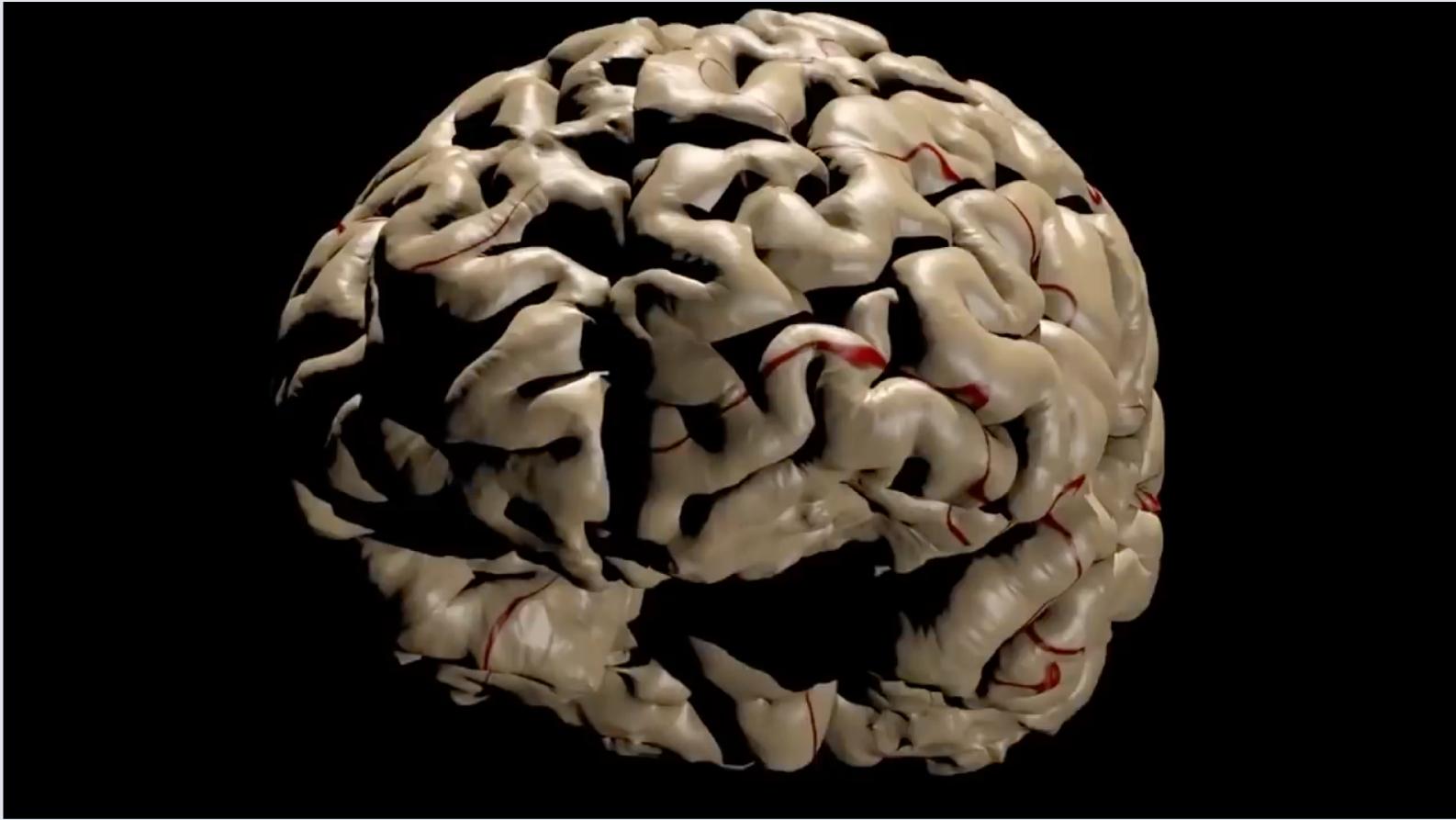


# Magical World of GANs

---

Krishna Balaga  
Developer Advocate

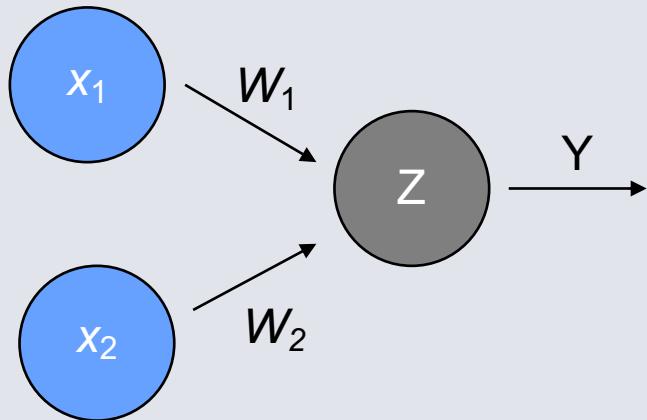
# Where it all Started



[youtu.be/31YqXxHjPll](https://youtu.be/31YqXxHjPll)

Search for : Neuron 3d structure

# The Inception!

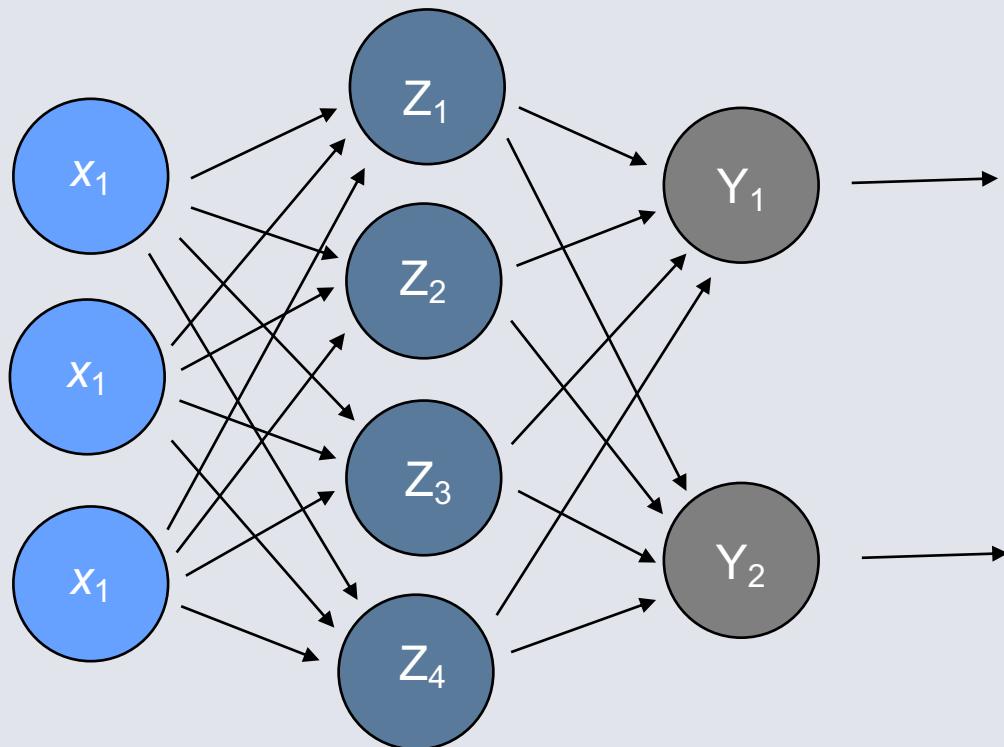


A detailed diagram showing the internal calculation of the output  $Y$ . The input  $x_1$  is multiplied by the weight  $W_1$  to produce a weighted sum. This process is repeated for input  $x_2$  with weight  $W_2$ . The resulting weighted sums are then summed up. This summation step is labeled "Weighted Sum". The final result of this summation is passed through an activation function  $g$  to produce the output  $Y$ . This final step is labeled "Activation".

