GENERATIVE ADVERSERIAL NEURAL NETWORKS

INTRODUCTION TO GANS

What's in store

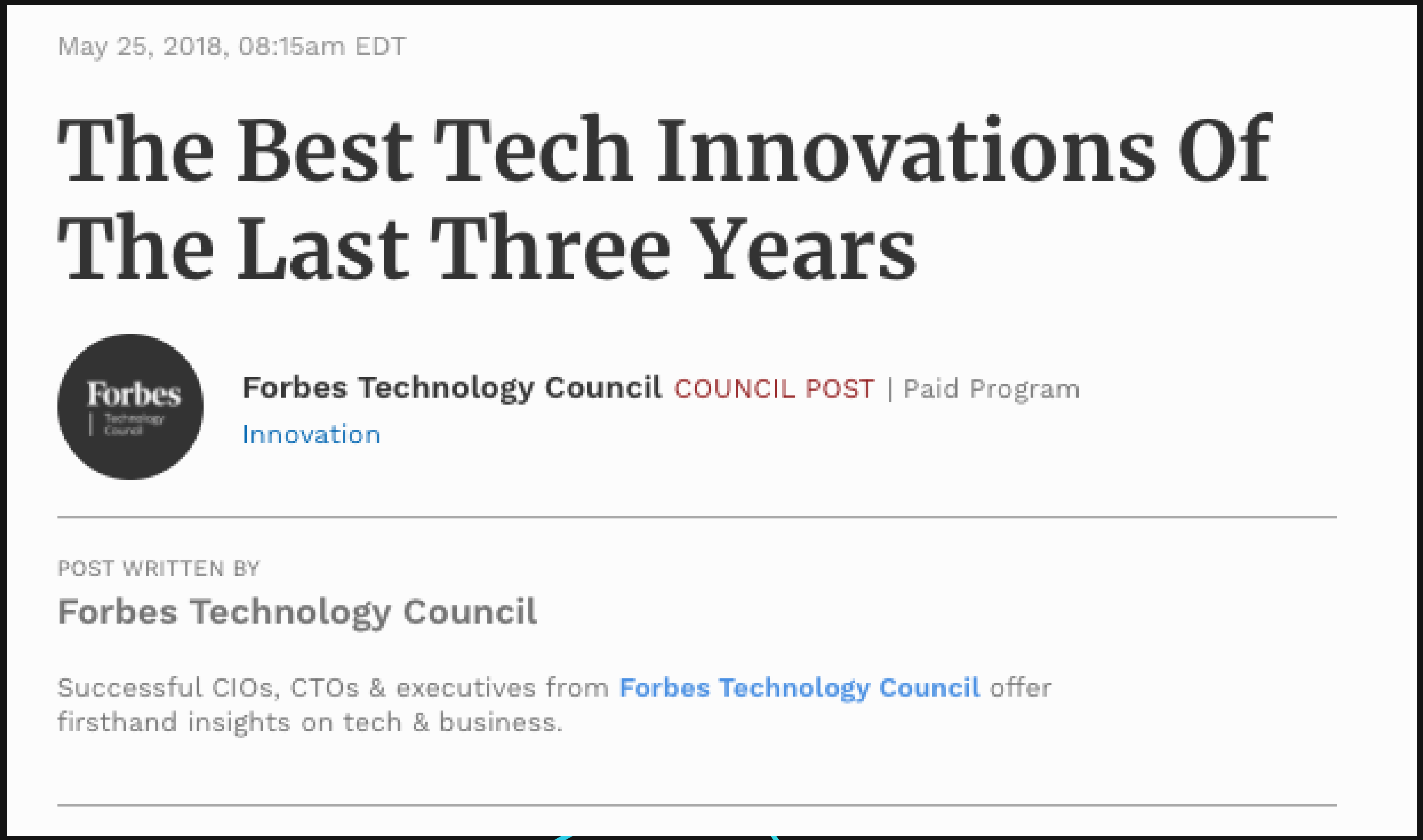
1. WHAT ARE GANS
2. EXAMPLES OF GANS
3. HOW GANS WORK
4. TRAINING GANS & CHALLANGES
5. PRACTICAL IMPLEMENTATIONS OF GANS

Face generation performed by GANs.
Taken from Progressive Growing of GANs for Improved Quality, Stability, and Variation, 2017.

Google's BigGAN

What Are GANs?

First Introduced in 2014 by Ian Goodfellow et al. GANs are a type of Neural Network that generates data that plausibly comes from an existing distribution of samples.



May 25, 2018, 08:15am EDT

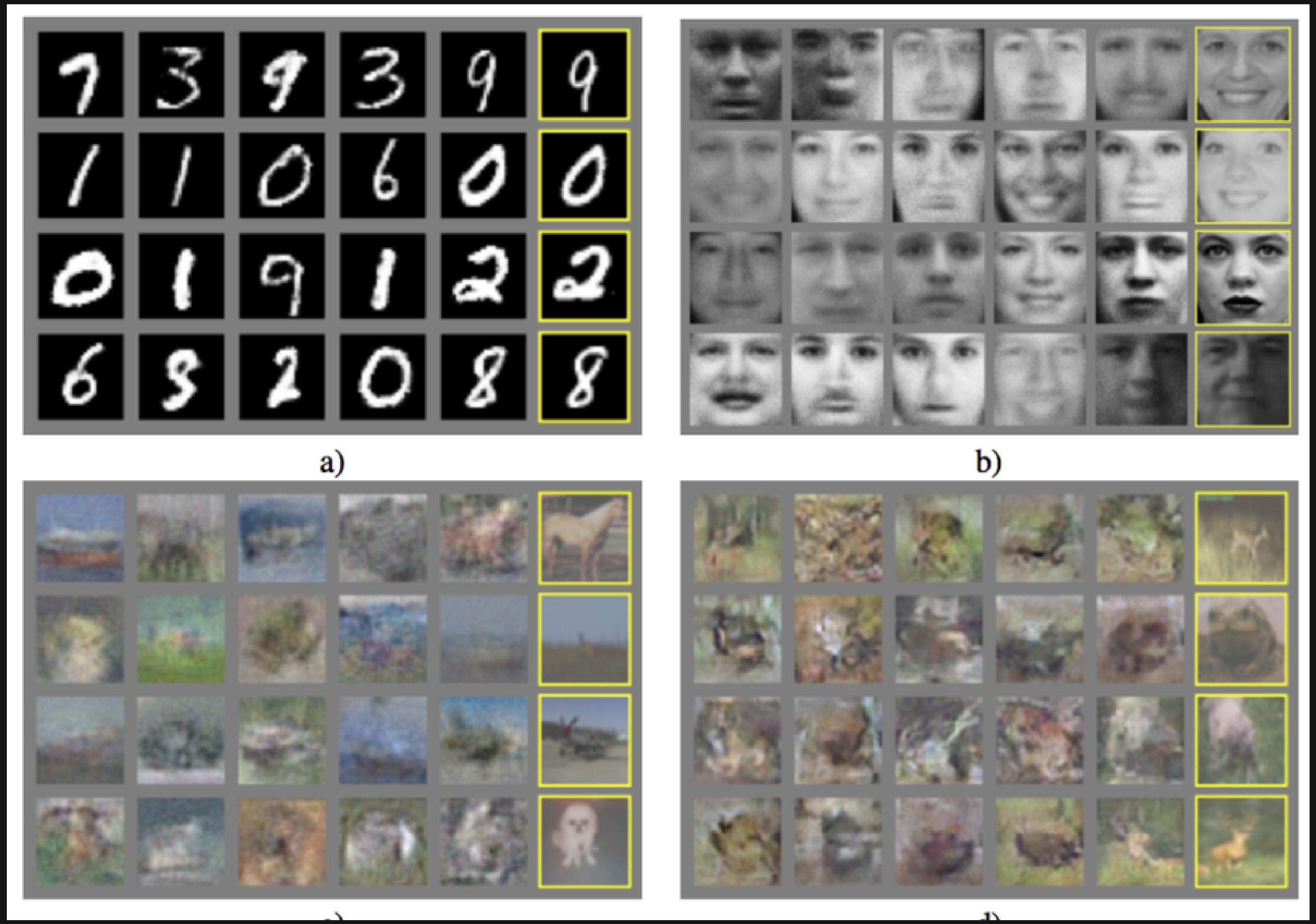
The Best Tech Innovations Of The Last Three Years

Forbes Technology Council COUNCIL POST | Paid Program
Innovation

POST WRITTEN BY
Forbes Technology Council

Successful CIOs, CTOs & executives from Forbes Technology Council offer firsthand insights on tech & business.

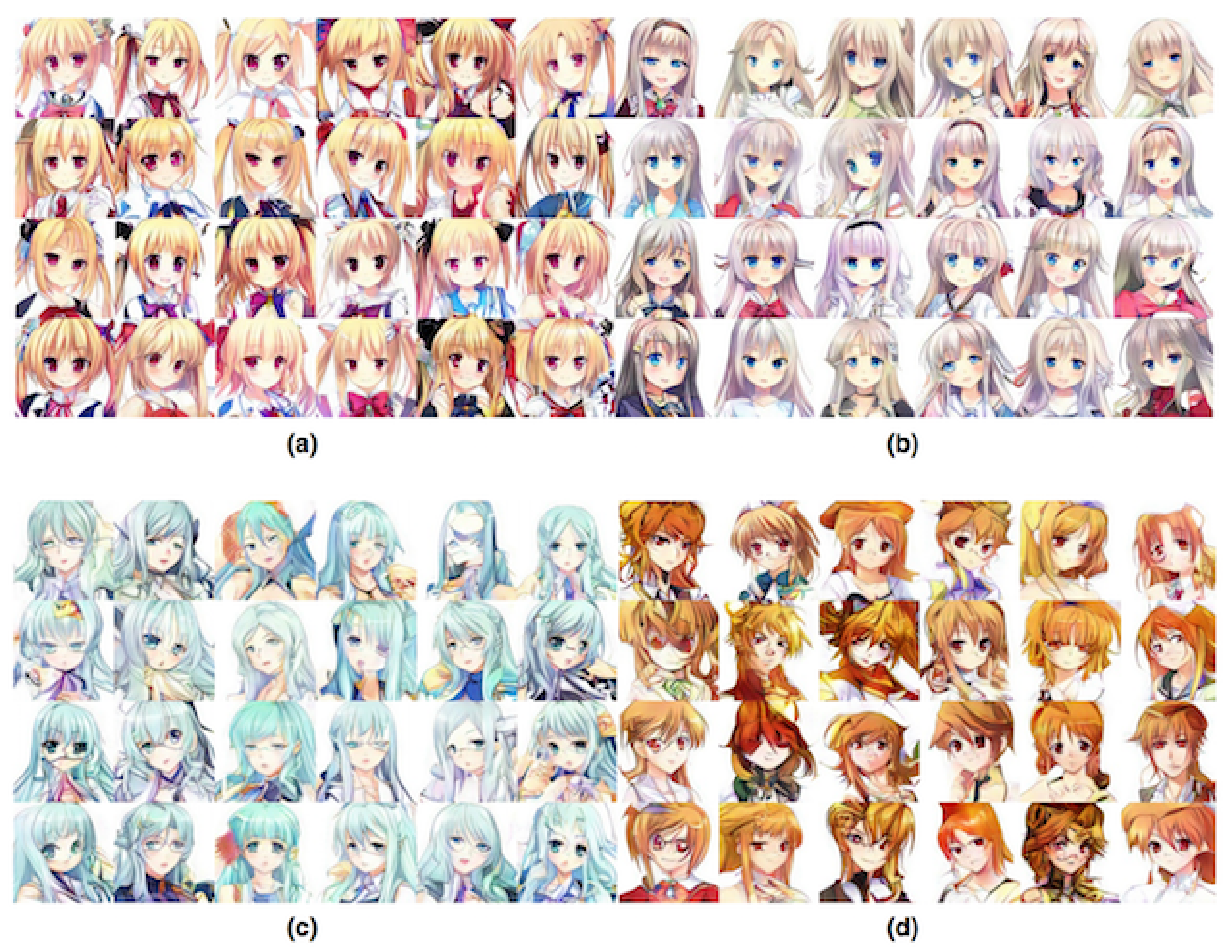
Examples of GANs used to Generate New Plausible Examples for Image Datasets.



