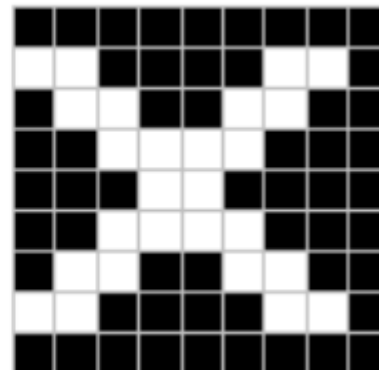
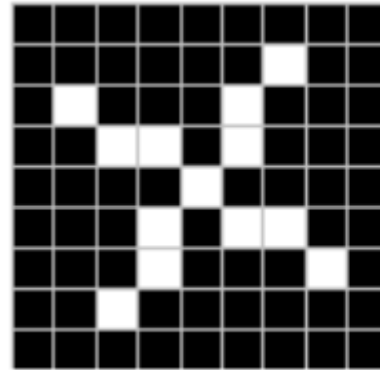
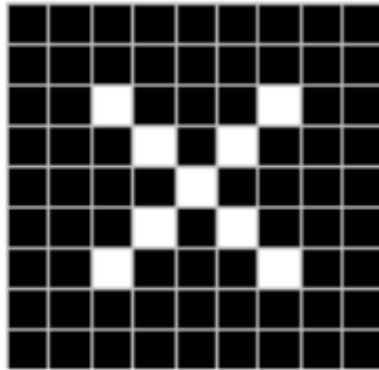
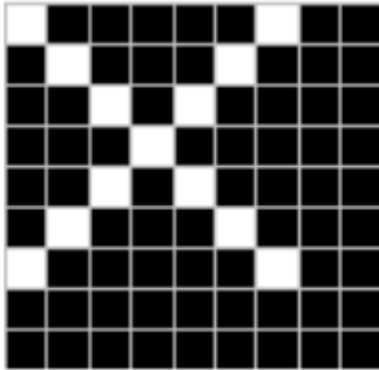
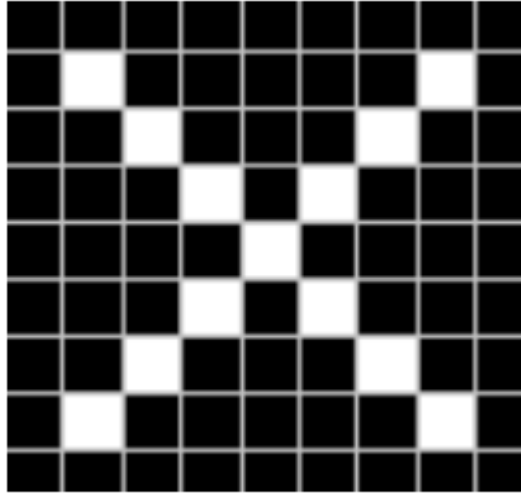


