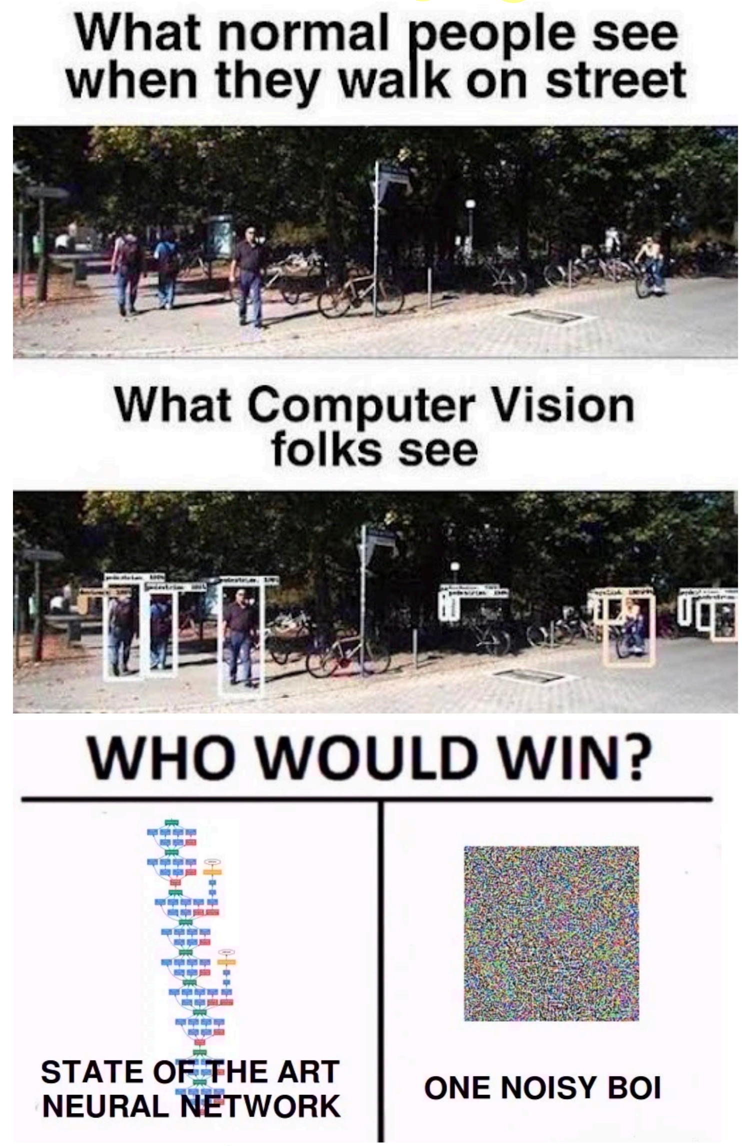
April 14, 2023

DSML: Computer Vision.

Generative models and GANs.

Class starts @ 9:05 pm.



Recap:

$$\begin{array}{cccc} 0.001 & 0.009 & 0.9 & 0.09 \\ 0 & 1 & 2 & 3 \end{array}$$

Image Classification

$f(x)$ → Dense → 0 0 0 } Classification output.

CNN. Image

outputs an embedding.

Object Detection

CNN → $\hat{c}$, (x_s, y_s, x_e, y_e)

Image Generation.

z noise vector → CNN-G → generator network.

Object Segmentation

CNN → $\hat{c}$, (x_s, y_s, x_e, y_e) → CNN → class segmentation masks.

Image $\rightarrow$ $f(\hat{x})$ $\rightarrow$ Embedding $\rightarrow$ Flatten

$\downarrow$

$[\quad]$

noise.

to

generator network.

z $\rightarrow$ CNN-G $\rightarrow$

z

noise vector

Image Generation is like trying to find an inverse for $f(x)$.

Discriminative vs. **Generative models**.

A model / algrithim which performs Image Classification.

$\{X, Y\}$ → Labelled Dataset.

x_i y_i

$$P(y \mid x_i)$$

what is learnt? Decision Boundary.

CNN, Decision trees, SVMs. Regression etc.

Tries to find the reverse mapping.

$$P(x \mid y).$$

Underlying probability distribution of the Image space.

Intuition behind training a generative model.

Think about counterfiet currency notes.

Natural Images.

Generato → Fake Currency notes.

GAN.