Examples of GANs
Realistic Photo Generation

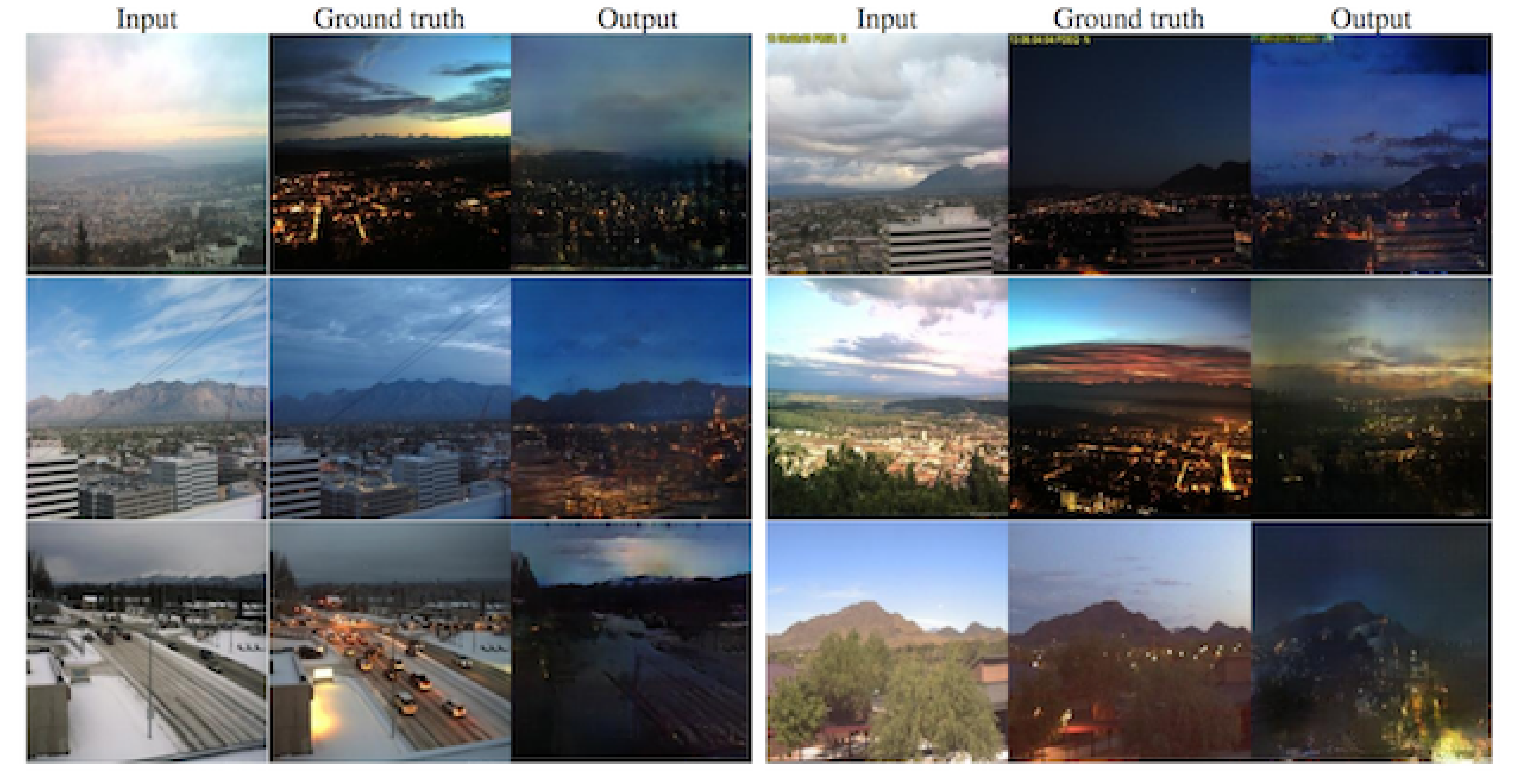
Example of Realistic Synthetic Photographs Generated with
BigGANTaken from Large Scale GAN Training for High Fidelity Natural Image Synthesis, 2018.

Generate Anime Cartoon Characters

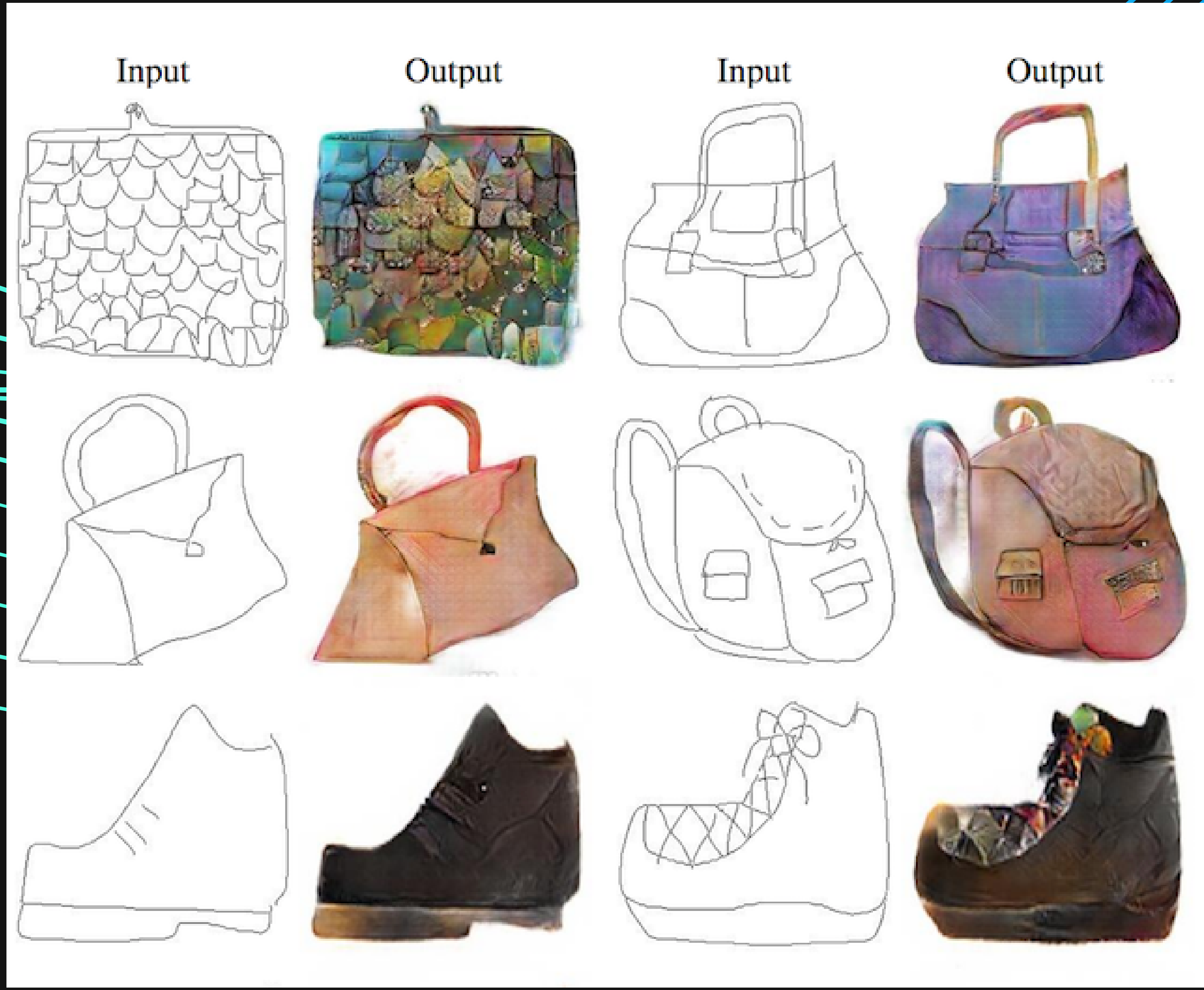


Example of GAN-Generated Anime Character Faces.Taken from T
owards the Automatic Anime Characters Creation with Generative Adversarial Networks, 2017

Image-to-Image Translation

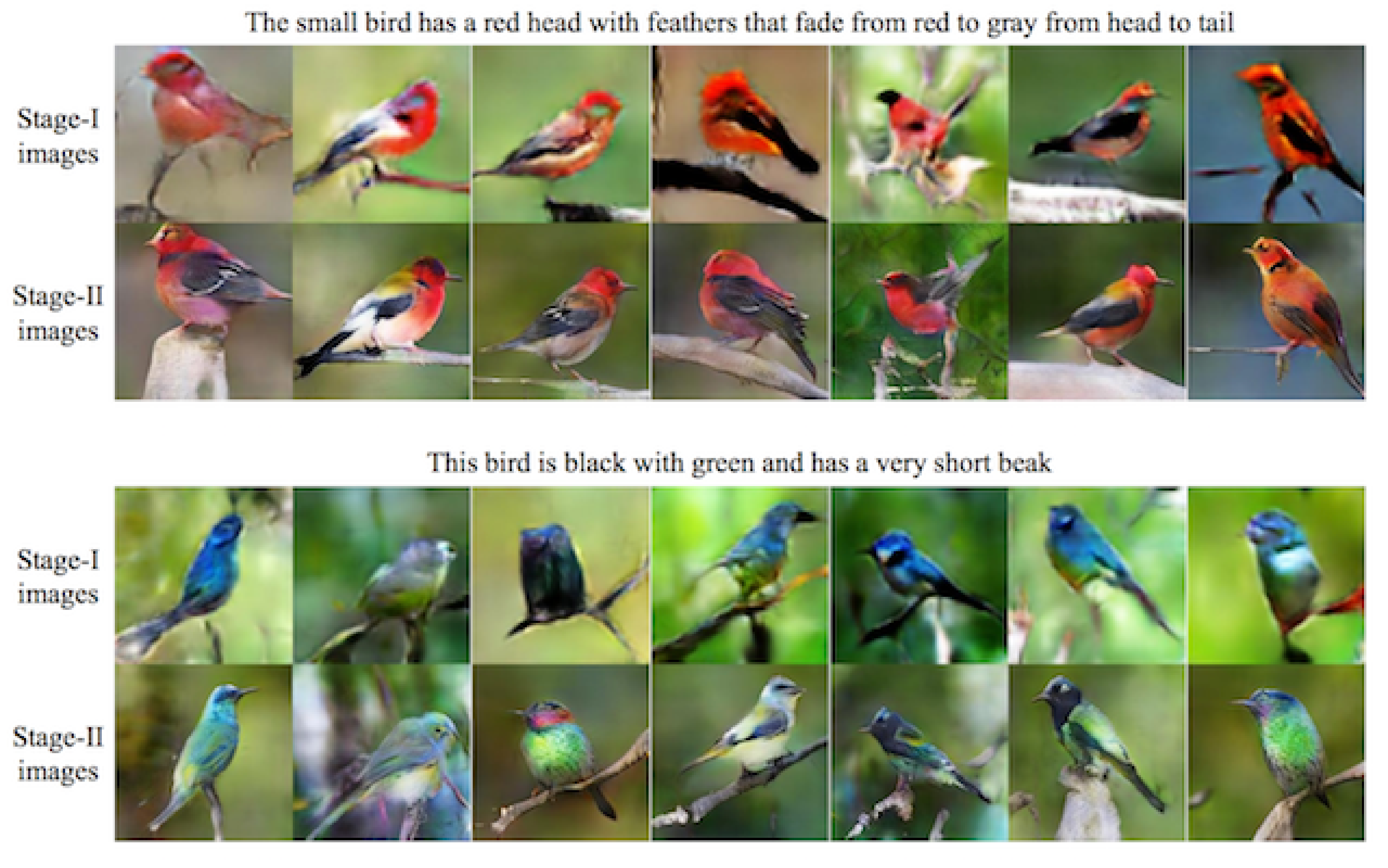


Example of Photographs of Daytime Cityscapes to Nighttime With pix2pix.
Taken from Image-to-Image Translation with Conditional Adversarial Networks, 2016.



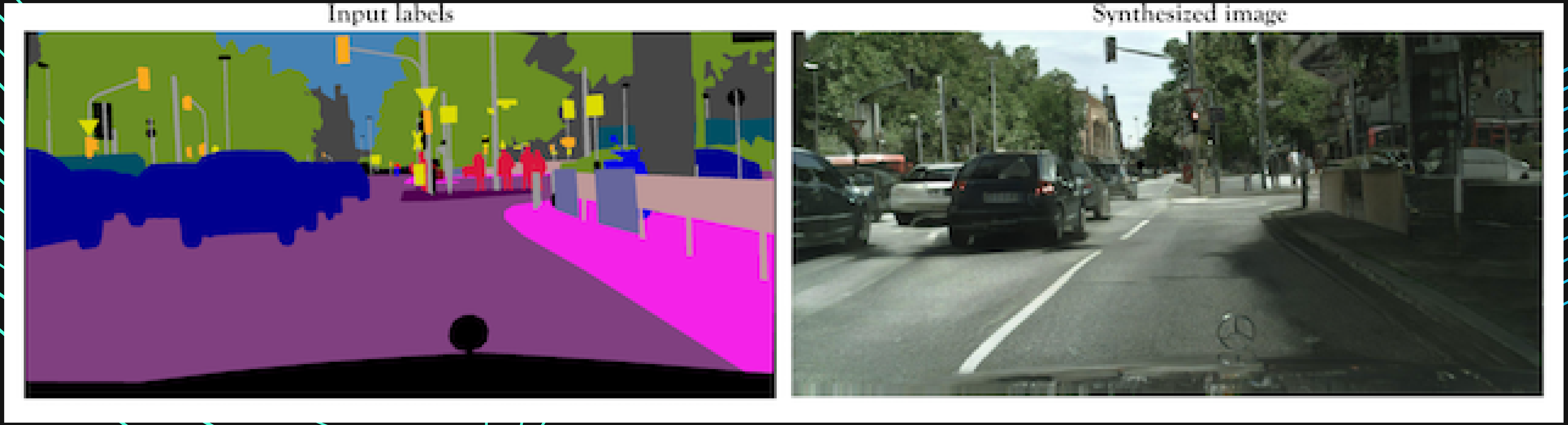
Example of Sketches to Color Photographs With pix2pix.Taken from Image-to-Image Translation with Conditional Adversarial Networks, 2016.

Text-to-Image Translations



Example of Textual Descriptions and GAN-Generated Photographs of BirdsTaken from StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks, 2016.

Semantic-Image-to-Photo-Translation



Example of Semantic Image and GAN-Generated Cityscape Photograph.Taken from High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs, 2017.

