

# Machine Learning



**RAJESH SHARMA**

Walt Disney Animation Studios



Thank you to ACM SIGGRAPH!



**Pol Jeremias-Vila** : SIGGRAPH 2021 Chair

**Tomasz Bednarz**: Frontiers Program Chair

**Alex Bryant**: Student Volunteers Chair

**Tim Hendrickson**: Digital Marketing Manager

Student Volunteers:

**Rogelio, Trinity, Aurora, Emily, Hunter & Kendra**



**SIGGRAPH 2021**

# Machine Learning

————— Rajesh Sharma —————

# Ashish Vaswani



## Research Scientist Google

**Ashish Vaswani** is a Senior Research Scientist in the [Brain group](#) within GoogleAI, where he works on machine learning with neural networks. His research has focussed on developing pure attention based models, such as the [Transformer](#) for generation and classification. Before joining Google, he was a PhD student, and later Research Scientist, in natural language processing at the University of Southern California Information Sciences Institute.

## Research Scientist

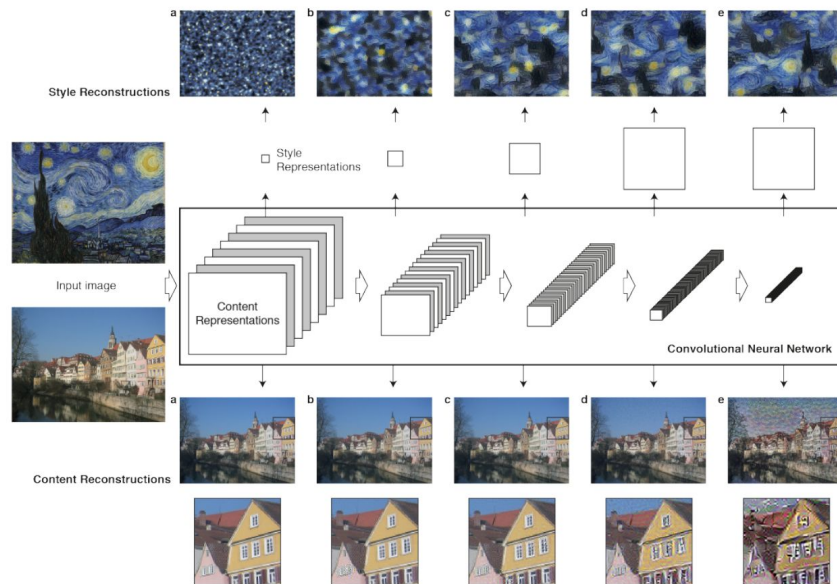
# Today

- Quick Recap: GAN, Facial Detection
- HW: CNN, GAN
- Variational Autoencoder
- Face Detection
- RNN, Transformers

# Hands-on

- ★ Log in to your google drive
- ★ Make a shortcut to: `https://bit.ly/3oKCVCh`
- ★ Make a copy of:
  - `Autoencoder.ipynb`
  - `dataPipeline.ipynb`
  - `denoiserCNN.ipynb`
  - `styleTransfer.ipynb`
  - `facialRecognition02.ipynb`
  - `TextRNN.ipynb`

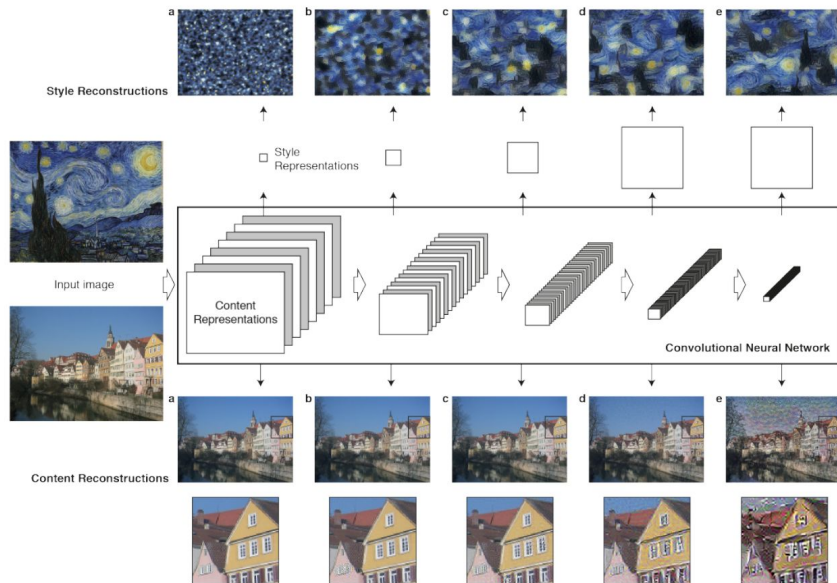
# Artistic Style Transfer



$$L_{\text{total}}(\check{c}, \check{s}, \check{x}) = \alpha L_{\text{content}}(\check{c}, \check{x}) + \beta L_{\text{style}}(\check{s}, \check{x})$$

# Artistic Style Transfer

styleTransfer.ipynb



$$L_{\text{total}}(\check{c}, \check{s}, \check{x}) = \alpha L_{\text{content}}(\check{c}, \check{x}) + \beta L_{\text{style}}(\check{s}, \check{x})$$