Generative Adversarial network.

Enforcer/Discriminator → "Fake" / "Real".

(Special Machine)

1 CNN for Generation.

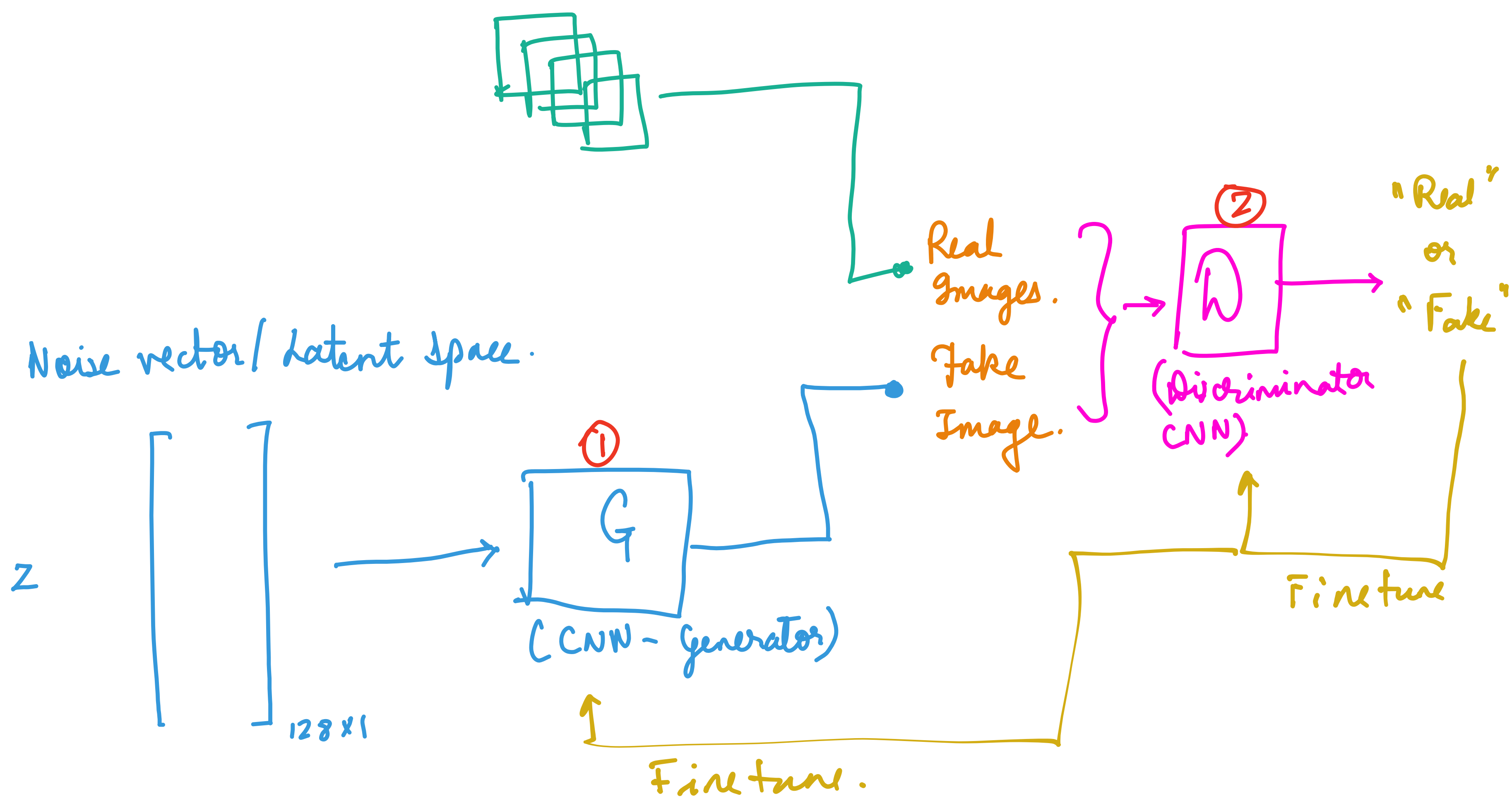
Noise → Image

→ "Generator network"