Super Resolution (SRGAN)

Figure 2: From left to right: bicubic interpolation, deep residual network optimized for MSE, deep residual generative adversarial network optimized for a loss more sensitive to human perception, original HR image. Corresponding PSNR and SSIM are shown in brackets. [4× upscaling]

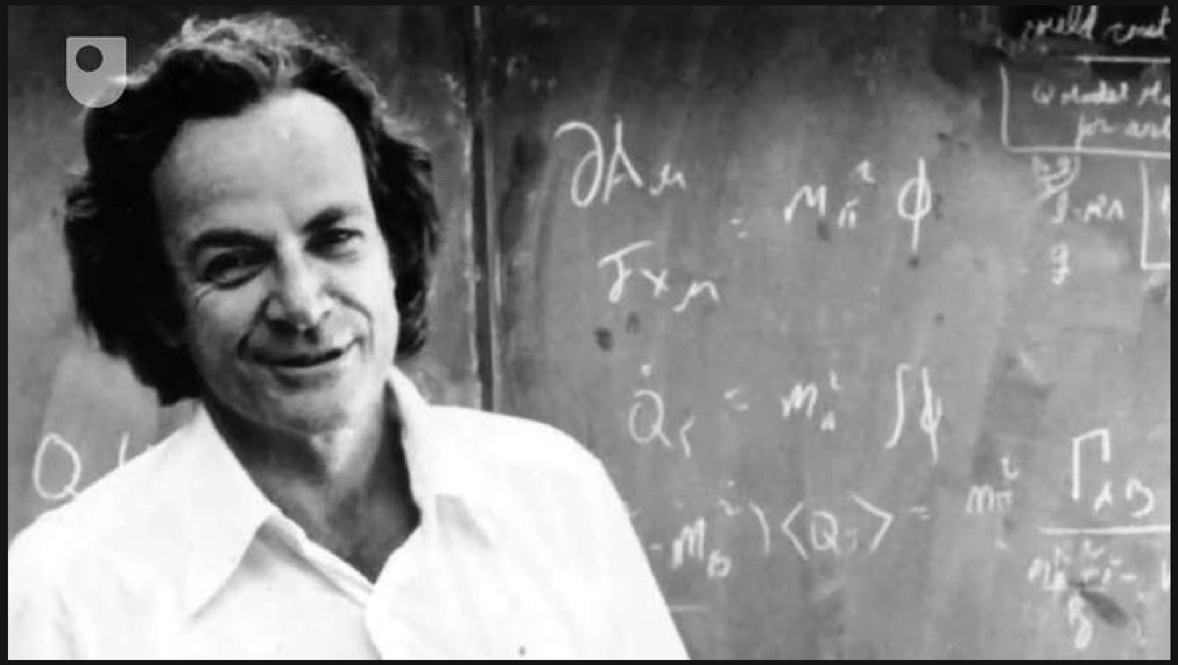
Next..

How do GANs Work?

HOW DO GANS WORK?

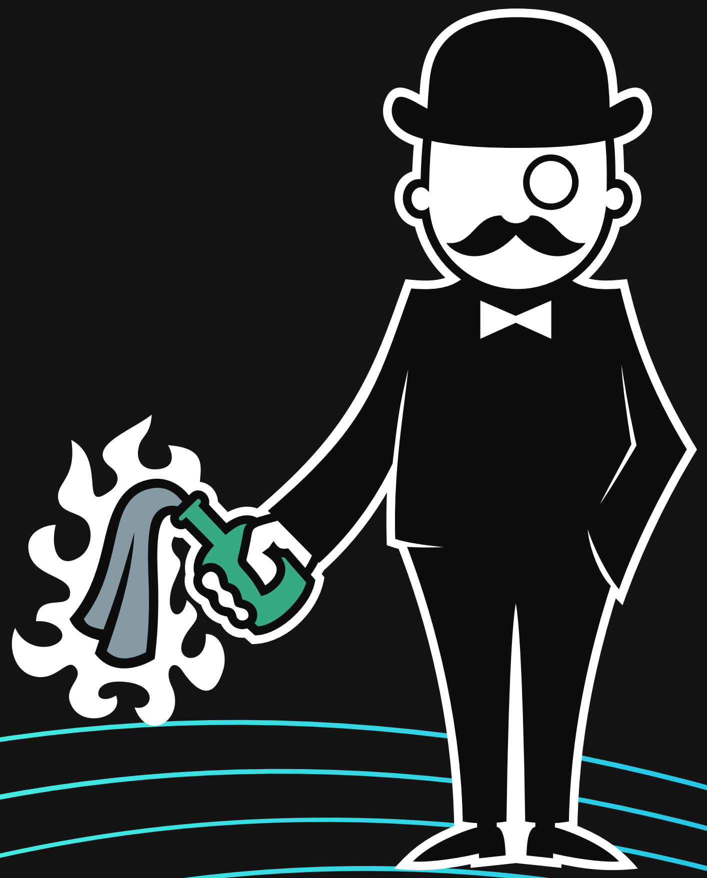
"What I cannot create, I do not understand"

– Richard Feynman (Theoretical Physicist)



Let's start with a Analogy

Imagine there's an ambitious young criminal who wants to counterfiet money and sell to a mobster who's an expert in counterfiets.

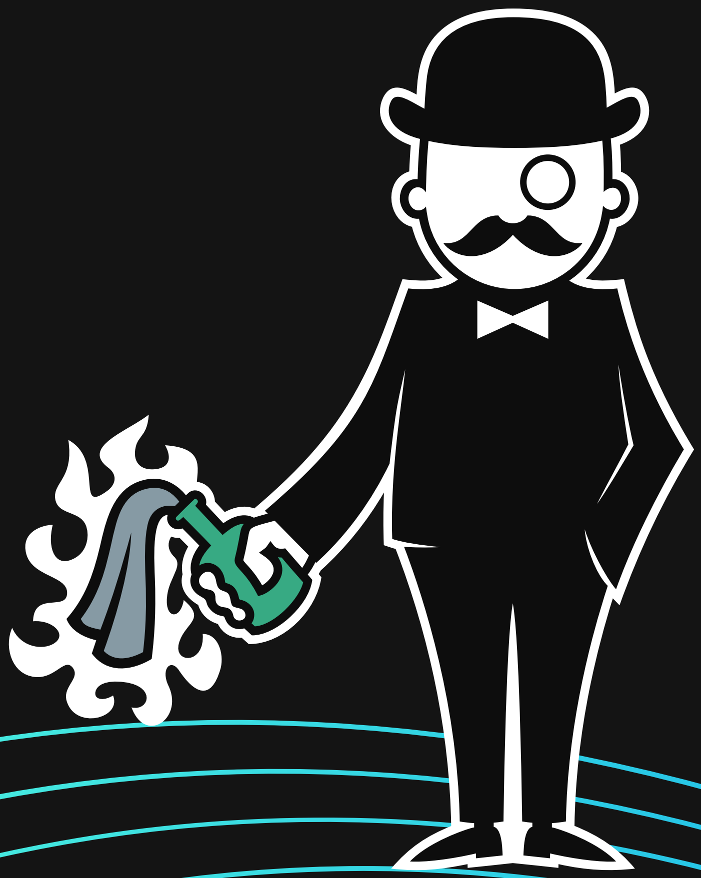


At first he's no good

Our counterfieter takes both his fake bills and some real ones and gives it to our expert Mobster who then labels which are fake and real.

I'LL TRY HARDER!

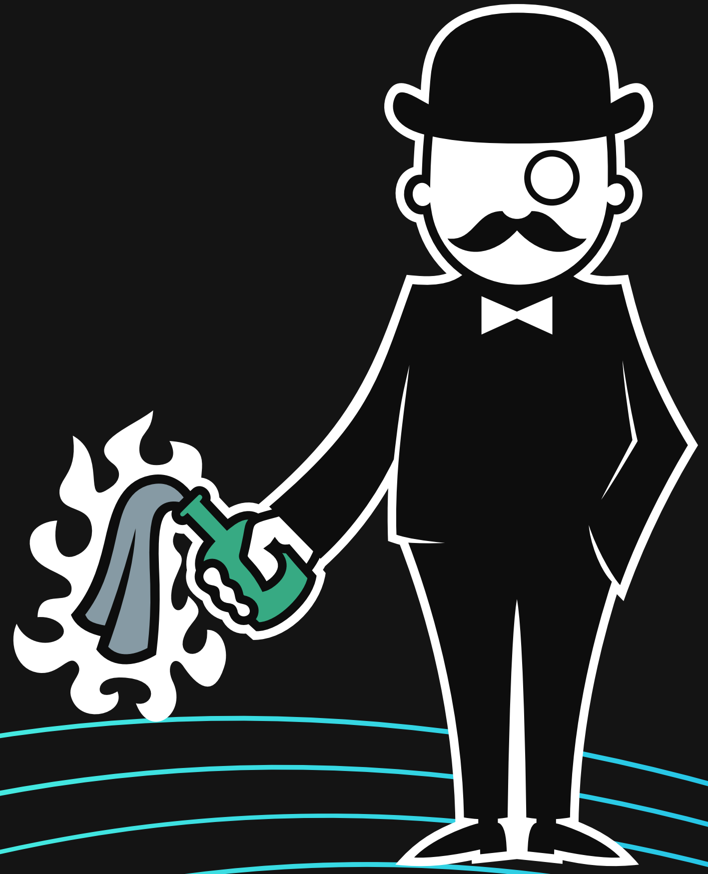
MAKE THIS OVER!!!



Then he slowly gets better...

I'LL KEEP IMPROVING

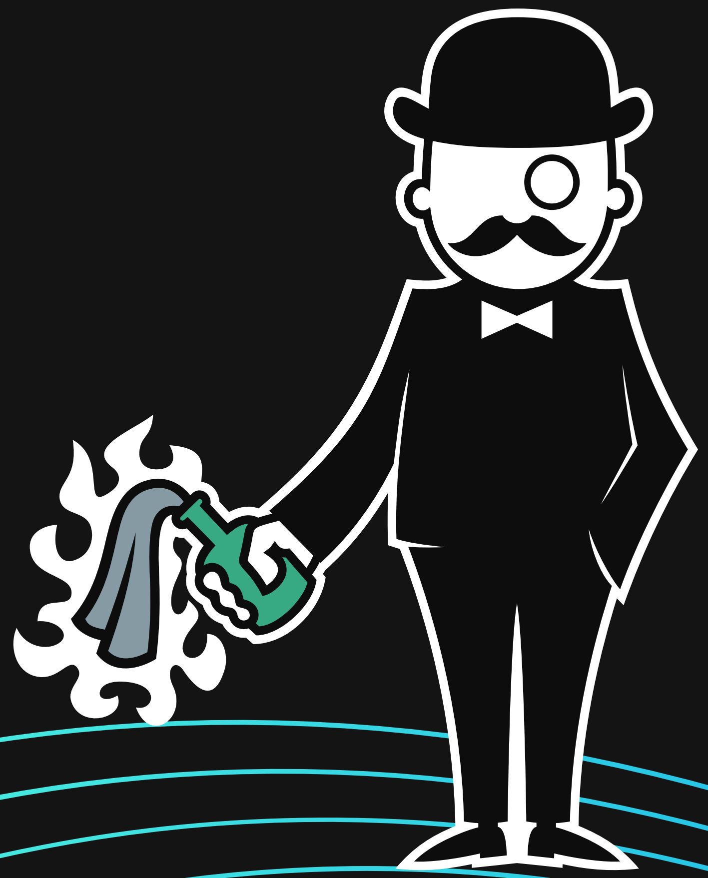
YOU'RE ALMOST FOOLING ME!



Both get better!

I'M AN EXPERT AT MAKING COUNTERFEITS

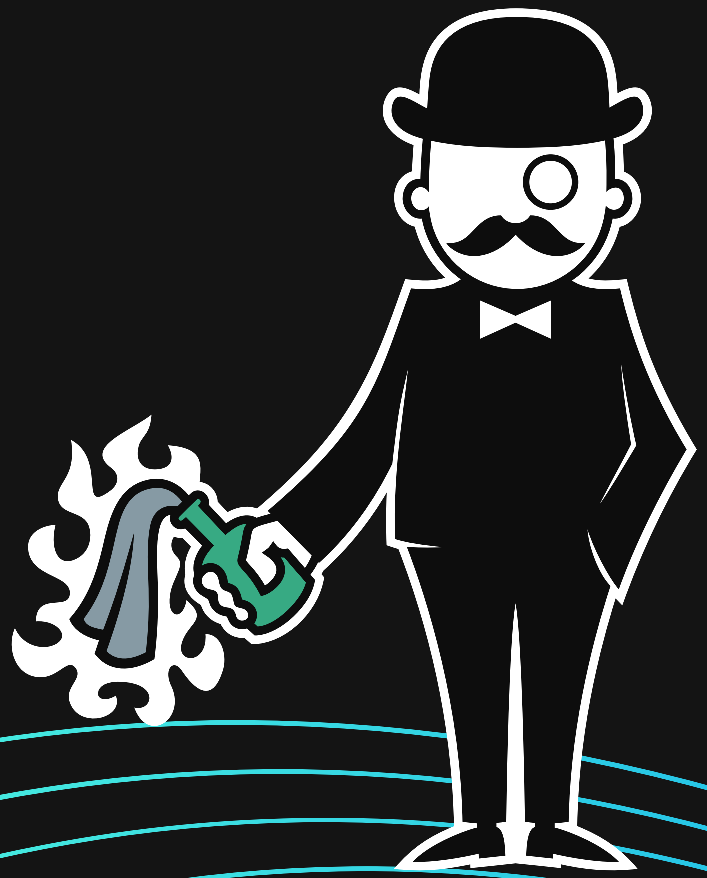
I'M AN EXPERT AT SPOTTING FAKES AND I CANNOT TELL YOUR FAKES FROM THE REAL THING



Two Components of GANs

In our analogy we have two antagonistic networks contesting against each other.