# Transfer Learning

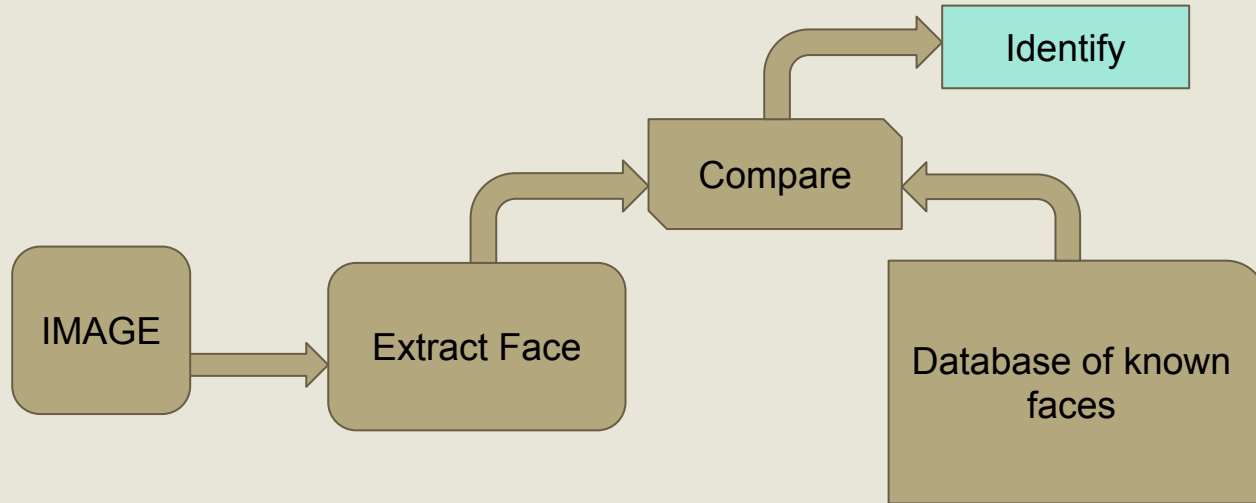
## Build a Facial Recognition System



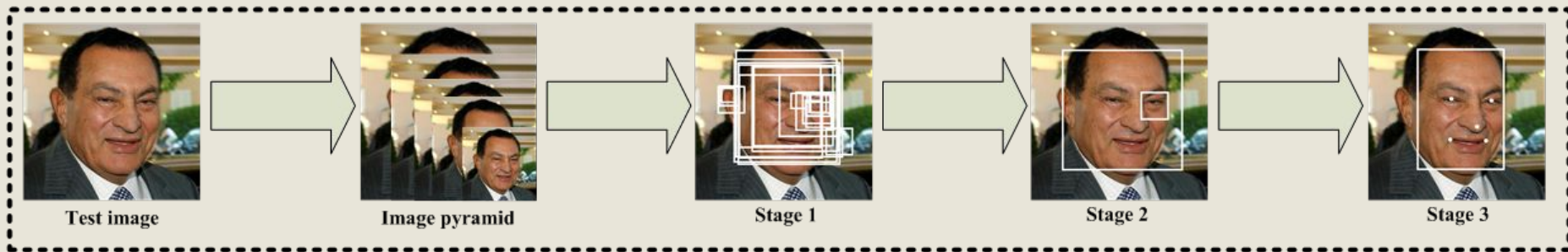
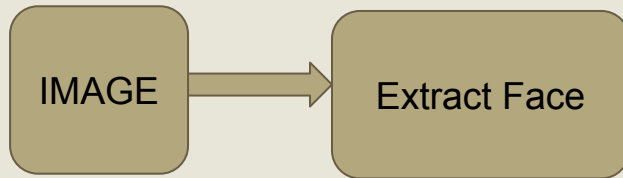
*Delta News Hub / CC BY (<https://creativecommons.org/licenses/by/2.0>)*