Discussion 12

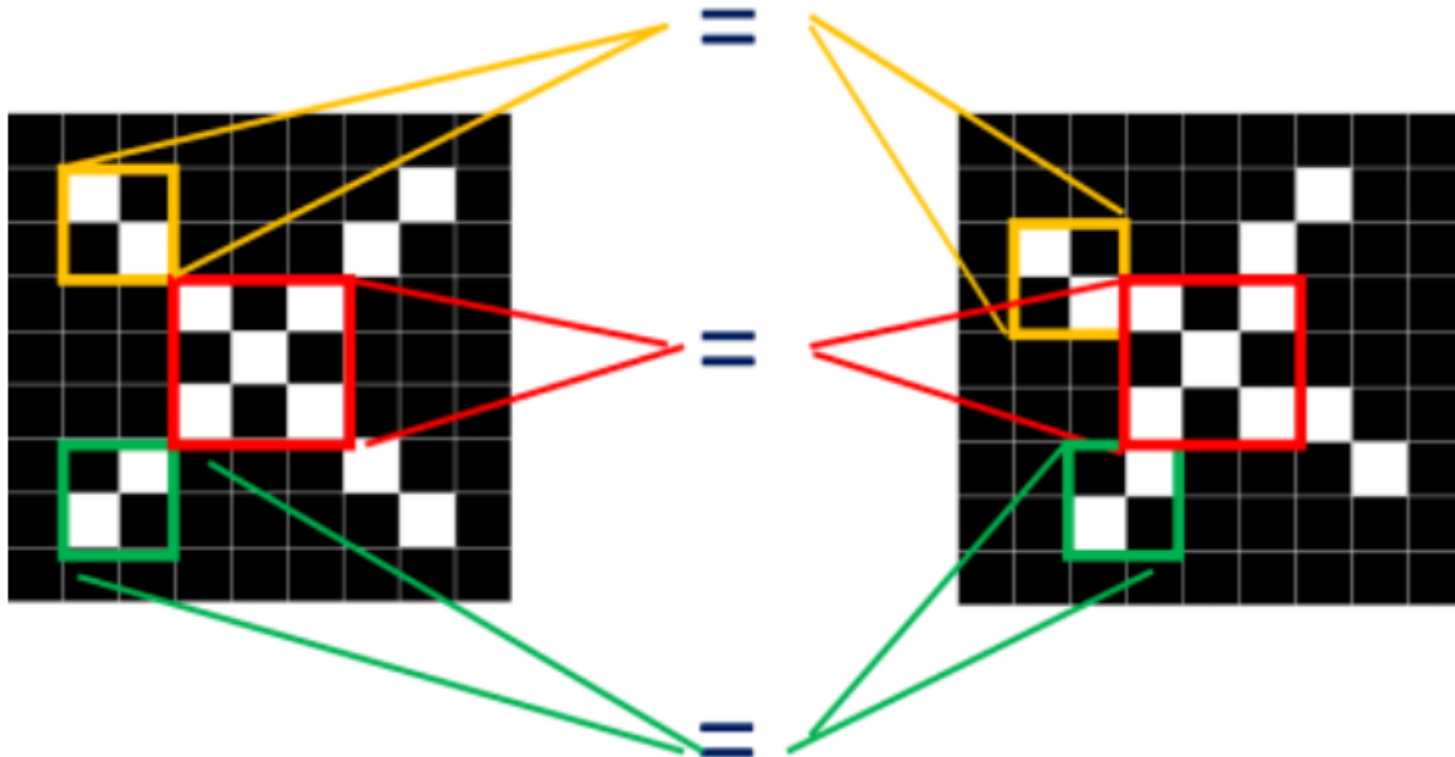
2022.06.09

- CNN
- RNN
- LSTM
- Autoencoder
- VAE
- GAN

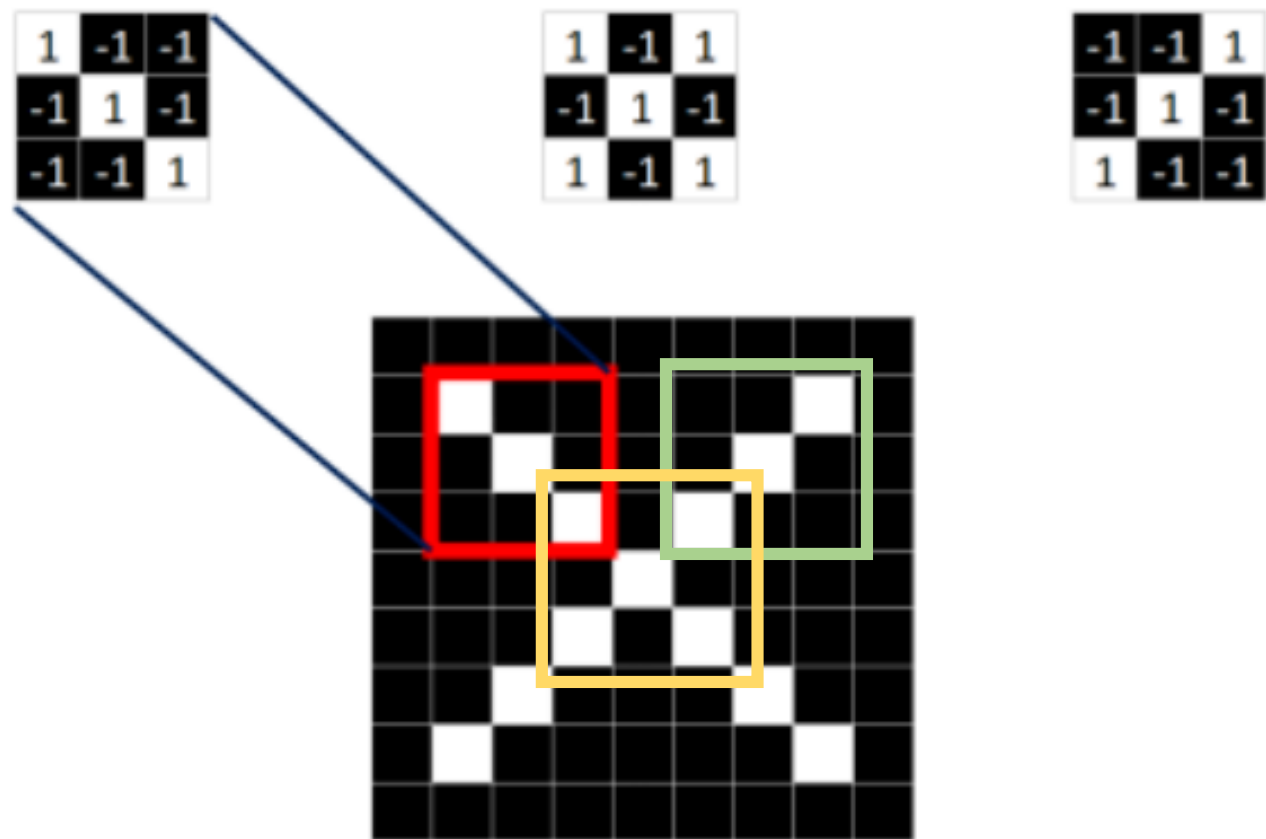
CNN



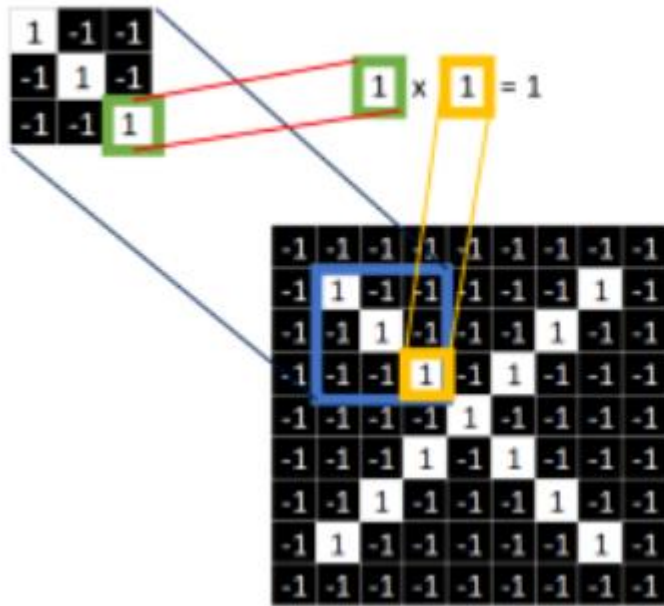
CNN



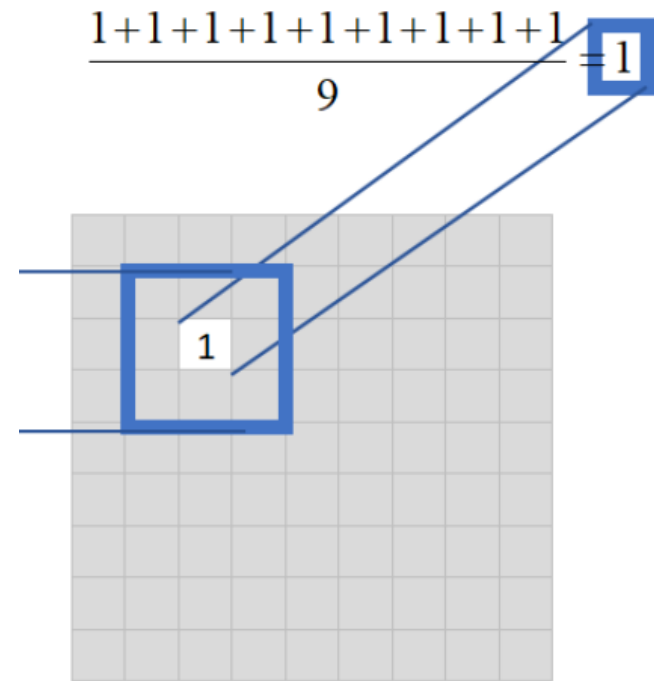
CNN/filter



CNN/convolution



1	1	1
1	1	1
1	1	1



CNN/convolution

1	-1	-1
-1	1	-1
-1	-1	1

-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	-1	-1	1	-1	-1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1

对图像运用该卷积核，
产生的结果



Feature map

0.77	-0.11	0.11	0.33	0.55	-0.11	0.33
-0.11	1.00	-0.11	0.33	-0.11	0.11	-0.11
0.11	-0.11	1.00	-0.33	0.11	-0.11	0.55
0.33	0.33	-0.33	0.55	-0.33	0.33	0.33
0.55	-0.11	0.11	-0.33	1.00	-0.11	0.11
-0.11	0.11	-0.11	0.33	-0.11	1.00	-0.11
0.33	-0.11	0.55	0.33	0.11	-0.11	0.77

CNN/convolution

-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	-1	-1	1	-1	-1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1



1	-1	-1
-1	1	-1
-1	-1	1

=

0.77	0.11	0.11	0.33	0.55	0.11	0.33
0.11	1.00	0.11	0.33	0.11	0.11	0.11
0.11	0.11	1.00	0.33	0.11	0.11	0.55
0.33	0.33	0.33	0.55	0.33	0.33	0.33
0.55	0.11	0.11	0.33	1.00	0.11	0.11
0.11	0.11	0.11	0.33	0.11	1.00	0.11
0.33	0.11	0.55	0.33	0.11	0.11	0.77

-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1



1	-1	1
-1	1	-1
1	-1	1

=

0.33	0.55	0.11	0.11	0.11	0.55	0.33
0.55	0.55	0.55	0.33	0.55	0.55	0.55
0.11	0.55	0.55	0.77	0.55	0.55	0.11
0.11	0.33	0.77	1.00	0.77	0.33	0.11
0.11	0.55	0.55	0.77	0.55	0.55	0.11
0.55	0.55	0.55	0.33	0.55	0.55	0.55
0.33	0.55	0.11	0.11	0.11	0.55	0.33

-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	-1	-1	1	-1	-1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1

