

Deep Learning (1470)

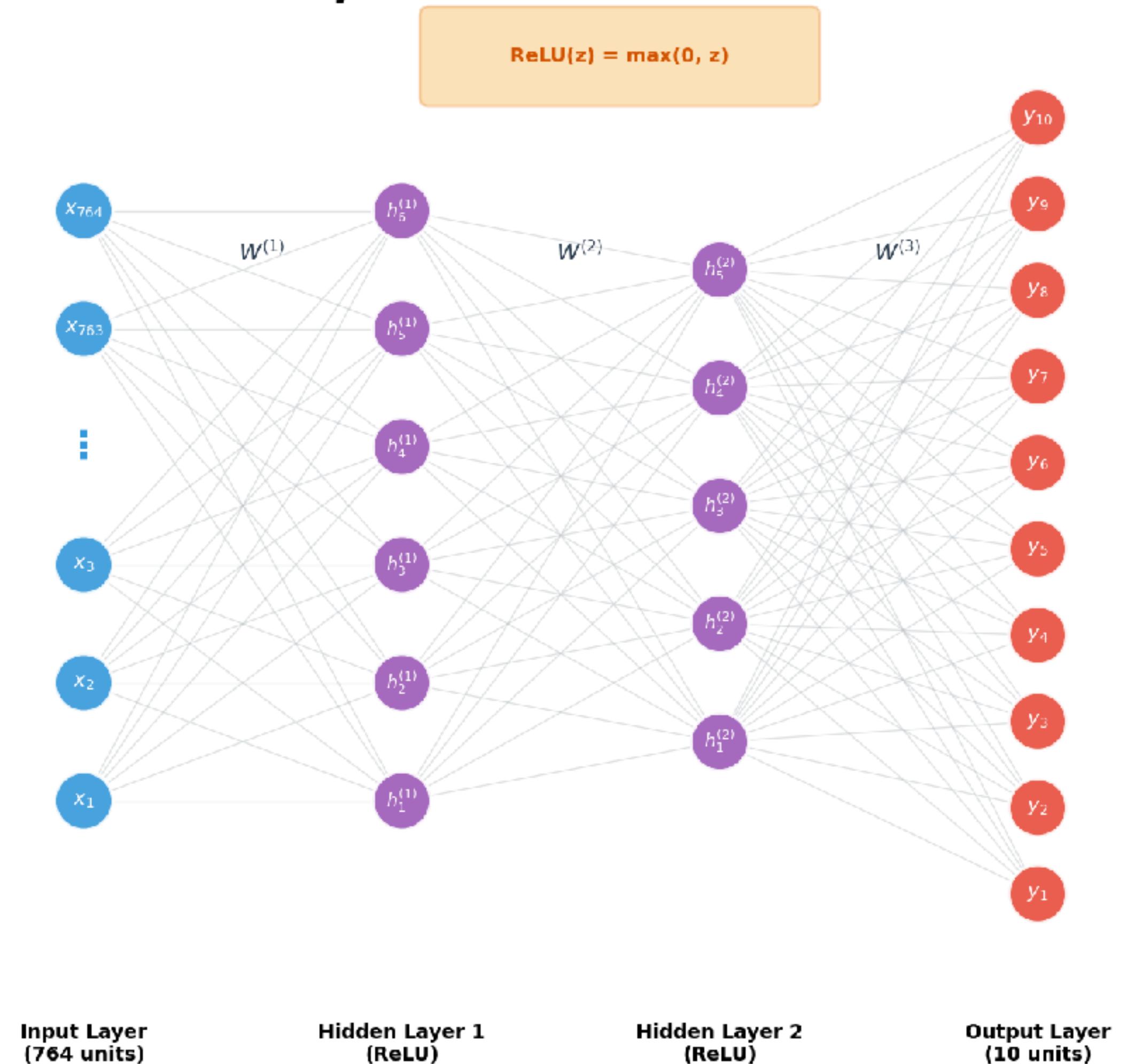
Randall Balestriero

Class 8: Convolutional Neural Networks

