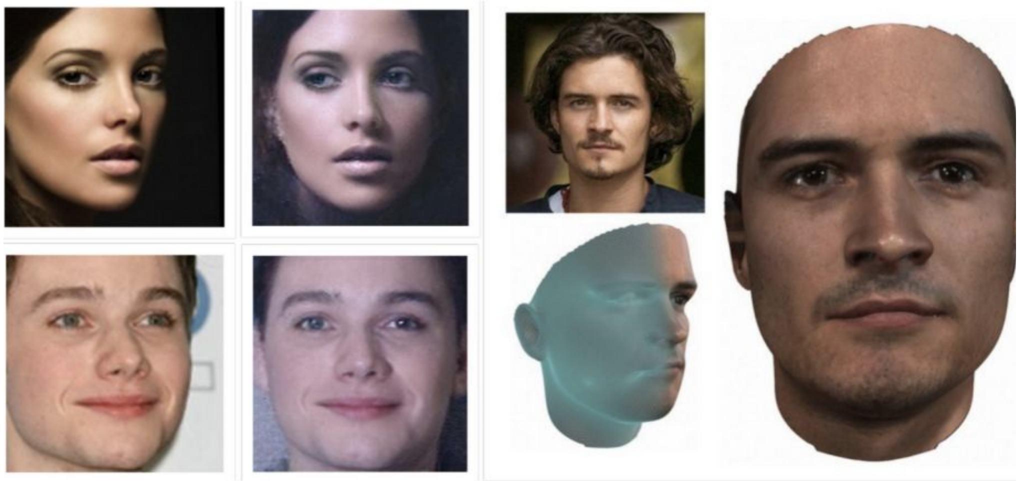
3D HUMAN BODY RECONSTRUCTION

Using Generative Adversarial Networks



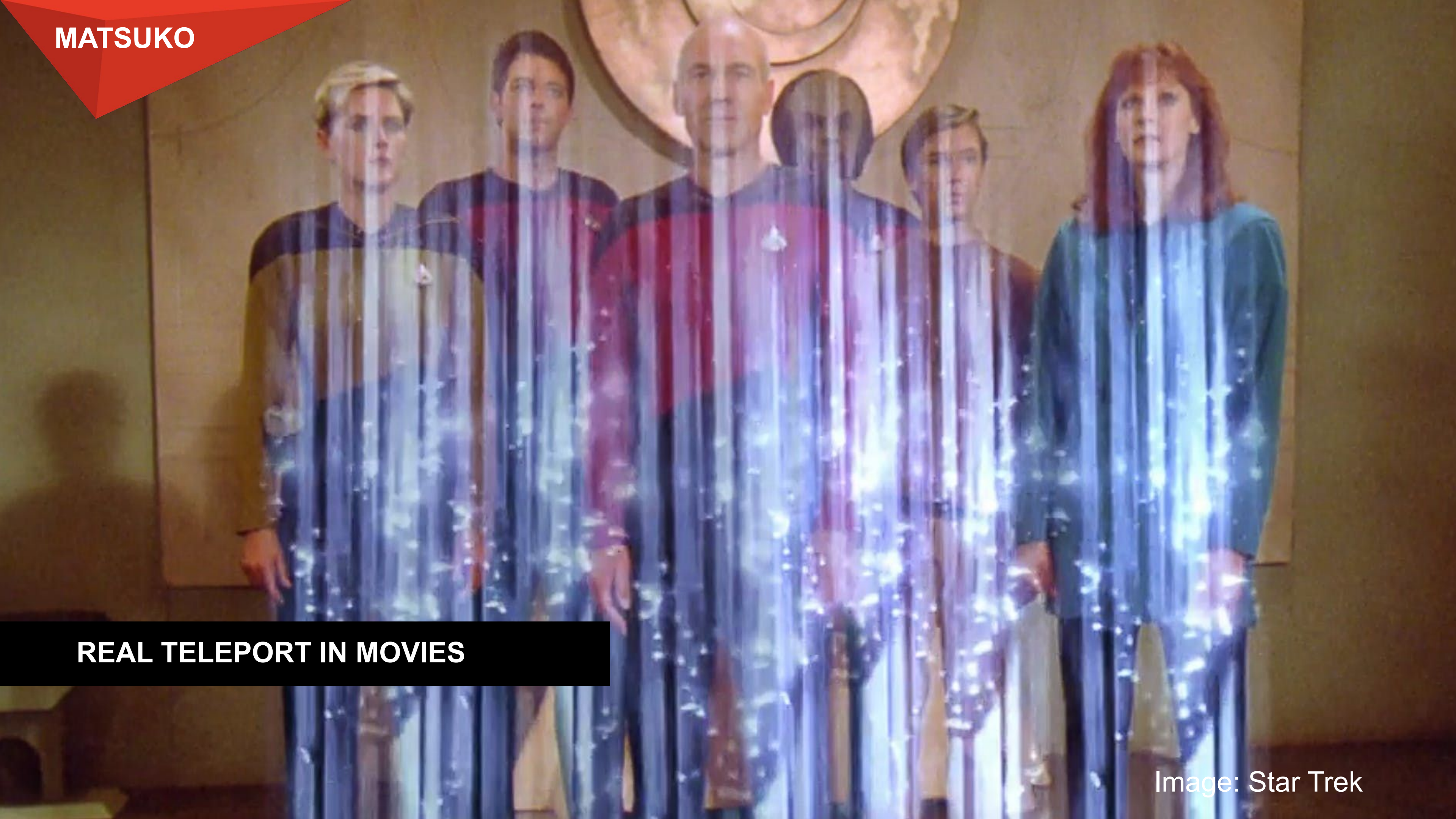
MATSUKO

1. Towards the 3D Telepresence
2. Why ML (unique selling point)
3. GANs
4. Examples of Use Cases
5. Matsuko's Real Case Studies
6. Deeper into GANs

1. Towards the 3D Telepresence

https://www.youtube.com/watch?time_continue=99&v=ptVhLn_CSBQ

MATSUKO


MATSUKO
REAL TELEPORT IN MOVIES
Image: Star Trek

HOLOGRAPHIC TELEPORT IN MIXED REALITY

Image: Spatial

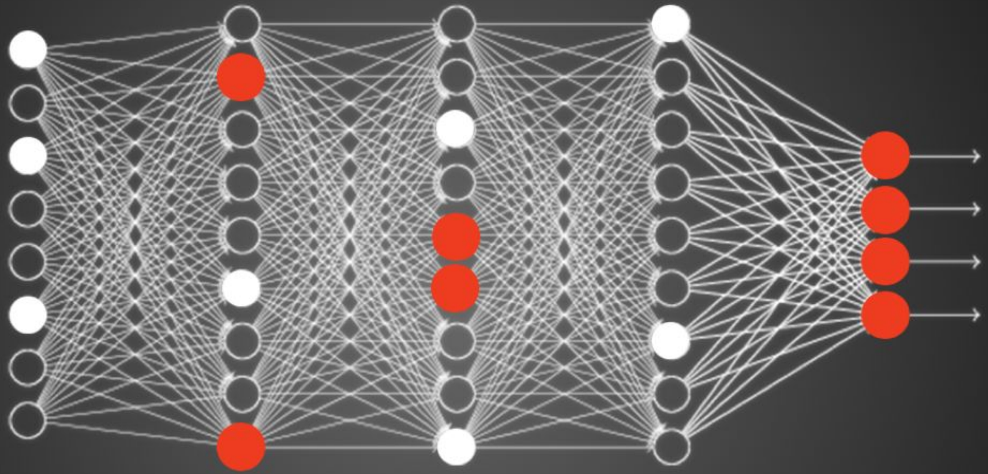
2. Why ML (unique selling point)

VOLUMETRIC VIDEO STUDIOS

Image: 8i

MATSUKO

We are transmitting people in 3D using only 1 camera.
Neural networks reconstruct the whole head / body only from one angle, even the non visible parts.



To do so, we are using mixed reality, artificial intelligence, 3D, computer vision, games technology to change the way people see and interact with virtual objects.