$$Y = g \left( \sum_{i=0}^n x_i W_i \right)$$

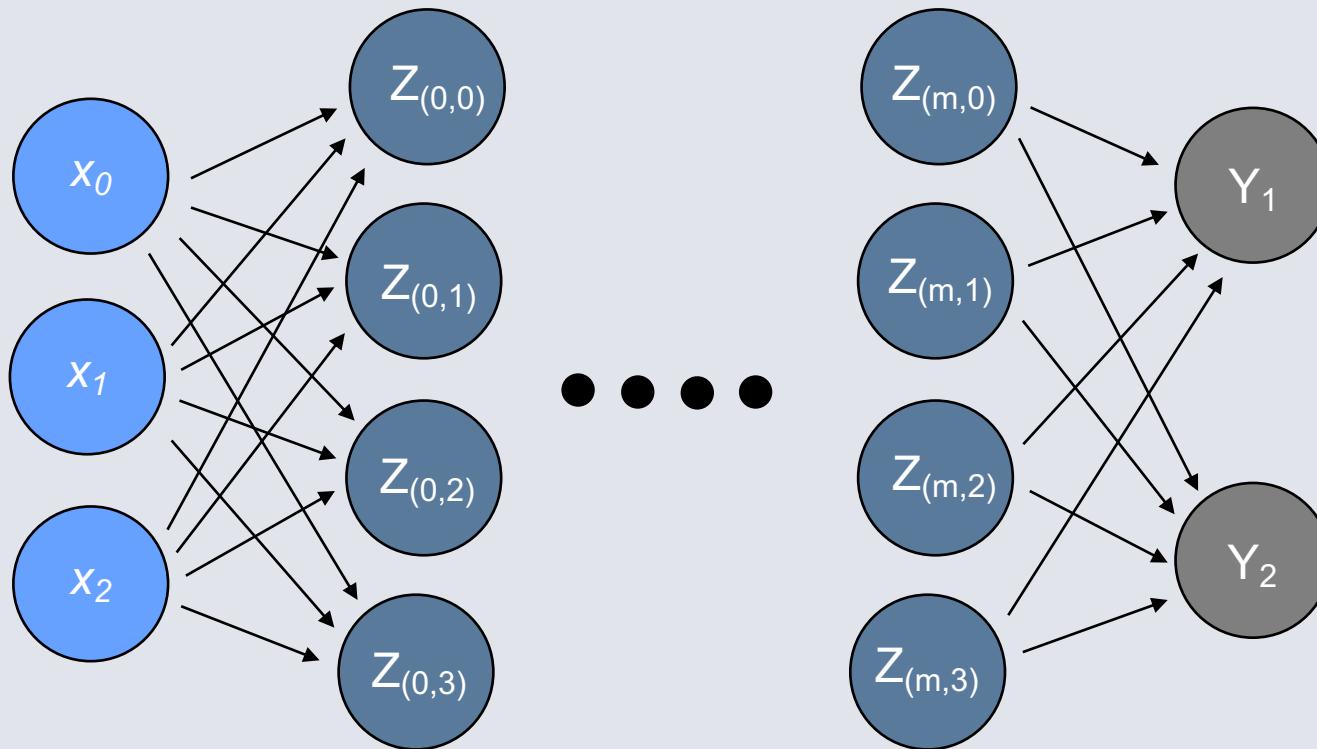
Building a simple perceptron

# The Inception!



Assembling a Neural Net

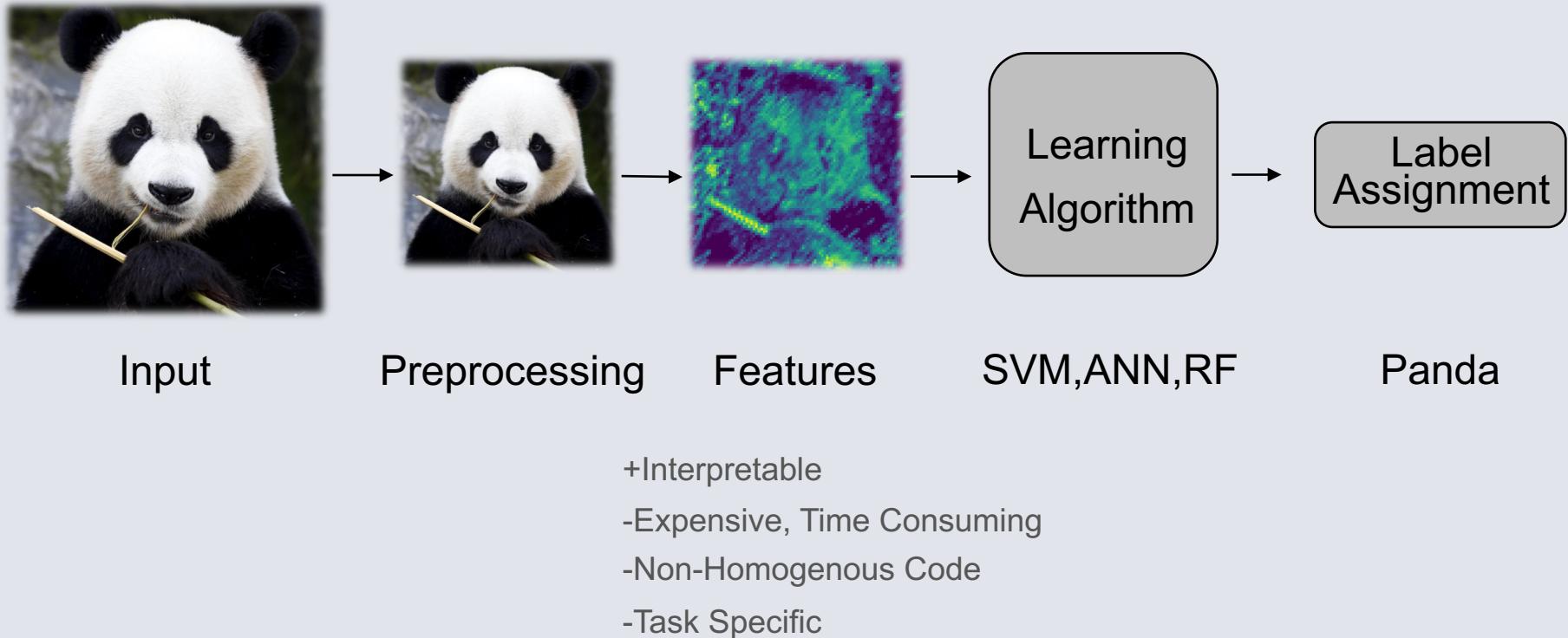
# The Inception!



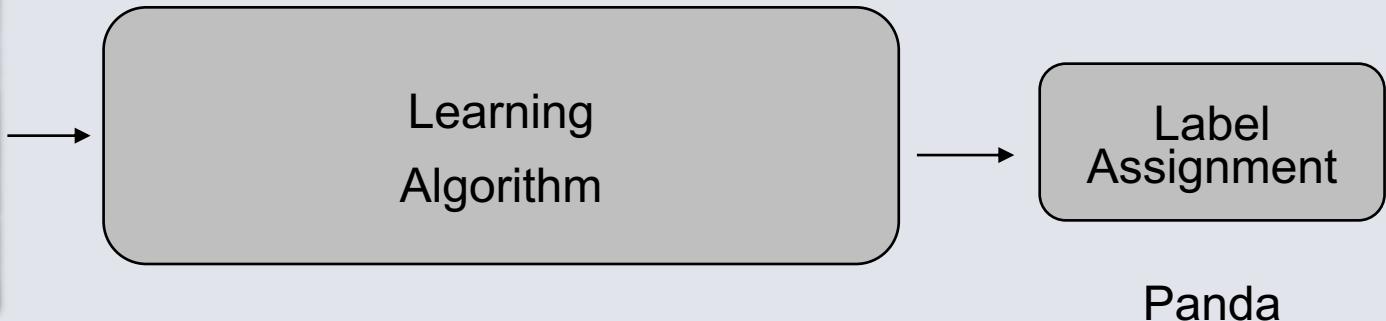
Making a Deep Neural Networks

# How are they Different?

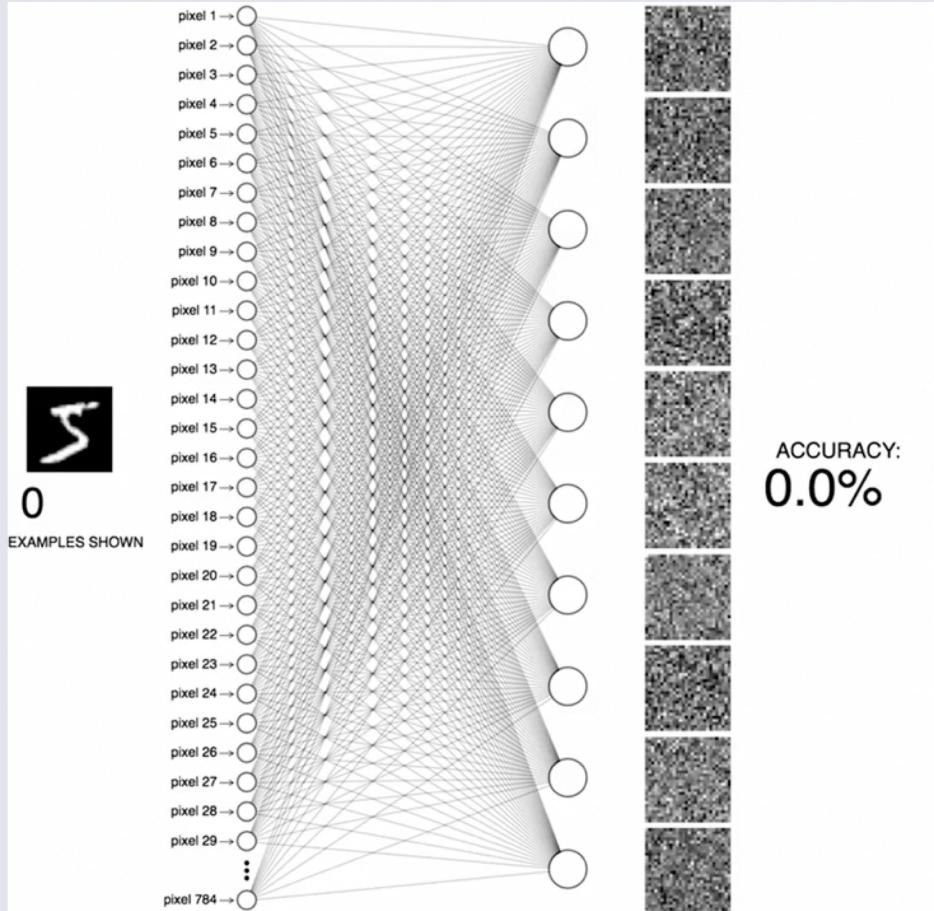
# Computer vision before Deep Learning



# Computer vision with Deep Learning



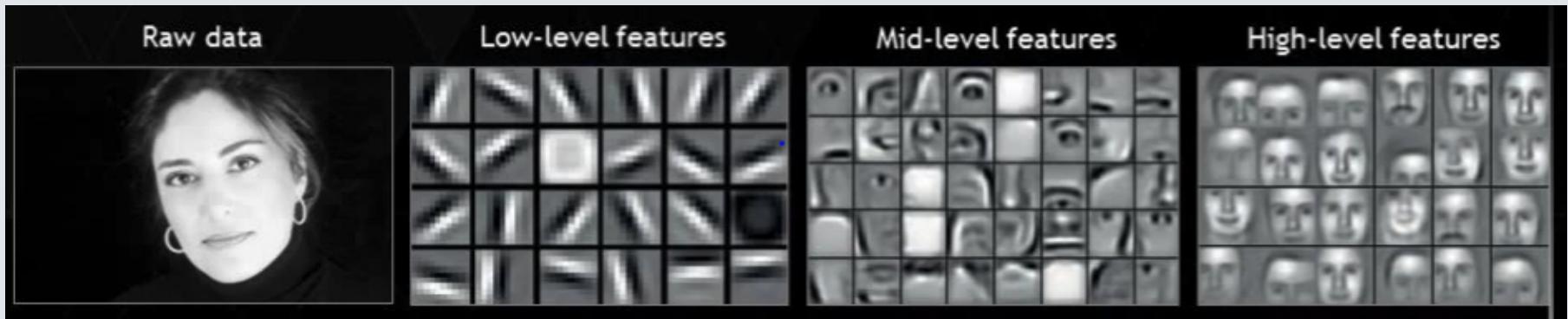
- +Trained End to End
- +Scalable, homogenous code
- +Generalizes well
- Resource and Data Intensive