Our young criminal was the Generator Network and our Descriminator was the mobster.

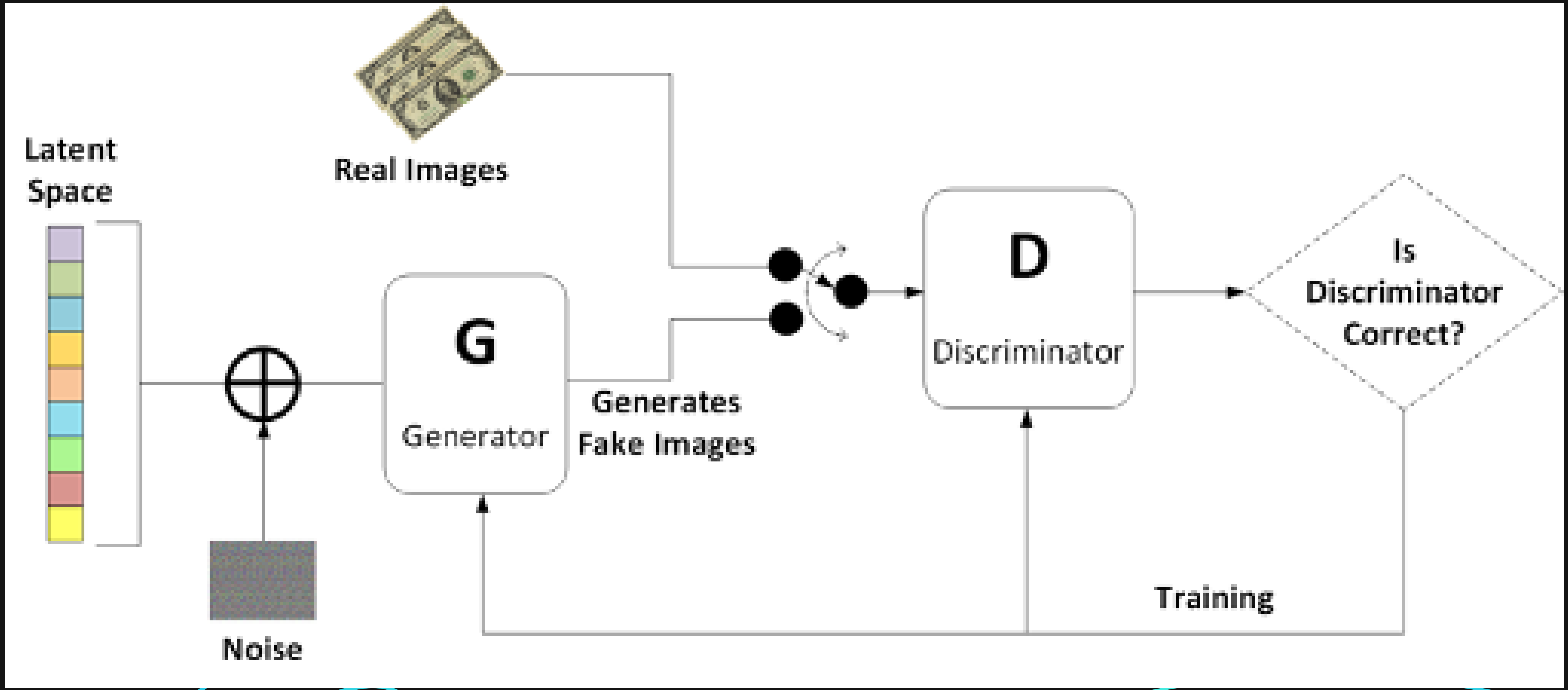


Generator & Descriminator Networks

The purpose of the Generator Network is take random data initializations and decode it into a synthetic sample.

The purpose of the Descriminator Network is to then take this input from our Generator and predict whether or not this sample came from the real dataset or not.

The Basic GAN Architecture



Adverserial Games

In GANs we have two adverserial networks (G & D) where they're both being trained to win over the other. The function below serves as our Joint Loss Function or Value Function.

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

$$V(D,G) = \mathbb{E}_{x \sim p_{data}(x)}[\log D(x)] + \mathbb{E}_{z \sim p_z(z)}[\log(1 - D(G(z)))]$$

It's effectively a minimax game where the Descriminator is trying maximize it's chances of winning and the Generator is trying to minimize it's chances of losing.

Next..

Training GANs

TRAINING GANS

Training GANs

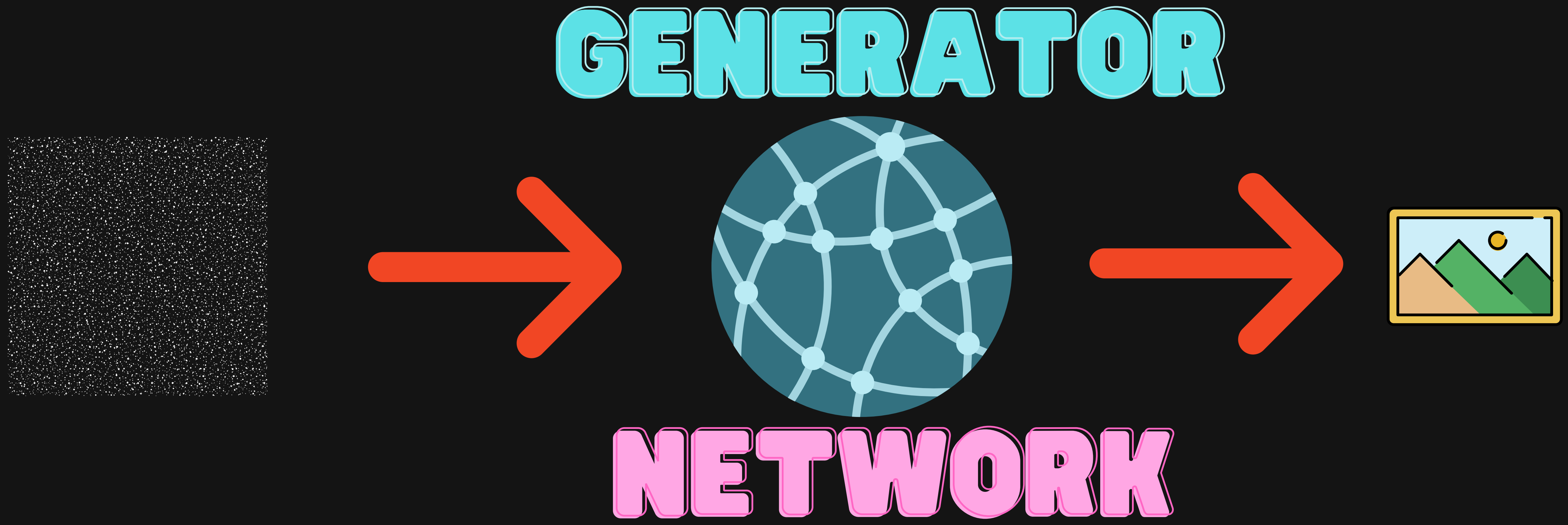
- Training GANs is notoriously difficult compared to Neural Networks we use gradient descent to change our weights to reduce our loss.
- In a GANs, every weight change can change the entire balance of our dynamic system.
- We are not seeking to minimize loss, but finding an equilibrium between our two opposing Networks.
- Training stops when the Descriminator (or you) cannot tell apart Real vs Fake Data

The Training Process

1. We randomly generate a noisy vector

The Training Process

1. We randomly generate a noisy vector
2. Input this noise into our Generator Network to generate some synthetic data



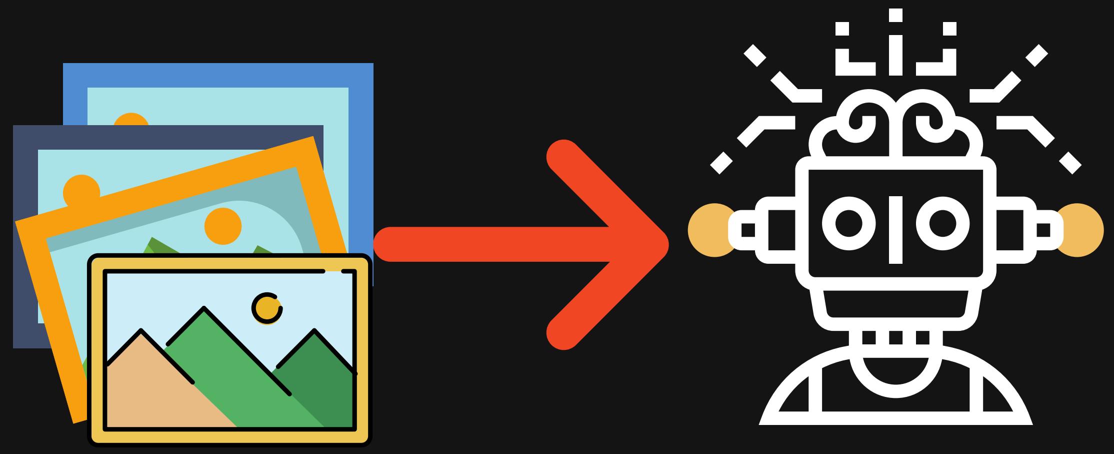
The Training Process

1. We randomly generate a noisy vector
2. Input this into our Generator Network to generate a sample data
3. We then take some sample data from our real data and mix it with some of our generated/synthetic data

Mixed Data

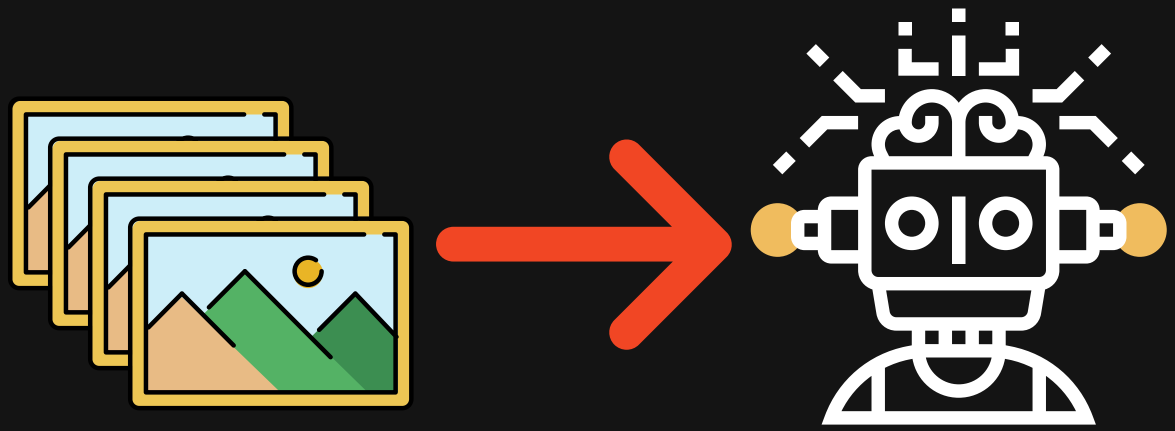
The Training Process

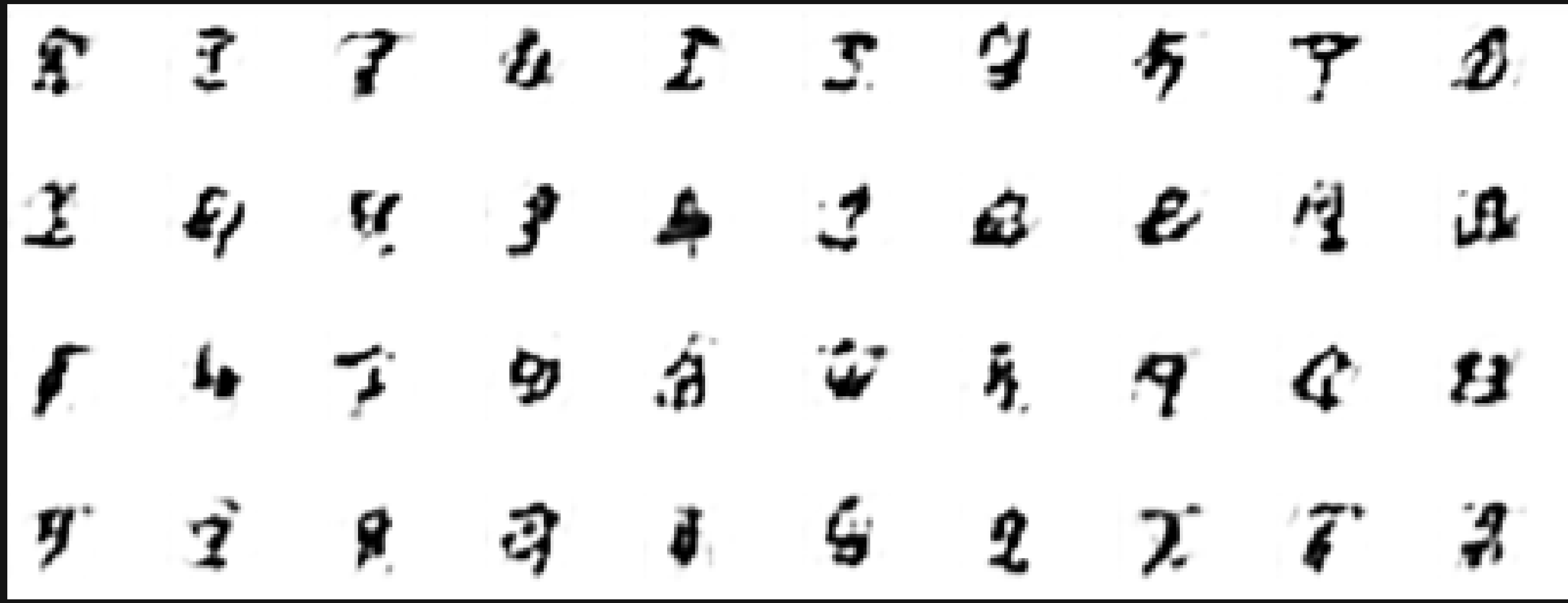
1. We randomly generate a noisy vector
2. Input this into our Generator Network to generate a sample data
3. We then take some sample data from our real data and mix it with some of our generated/synthetic data
4. Train our descriminator to classify this mixed dataset and thus update it's weights accordingly