2D

DEEP LEARNING

3D

MATSUKO

Demo with photo

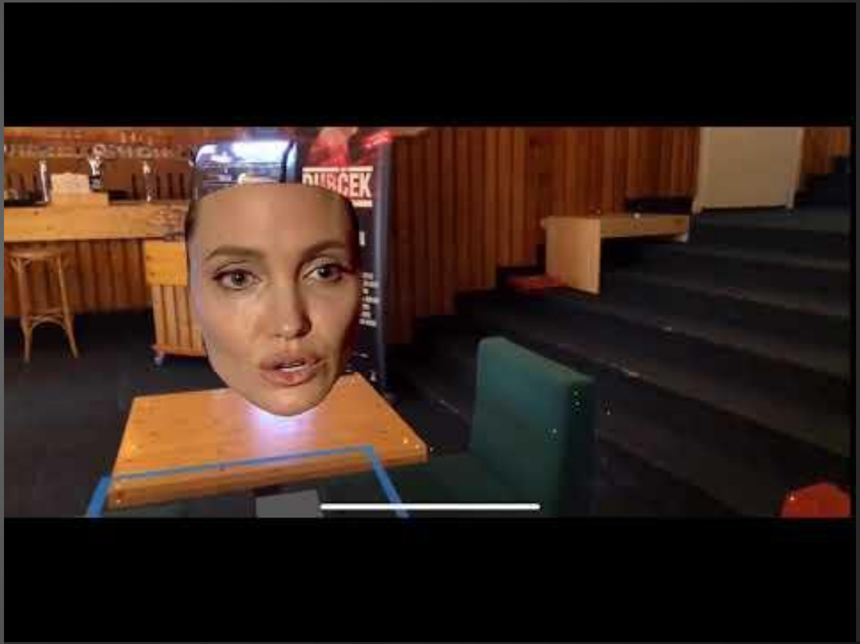
2D photo

MATSUKO static 3D model

MATSUKO

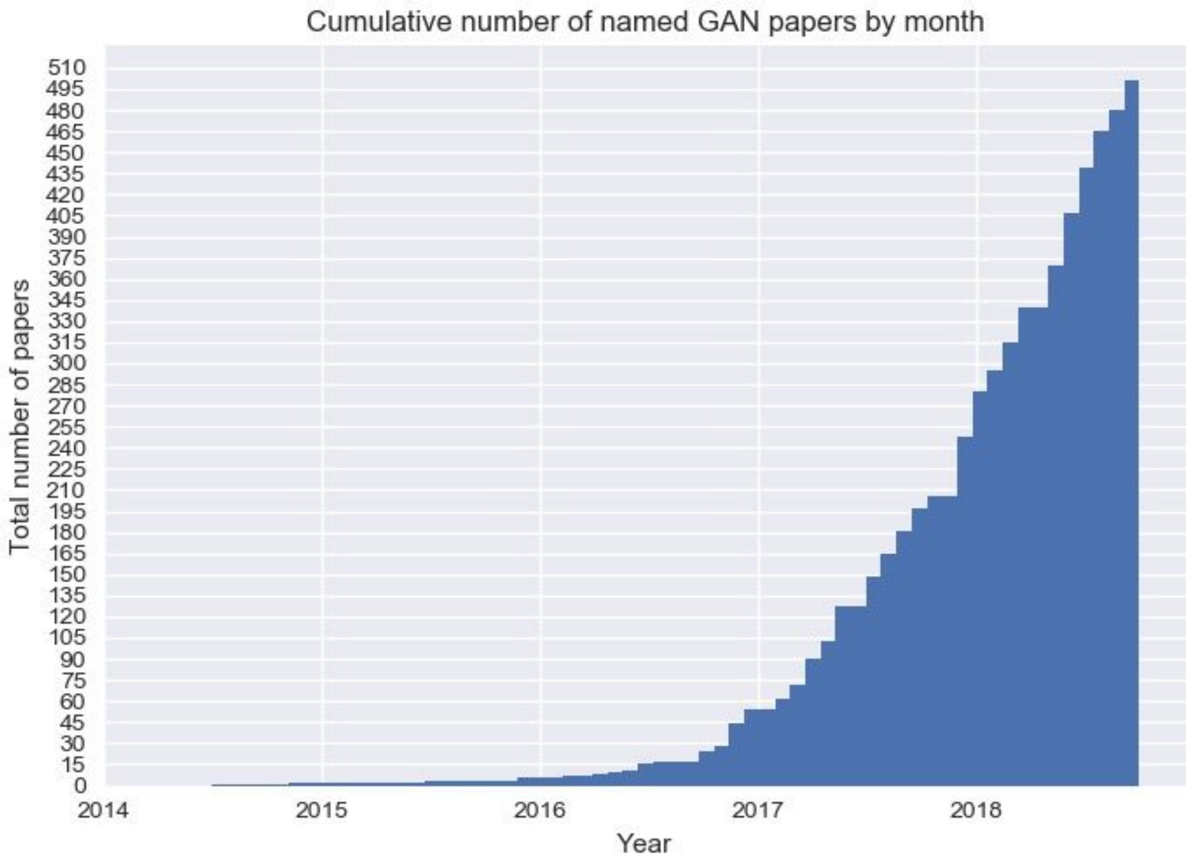
Demo with video

10” original video



MATSUKO 3D realistic reconstruction

MATSUKO



Cumulative number of named GAN papers by month
Total number of papers
510
495
480
465
450
435
420
405
390
375
360
345
330
315
300
285
270
255
240
225
210
195
180
165
150
135
120
105
90
75
60
45
30
15
0
2014
2015
2016
2017
2018
Year

3. GANs

MATSUKO

Real or GAN generated?

Image: https://thispersondoesnotexist.com/

It all started with a … beer

The GANfather

“the coolest idea in deep learning in the last 20 years” - Yann LeCun (Facebook’s chief AI scientist)

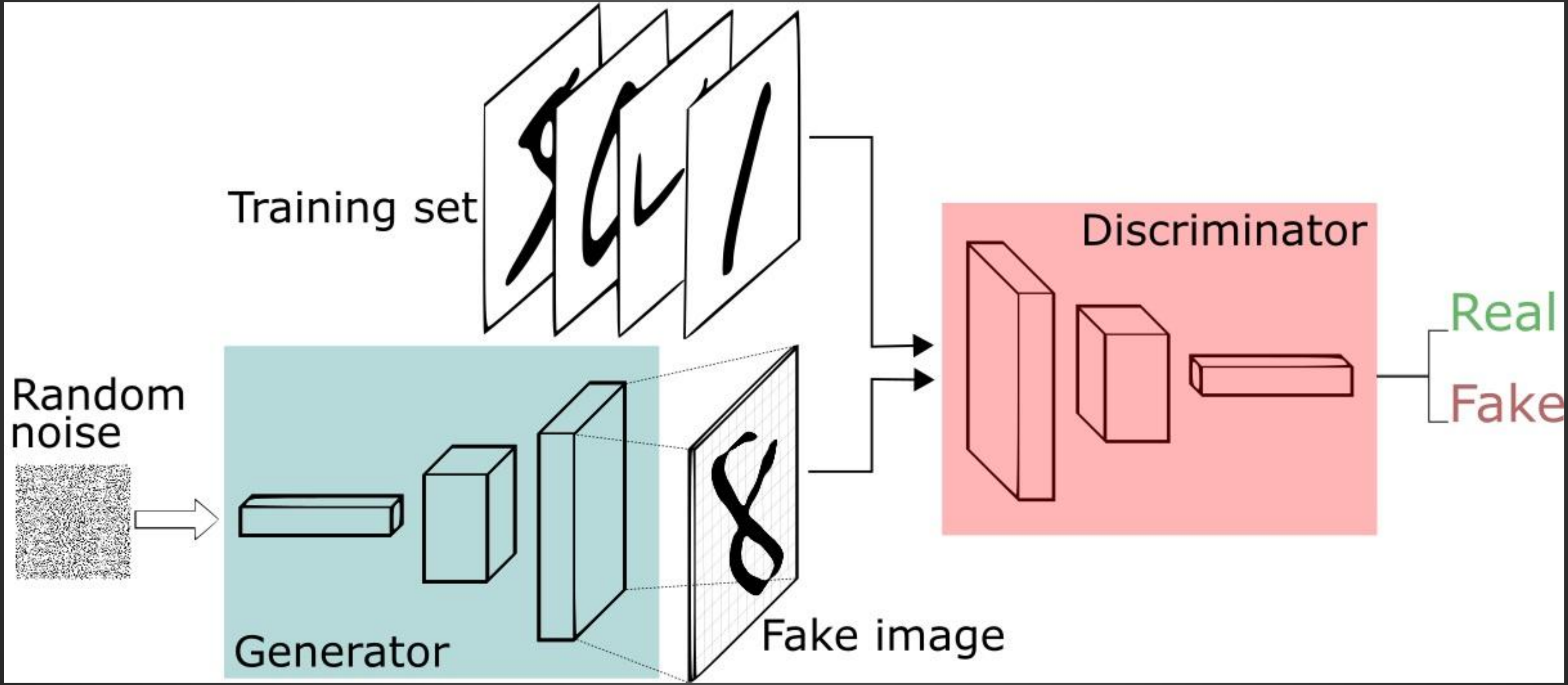
“a significant and fundamental advance” - Andrew Ng (former chief scientist of China’s Baidu)