During training, the network starts finding out a good set of weights for the entire network that maps the input with output

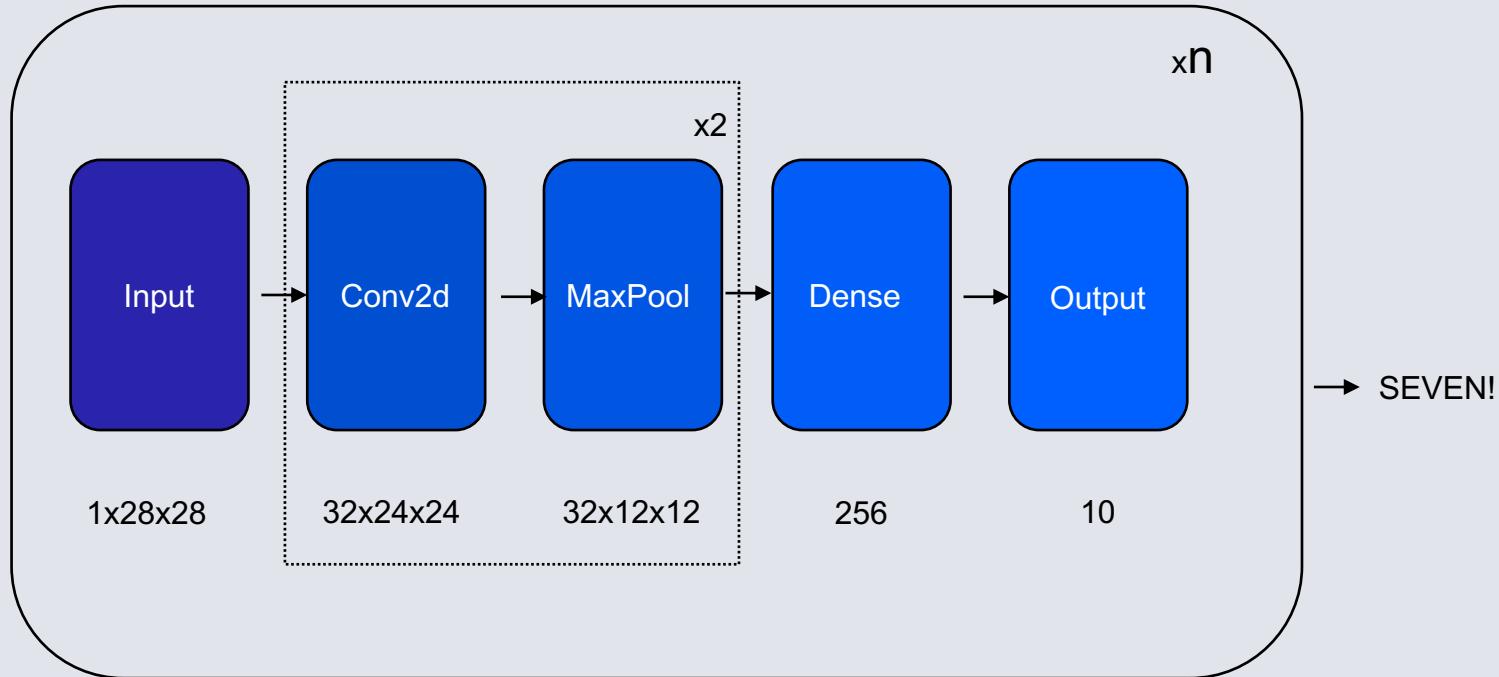
# Essentially,

With Deep learning, it is possible to learn the  
**underlying features** directly from data