# Transfer Learning

## Build a Facial Recognition System



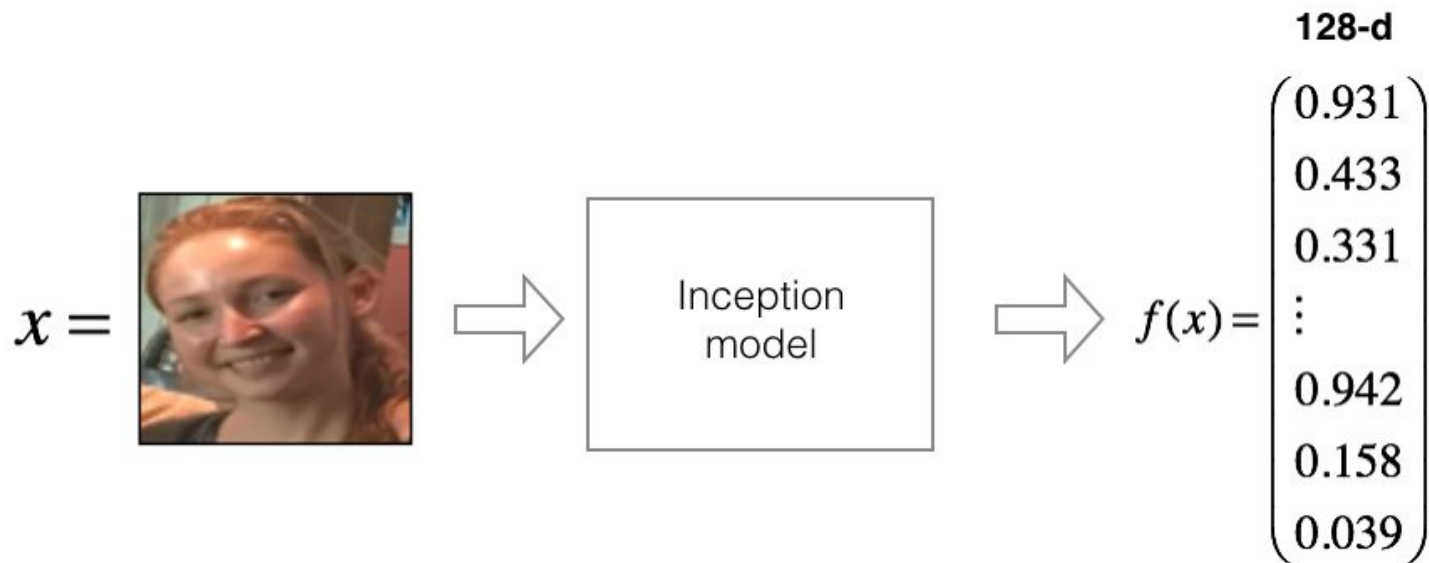
# Extracting Faces -- MT-CNN



# End to End System - Transfer Learning

1. Get faces from ground-truth images (MTCNN)
2. Encode ground-truth images (FACENET)
3. Read Camera Input Stream
4. Isolate Faces (MTCNN)
5. Encode input face (FACENET)
6. Compare encoding with stored ground-truth
7. Identify person

# Facenet - triplet loss: [Paper](#)



$$Loss = \sum_{i=1}^N \left[ \|f_i^a - f_i^p\|_2^2 - \|f_i^a - f_i^n\|_2^2 + \alpha \right]_+$$

# Hands-on

★ ***facialRecognition02.ipynb***

# Homework:

Colorization: `tf.image.adjust_saturation`

## Up-resing:

```
tf.image.resize(image, size=[256,256], method=tf.image.ResizeMethod.NEAREST_NEIGHBOR)
```

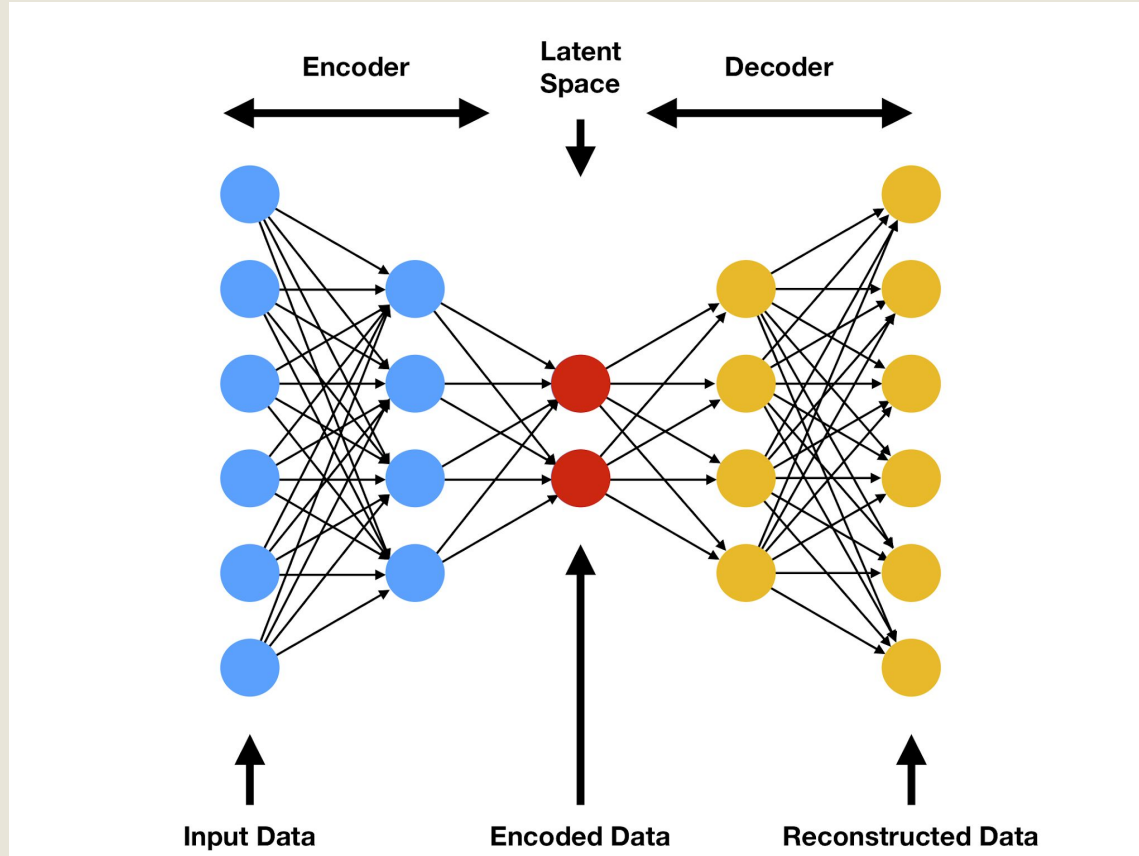
## In-Painting:

```
mask = np.ones((PATCH_WIDTH, PATCH_HEIGHT), dtype=np.float32)
scale = 0.25
low, upper = int(PATCH_WIDTH * scale), int(PATCH_HEIGHT * (1.0 - scale))
mask[:, low:upper, low:upper] = 0.
tf.multiply(patch, mask)
```