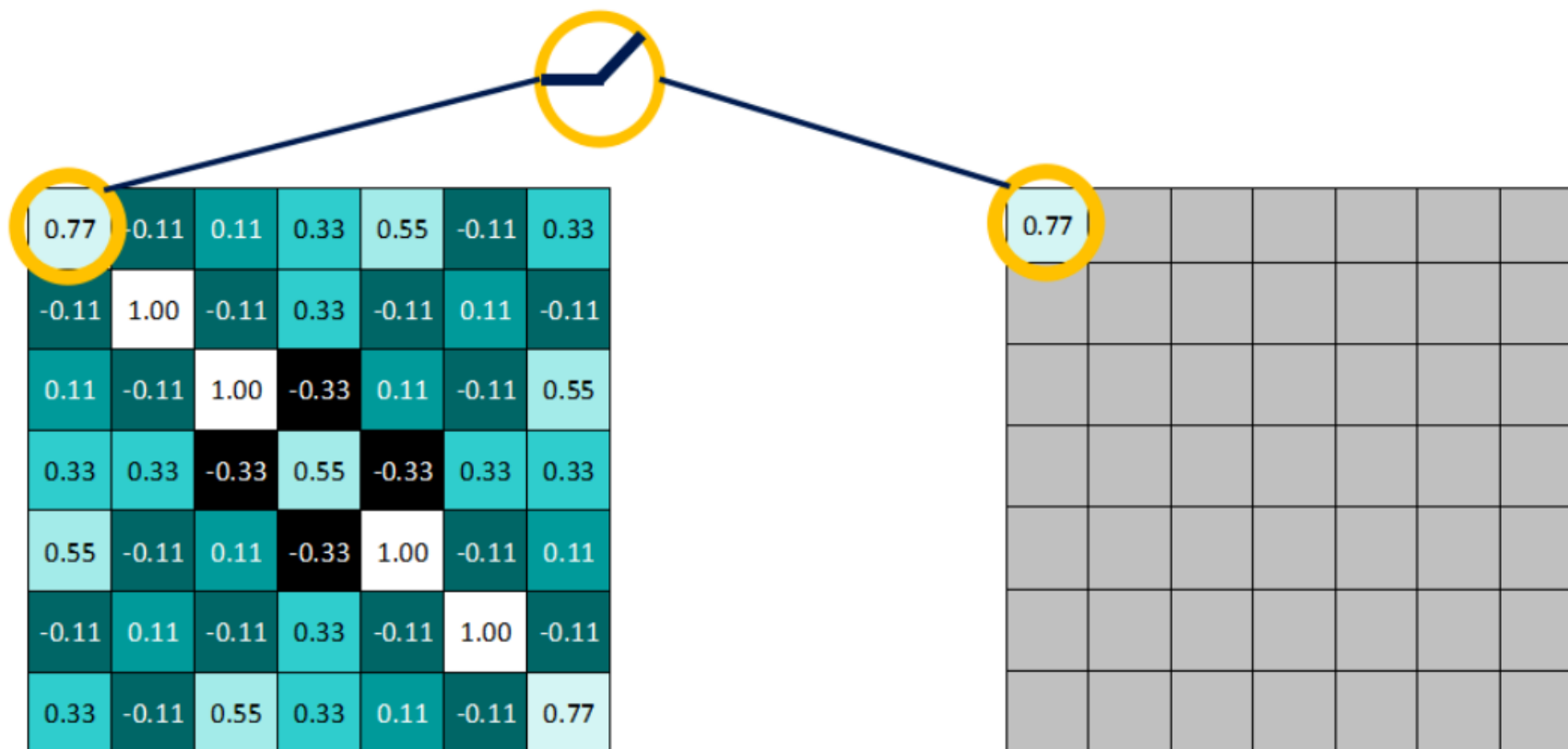


-1	-1	1
-1	1	-1
1	-1	-1

=

0.33	0.11	0.55	0.33	0.11	0.11	0.77
0.11	0.11	0.11	0.33	0.11	1.00	0.11
0.55	0.11	0.11	0.33	1.00	0.11	0.11
0.33	0.33	0.33	0.55	0.33	0.33	0.33
0.11	0.11	1.00	0.33	0.11	0.11	0.55
0.11	1.00	0.11	0.33	0.11	0.11	0.11
0.77	0.11	0.11	0.33	0.55	0.11	0.33

CNN/Relu



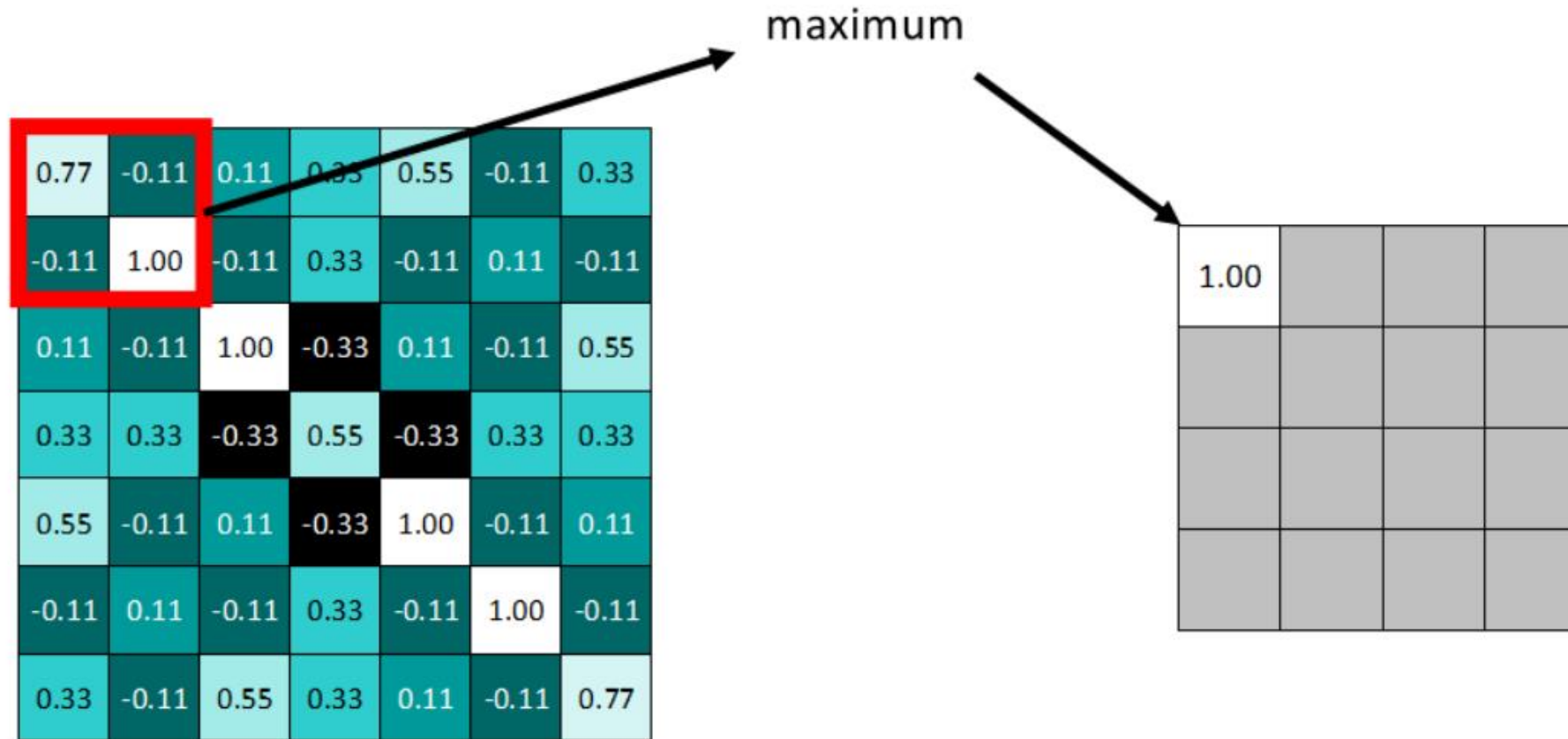
CNN/Relu

0.77	-0.11	0.11	0.33	0.55	-0.11	0.33
-0.11	1.00	-0.11	0.33	-0.11	0.11	-0.11
0.11	-0.11	1.00	-0.33	0.11	-0.11	0.55
0.33	0.33	-0.33	0.55	-0.33	0.33	0.33
0.55	-0.11	0.11	-0.33	1.00	-0.11	0.11
-0.11	0.11	-0.11	0.33	-0.11	1.00	-0.11
0.33	-0.11	0.55	0.33	0.11	-0.11	0.77



0.77	0	0.11	0.33	0.55	0	0.33
0	1.00	0	0.33	0	0.11	0
0.11	0	1.00	0	0.11	0	0.55
0.33	0.33	0	0.55	0	0.33	0.33
0.55	0	0.11	0	1.00	0	0.11
0	0.11	0	0.33	0	1.00	0
0.33	0	0.55	0.33	0.11	0	0.77