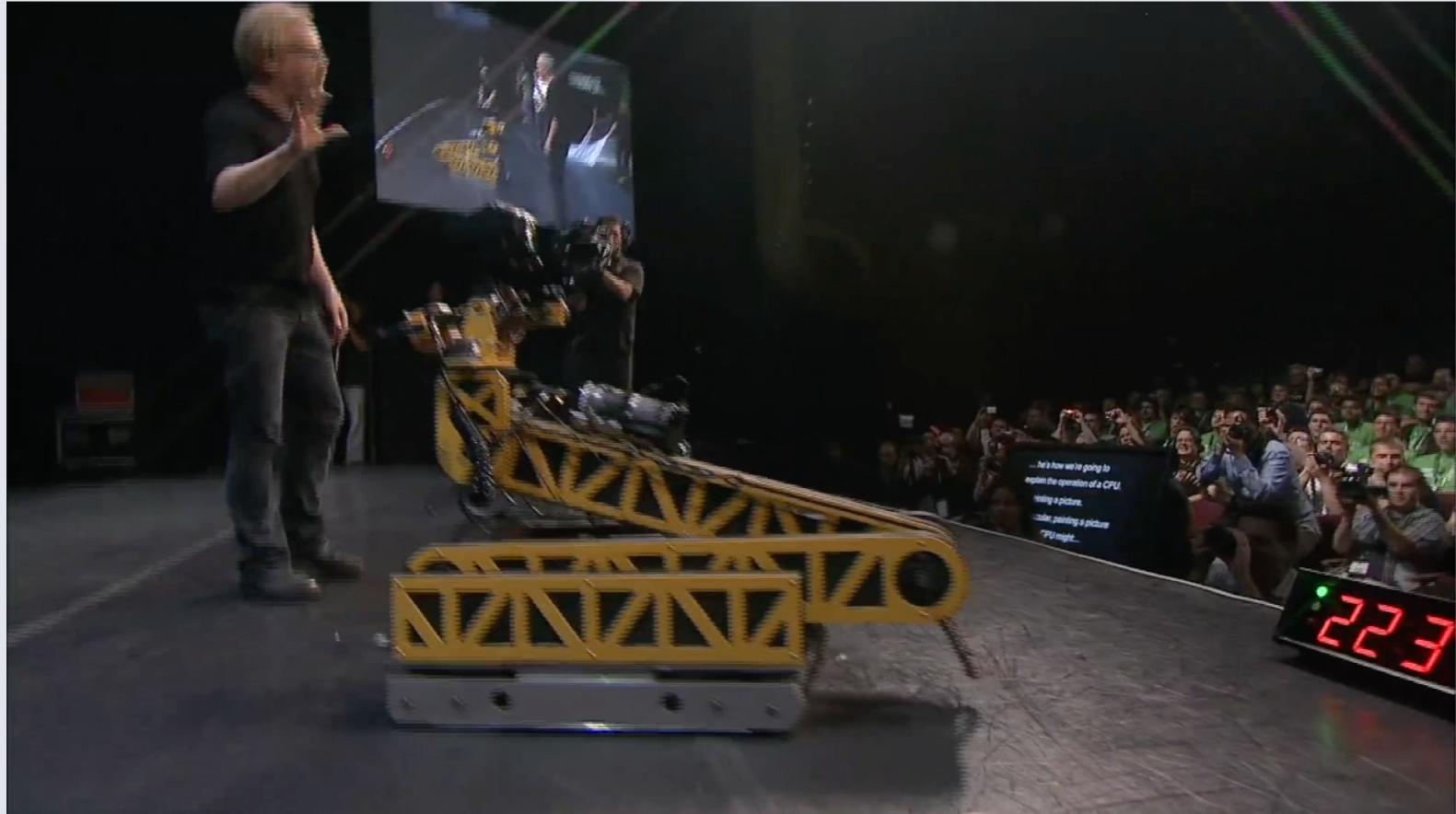
28x28x1  
60,000 samples



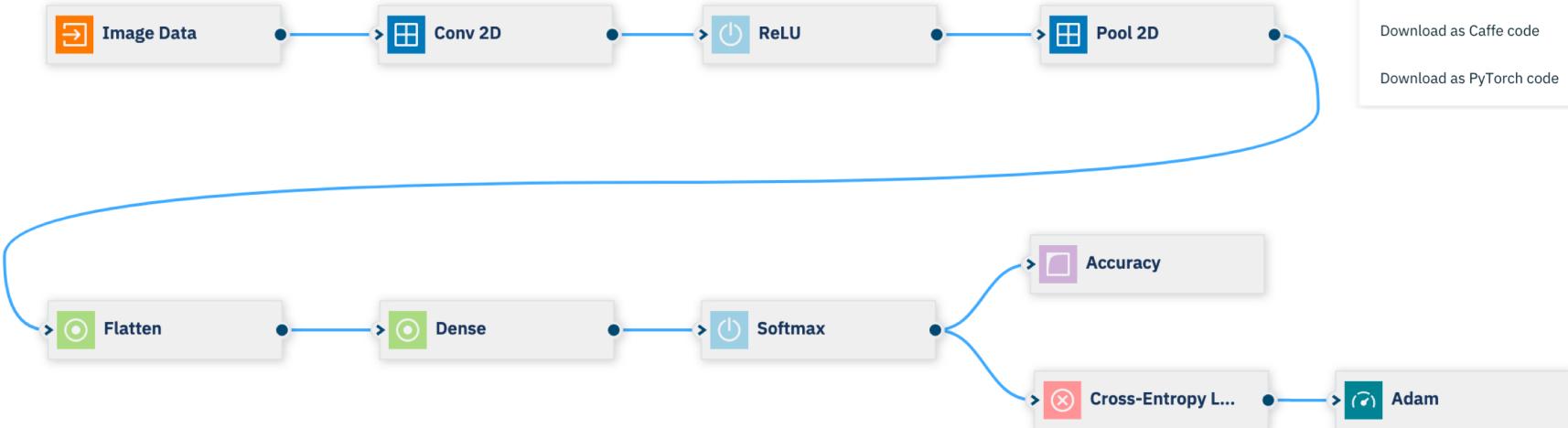
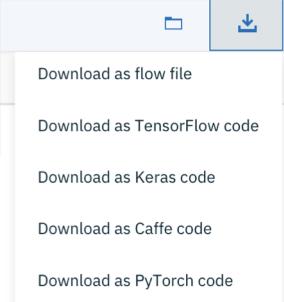
---

217,706

# CPU vs GPU



# Watson Modeler Flows



# Competitor vs IBM

```
Time (s) to convolve 32x7x3  
CPU (s):  
9.5946969986  
GPU (s):  
2.33175206184  
GPU speedup over CPU: 4x
```

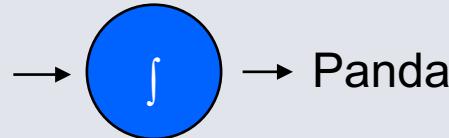
k80

```
Time (s) to convolve 32x7x7x3  
CPU (s):  
3.8780993139371276  
GPU (s):  
0.029320250963792205  
GPU speedup over CPU: 132x
```

v100

Free offerings

# UseCases



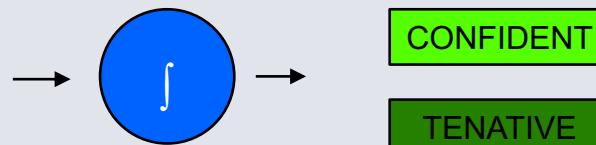
## Image Classification

Food and water ran out four days ago.

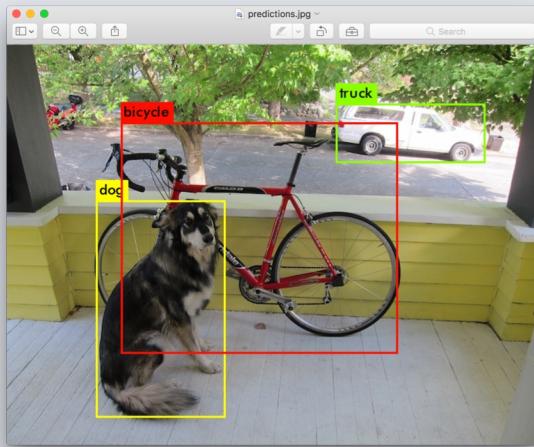
Oxygen will run out tomorrow morning.

When I drift off, I will dream about you.

It's always you

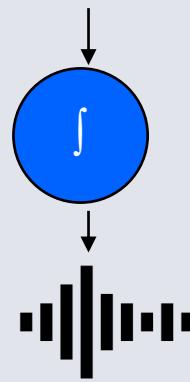


## sentiment analysis

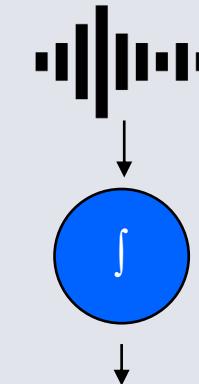


Object detection

Hello there!