## Frame interpolation:

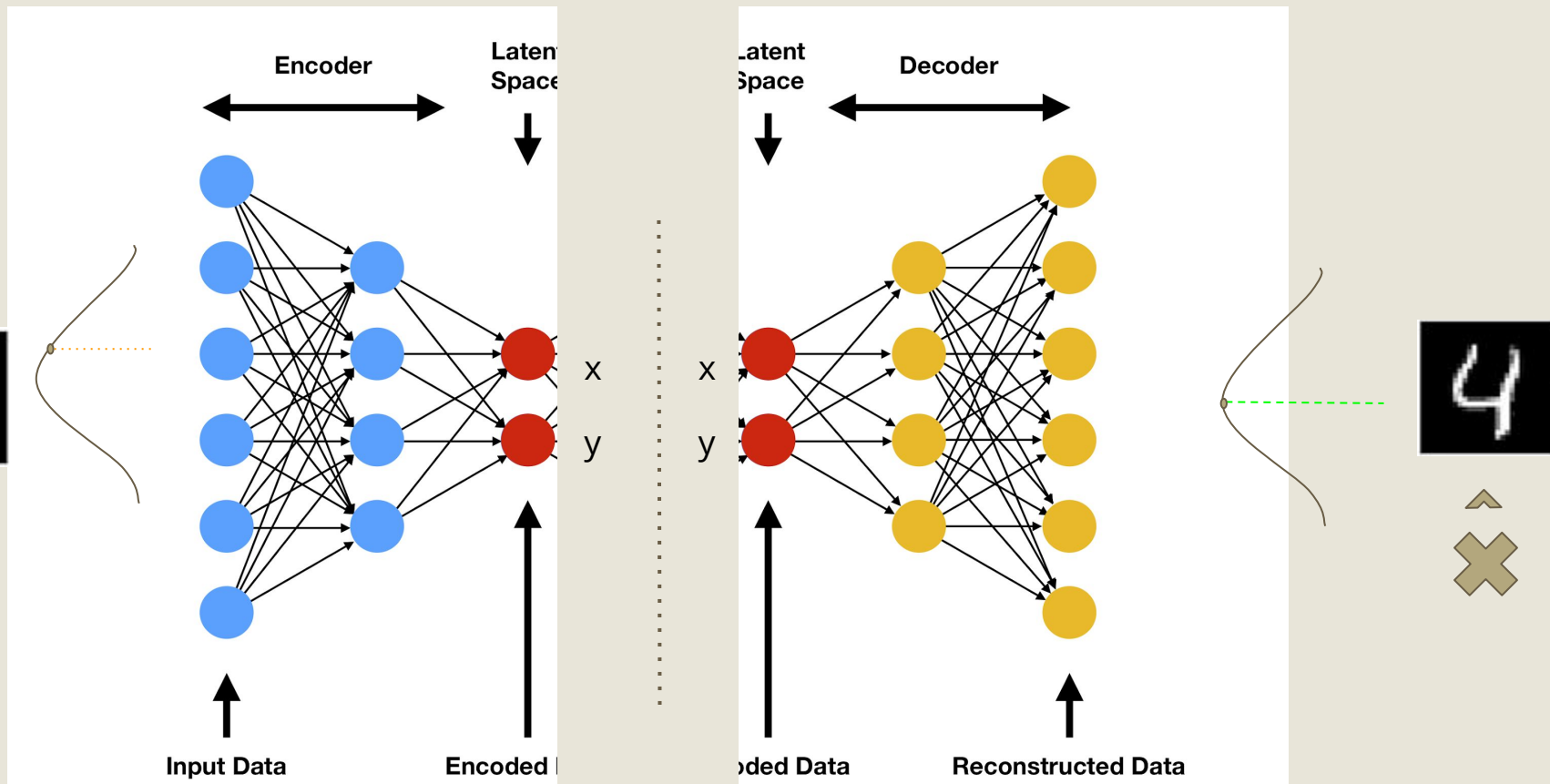
```
stacked = tf.concat([frame1, frame3], axis=-1)
```