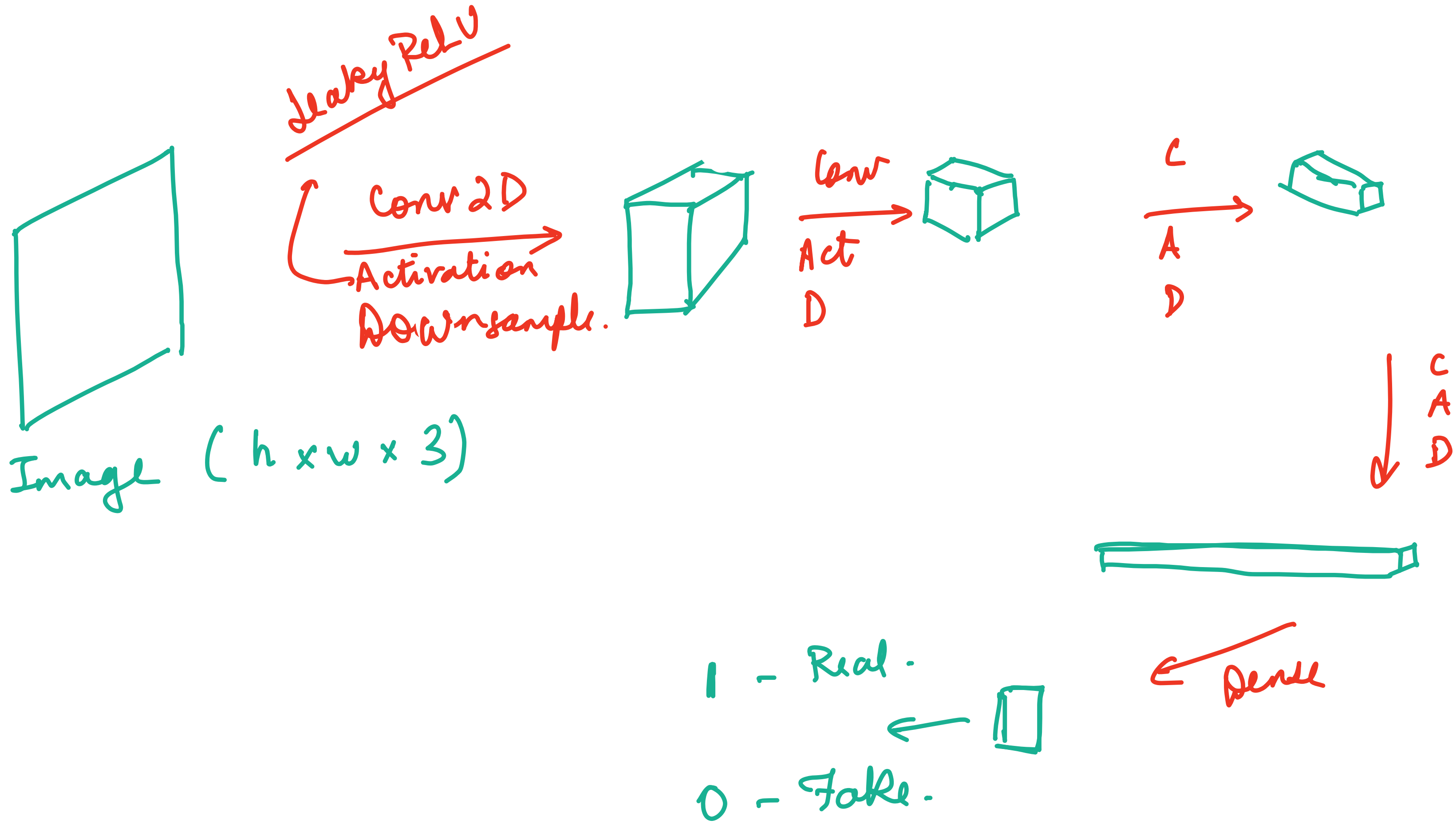
1 CNN for Discrimination.

→ "Discriminator network."