CNN/pooling



CNN/pooling

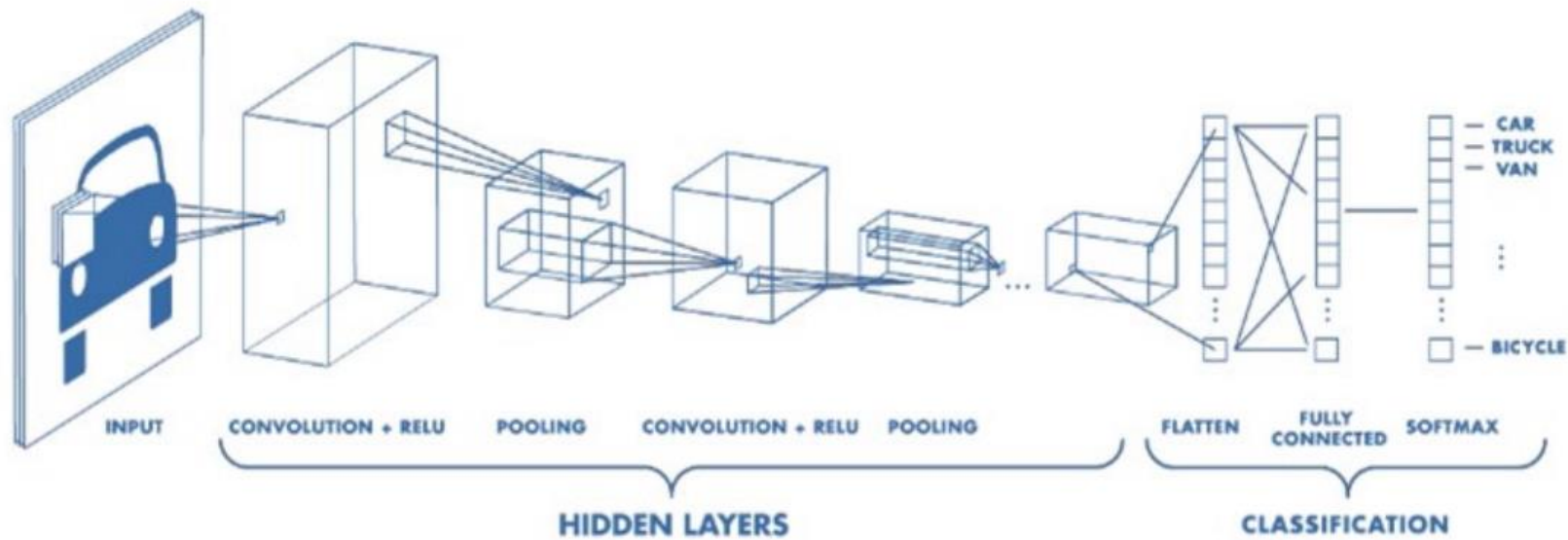
0.77	-0.11	0.11	0.33	0.55	-0.11	0.33
-0.11	1.00	-0.11	0.33	-0.11	0.11	-0.11
0.11	-0.11	1.00	-0.33	0.11	-0.11	0.55
0.33	0.33	-0.33	0.55	-0.33	0.33	0.33
0.55	-0.11	0.11	-0.33	1.00	-0.11	0.11
-0.11	0.11	-0.11	0.33	-0.11	1.00	-0.11
0.33	-0.11	0.55	0.33	0.11	-0.11	0.77