Theory of GAN

- Ian Goodfellow et al. - **GENERATIVE ADVERSARIAL NETS** (2014)
- simultaneously training two models
- a generative model ***G*** that captures the data distribution
- a discriminative model ***D*** that estimates the probability that a sample came from the training data rather than ***G***
- the training procedure for ***G*** is to maximize the probability of ***D*** making a mistake
- in the space of arbitrary functions ***G*** and ***D***, a unique solution exists, with ***G*** recovering the training data distribution and ***D*** equal to 0.5 everywhere
- in the case where ***G*** and ***D*** are defined by MLP, the system can be trained with backpropagation

MATSUKO

Structure of GAN



Theory of GAN

- the adversarial modeling framework is most straightforward to apply when the models are both MLPs
- to learn the generator's distribution $\boldsymbol{p_g}$ over data $\boldsymbol{x}$, an input noise variable $\boldsymbol{p_z(z)}$ is defined, then a mapping to data space is represented by $\boldsymbol{G(z; \theta_g)}$
- another MLP $\boldsymbol{D(z; \theta_d)}$ is defined which outputs a single scalar, where $\boldsymbol{D(x)}$ represents the probability that $\boldsymbol{x}$ came from the data rather than $\boldsymbol{p_g}$.
- $\boldsymbol{D}$ is trained to maximize the probability of assigning the correct label to both training examples and samples from $\boldsymbol{G}$
- $\boldsymbol{D}$ and $\boldsymbol{G}$ play a two-player minimax game with a value function $\boldsymbol{V(G, D)}$

$$E_x[log(D(x))] + E_z[log(1 - D(G(z)))]$$

MATSUKO

How to train a GAN - theory

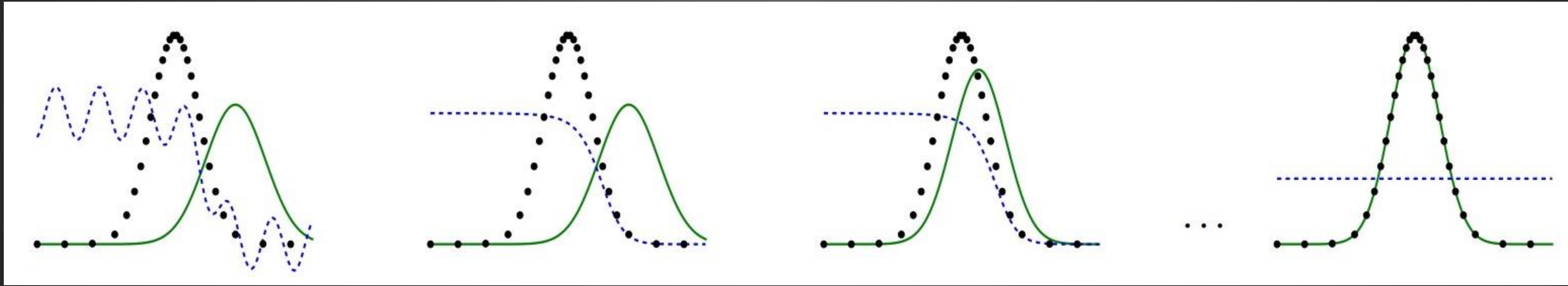


Image: Generative Adversarial Nets

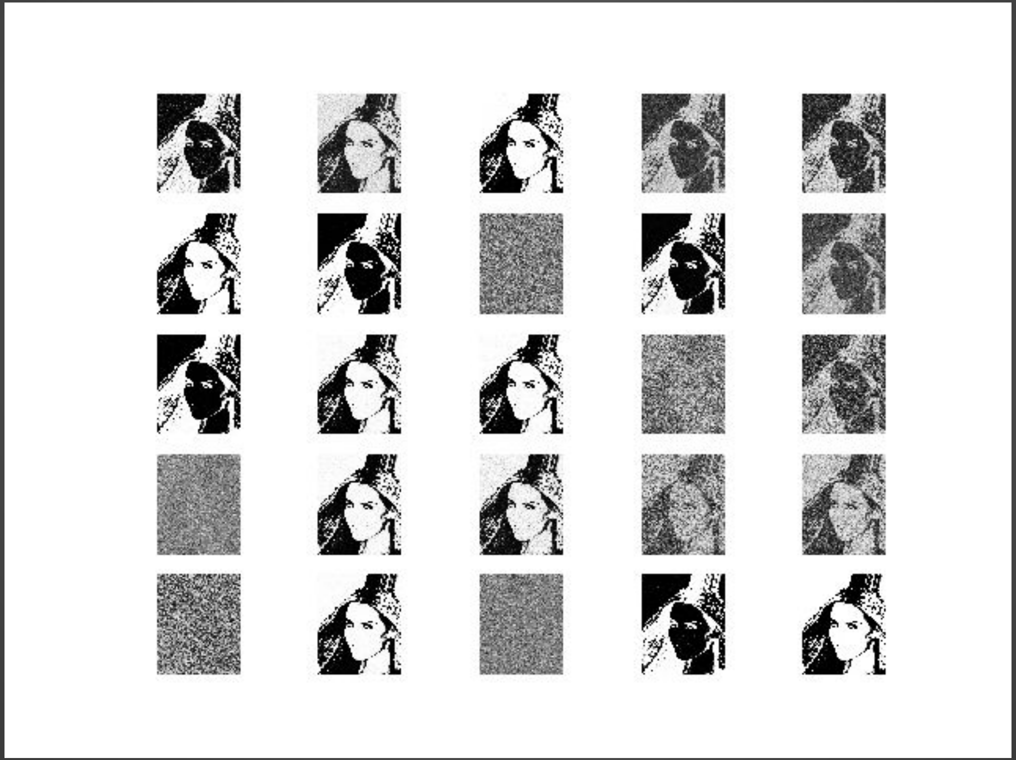
MATSUKO

How to train a GAN - theory

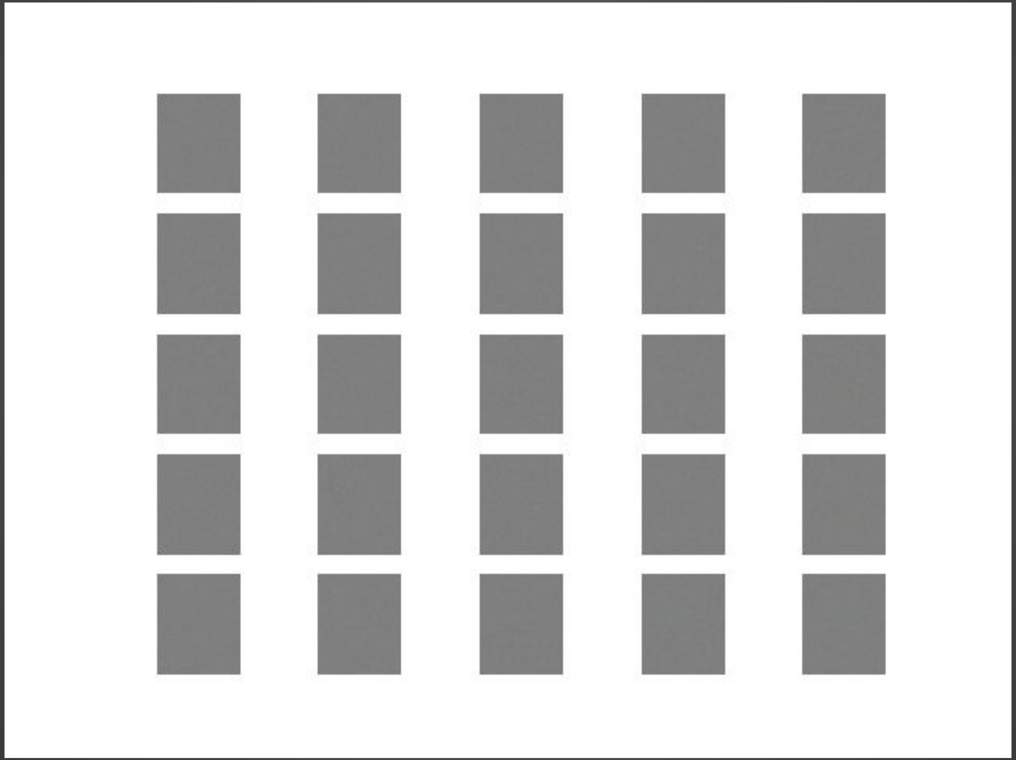
- https://poloclub.github.io/ganlab/

MATSUKO

How to train a GAN - reality



How to train a GAN - reality



Advantages and disadvantages

- there is no explicit representation of $\boldsymbol{p_g(x)}$
- ***D*** must be synchronized well with ***G*** during training (in particular, ***G*** must not be trained too much without updating ***D***, in order to avoid the ‘Helvetica scenario’ in which ***G*** collapses too many values of ***z*** to the same value of ***x*** to have enough diversity to model $\boldsymbol{p_{data}}$)