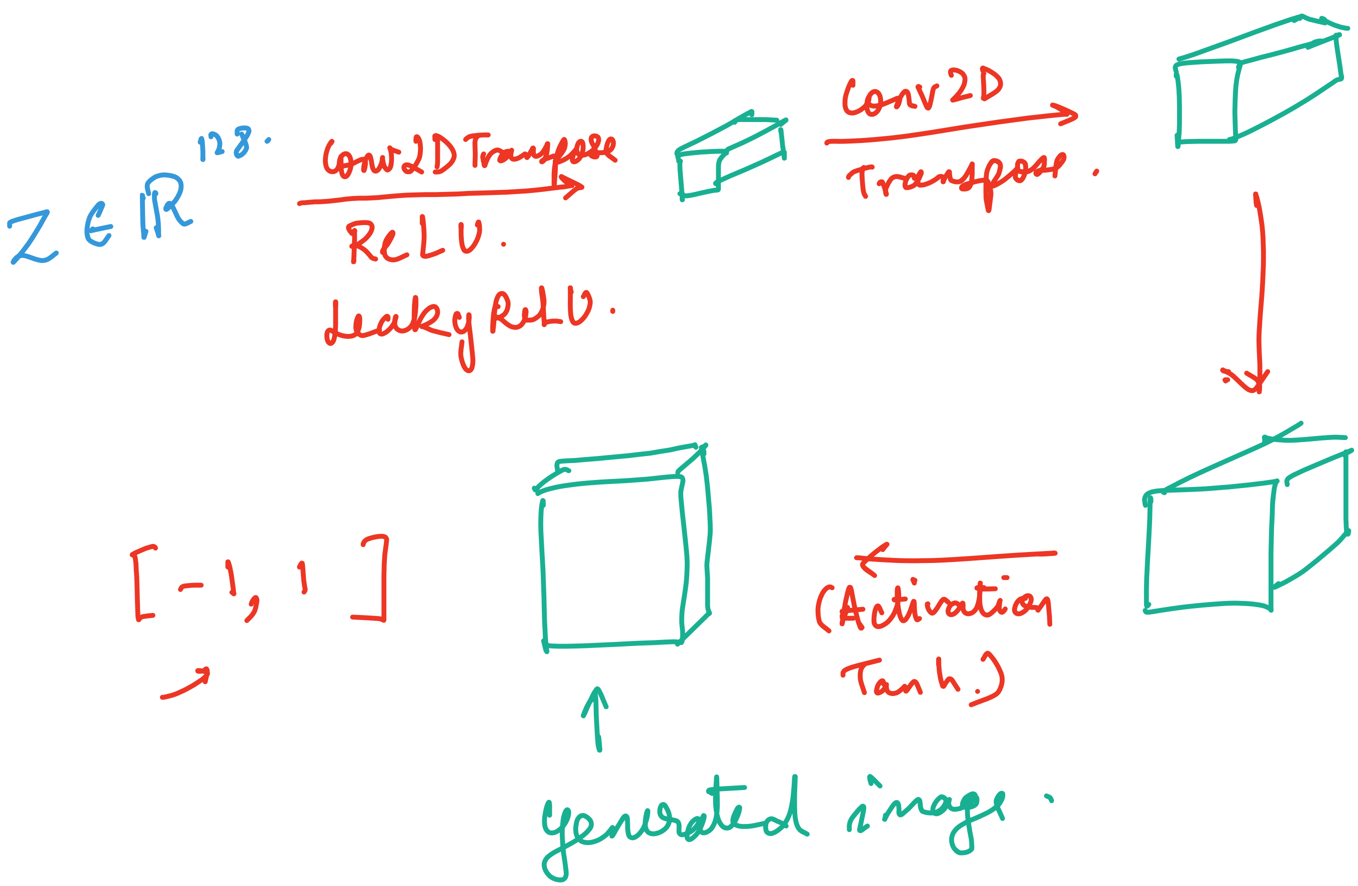
Architecture of a generative adversarial network.



Discriminator Architecture.