max pooling

1.00	0.33	0.55	0.33
0.33	1.00	0.33	0.55
0.55	0.33	1.00	0.11
0.33	0.55	0.11	0.77

CNN

1. Convolution Layer(s)
2. Pooling Layer(s)
3. Classification - Fully Connected Layer



RNN

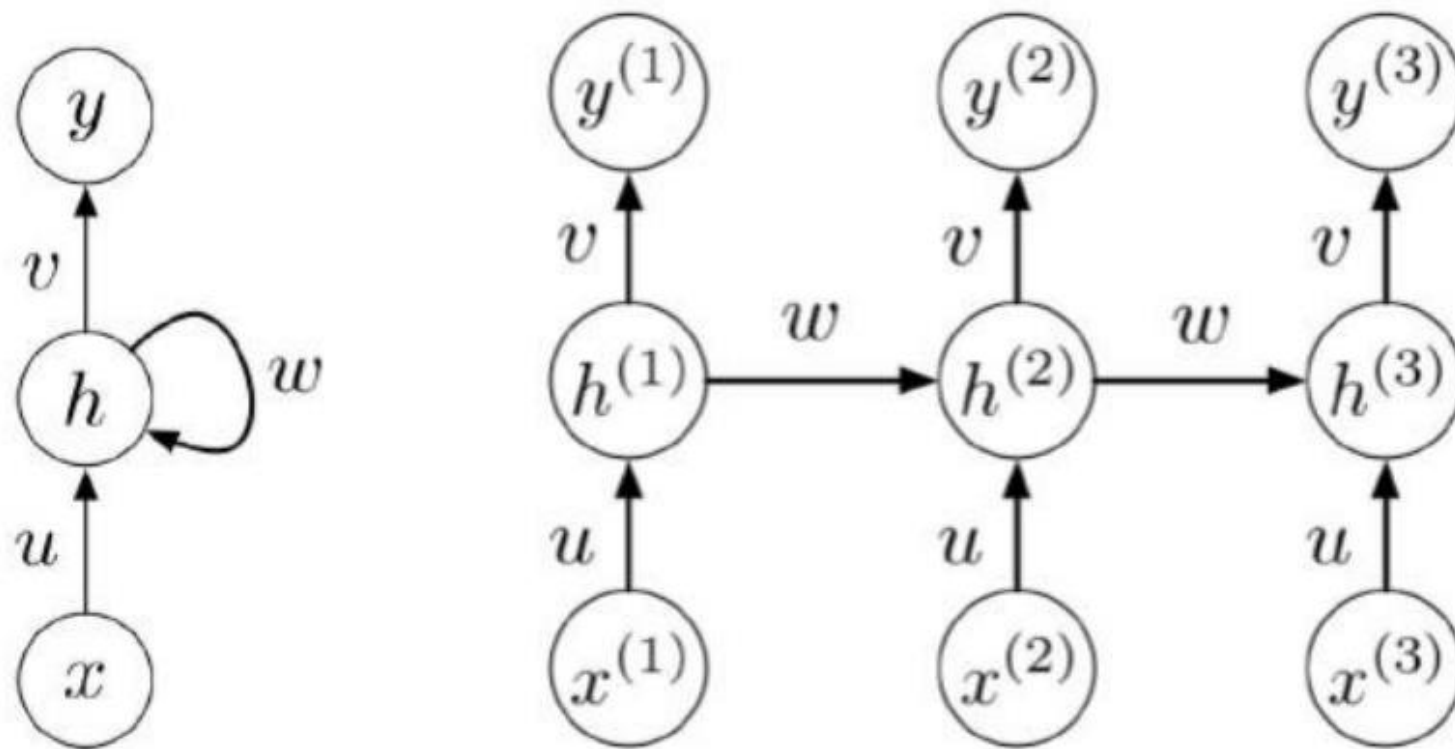
apple



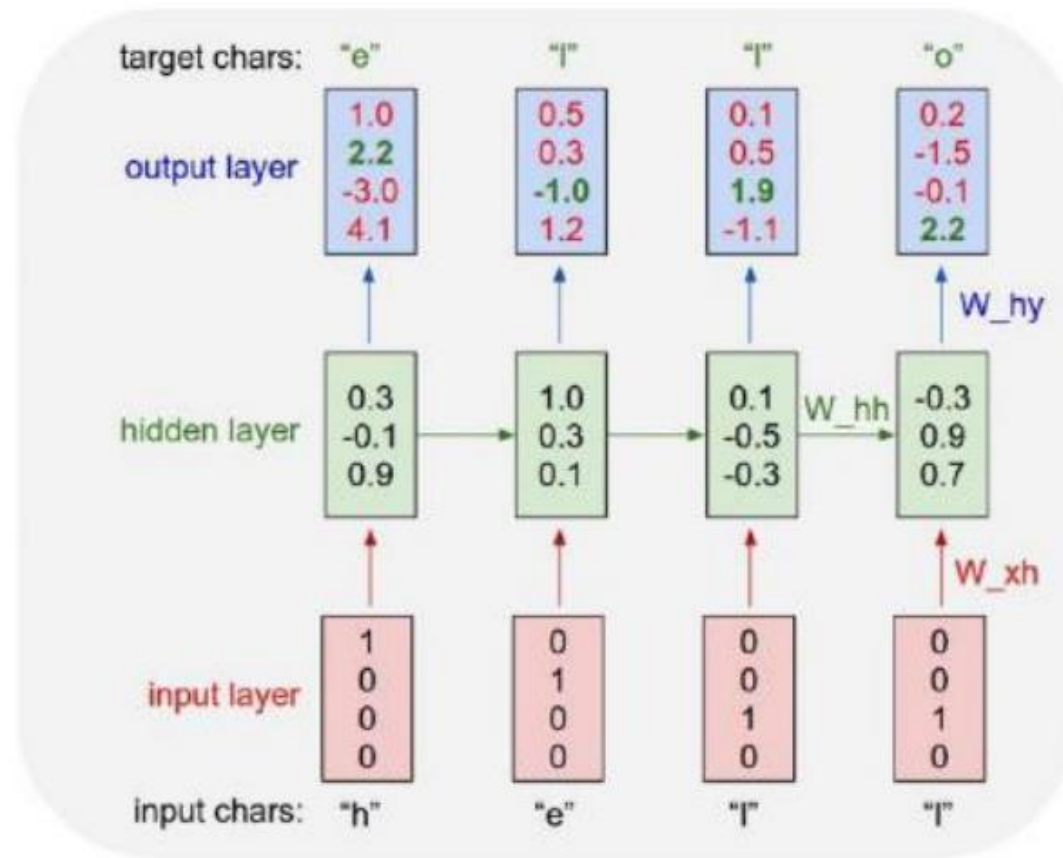
I ate an apple.



RNN



RNN



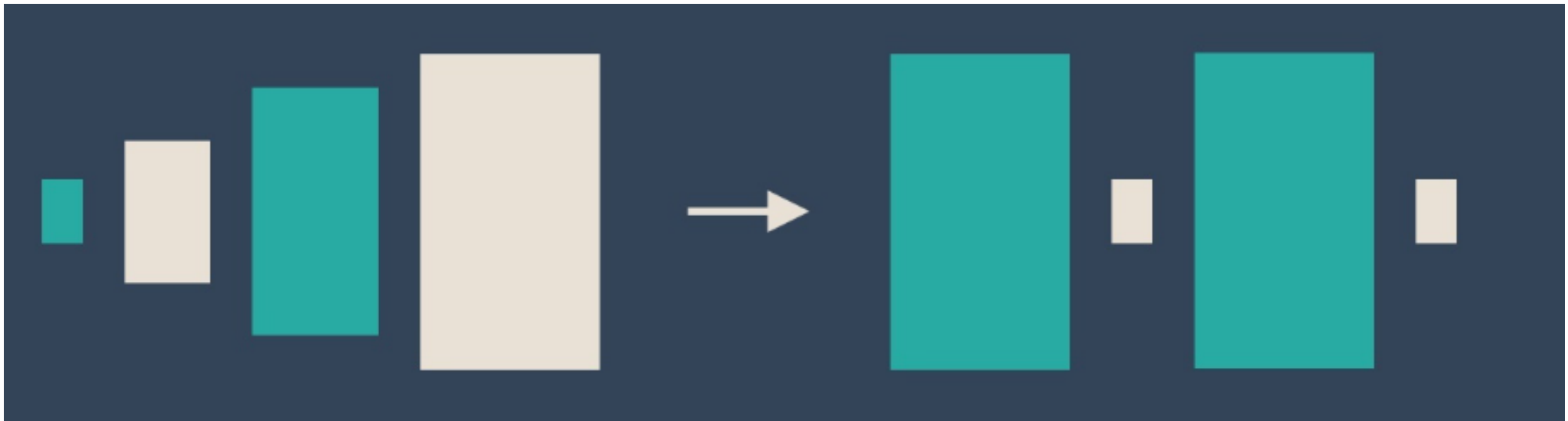
$$\boxed{h_t} = \boxed{f_W}(\boxed{h_{t-1}}, \boxed{x_t})$$

new state / some function with parameters W old state input vector at some time step

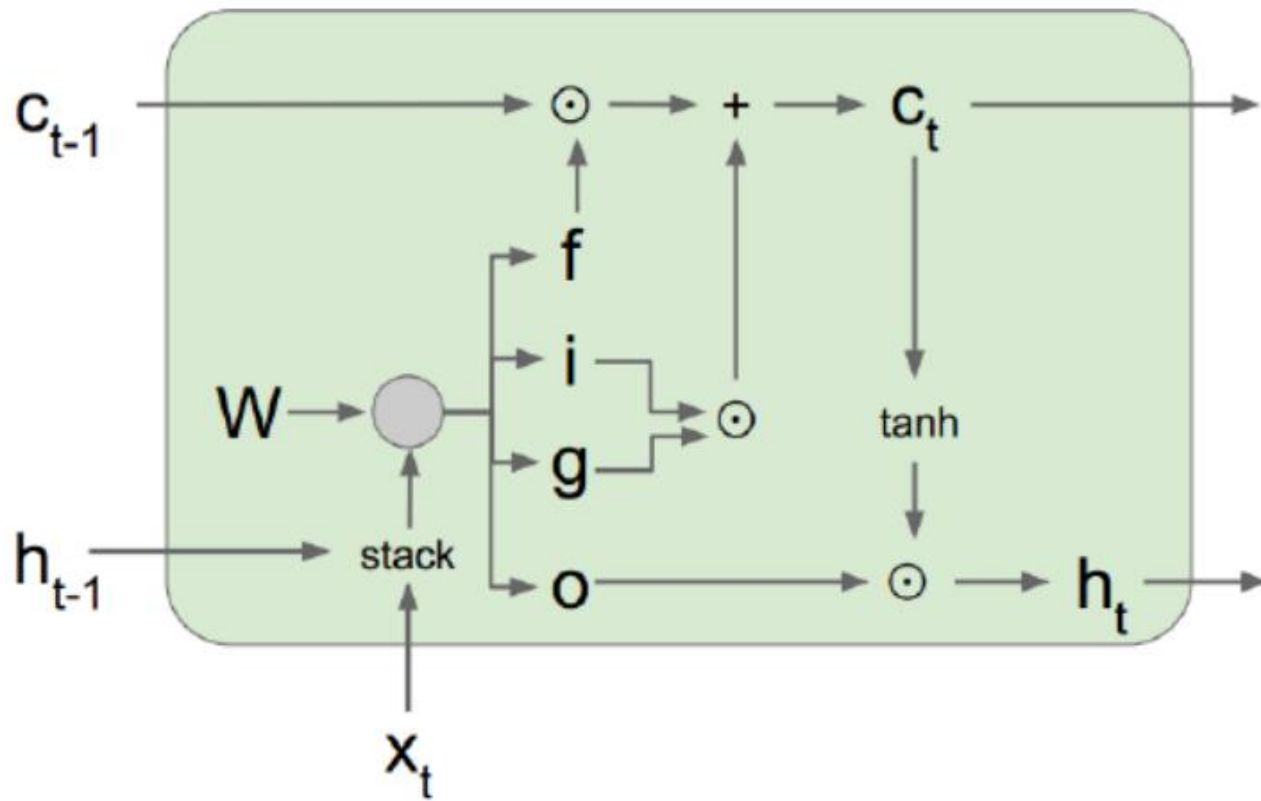
$$y_t = W_{hy}h_t$$

LSTM (Long short-term memory)