# Autoencoder

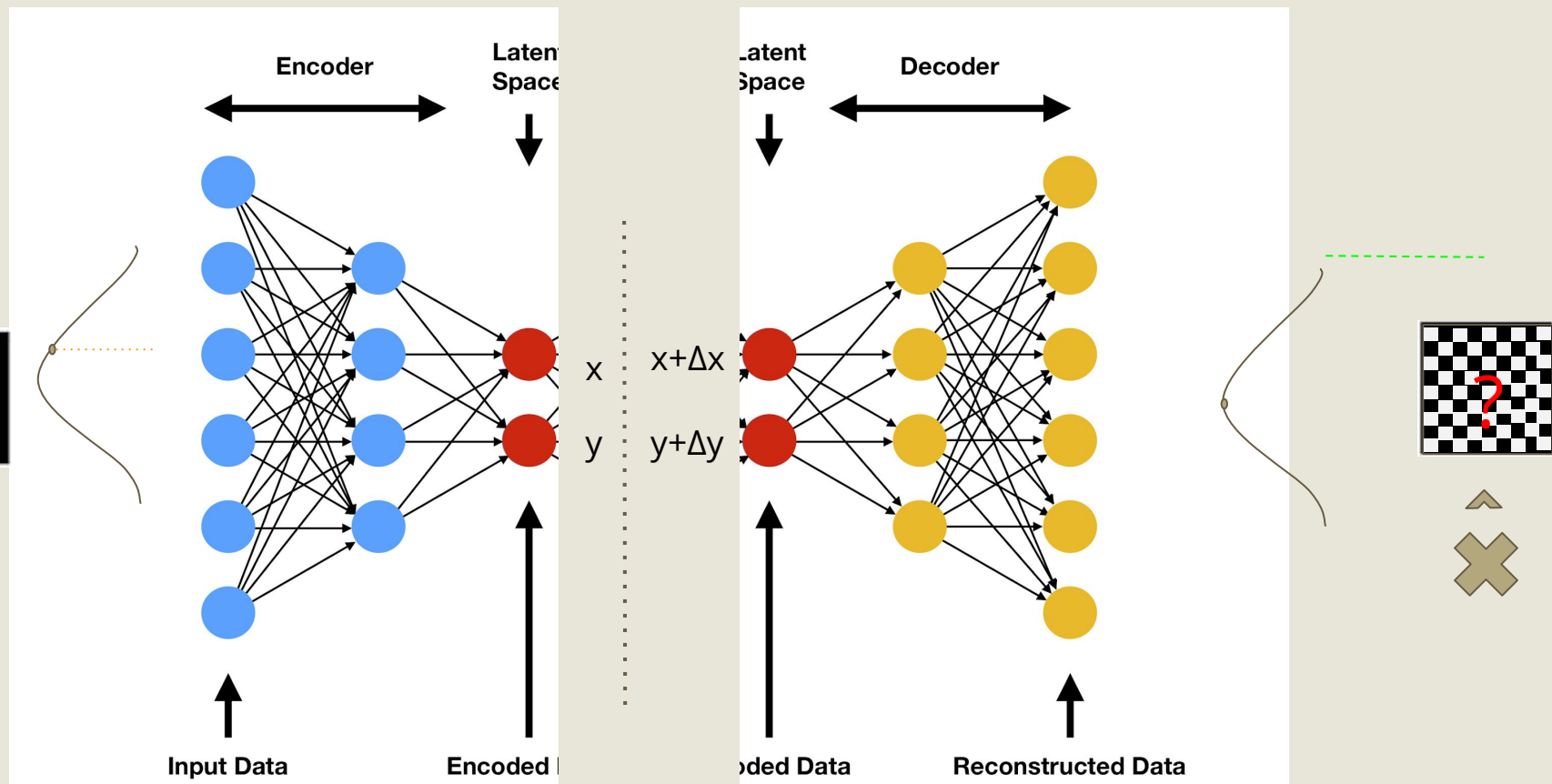




# Autoencoder - break it up after training



# Autoencoder - A variation



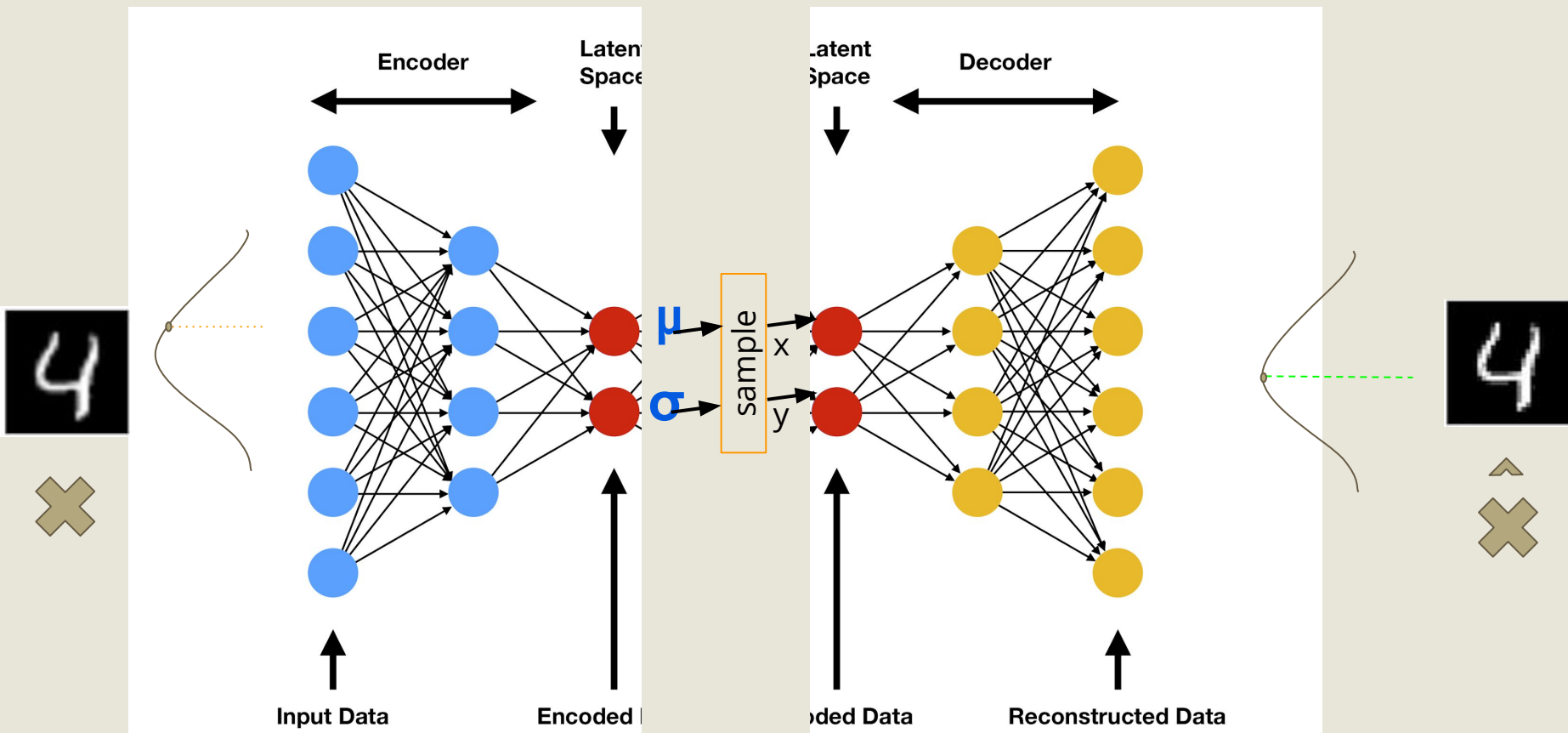