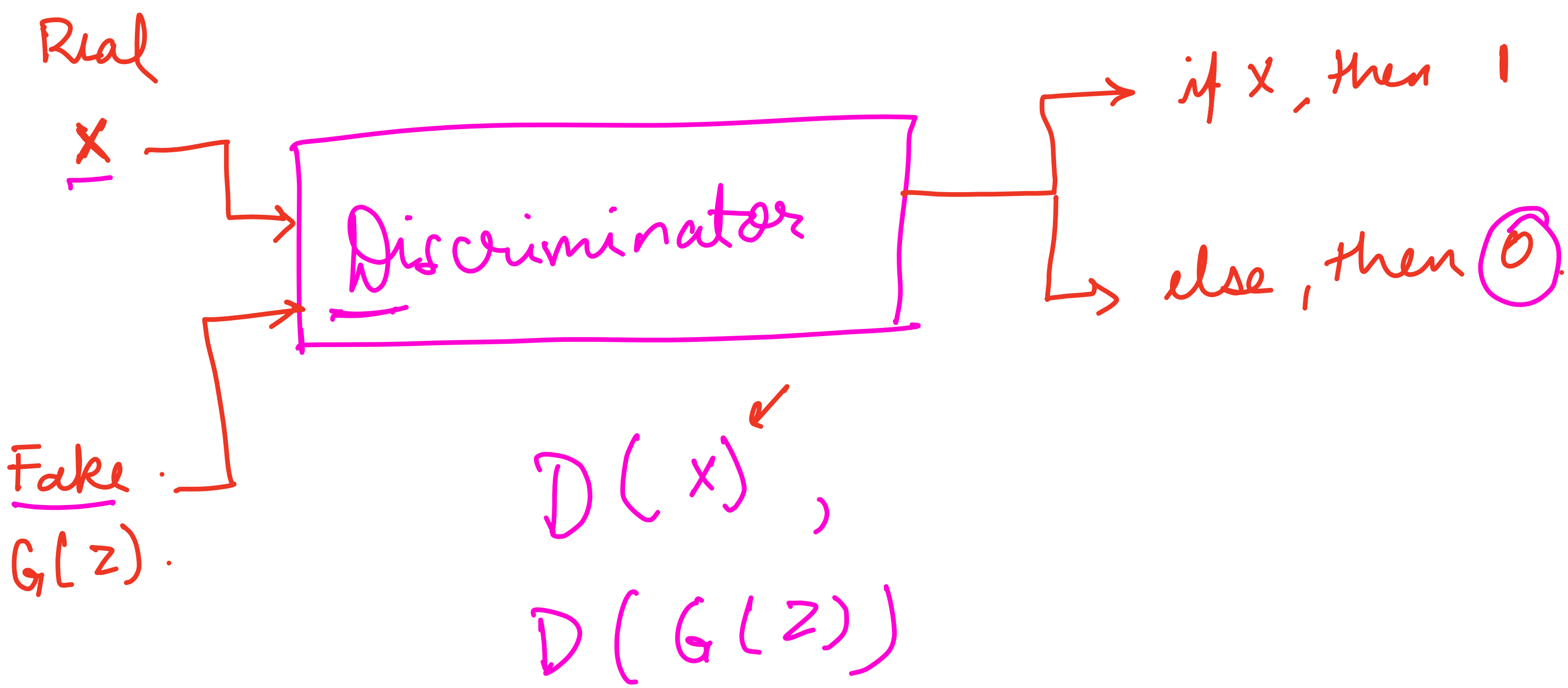
Generator Architecture.