Decide if something is important or not



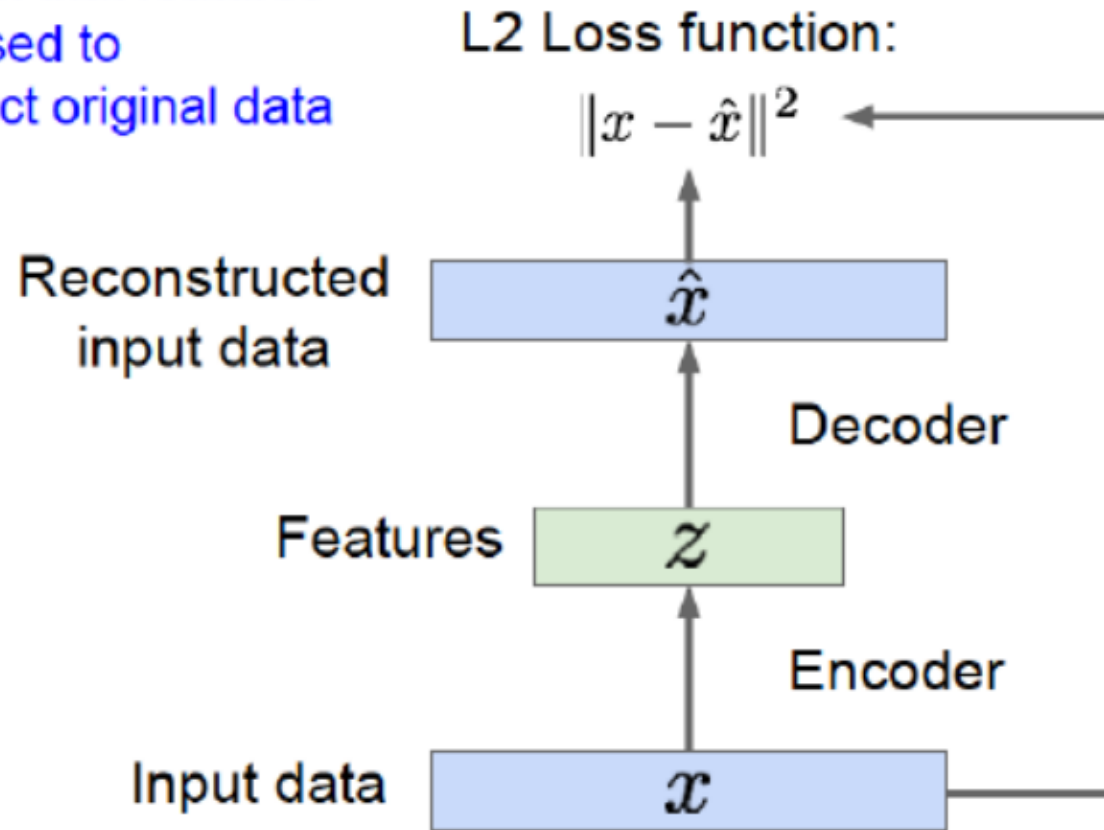
LSTM



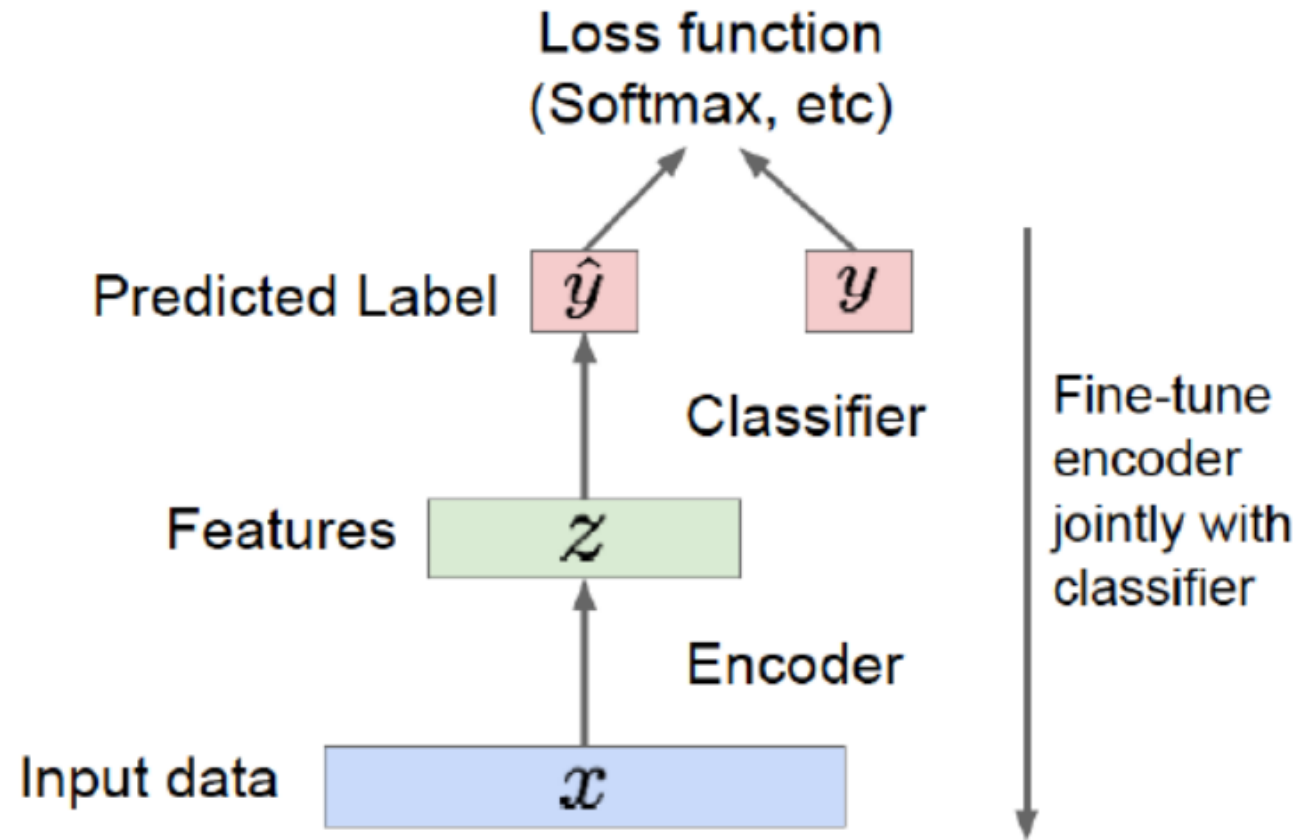
$$\begin{pmatrix} i \\ f \\ o \\ g \end{pmatrix} = \begin{pmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{pmatrix} W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix}$$
$$c_t = f \odot c_{t-1} + i \odot g$$
$$h_t = o \odot \tanh(c_t)$$

Autoencoder

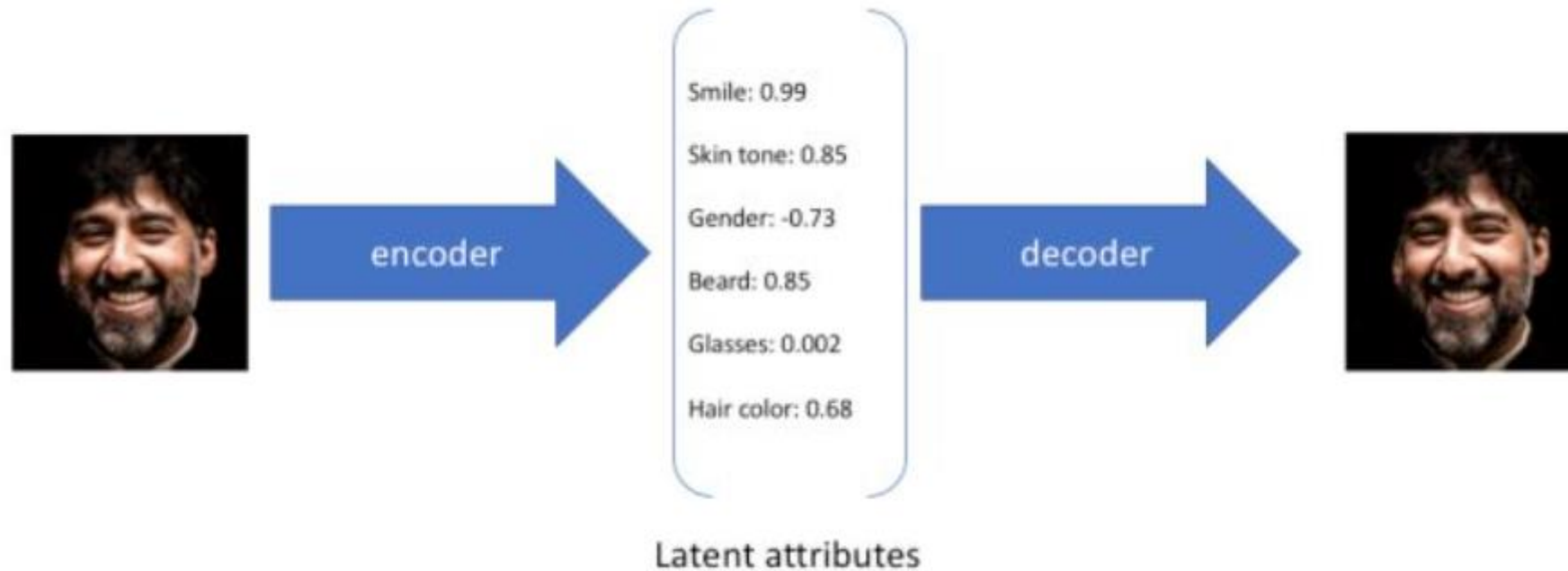
Train such that features
can be used to
reconstruct original data