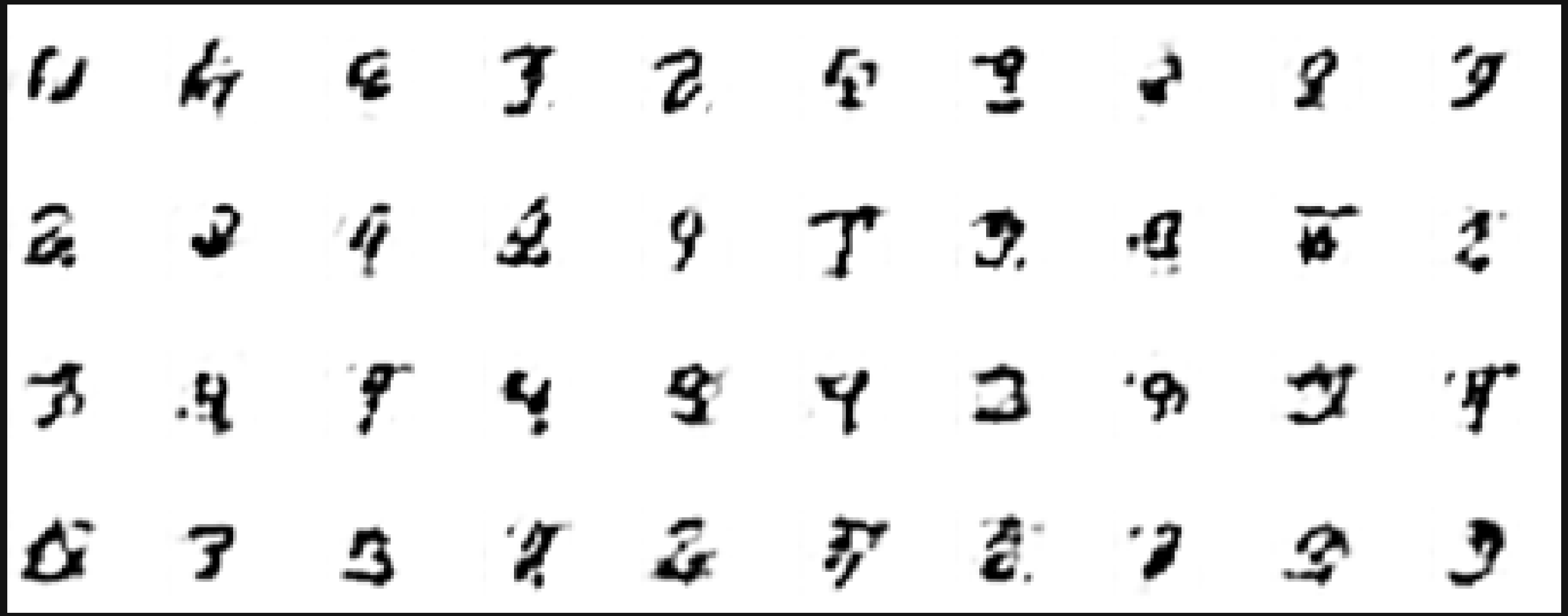
The Training Process

1. We randomly generate a noisy vector
2. Input this into our Generator Network to generate a sample data
3. We then take some sample data from our real data and mix it with some of our generated data
4. Train our descriminator to classify this mixed dataset and thus update it's weights accordingly
5. Now we then train the Generator. We make more random noisy vectors and create synthetic data.With the weights of the Descriminator frozen, we use the feedback from the discriminator to now update the weights of the generator. This is how we teach both our Generator (to make better synthetic images) and Discriminator to get better at spotting fakes.

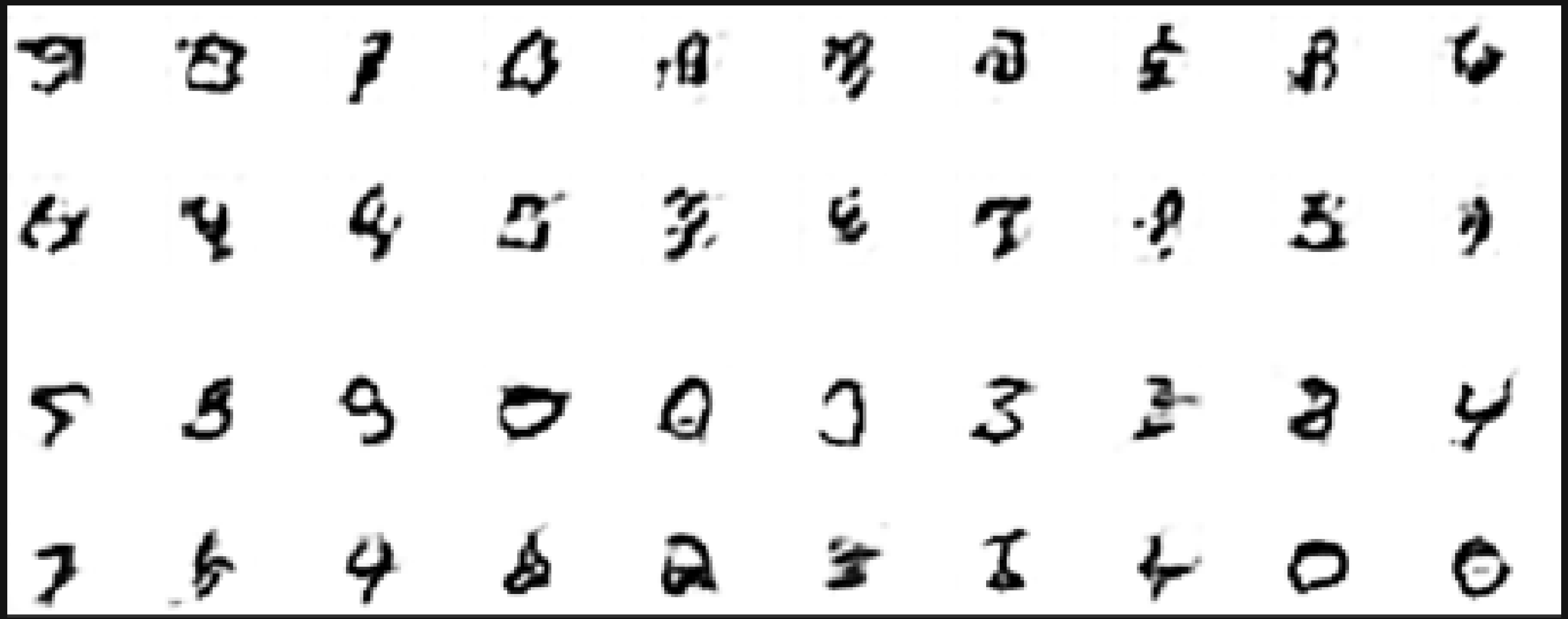




1

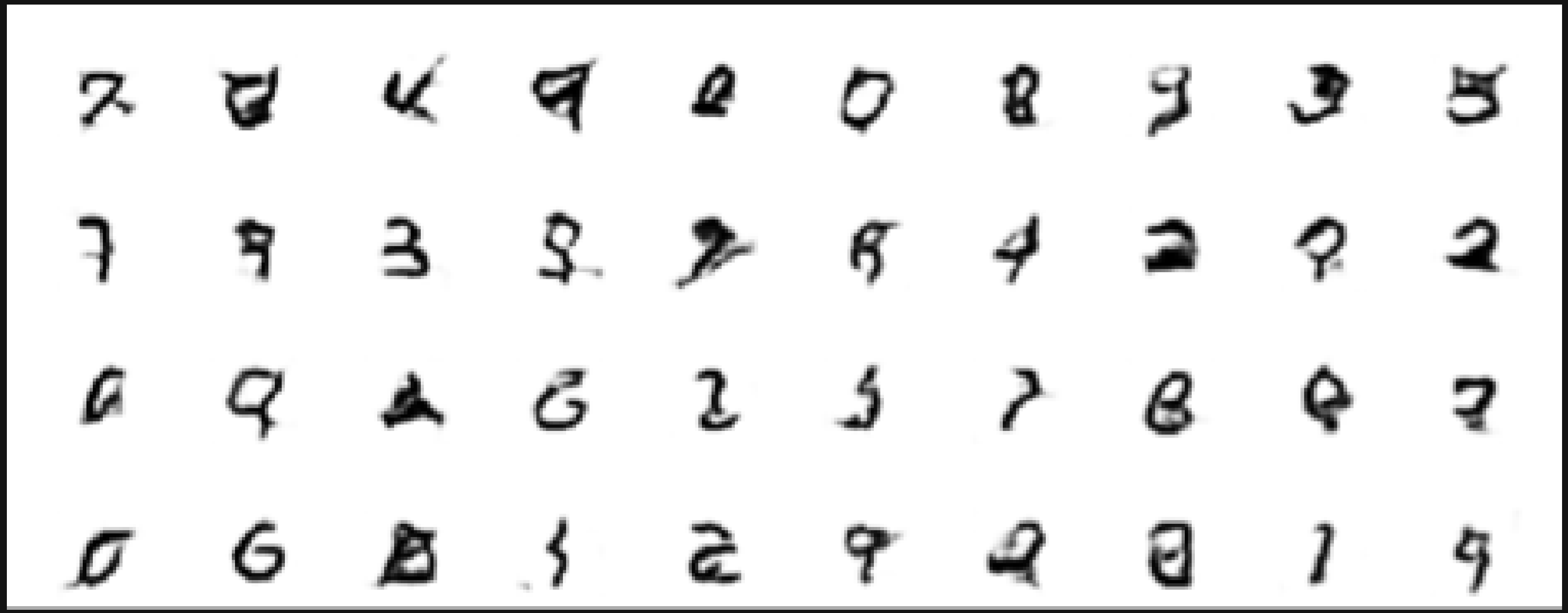


2



3

4



5

Are There Any Issues With GANs?

Are There Any Issues With GANs?

Challanges in Training

- Achieving Equilibrium - Due to it being a dynamic system
- Time - Training GANs is computationally demanding and necessitates tweaking of Hyperparameters such as initializations, altering hidden layers, different activations, using BN or Dropout.
- Bad Initializations causing the descriminator loss to go close to zero (not what you want, ideally you want this to ~0.5 where synethic and real images are indestinguisable). One solution is two purposely make your Generator learn slower than the Descriminator, however this too can be hard to balance as a good Descriminator depends on a good Generator.
- Mode Collapse - happens when regardless of the noise input fed into your generator, the generated output varies very little. It occurs when a small set of images look good to the descriminator and get scored better than other images. The GAN simple learns to reproduce those images over and over (analgous to overfitting).

Next..