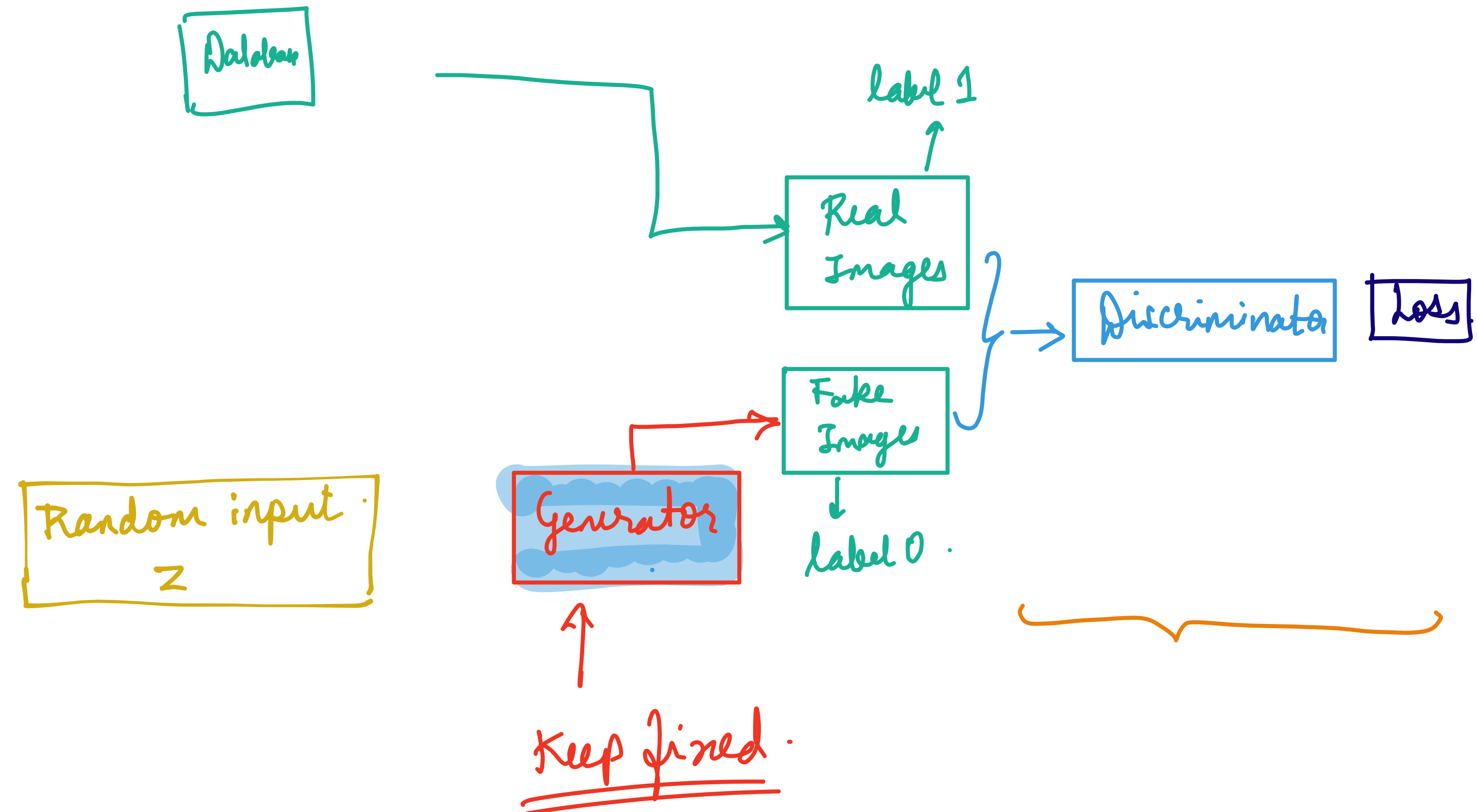


Training a GAN: The GAN Loss function.

Binary Cross-Entropy: $N \rightarrow$ size of dataset.

$$-\frac{1}{N}\sum_{i=1}^{N}\left[y_i \cdot \log\left(D(x)\right) + (1-y_i)\log\left(1-D(G(z))\right)\right]$$