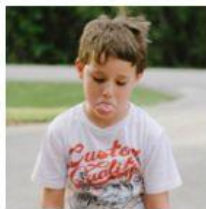
Autoencoder



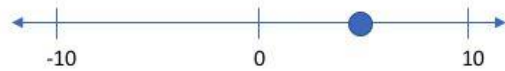
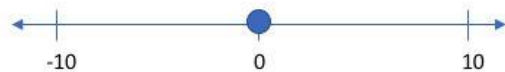
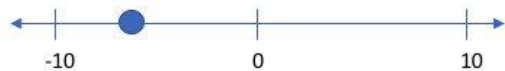
Autoencoder



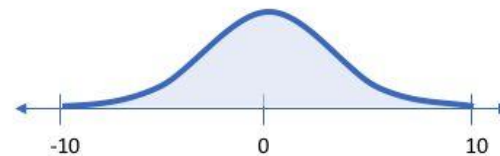
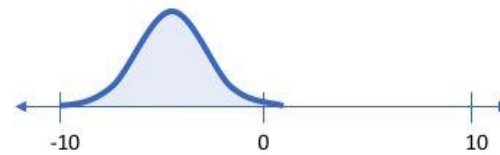
VAE



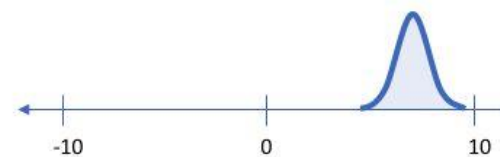
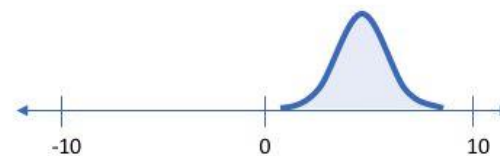
Smile (discrete value)



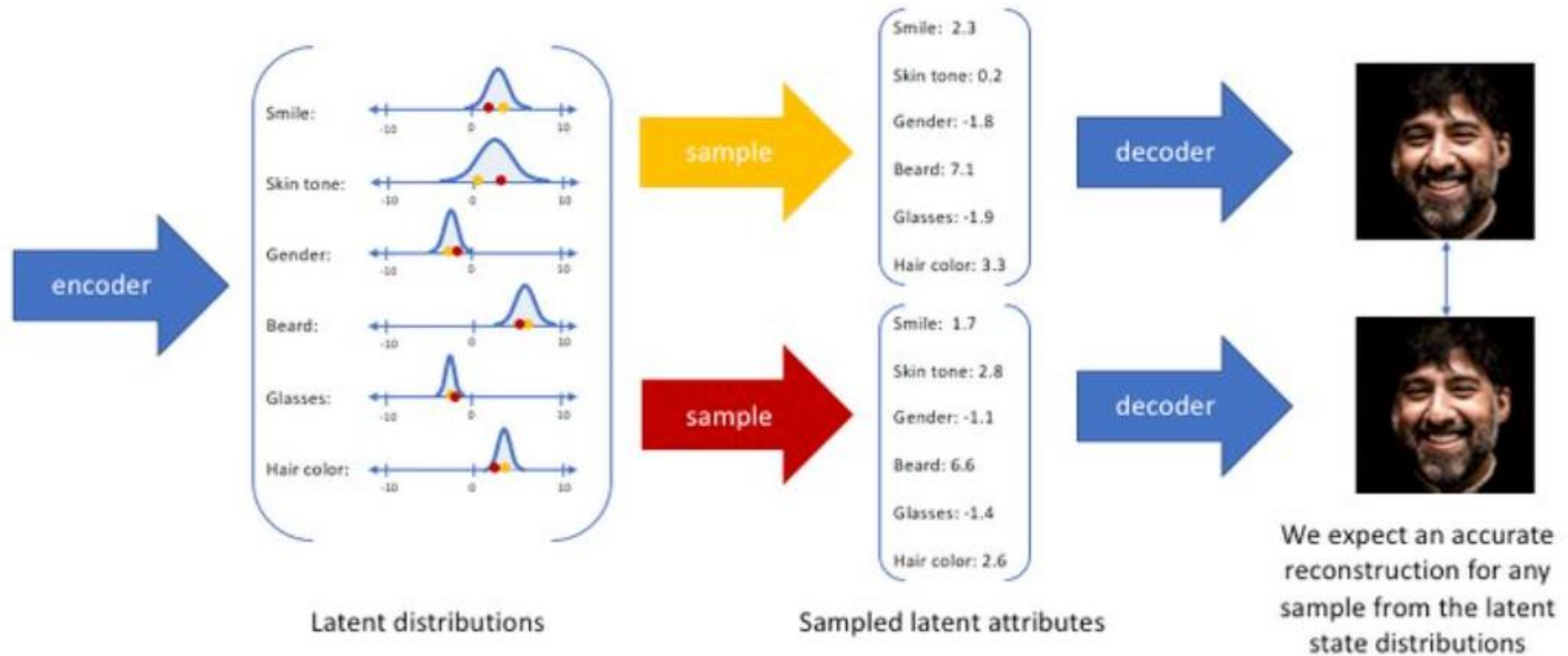
Smile (probability distribution)



vs.