# You don't because...

The latent space and the input distributions  
are *different*!

But there is a way:

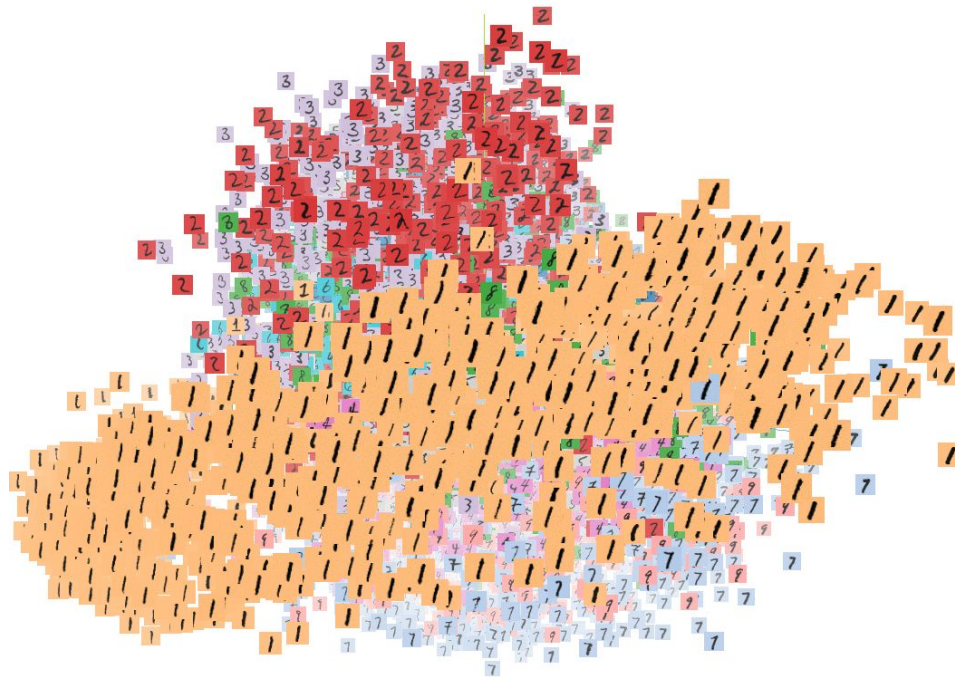
Treat encoder output as  $\mu$  and  $\sigma$  of a distribution