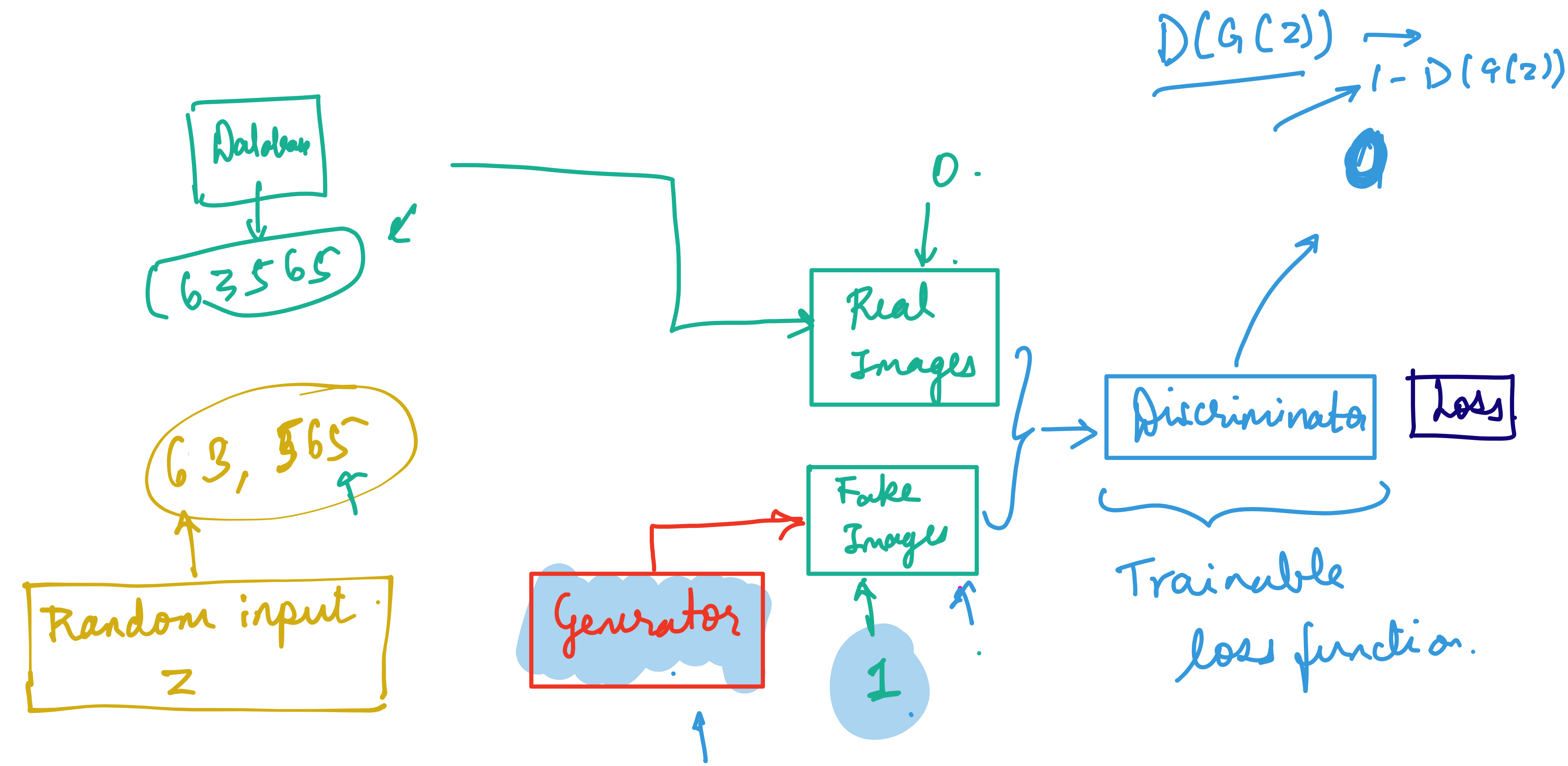




If the input is real, loss : $\max\limits_{D} \log(D(x))$

If the input is fake, loss : $\max\limits_{D} \log(1 - D(G(z)))$

Overall :



$$\min_{G} L\left(D(G(z)), 1 \right) = 1 \cdot \log\left(1 - D(G(z))\right)$$

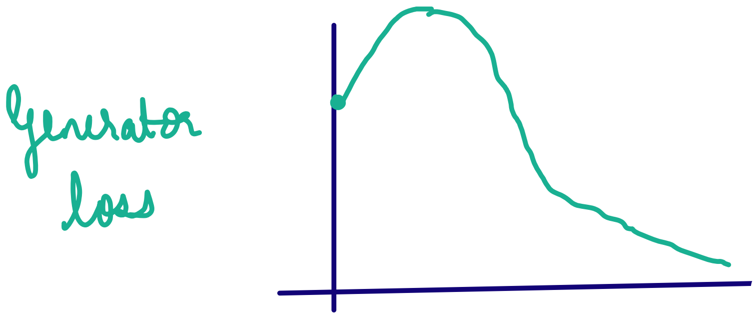
$$(1-1) \cdot \log\left(D(G(z)) \right)$$

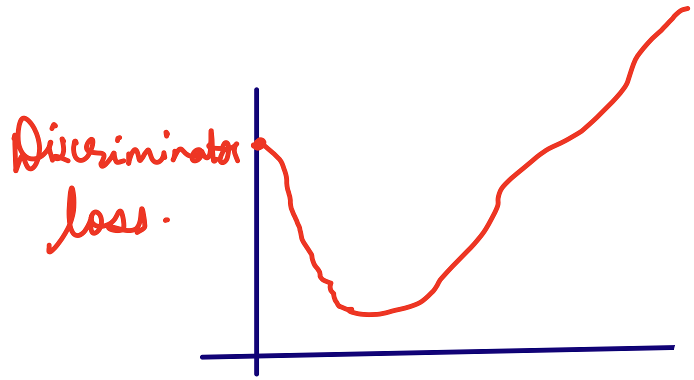
$$\min_{G} \left\{ \log\left(1 - D(G(z))\right) \right\}$$

Fixed mapping

$$\begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$$

Good GAN training





Bad GAN training.