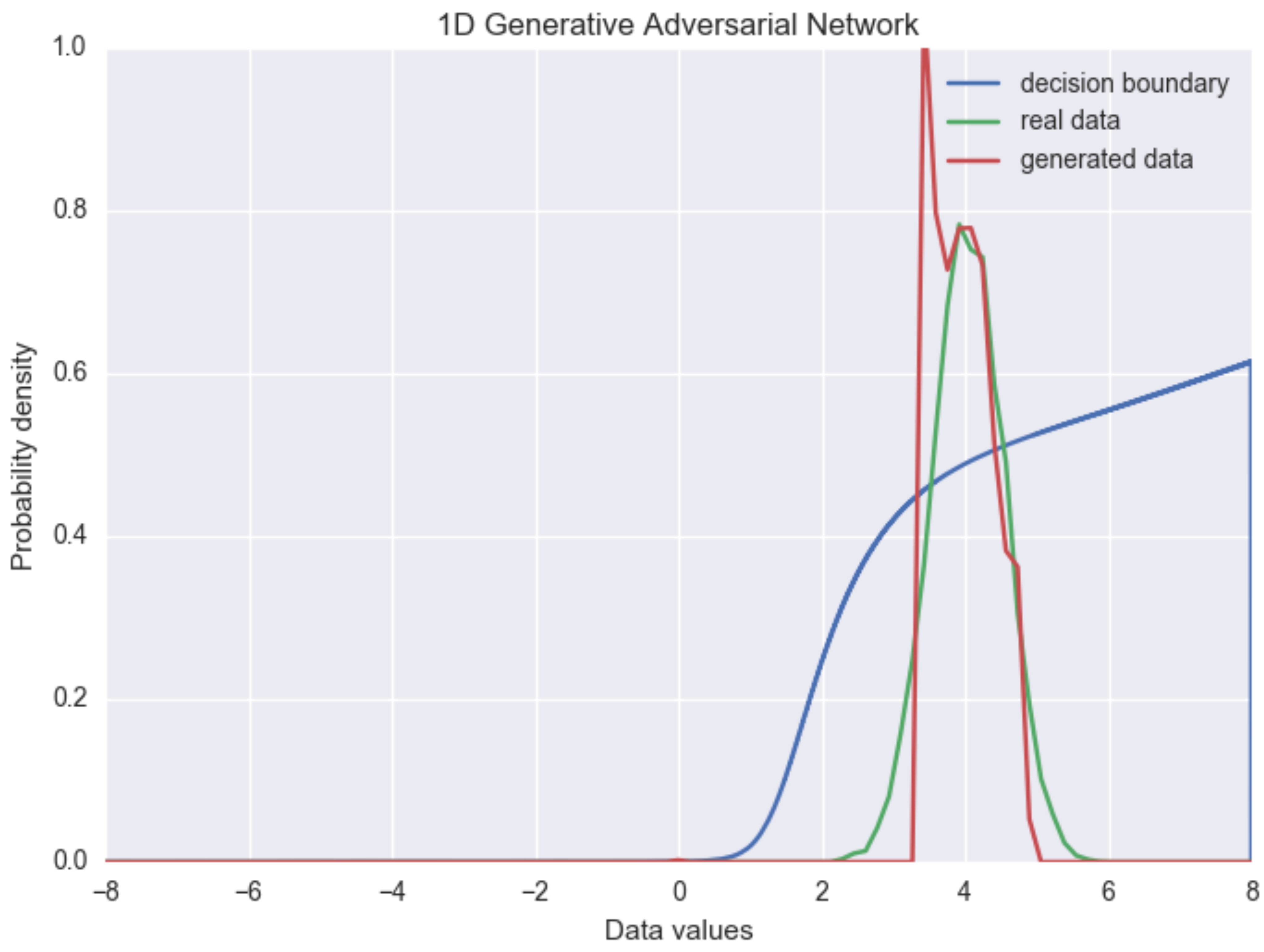

1D Generative Adversarial Network
decision boundary
real data
generated data
Probability density
Data values
1.0
0.8
0.6
0.4
0.2
0.0
−8
−6
−4
−2
0
2
4
6
8

MINIBATCH DISCRIMINATION

- In the paper, mini-batch discrimination is defined to be any method where the discriminator is able to look at an entire batch of samples in order to decide whether they come from the generator or the real data.

- They also present a more specific algorithm which works by modelling the distance between a given sample and all other samples in the same batch.

- These distances are then combined with the original sample and passed through the discriminator, so it has the option to use the distance measures as well as the sample values during classification.


1D Generative Adversarial Network
decision boundary
real data
generated data
Probability density
Data values
1.0
0.8
0.6
0.4
0.2
0.0
−8
−6
−4
−2
0
2
4
6
8

Limitations of GANs

Mode collapse

- Generator keeps generating highly similar looking images

- Happens when the generator is optimized while keeping the discriminator constant for many iterations

Source: Artificia

Difficulty in reaching convergence

- Generator and discriminator losses keep oscillating
- Network does not converge to an optimal solution

Source: Artificia

Relative strength of the two networks

- Either of the two networks becoming extremely strong relative to the other

- Network never learns beyond this point

Source: [Artificia]()

Dealing with these issues - GAN Hacks

Normalizing images

- Standard practice of normalizing images by mean normalizing and scaling by stddev should work

- Make sure that the images are normalized to values between -1 and +1

- Paper explaining intuition [Sampling Generative Networks]

Source: [Artificia]

Inverting labels

- While training generator, flip labels sent to the discriminator

- Label of real image is made fake(say class 0) and that of generated image is made real(class 1)

Source: Artificia

Schedule for training G and D

- People tend to find a fixed (nIter_Gen/nIter_Disc) training schedule

- Better to train G or D based on a loss threshold

- if loss_G > A:
      train_G()
  If loss_D > B:
      train_D()

Source: Artificia

Noisy D input

- Add noise to inputs of the discriminator
- Label of real image is made fake(say class 0) and that of generated image is made real(class 1)
- Paper explaining intuition: Towards Principled Methods for Training Generative Adversarial Networks

Source: Artificia

Feature matching

- Modify the loss function of the generator to include intermediate feature activations rather than output

- Forces the generator to generate data that match the statistics of real data

Source: [Artificia]

One sided label smoothing

- Technique proposed in the 1980’s
- Instead of having hard labels like 0 and 1, smoothen the labels by making them close to 0 and 1
- For example, 0,1 -> 0.1,0.9

Source: [Artificia]()

THANK YOU!

@BOOKWORMENGR