Practical use cases for GANs

PRACTICAL USE CASES FOR GANS

Practical Use Case of GANs

- Creative Industries - Art, Music and Design
- Deep Fakes - Replicating facial style in video
- Security - Privacy Preserving, enhancing poor CCTV feeds
- Medical Applications - Data Augumentation and Drug Discovery
- Photography- FaceApp, Smartphone Cameras scene enhancement
- Video and Image Effects - Special Effects Industries
- Video Games - Graphic enhancements using DLSS
- Video Compression - NVIDIA's Maxine
- Marketing Materials - Virtual Try-ons
- Autonomous Vehicles
- Space & Physics - improve astronomical images

Creative Industries

- Art created by GANs have sold for thousands
- Music
  - MidiNet: A Convolutional Generative Adversarial Network for Symbolic-domain Music Generation https://arxiv.org/abs/1703.10847
- GANSynth - Adversarial Neural Audio Synthesis
- WaveGAN and WaveNet- Voice Synthesis for text to speech
- Logos - LoGAN: Generating Logos with a Generative Adversarial Neural Network Conditioned on color (https://arxiv.org/pdf/1810.10395.pdf)

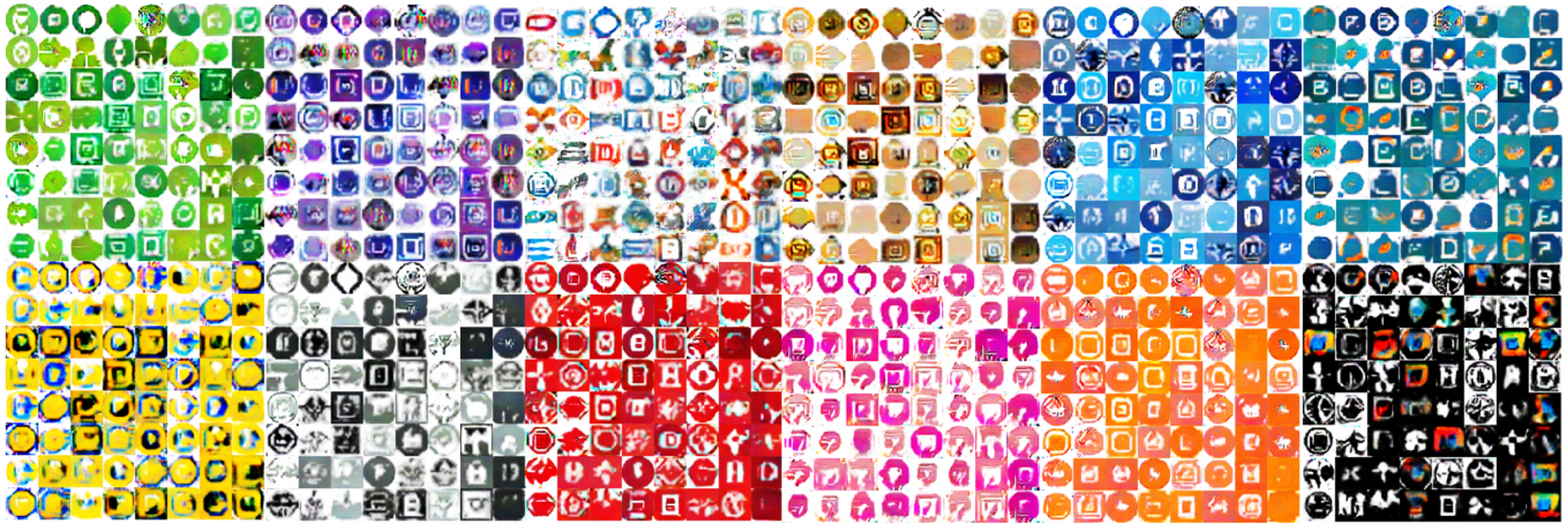
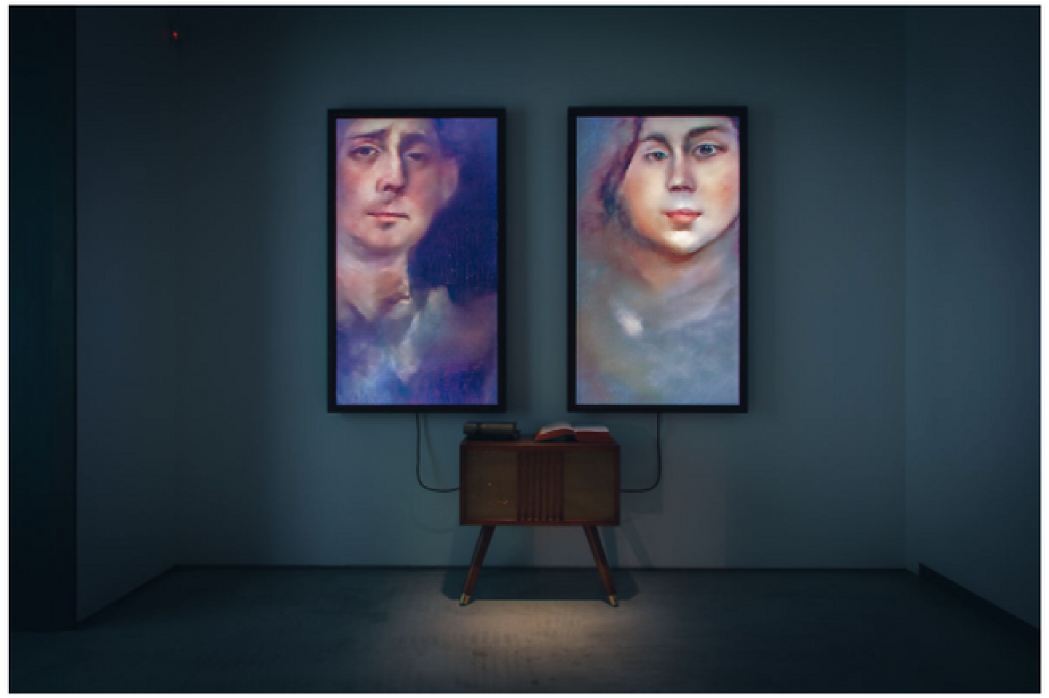
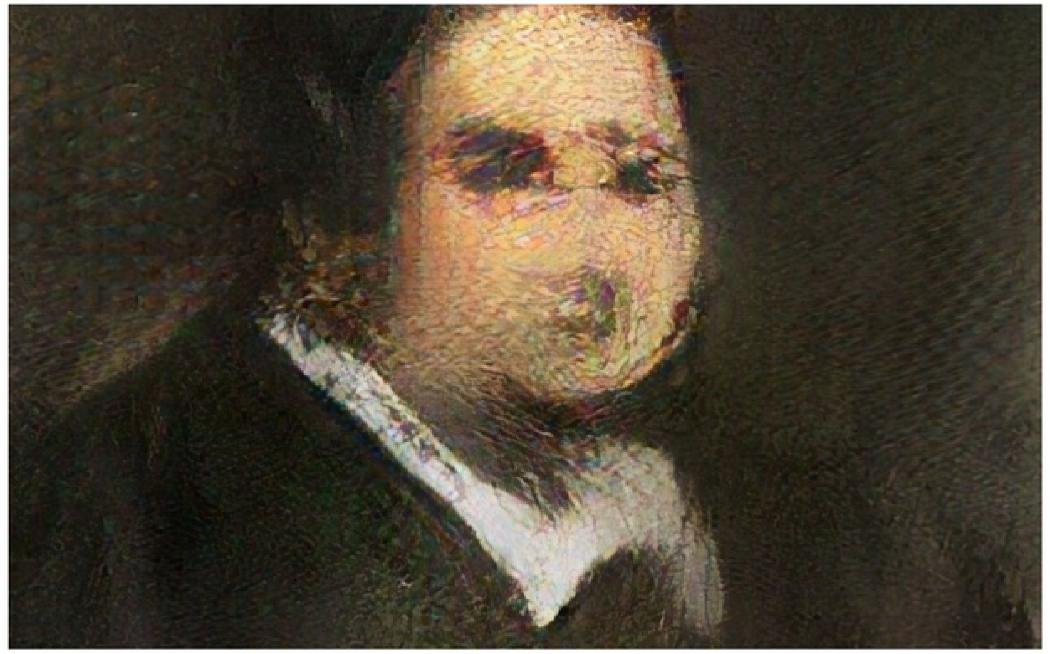


Fig. 6. Results from the generation of 64 logos per class after 400 epochs of training. Classes from left to right top to bottom: green, purple, white, brown, blue, cyan, yellow, gray, red, pink, orange, black.

An artwork created by AI sold for £40,000 at Sotheby's, failing to generate the fervor that propelled another AI work to sell for 40 times its estimate last year.



Mario Klingemann, *Memories of Passersby I*, 2018. Sold for £40,000. Courtesy Sotheby's.

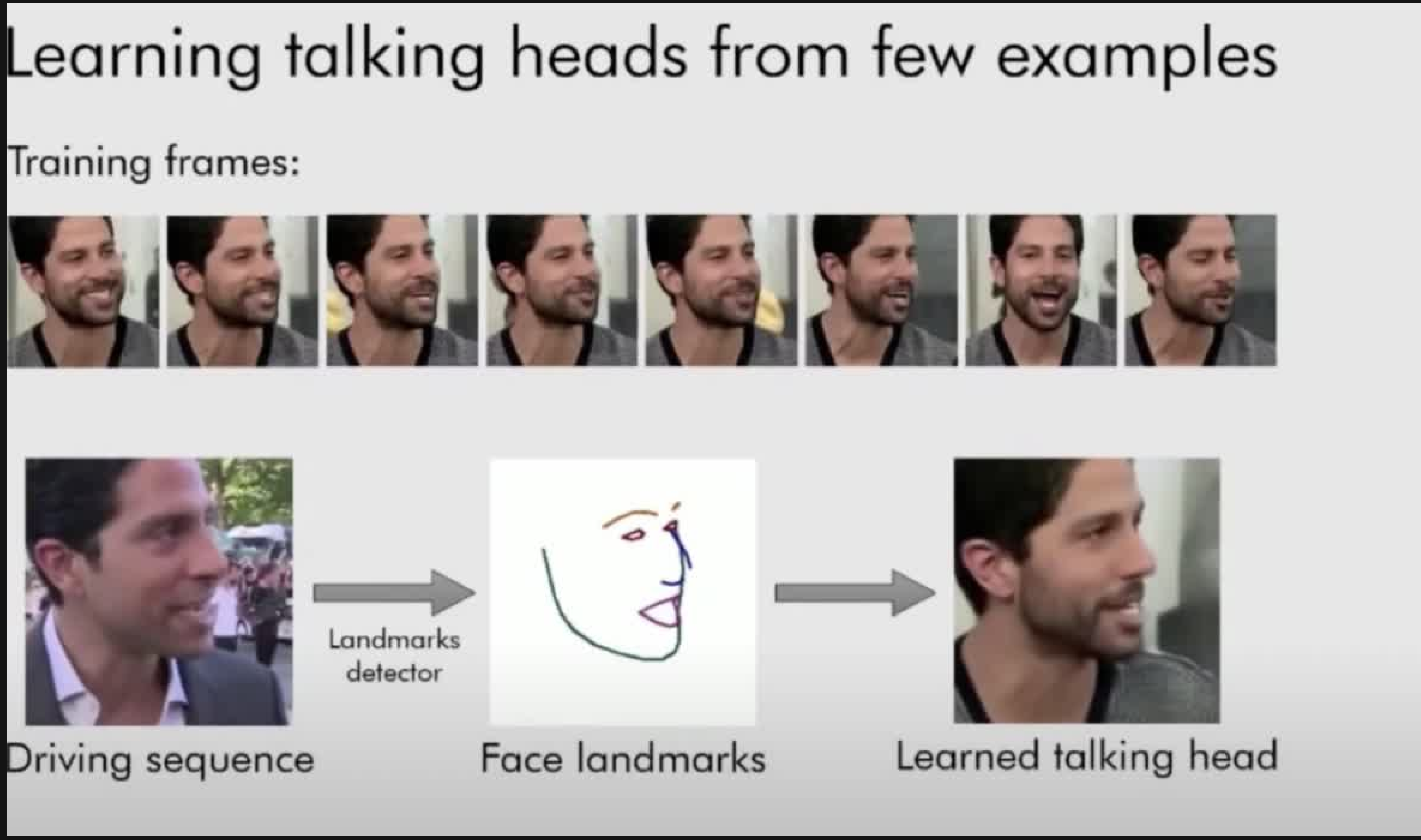


Is artificial intelligence set to become art's next medium?

AI artwork sells for $432,500 — nearly 45 times its high estimate — as Christie's becomes the first auction house to offer a work of art created by an algorithm

Deep Fakes

- Few-Shot Adversarial Learning of Realistic Neural Talking Head Models
- ArcaneGAN (Converts images to Netflix's Arcane's style!)