\+ a wide variety of functions can be incorporated into the model

4. Examples of Use Cases

3D human reconstruction

MATSUKO

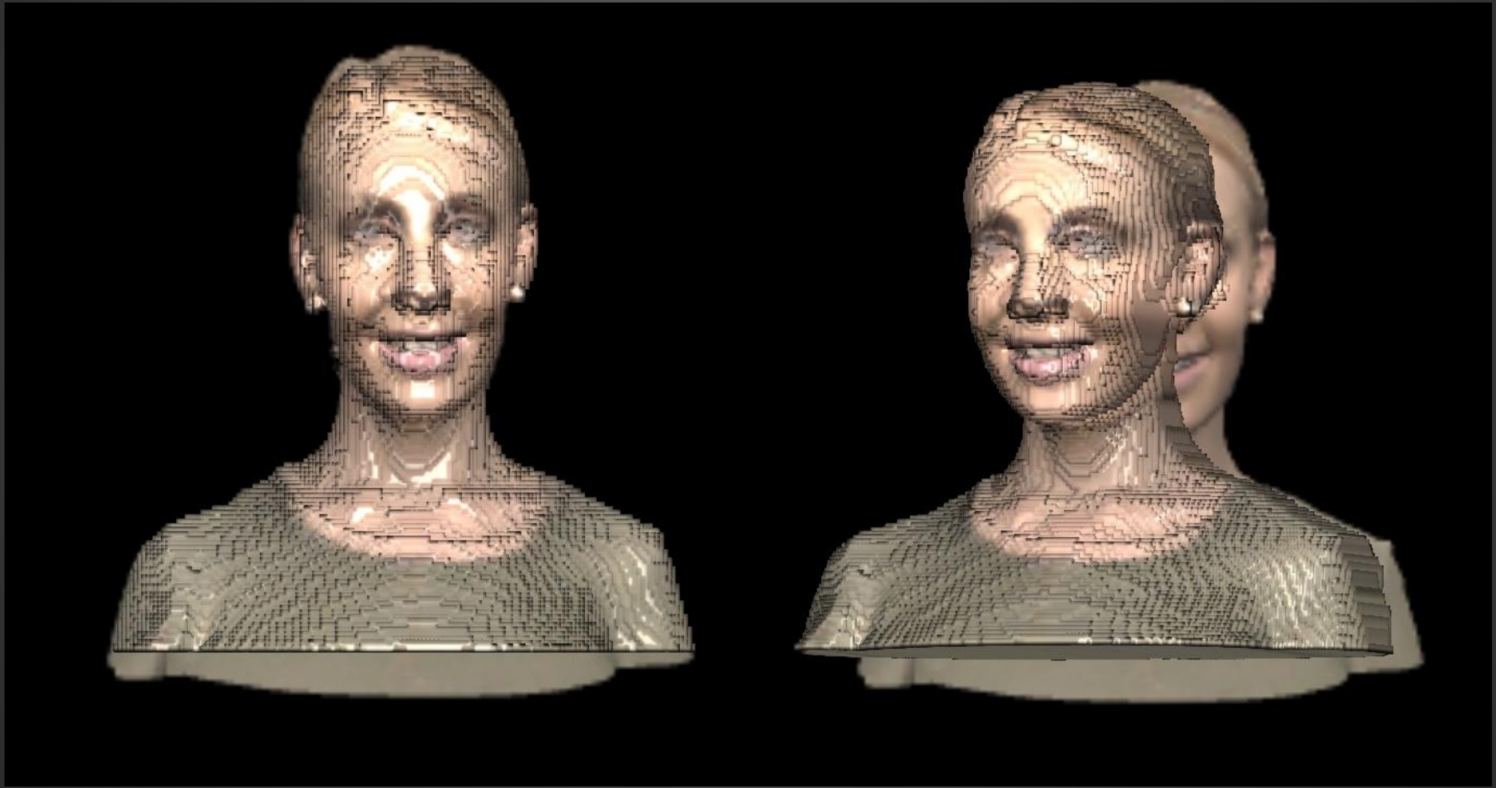
3D human reconstruction

MATSUKO

3D human reconstruction

MATSUKO

3D human reconstruction



MATSUKO

Generate photographs of human faces

Image: The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation

MATSUKO

Text to photo-realistic image

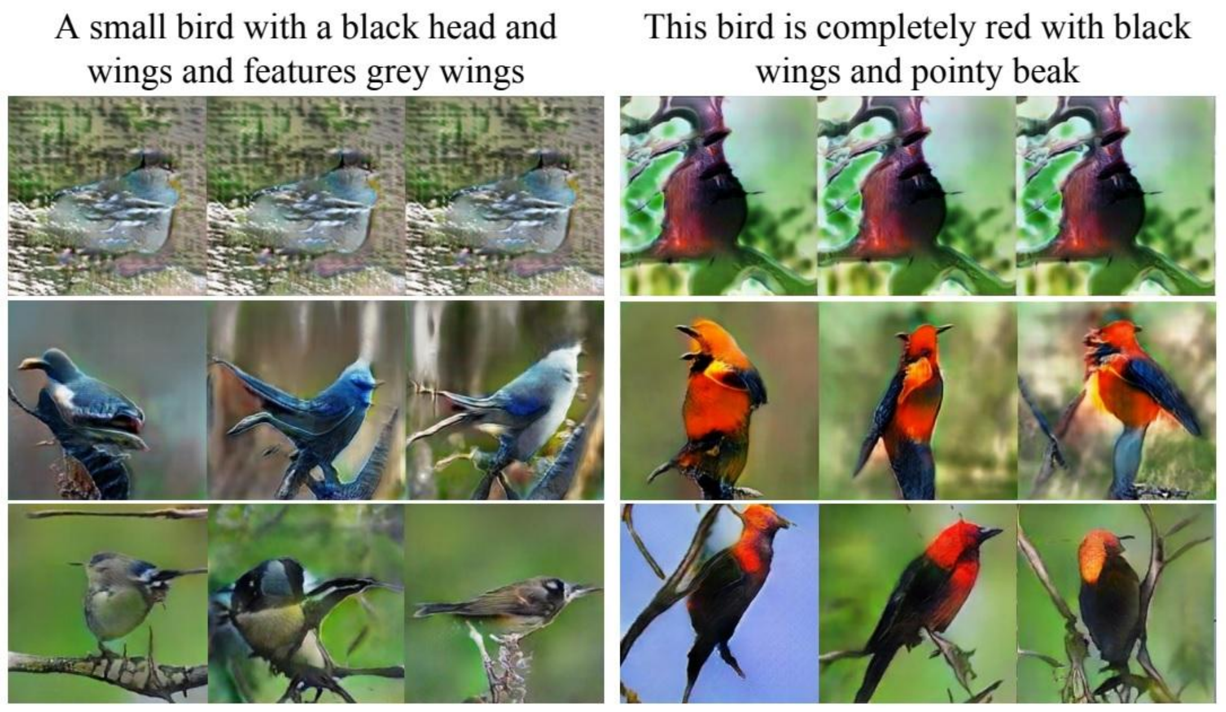


Image: StackGAN: Text to Photo-realistic Image Synthesis with Stacked GANs

Face frontal view generation

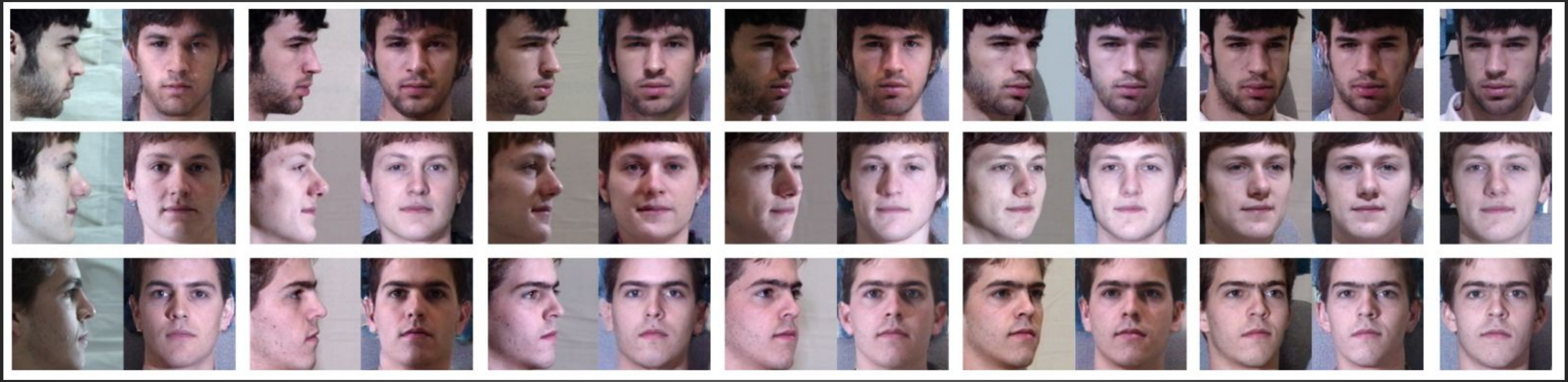


Image: Beyond Face Rotation: Global and Local Perception GAN for Photorealistic and Identity Preserving Frontal View Synthesis

Face frontal view generation

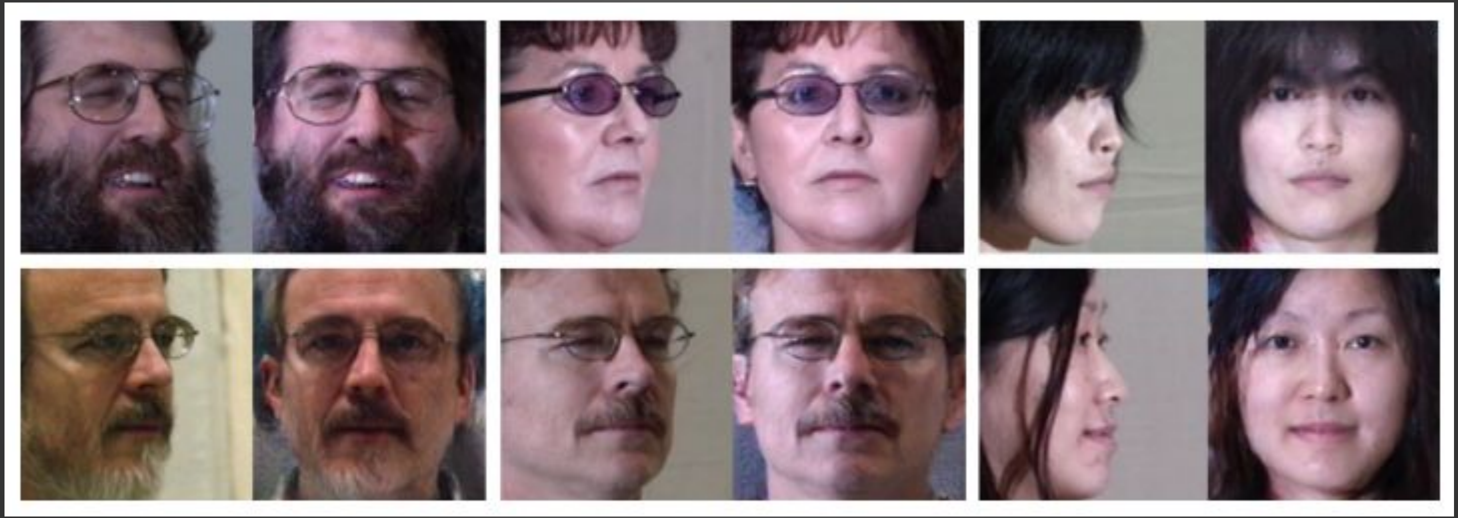


Image: Beyond Face Rotation: Global and Local Perception GAN for Photorealistic and Identity Preserving Frontal View Synthesis