Text to speech

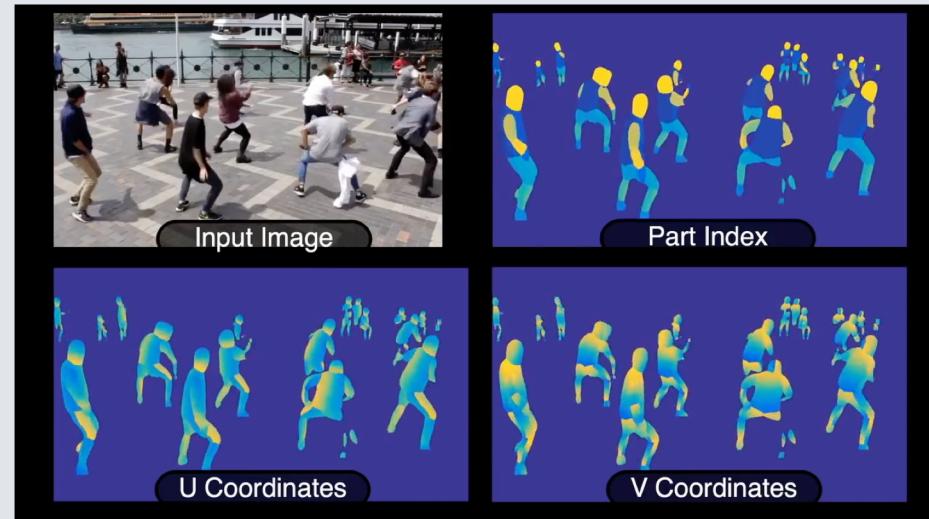


hope you are  
having fun

Speech to text



Markerless full body pose estimation

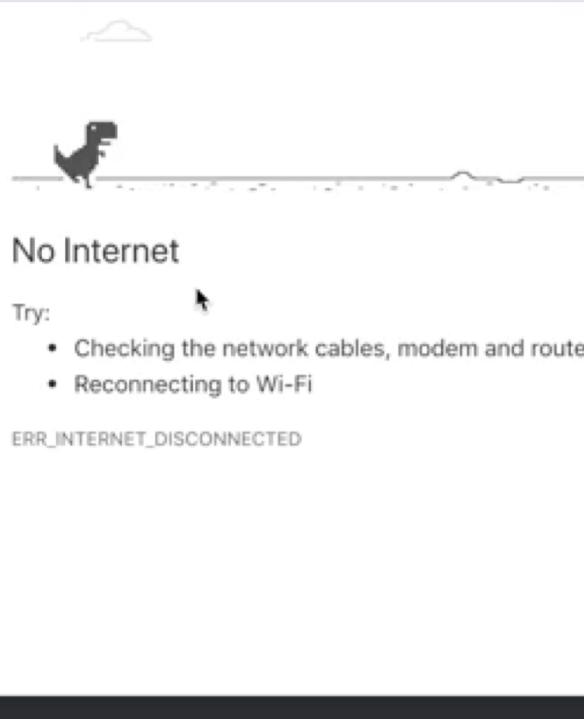


Dense Pose



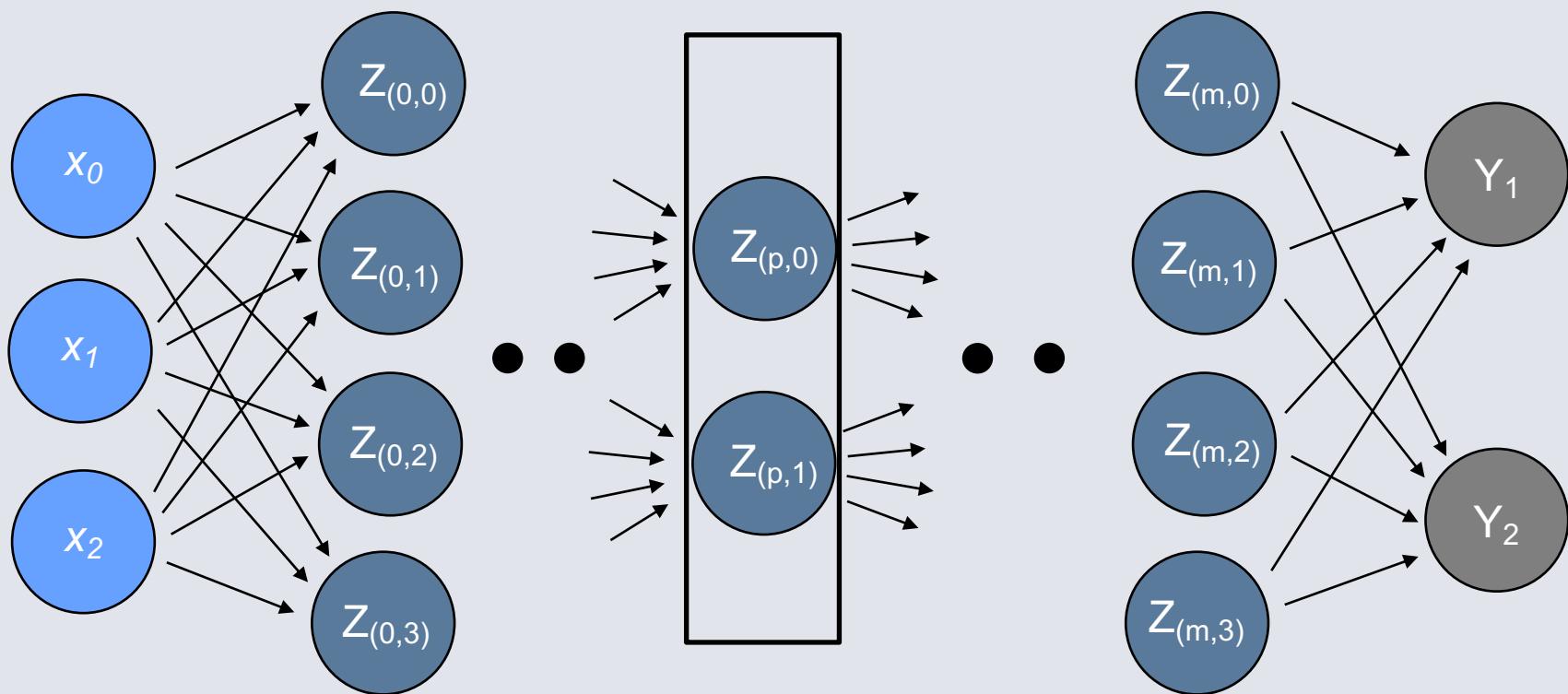
## Semantic Segmentation

# FACIAL LANDMARKING



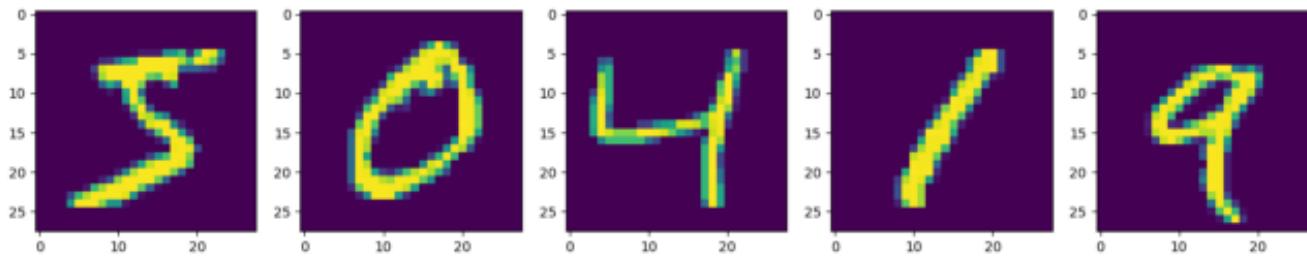
# Auto Encoders

## The Bottle Neck!

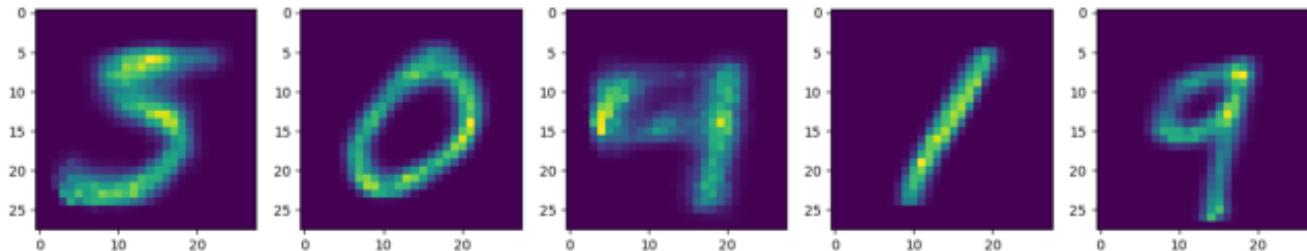


Making a Deep Neural Networks

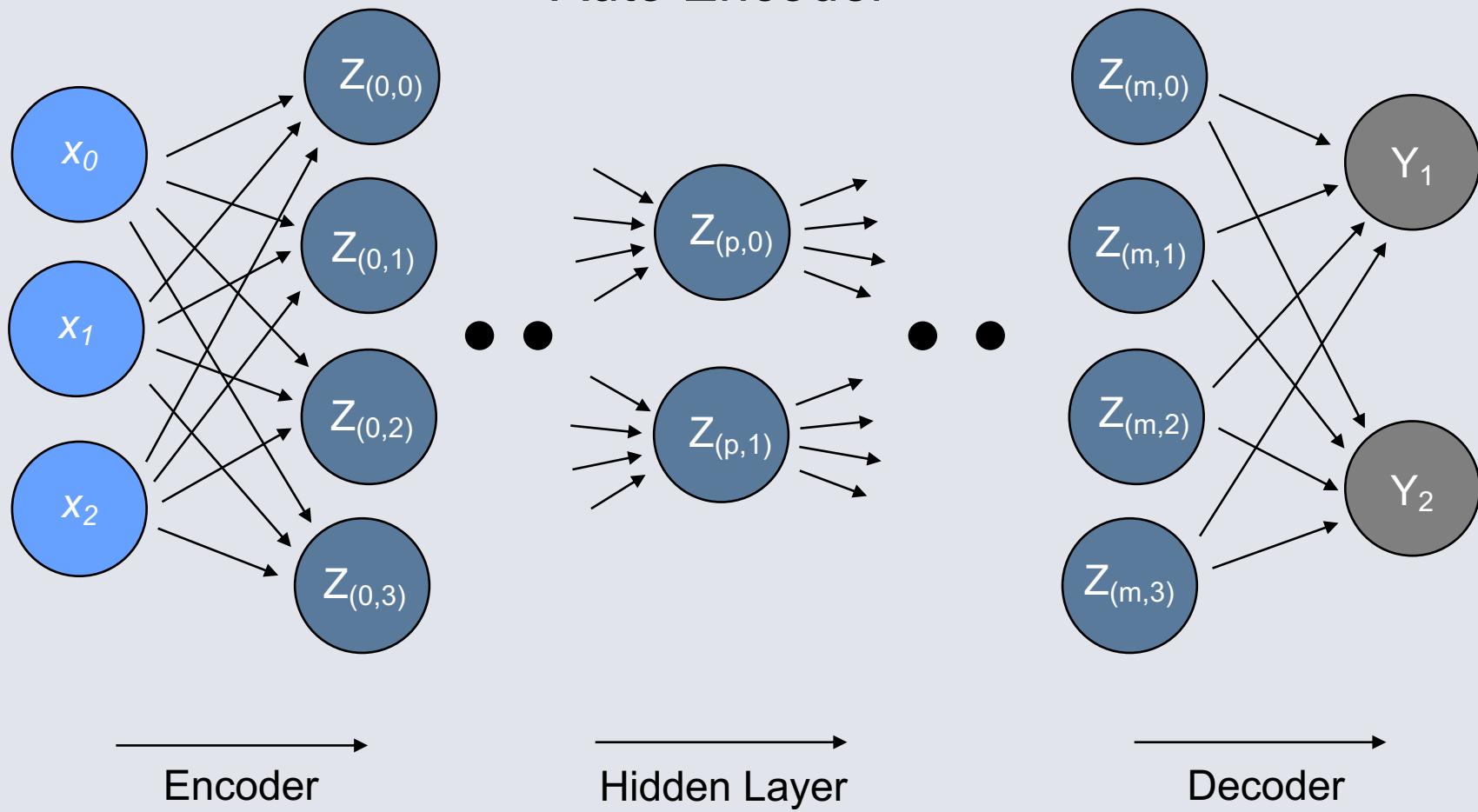
Input



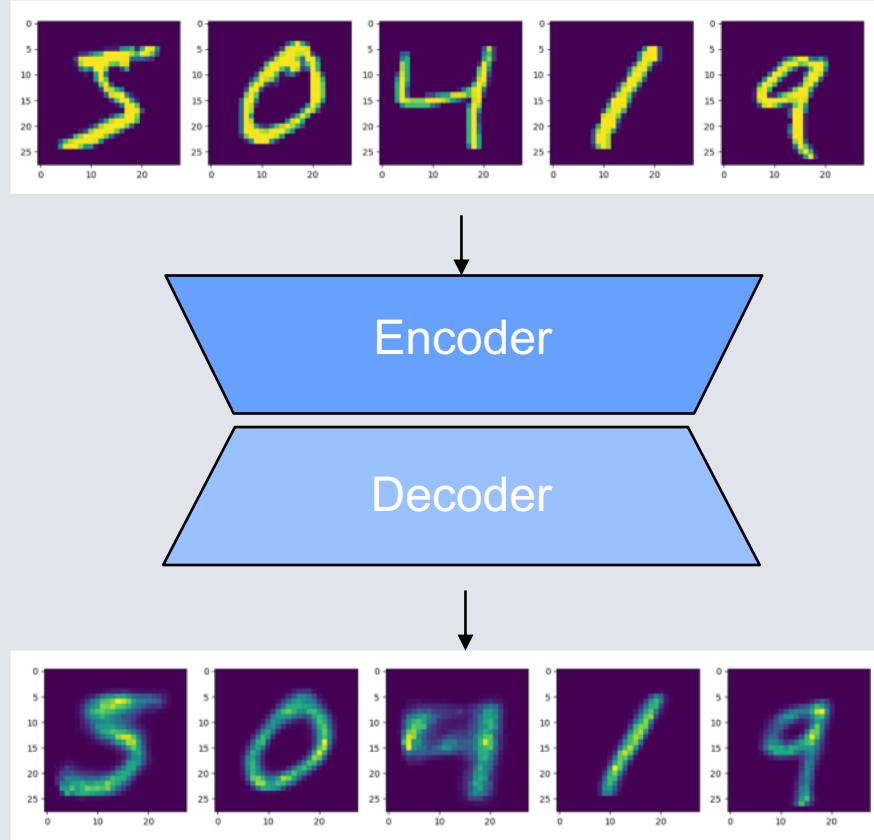
Output



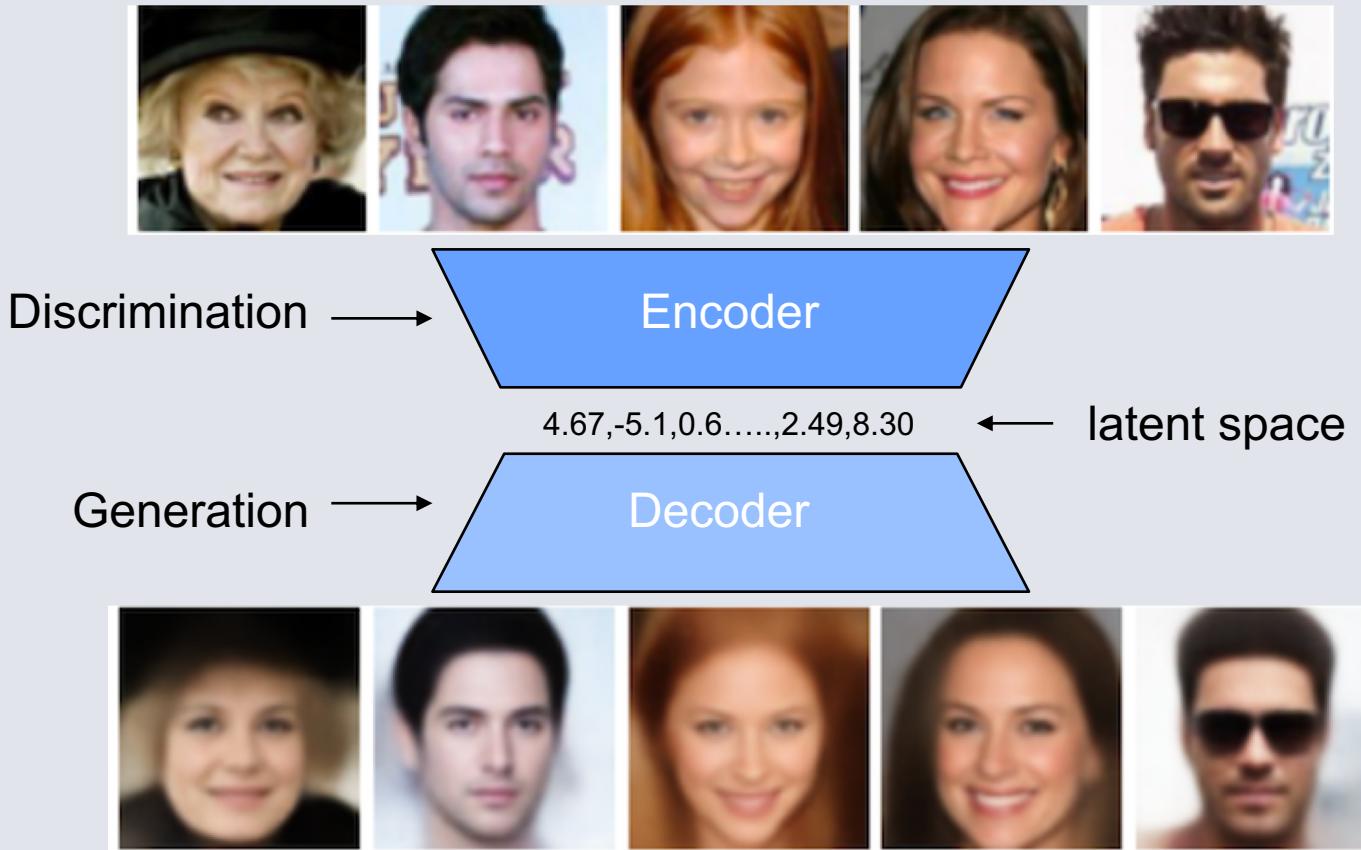
# Auto Encoder



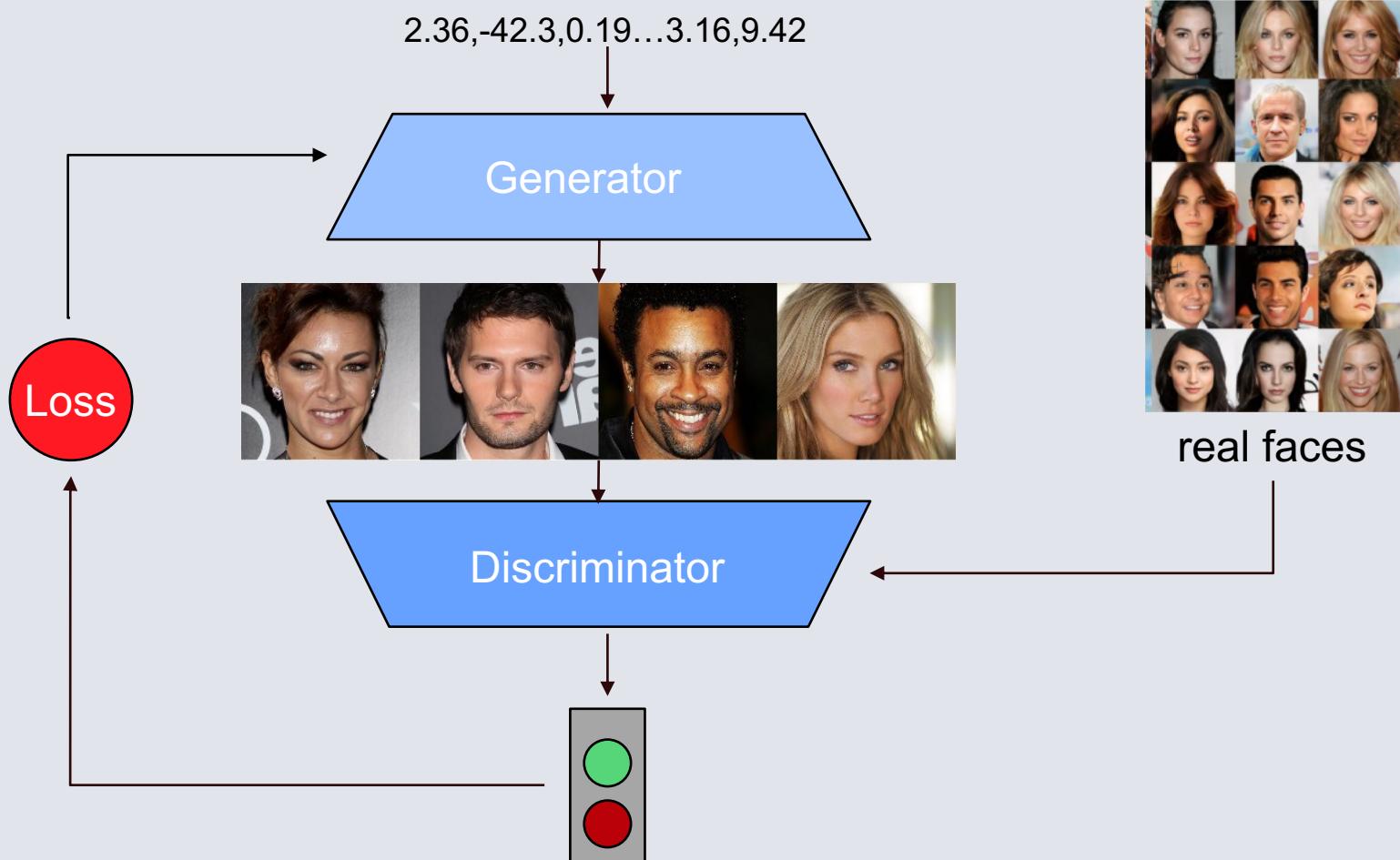
# Auto Encoder



# VARIATIONAL AUTO ENCODERS

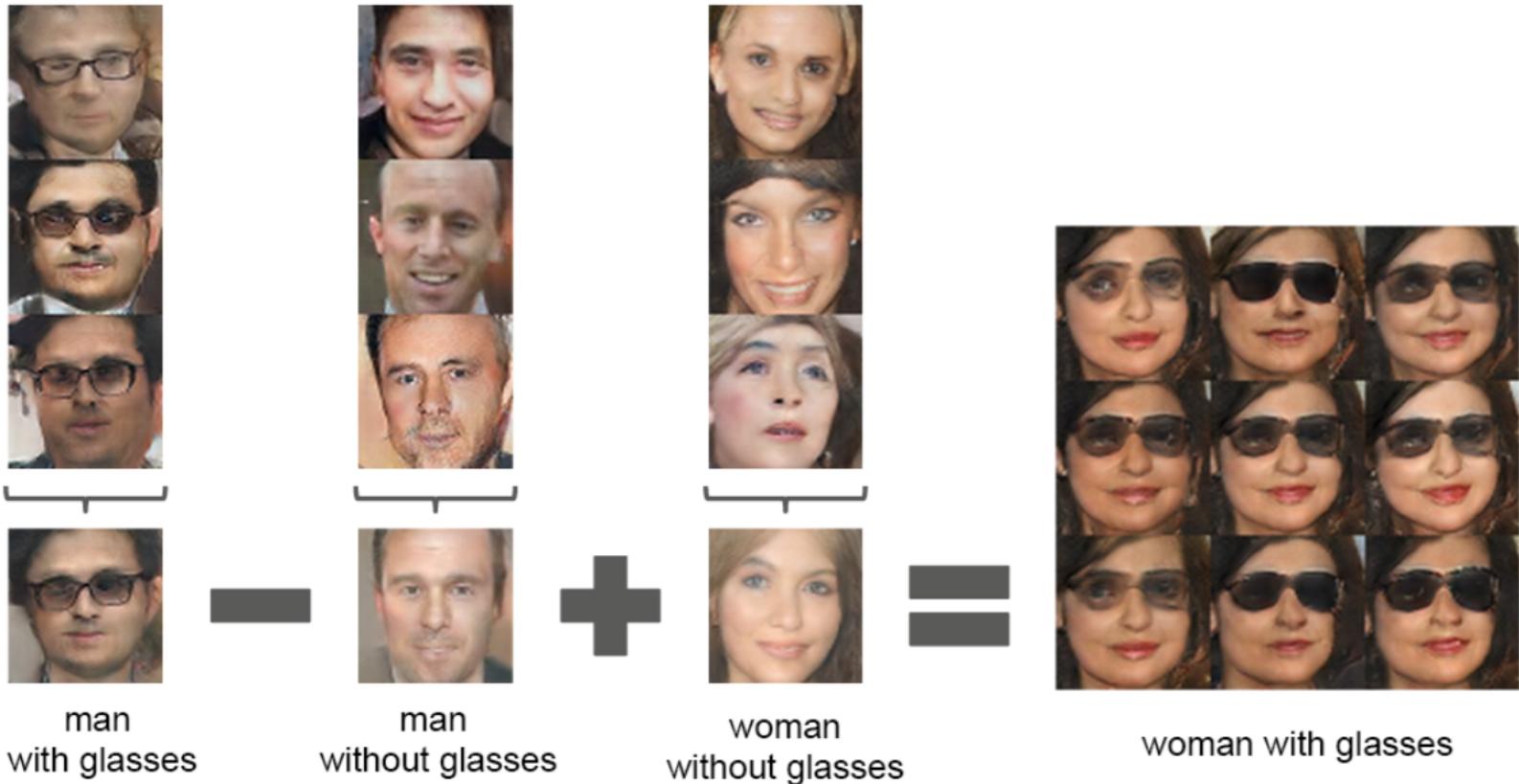


# Generative Adversarial Networks



# THE GENERATIVE WORLD!

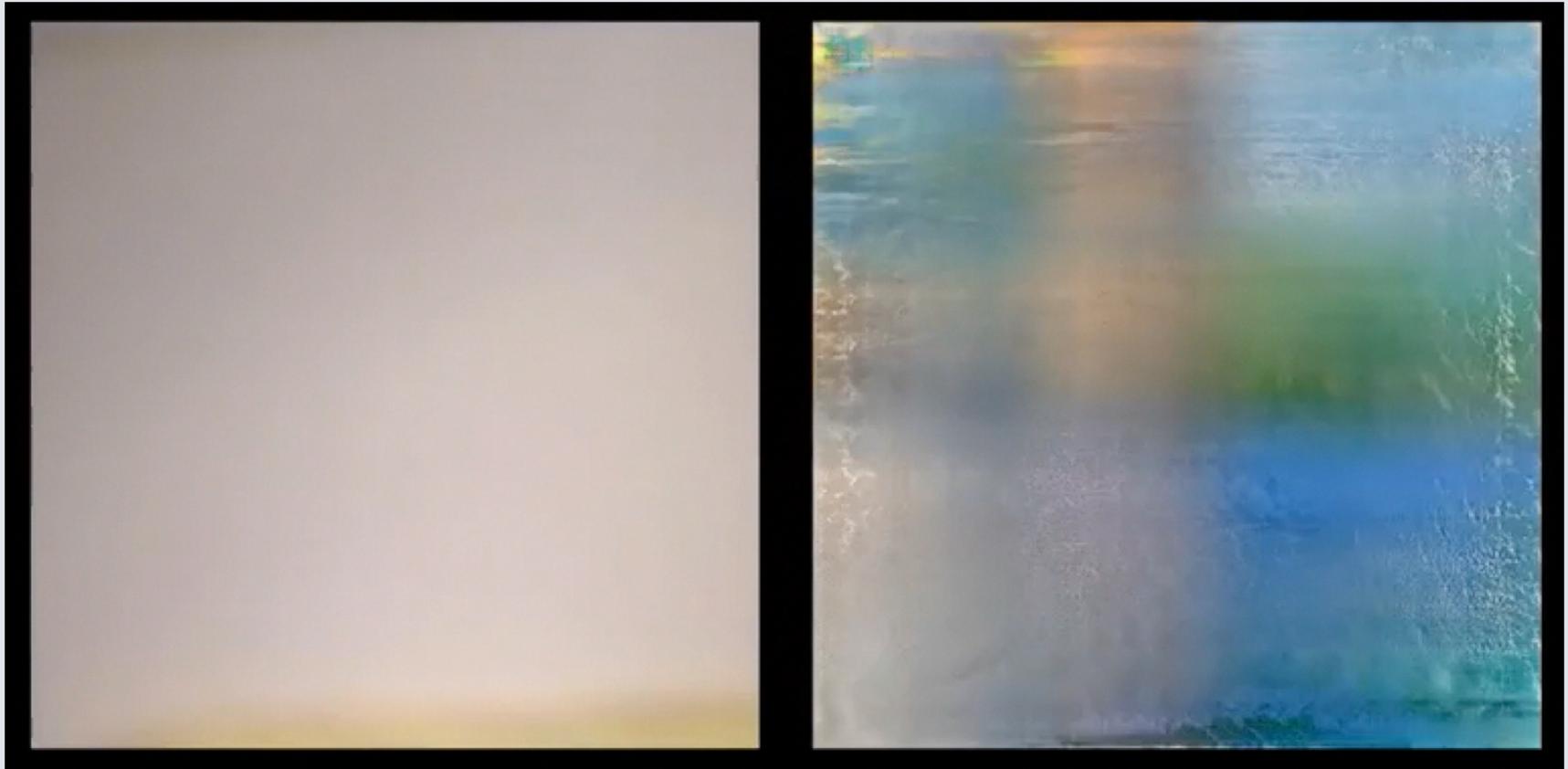
# Vector arithmetic



# Pix2Pix



## PIX TO PIX CANNY EDGE DETECTOR



# Face to Face



Real-time Reenactment