# Variational Autoencoder

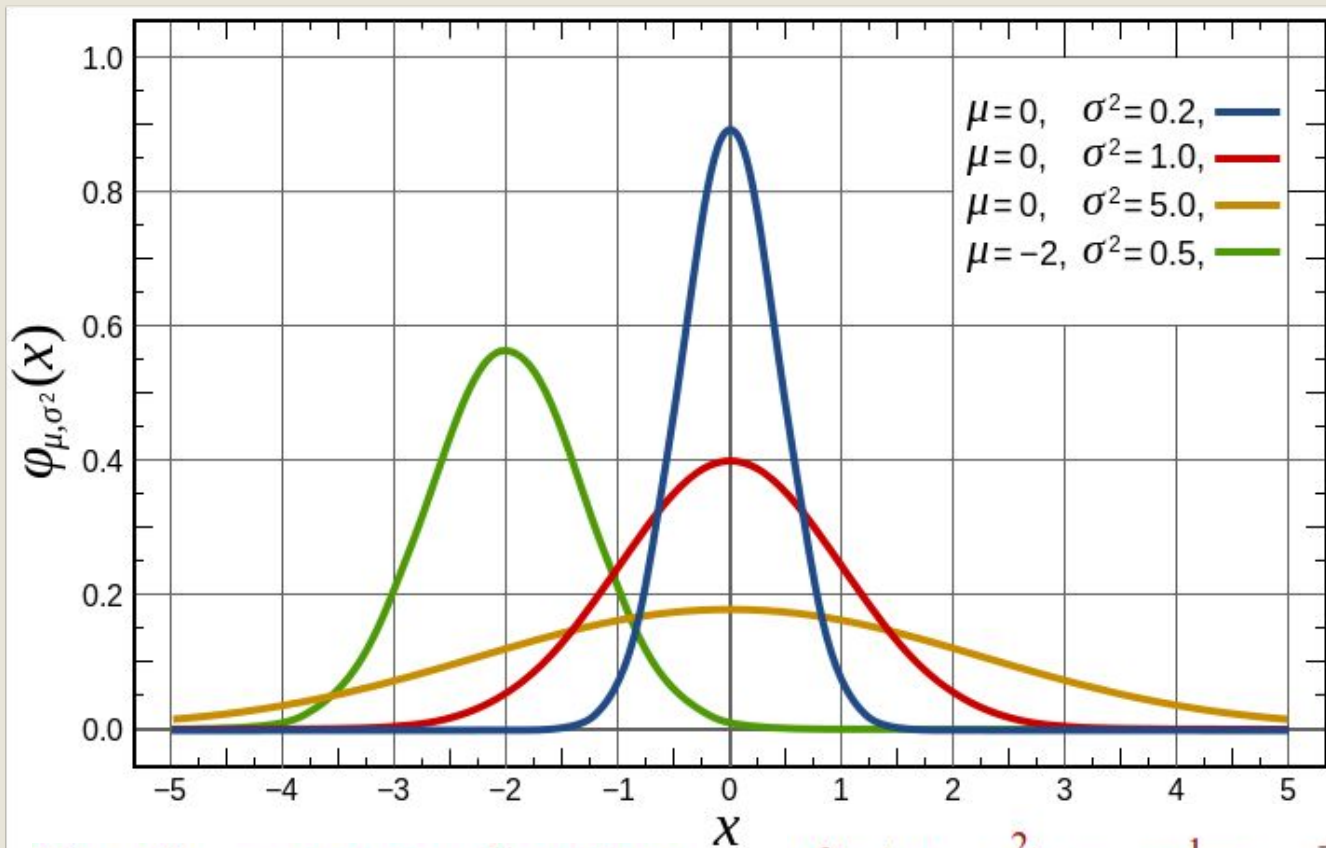


# Latent Spaces and Embeddings

<https://projector.tensorflow.org>



# You get nice continuous distribution for each input

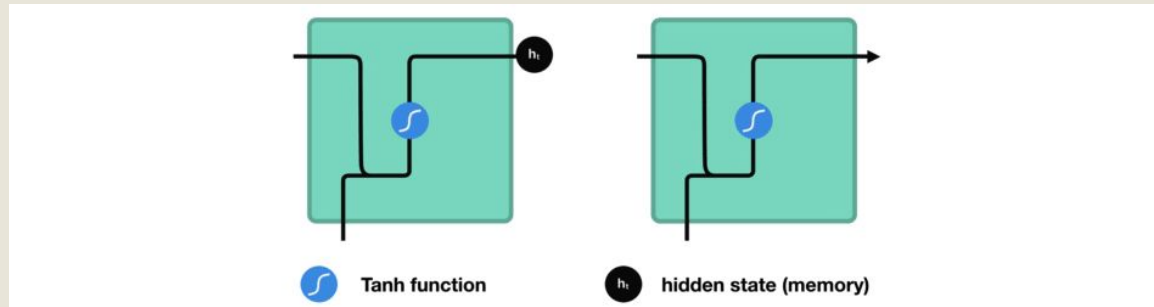
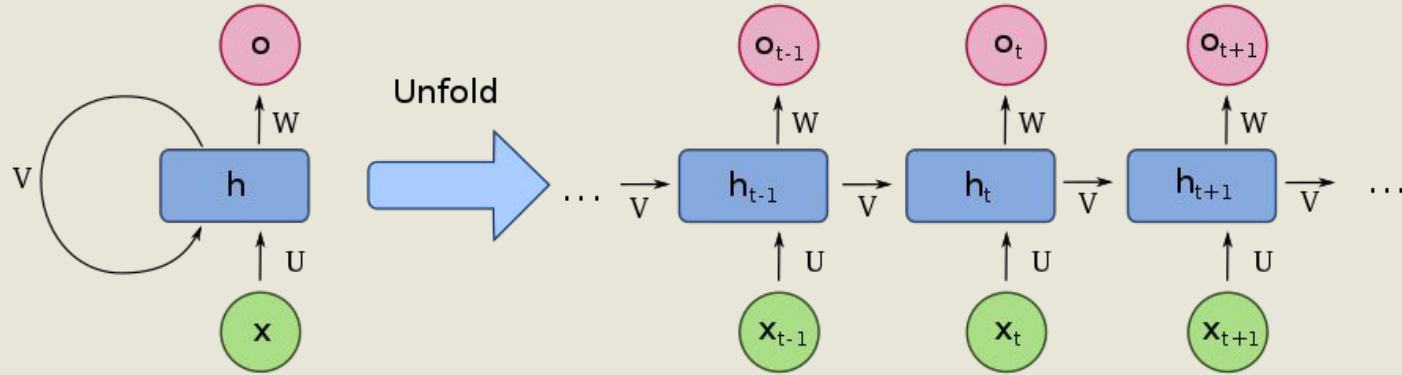