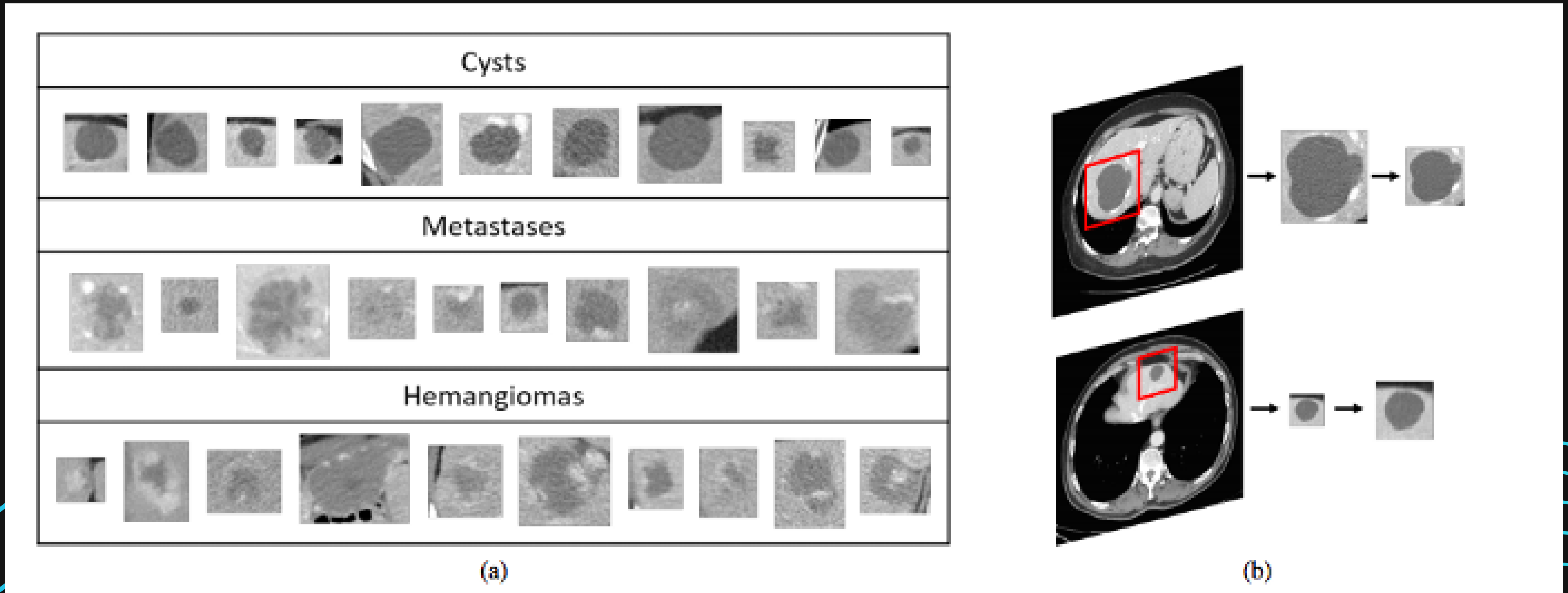


Security

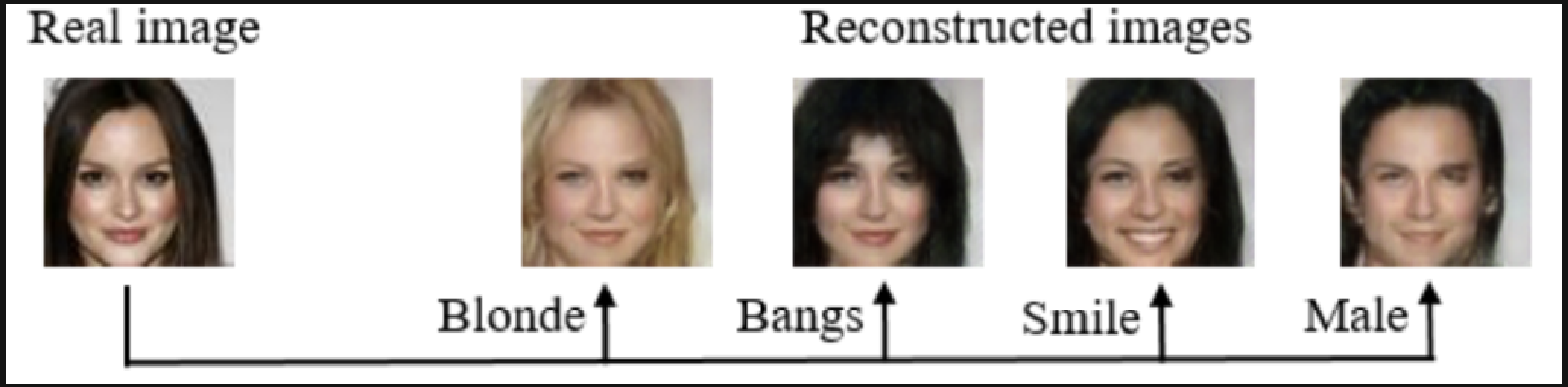
- Privacy Maintence: Instead of sharing real data we can share synthetic data that is indestinguisable from the real thing
- Cryptography - Learning to Protect Communications with Adversarial Neural Cryptography
    - https://arxiv.org/abs/1610.06918
- CCTV Footage Enhancement - SNIDER: Single Noisy Image Denoising and Rectification for
- Improving License Plate Recognition

Medical Applications

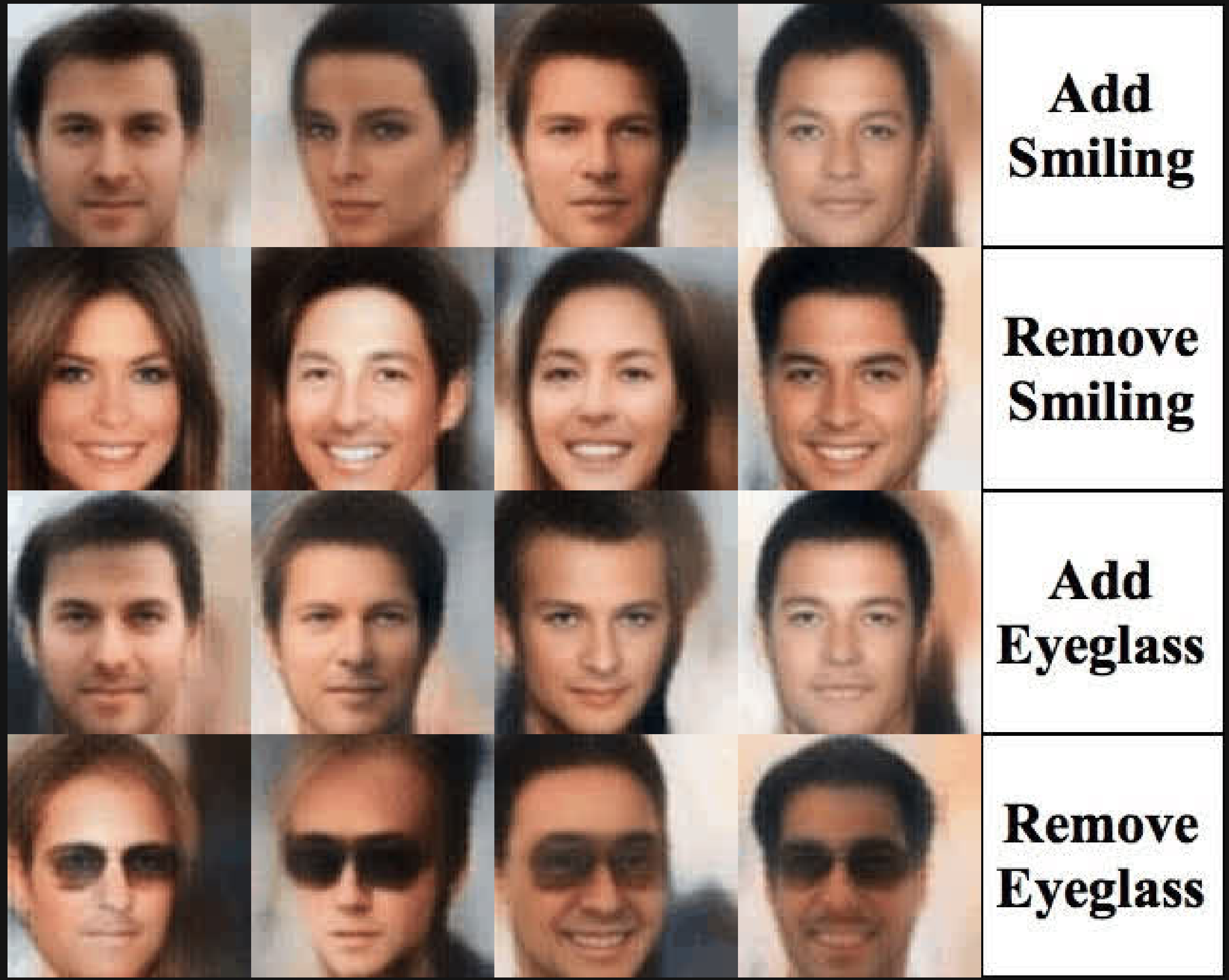
- Drug Discovery: GANs can quickly generate novel biological components to test hypothesis simultaneously - https://www.eurekalert.org/pub_releases/2018-10/imi-aid101018.php
- Data Augmentation - GANs were used to augment medical brain scan CT images which improved the sensitivity and specifity of their Brain Disease classifier to 85.7% and 92.4% respectively.
    - https://arxiv.org/pdf/1803.01229.pdf

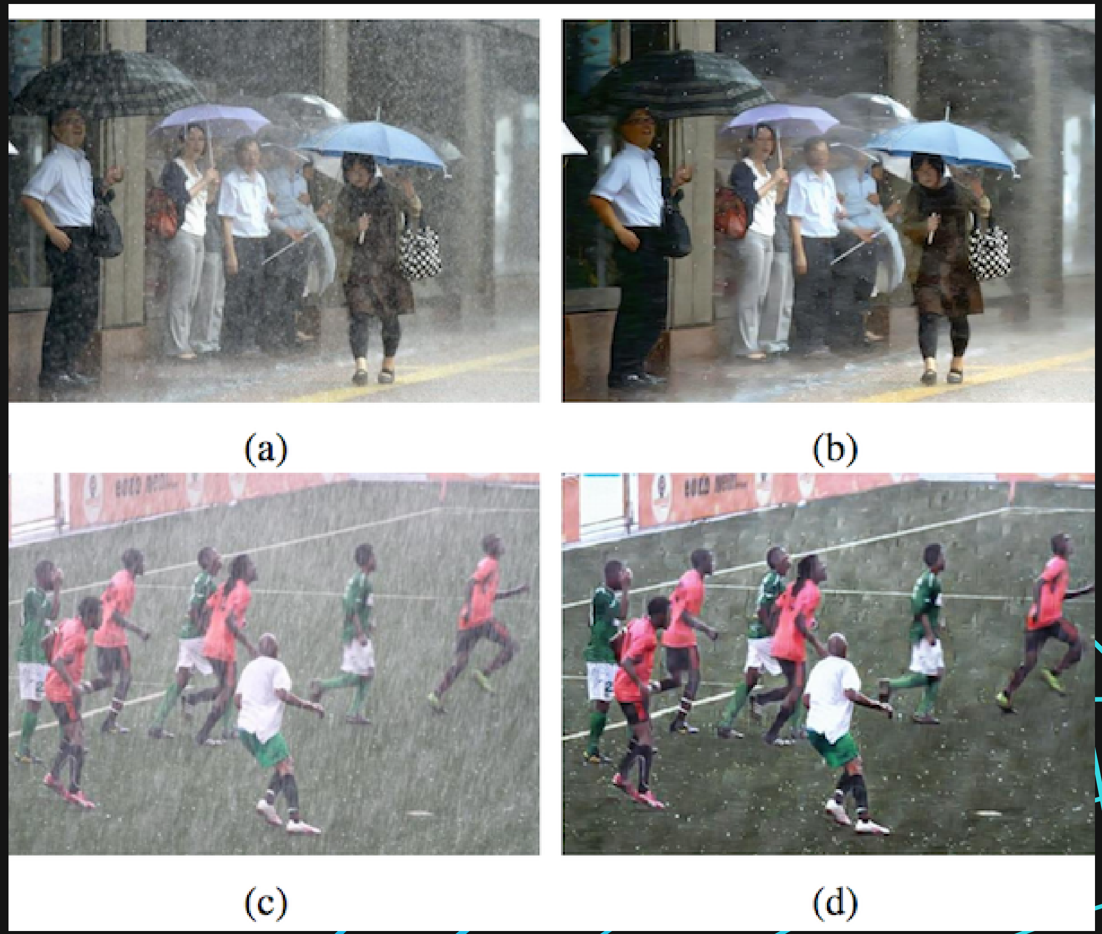


Photography



"Invertible Conditional GANs for Image Editing"
https://arxiv.org/abs/1611.06355





“Image De-raining Using a Conditional Generative Adversarial Network"
https://arxiv.org/abs/1701.05957