Reenactment Result



# Everybody dance now!

Source Subject



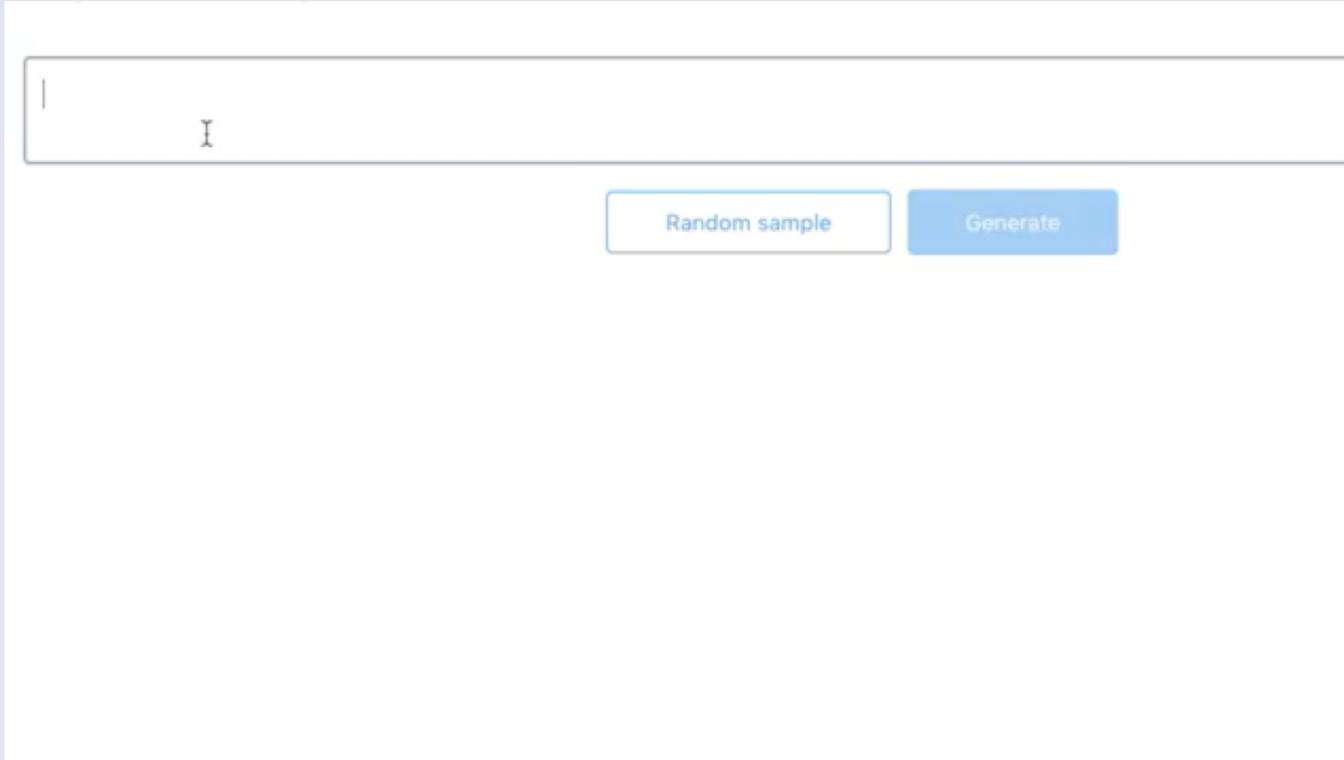
# Vid2Vid



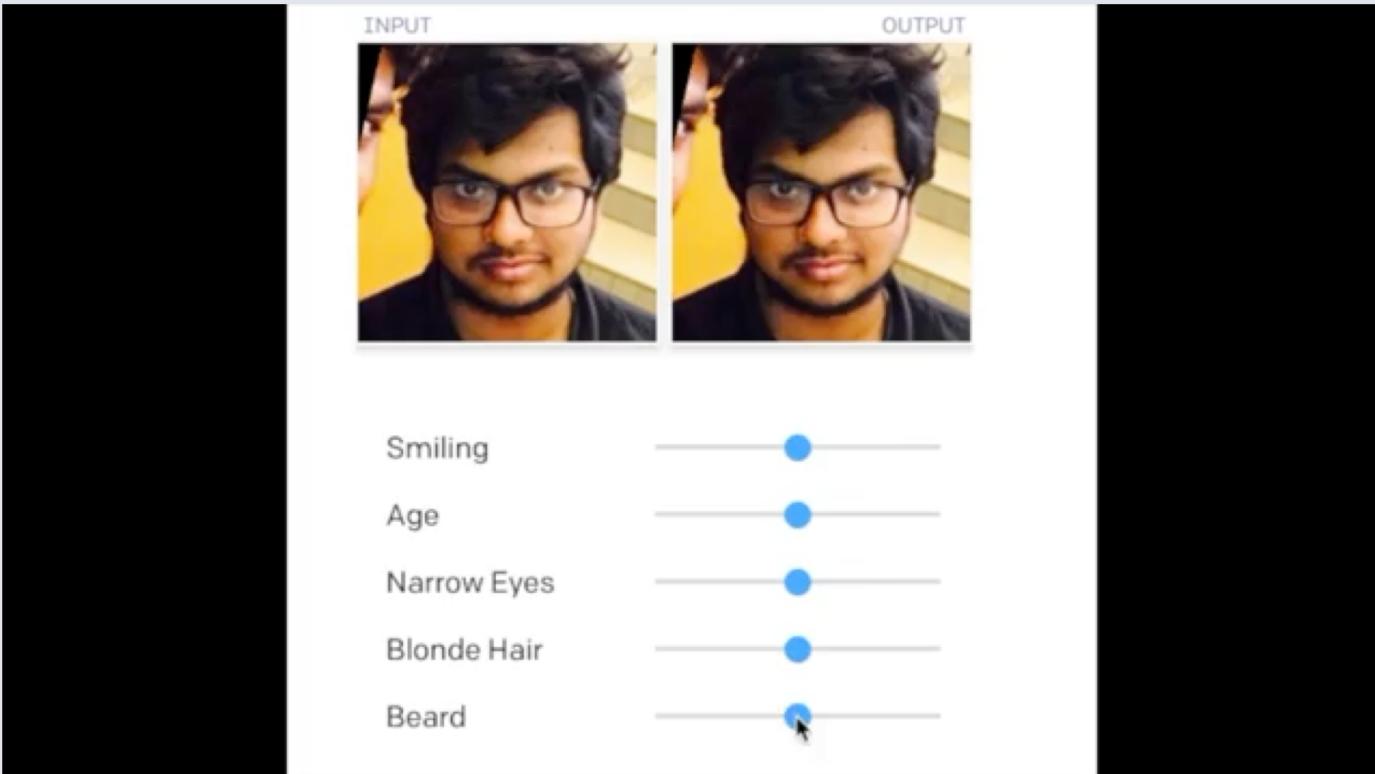
# CycleGan



# Voice Modelling



# Facial Attribute Manipulation



# 2D to 3D Facial Reconstruction

**3D Face Reconstruction from a Single Image**

Drag the 3D model around with your mouse. You can checkout the paper and code [here](#).

If it is blank, it may be an issue with WebGL or browser. Sometimes this happens with Chromium, but not Google Chrome, as far as I know.

Download Wavefront OBJ File  
(colours are stored per-vertex)

Try another image

Z Translate: ——○————

Show background image

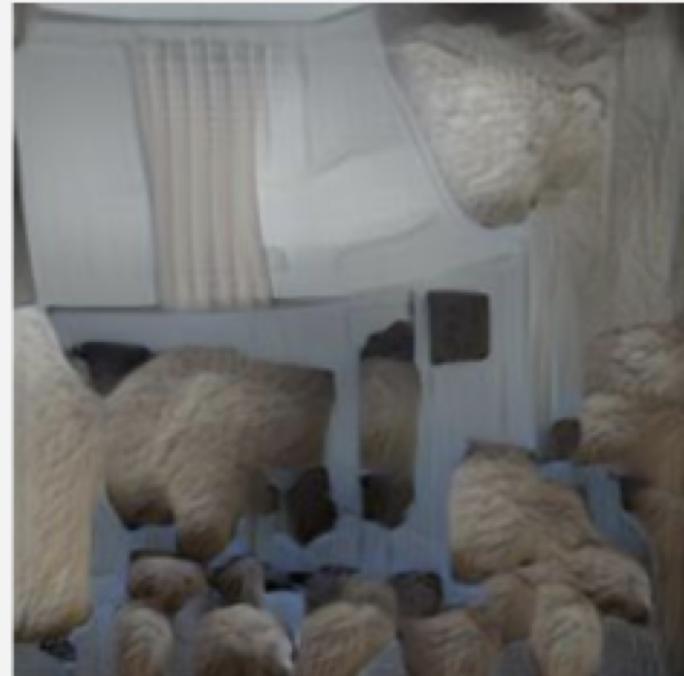
Please share and spread the word!

[Twitter](#) [Star](#) 3,163

aaron.jackson@nottingham.ac.uk ([website](#)) • WebGL rendering done with the Three.js framework.

# text to image generation

The sheep  
are in my  
bedroom  
jumping on  