MATSUKO

Face aging with CGANs

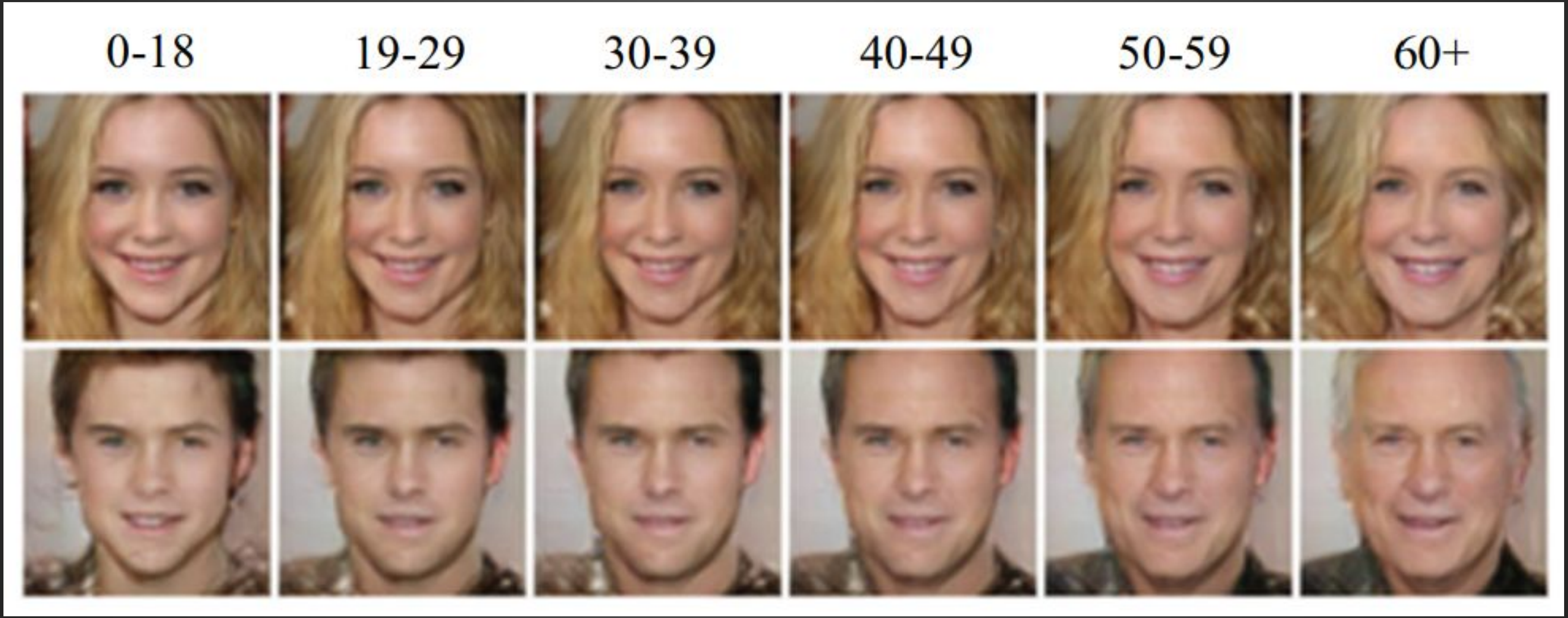


Image: Face Aging with Conditional Generative Adversarial Networks

Super resolution

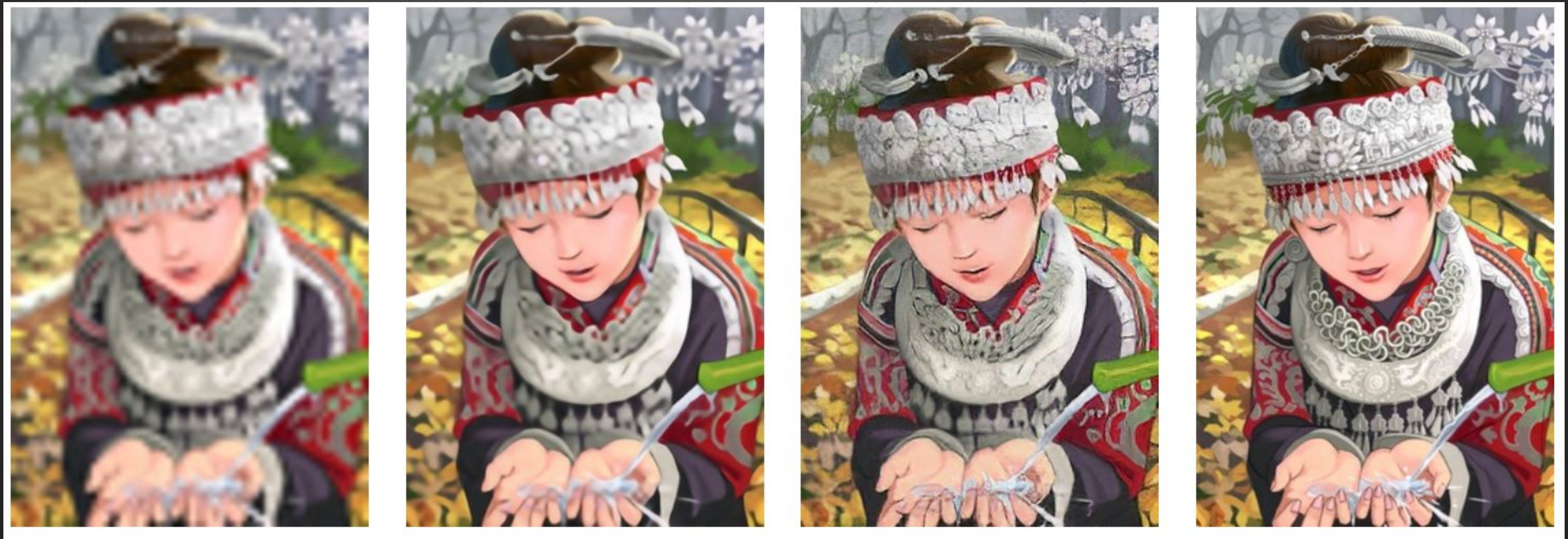


Image: Photo-Realistic Single Image Super-Resolution Using a GAN

MATSUKO

Photo inpainting

Image: Context Encoders: Feature Learning by Inpainting

MATSUKO

Face completion

Image: Generative Face Completion

MATSUKO

Transform images

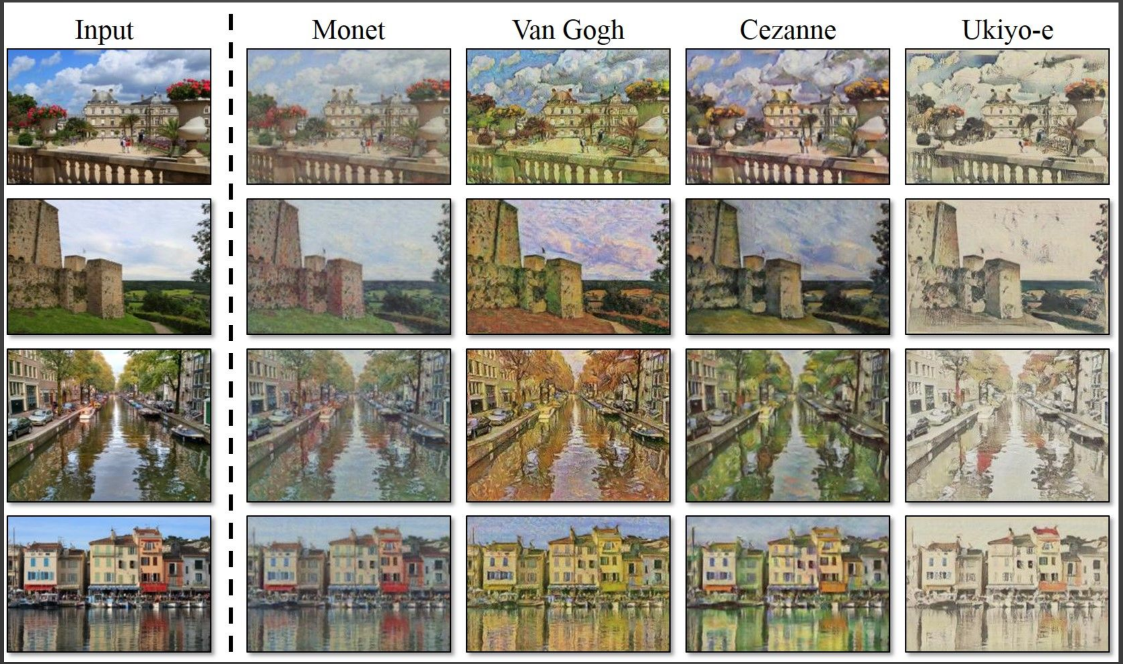


Image: Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks

Transform images

winter Yosemite → summer Yosemite

summer Yosemite → winter Yosemite

Image: Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks

Transform images

Image: Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks

MATSUKO

Data manipulation

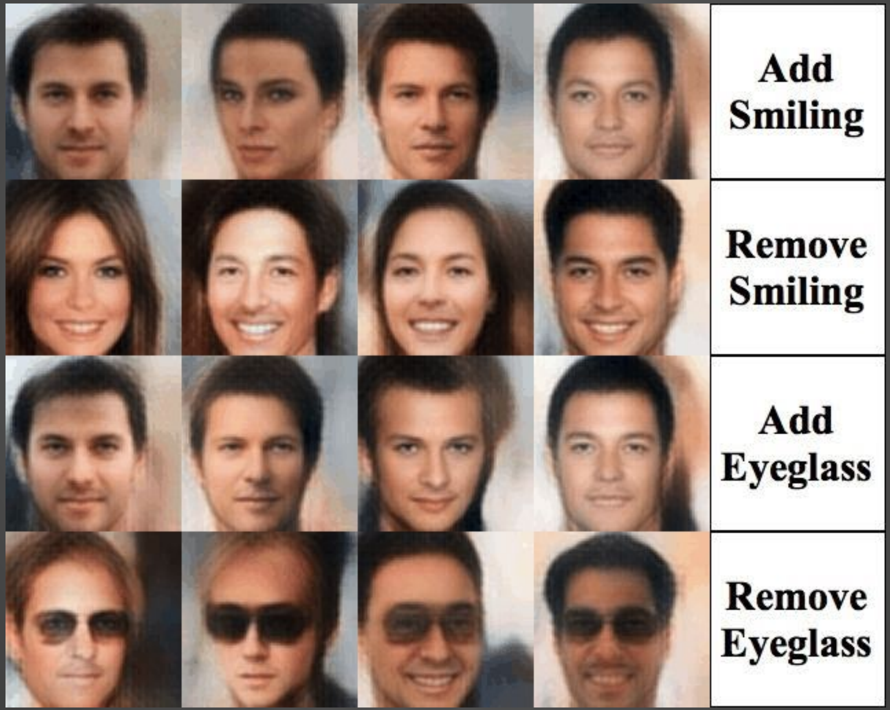


Image: Deep Feature Consistent Variational Autoencoder

Fight against deepfakes

- modern generative models are capable of synthesizing hyperrealistic images, speech, music, and even video