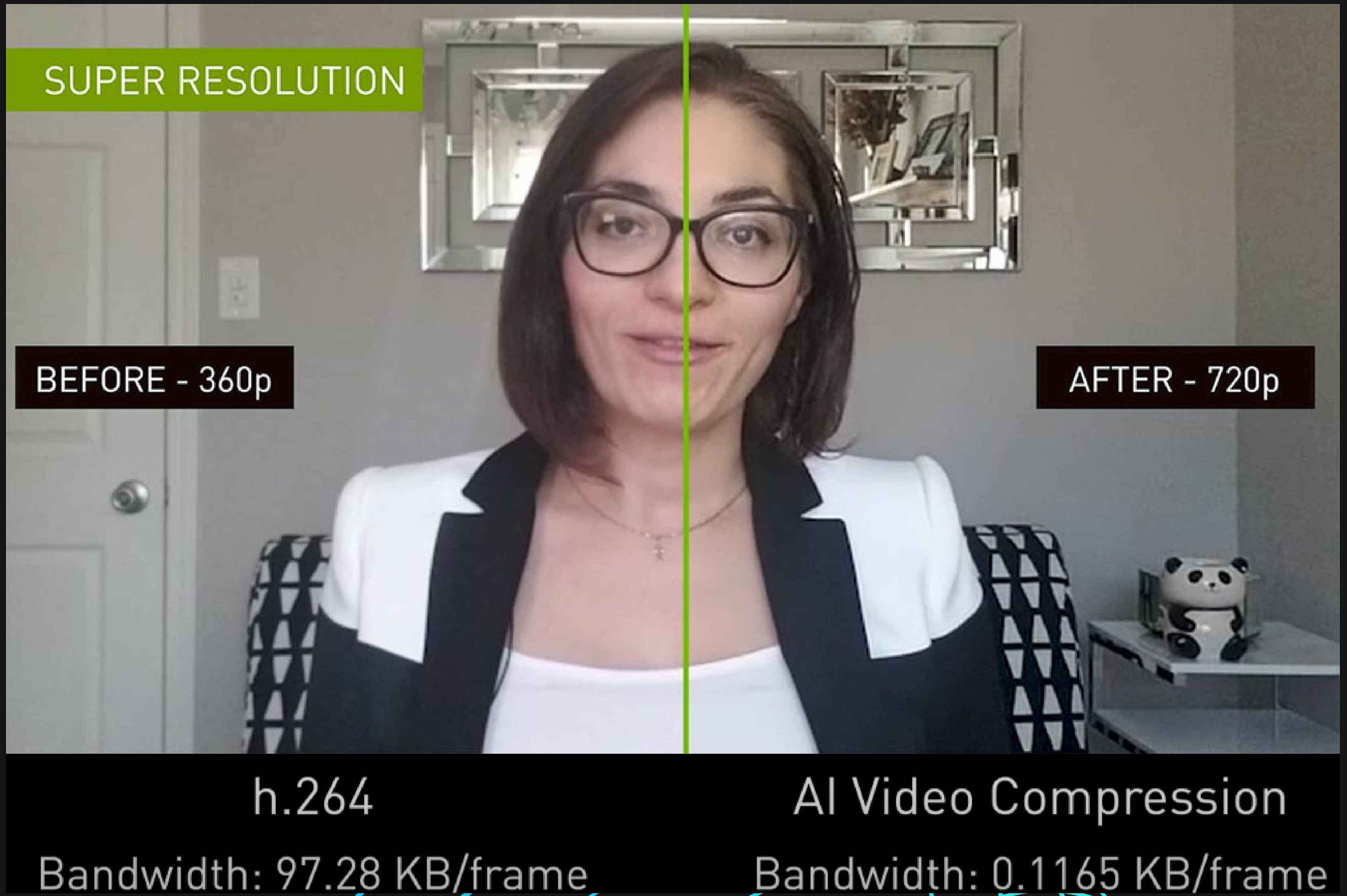
Videogame Graphics

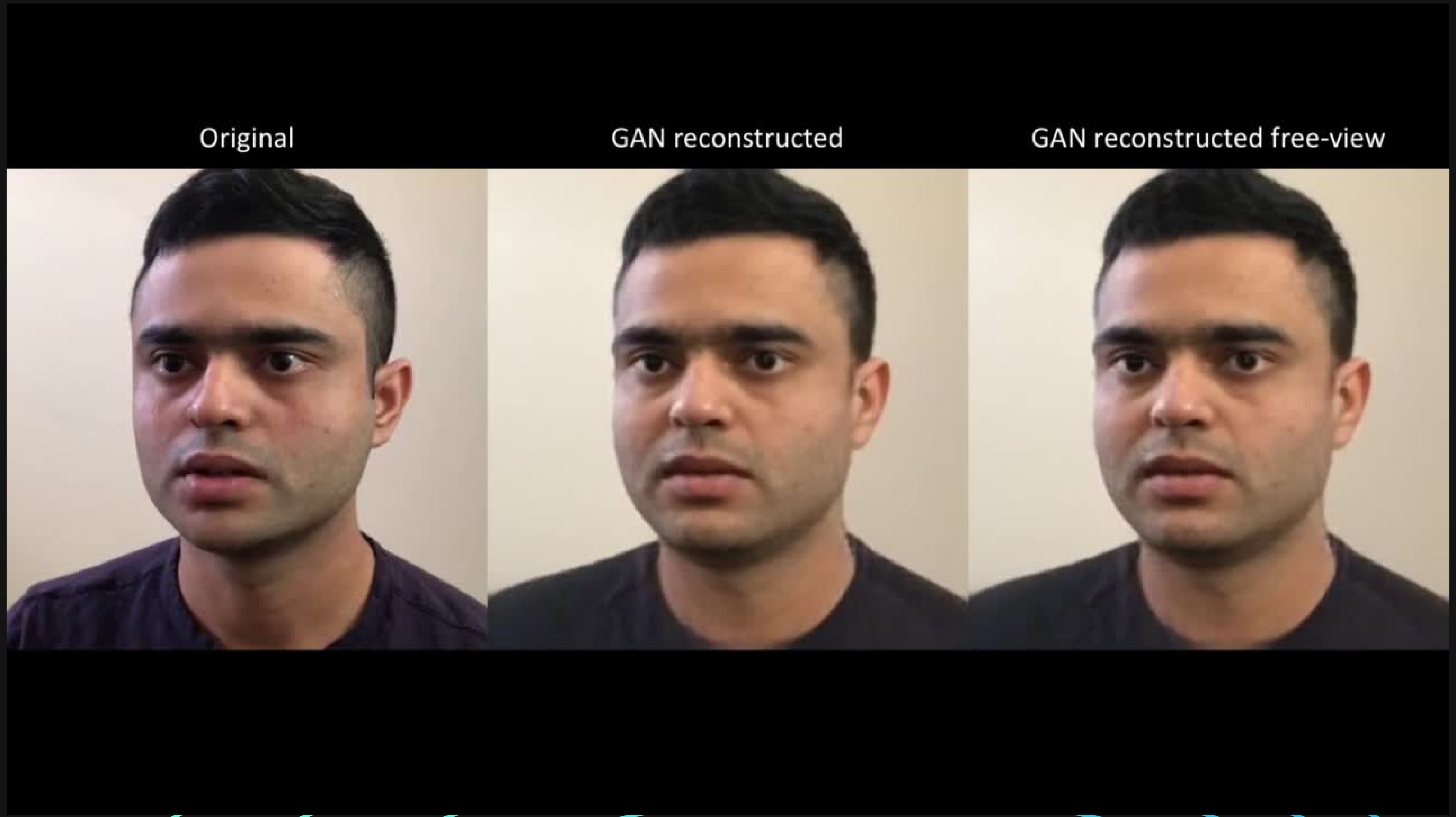
- Deep Learning Super Sampling (DLSS): It enables the upscaling of lower-resolution images to a higher-resolution for display on higher-resolution displays (4K), while rendering natively at lower resolutions.

https://www.nvidia.com/en-us/geforce/news/graphics-reinvented-new-technologies-in-rtx-graphics-cards/#dlss

Video Compression - NVIDIA's Maxine

A generative adversarial network on the receiver's side uses the initial image and the facial key points to reconstruct subsequent images on a local GPU.





Video and Image Effects

Sky Replacement - "Castle in the Sky: Dynamic Sky Replacement and Harmonization in Videos", in arXiv:2010.11800. Built using pytorch-CycleGAN-and-pix2pix

Marketing Materials

Pose Guided Person Image Generation - https://arxiv.org/abs/1705.09368

Autonomous Vehicles

- Simulating Environments (GraspGAN)
- SeGAN - Inferring Occluded Objects
- SRGAN (Super Resolution) for enhancing clarity of images
- Inpainting - For reducing noise in sensors

GAN-Generated Photograph Inpainting Using Context Encoders.Taken from Context Encoders: Feature Learning by Inpainting describe the use of GANs, specifically Context Encoders, 2016

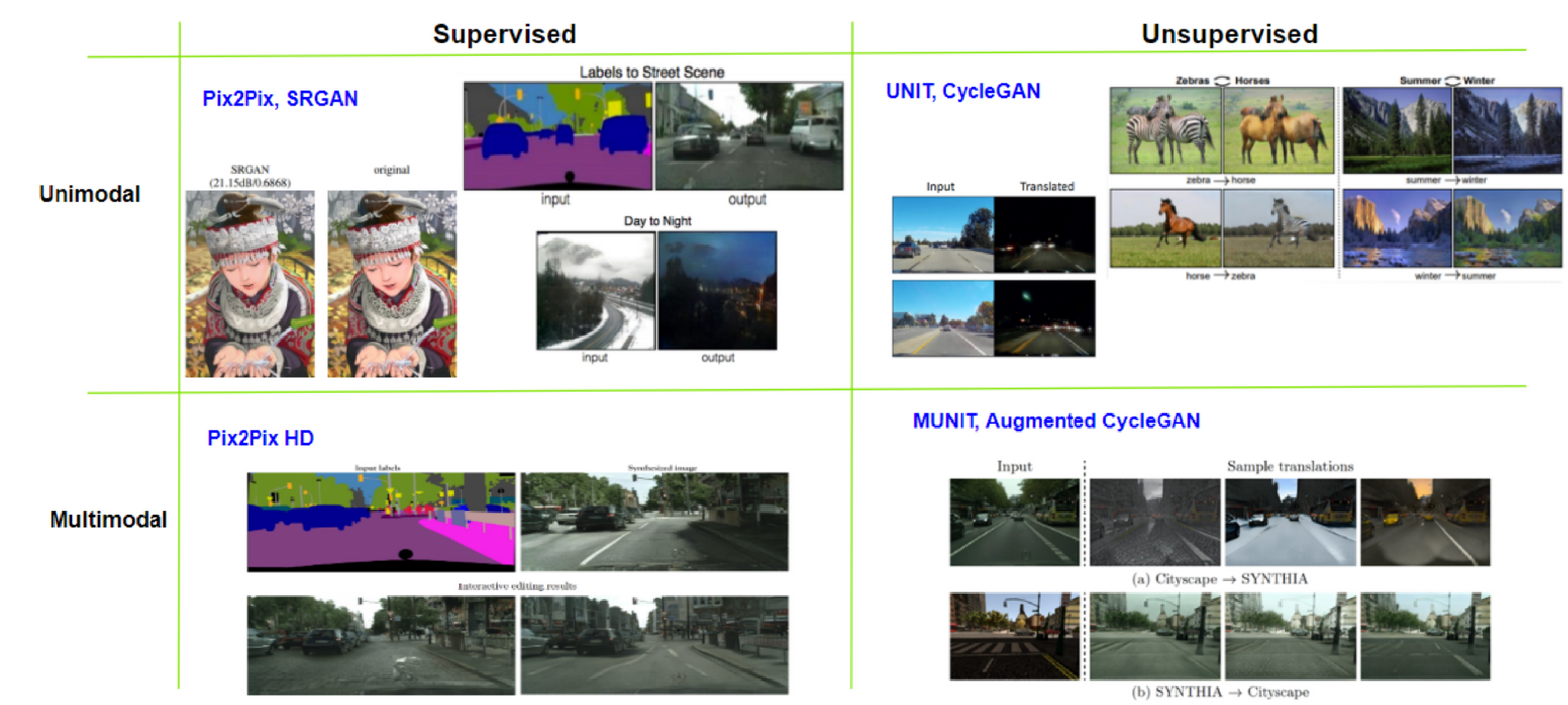
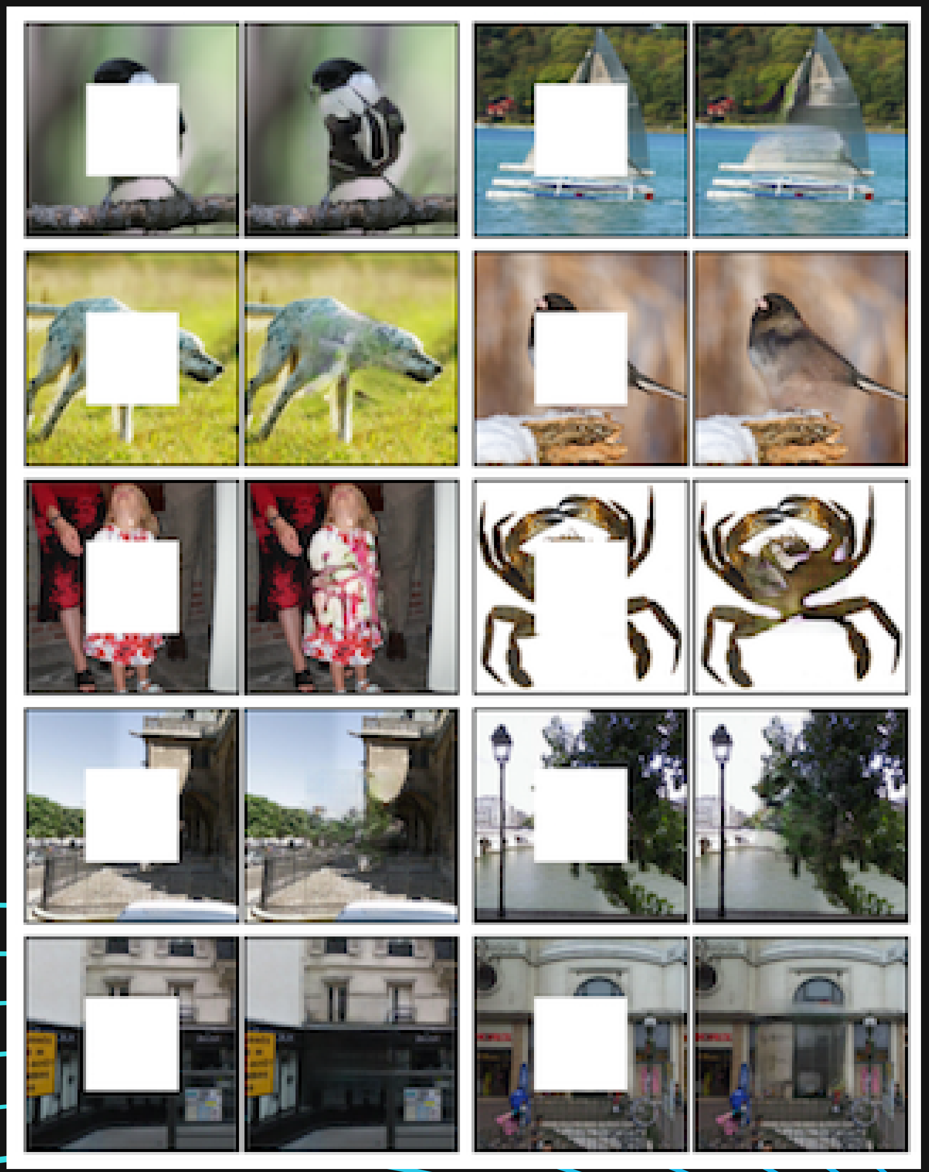


Figure 4: *Successful application of GAN for autonomous driving - Image to Image Translation*



Space & Phyiscs

CosmoGAN: Training a neural network to study dark matter - GANs can improve astronomical images and simulate gravitational lensing for dark matter research.