$N(\mu, \sigma^2)$  ;  $\mu$  - mean ,  $\sigma^2$  - variance

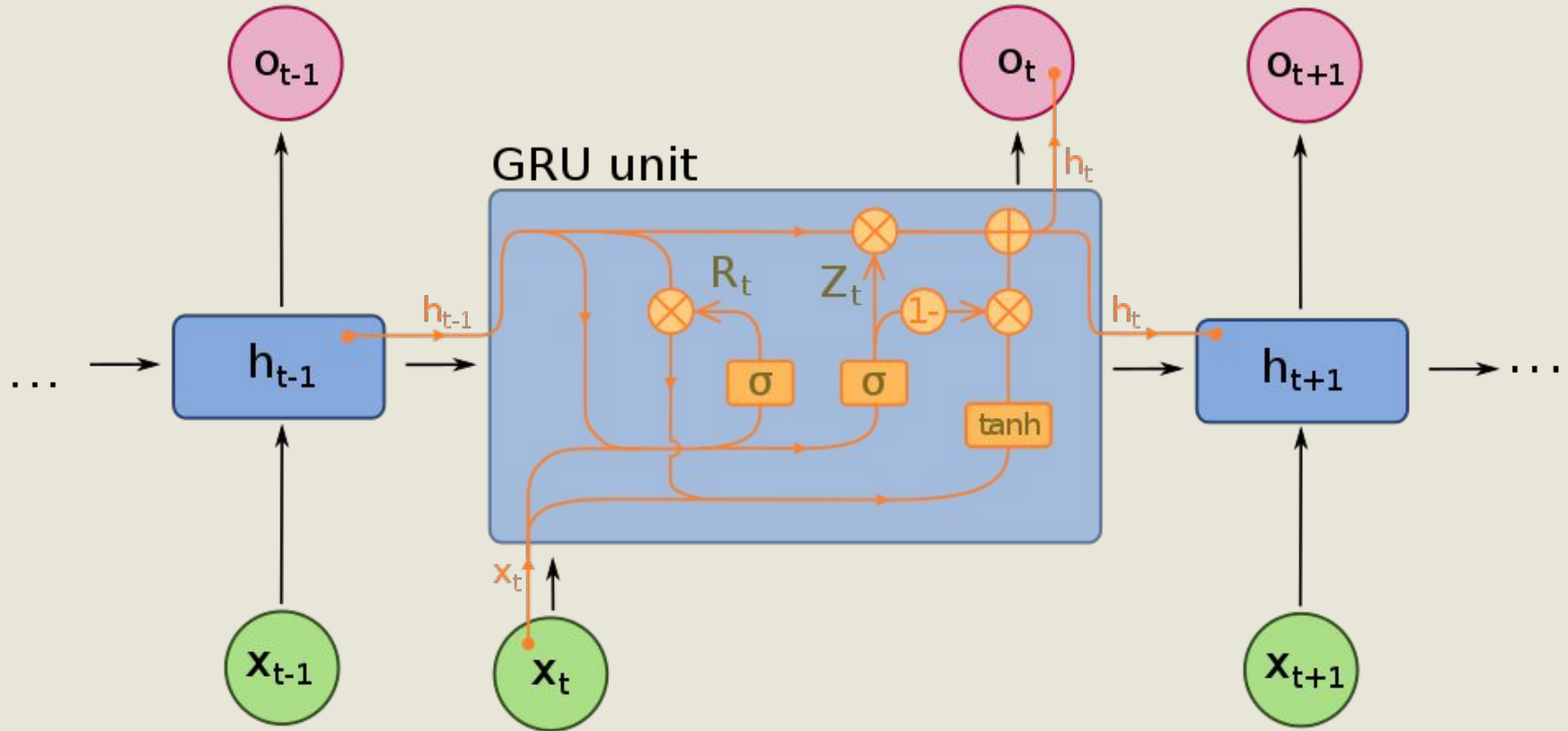
$$f(x | \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

# What are RNNs?

Information from previous timestep is passed forward



# GRU - RNN with memory/state





# Hands on...Text Generation

Given string of characters, what is the most probable next character?

Example:

Input:           Machin

Output:         achine

- Pick a sequence length
- Divide text into sequences
- Train on (sequence, sequence+1) pairs over the entire text
- Important: Convert all text to numbers first!

# Hands on...

Find and open:  
TextRNN.ipynb

# Next Class

- Guest: Andrew Glassner
- Reinforcement Learning
- Wrap up and where to go from here
- Homework:
- RNN: auto generated tweets
- Facial Recog: Closest match to a celebrity
- @xarmalarma, #siggraph2021