- deepfakes are produced by deep generative models that can manipulate video and audio clips
- Facebook, Microsoft in partnership with various universities are coming together to build the Deepfake Detection Challenge
- the full data release and the DFDC launch will happen at the Conference on Neural Information Processing Systems (NeurIPS)

MATSUKO

Fight against deepfakes

- Deep Fake Detection Dataset (Google & Jigsaw) - contains over 3000 manipulated videos from 28 actors in various scenes.

5. Case Studies

MATSUKO

Games

We are able to transform a raw scan into a fully functional and playable avatar for virtual worlds (gaming, VR/AR, pro simulation).

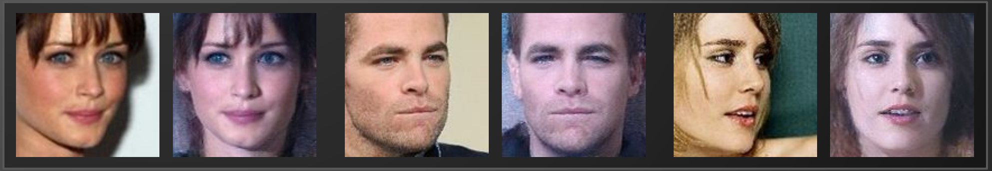
MATSUKO has a unique technology for human 3D video capture which enables not only to see the player’s face on the virtual character but also to capture the real expressions.

MATSUKO

Security

The company X is providing the industry leading fingerprint and facial biometrics solutions thanks to the award-winning Innovatrics algorithms. To date, they completed 500 projects in 70 countries.