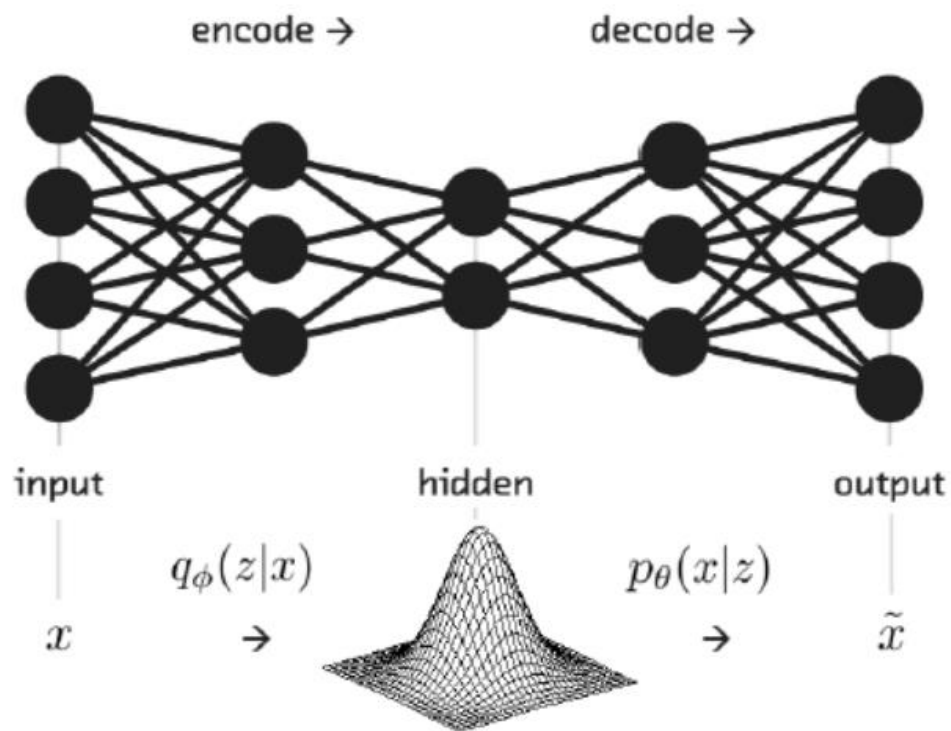


VAE

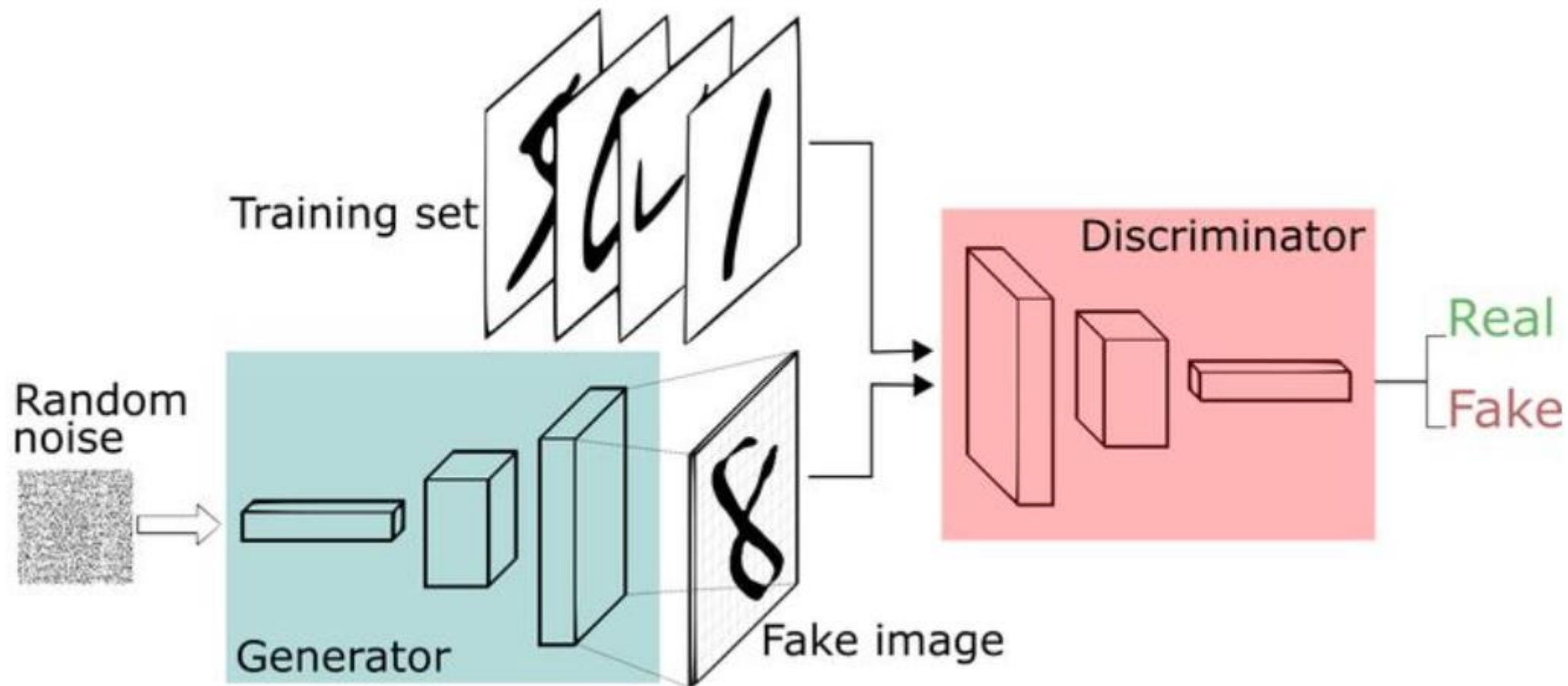


VAE

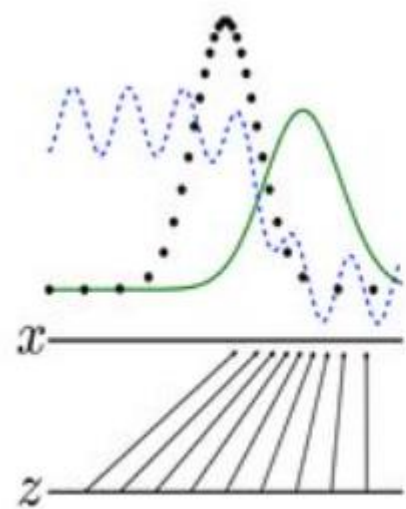
- Objective $\mathcal{L}(x, \phi, \theta) = \underbrace{-D_{KL}(q_{\phi}(z|x)||p_{\theta}(z))}_{\text{Regularization term}} + \underbrace{E_{q_{\phi}(z|x)}[\log p_{\theta}(x|z)]}_{\text{Reconstruction term}}$



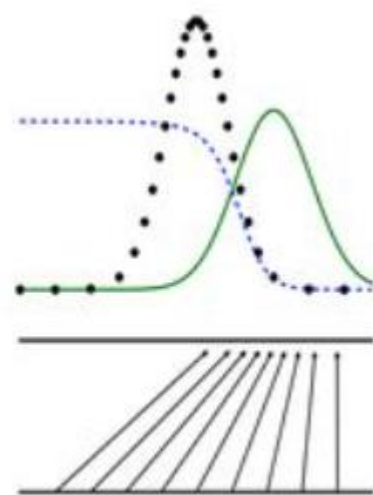
GAN



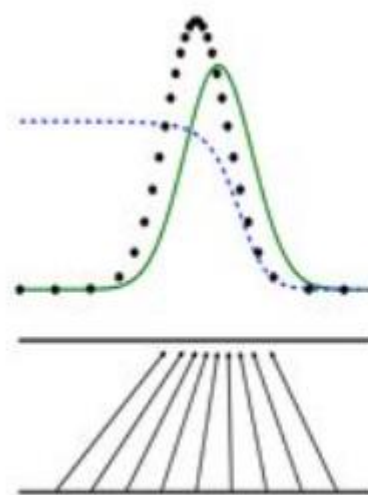
GAN



(a)

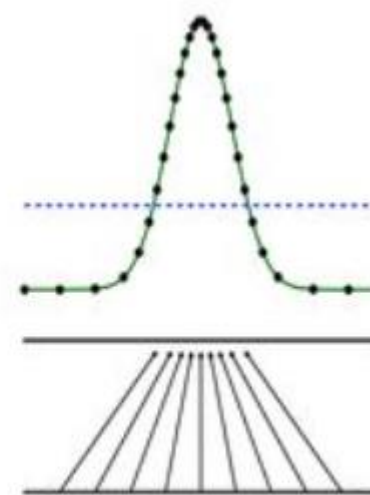


(b)



(c)

...



(d)

Reference

- <https://zhuanlan.zhihu.com/p/27908027>
- https://www.bilibili.com/video/BV1qM4y1M7Nv?p=3&vd_source=4e16ec5d64dfcf4e3da2a8cae0e32d0b
- <https://zhuanlan.zhihu.com/p/72228021>
- <https://www.zhihu.com/search?type=content&q=%E5%BE%AA%E7%8E%AF%E7%A5%9E%E7%BB%8F%E7%BD%91%E7%BB%9C>
- <https://www.zhihu.com/search?type=content&q=%E7%94%9F%E6%88%90%E5%AF%B9%E6%8A%97%E7%BD%91%E7%BB%9C>
- https://www.zhihu.com/search?q=%E8%87%AA%E7%BC%96%E7%A0%81%E5%99%A8%20VAE&utm_content=search_history&type=content
- <https://zhuanlan.zhihu.com/p/33752313>

THANKS