the bed



# ProGAN Landscape generation



# Image to Sound



Press **esc** to exit full screen

[ 17355 ]  
17355

0:13 / 0:25

▶ 🔍 🔊 ⏪ ⏴

A painting depicting a lively street scene, likely from the 18th century. In the foreground, several figures in traditional European attire are engaged in what appears to be a dance or a social gathering. One man in a dark coat and white breeches is prominent, while others in colorful dresses and bonnets look on. The background shows buildings with thatched roofs under a bright sky. A small text box in the upper right corner of the image area says "Press esc to exit full screen". Below the image, the number "[ 17355 ]" is followed by "17355". At the bottom, there is a control bar with a play button, a timestamp "0:13 / 0:25", a magnifying glass icon, a volume icon, a left arrow, and a download icon.

# Progressive GANS



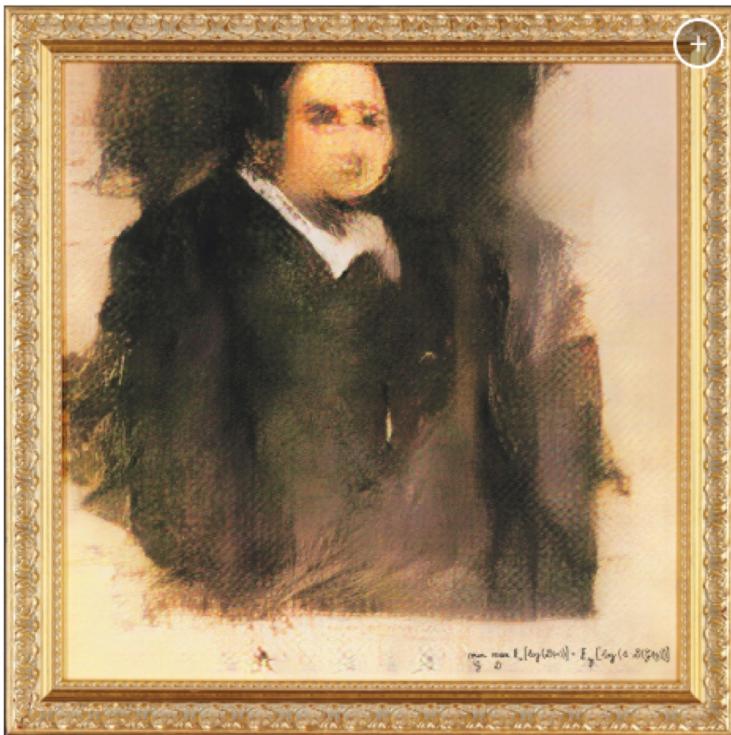
# Writing the script for an ad

LEXUS PRESENTS  
A FILM WRITTEN BY ARTIFICIAL INTELLIGENCE

# Wavenets- Generating Music



# AI GENERATED ART SOLD FOR \$432,500 (NEARLY 3CRORE INR)



*Portrait of Edmond Belamy*, 2018, created by GAN (Generative Adversarial Network). Sold for \$432,500  
on 25 October at Christie's in New York. Image © Obvious

# The Future: Neural nets and Reinforcement Learning



Boston Dynamics

[youtu.be/kHBcVlqpvZ8](https://youtu.be/kHBcVlqpvZ8)  
[youtu.be/M8YjvHYbZ9w](https://youtu.be/M8YjvHYbZ9w)

Search for : spot boston dynamics

# Checklist for Making a Movie

- ✓ WRITING THE SCRIPT
- ✓ CASTING
- ✓ CREATING THE VISUAL GRANDEUR
- ✓ GENERATING MUSIC

Any Questions?

# Any Questions?

If AI actually makes a Blockbuster movie,  
who would rightfully own the Profits?



Free GPU Access

Slides

Krishna Balaga  
krbalaga@in.ibm.com

@Krishnac71

# Thank you.

Krishn Balaga

Developer Advocate

—

Krbalaga@in.ibm.com

Github.com/krishnac7

ibm.com