MATSUKO has the technology for 3d face reconstruction enabling to frontalize (rotate) the face from various pictures in order to improve the accuracy of the face recognition systems.



MATSUKO

Cosmetics

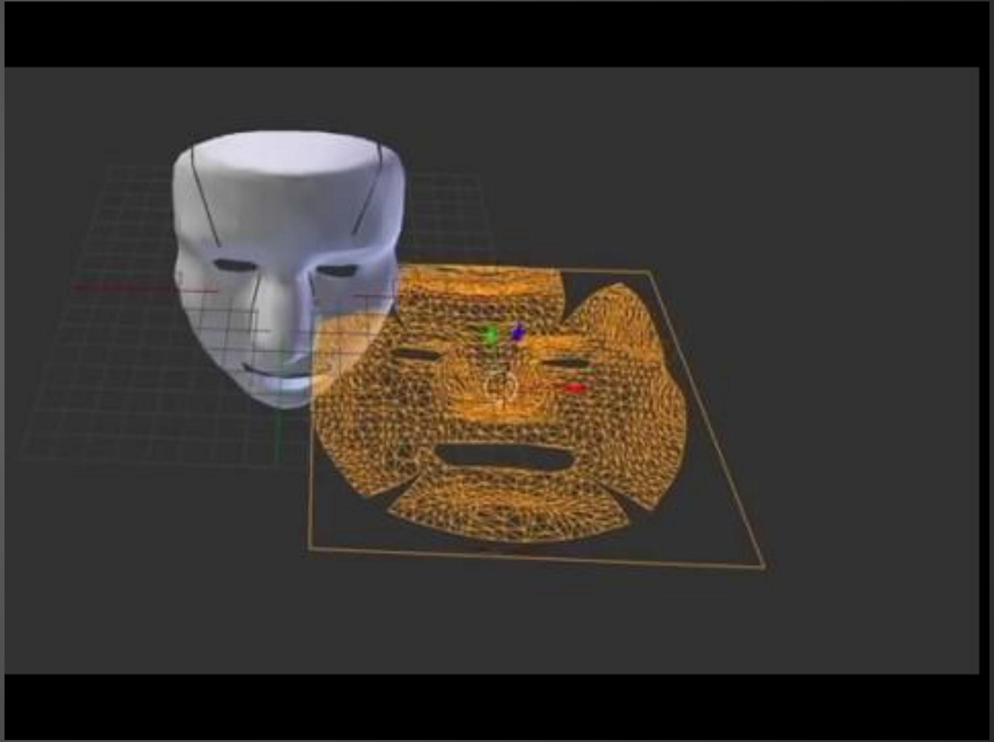
SYMRISE is one of the world's top suppliers in fragrances, flavorings, cosmetic base materials and substances, with up to $3 billion in revenues and market share of 11 % (2017).

3D SymSkin Analyzer developed by MATSUKO is the first AI-powered B2B skincare tool, used by Symrise. Neural network and computer vision analyse user's skin health.

Skincare

BEIERSDORF is a leading provider of innovative, high-quality skin care products and has over 130 years of experience in this market, with up to $7 billion in revenues.

MATSUKO is developing the world’s first skin 3d mask for personalized skin care. From a single photo the AI-powered 3d reconstruction creates 3d face mesh to target the areas of the user’s face.



MATSUKO

Education

Virtual trainings - our ongoing project

6. Deeper into GANs

How to Evaluate GANs
What type of GANs to choose
Where to start

MATSUKO

How to evaluate GANs?

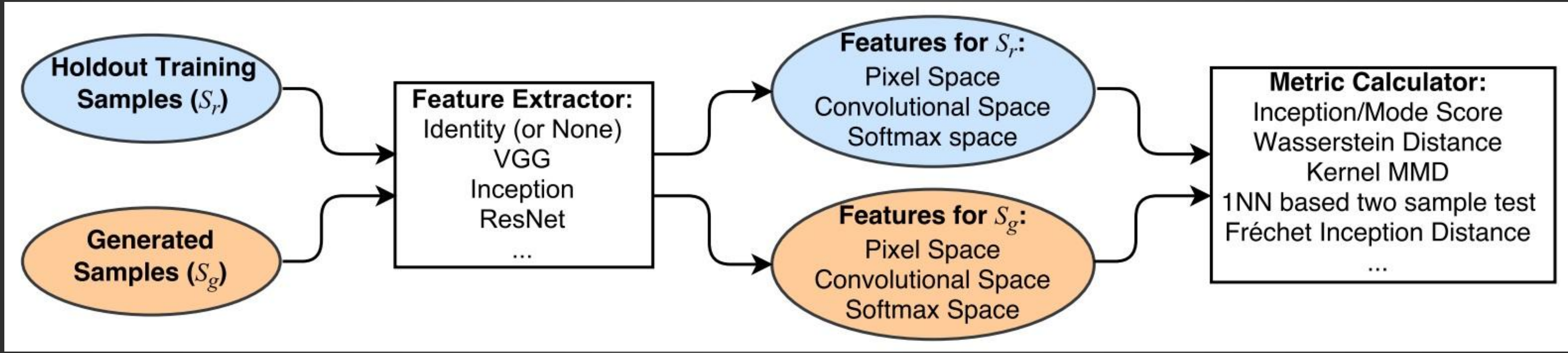


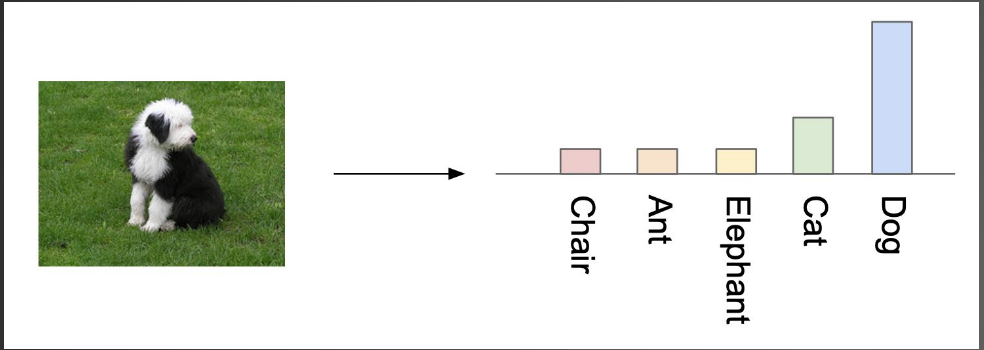
Image: An empirical study on evaluation metrics of generative adversarial networks

How to evaluate GANs?

- GANs besides producing surprisingly plausible images, have been innovatively applied in various fields
- their evaluation is still predominantly qualitative, very often resorting to manual inspection of the visual fidelity of generated images
- given the limitations of qualitative evaluations, proper quantitative metrics are crucial for the development of GANs to guide the design of better models
- two evaluation metrics were proposed to quantitatively access the performance of GANs, both assume access to a pre-trained classifier: Inception Score, Frechet Inception Score

MATSUKO

Inception score



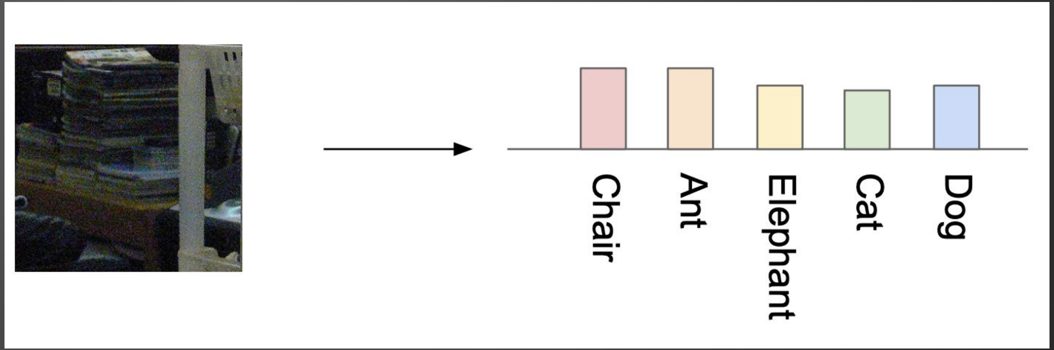


Image: A simple explanation of the Inception Score

Inception score

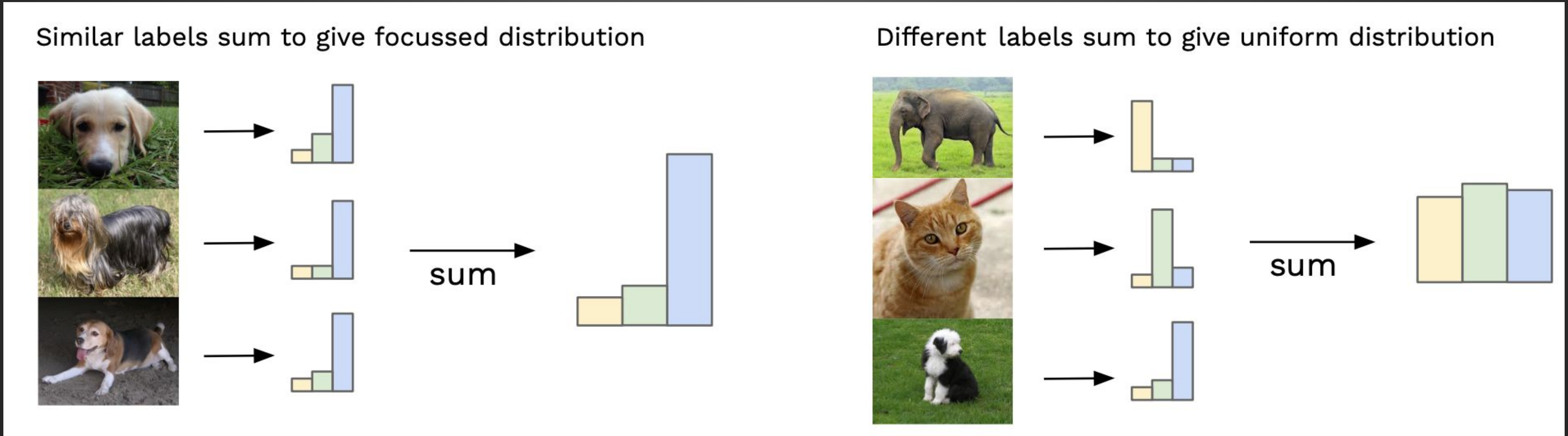


Image: A simple explanation of the Inception Score

Inception score

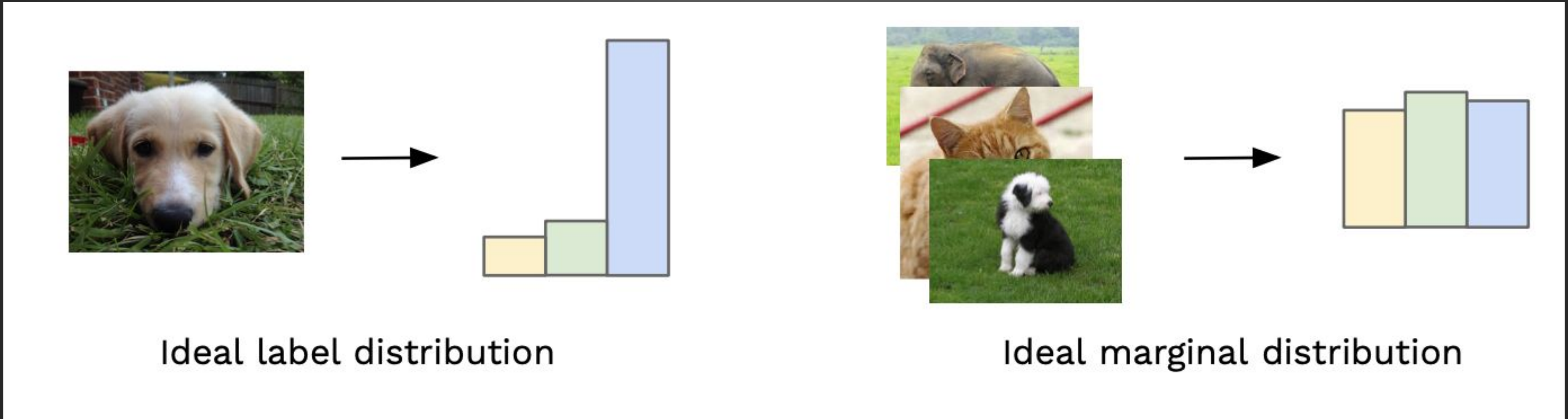


Image: A simple explanation of the Inception Score

Frechet Inception score

- is computed by considering the difference in embedding of true and fake data. Assuming that the coding layer follows a multivariate Gaussian distribution, the distance between the distributions is reduced to the Frechet distance between the corresponding Gaussians. It was shown that the score is consistent with human judgement and more robust to noise than Inception Score.
- lower FID is better, corresponding to more similar real and generated samples as measured by the distance between their activation distributions.

How to evaluate GANs?

- ***precision and recall*** - given a reference distribution ***P*** and a learned distribution ***Q***, precision intuitively measures the quality of samples from ***Q***, while recall measures the proportion of ***P*** that is covered by ***Q***.
- GANs often produce ‘sharper’ images, but can suffer from mode collapse (high precision, low recall), while e.g. variational autoencoders produce ‘blurry’ images, but cover more modes of the distribution (low precision, high recall).

MATSUKO

How to evaluate GANs?

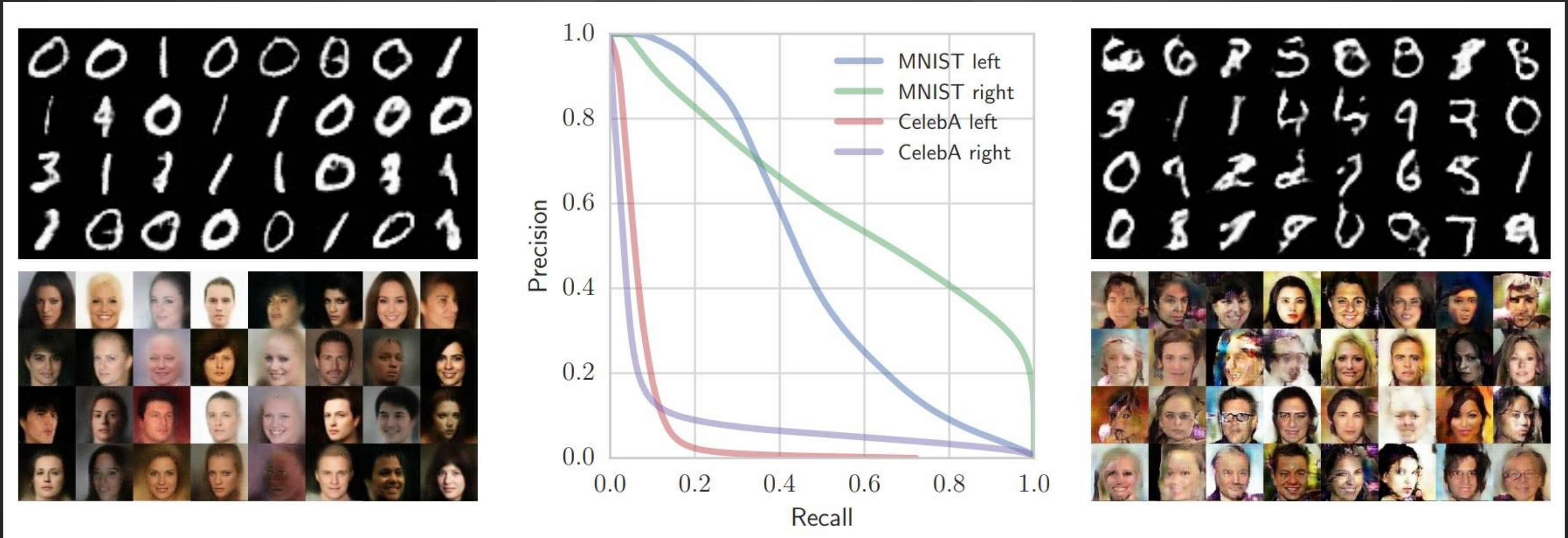


Image: Assessing Generative Models via Precision and Recall

But which type of GAN to choose?

DCGAN
Wasserstein GAN
adaGAN
Bayesian GAN
vanilla GAN
cycle GAN
style GAN
VEEGAN
SAGAN
TequilaGAN

A study conducted at the Google Brain in 2018 empirically demonstrates that nearly all of them can reach similar values of FID, given a high enough computational budget.

Where to start?

- Generative Adversarial Networks for beginners - link
- Ian Goodfellow's GAN tutorial at NIPS 2016 - link
- Various GAN implementations in Keras - link
- GAN Lab - play with GANs in your browser - link
- GAN Zoo - a list of all named GANs - link


MATSUKO
hello@matsuko.com