Multilayer Perceptrons

And why we need better

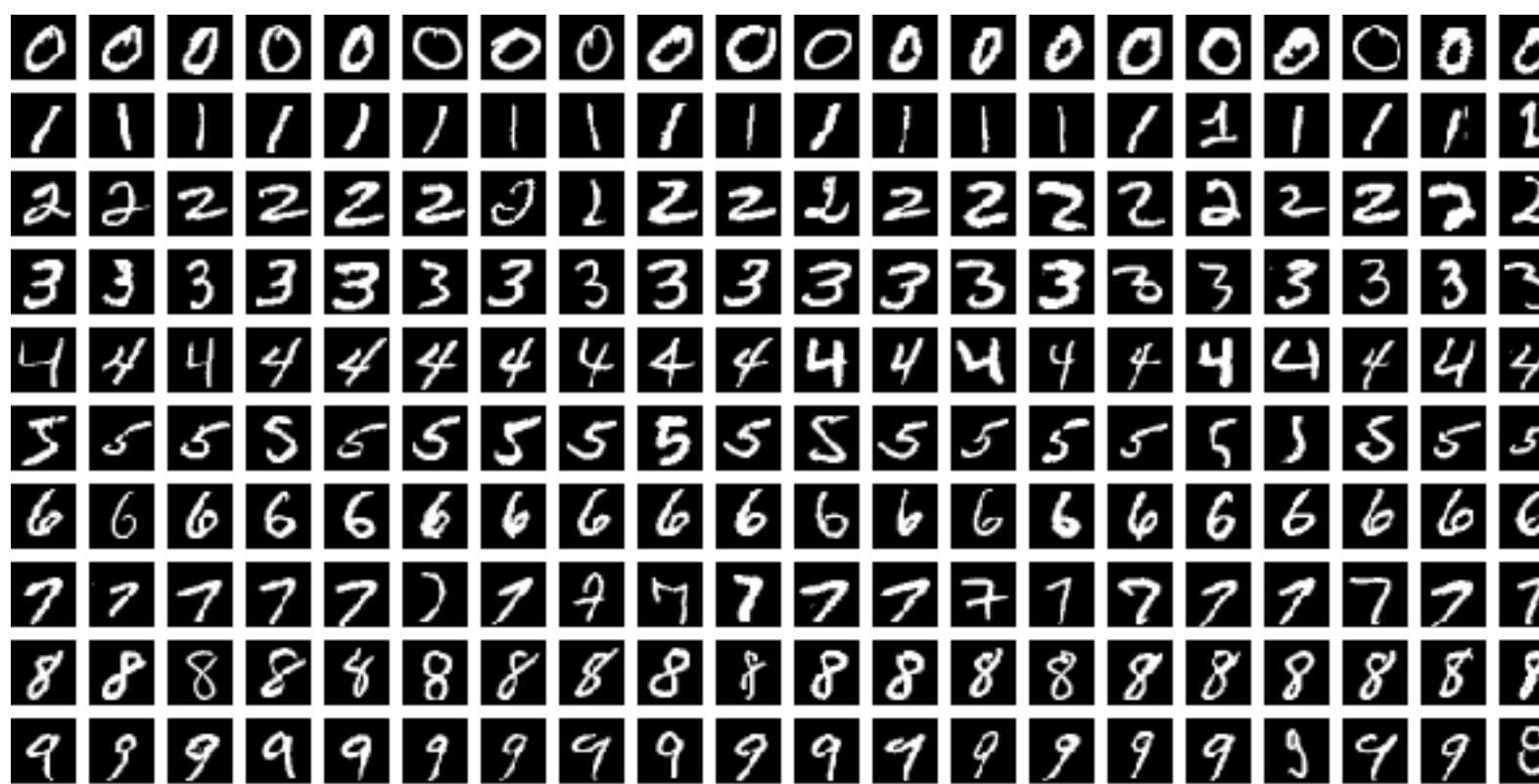
3-Layer MLP with ReLU Activations



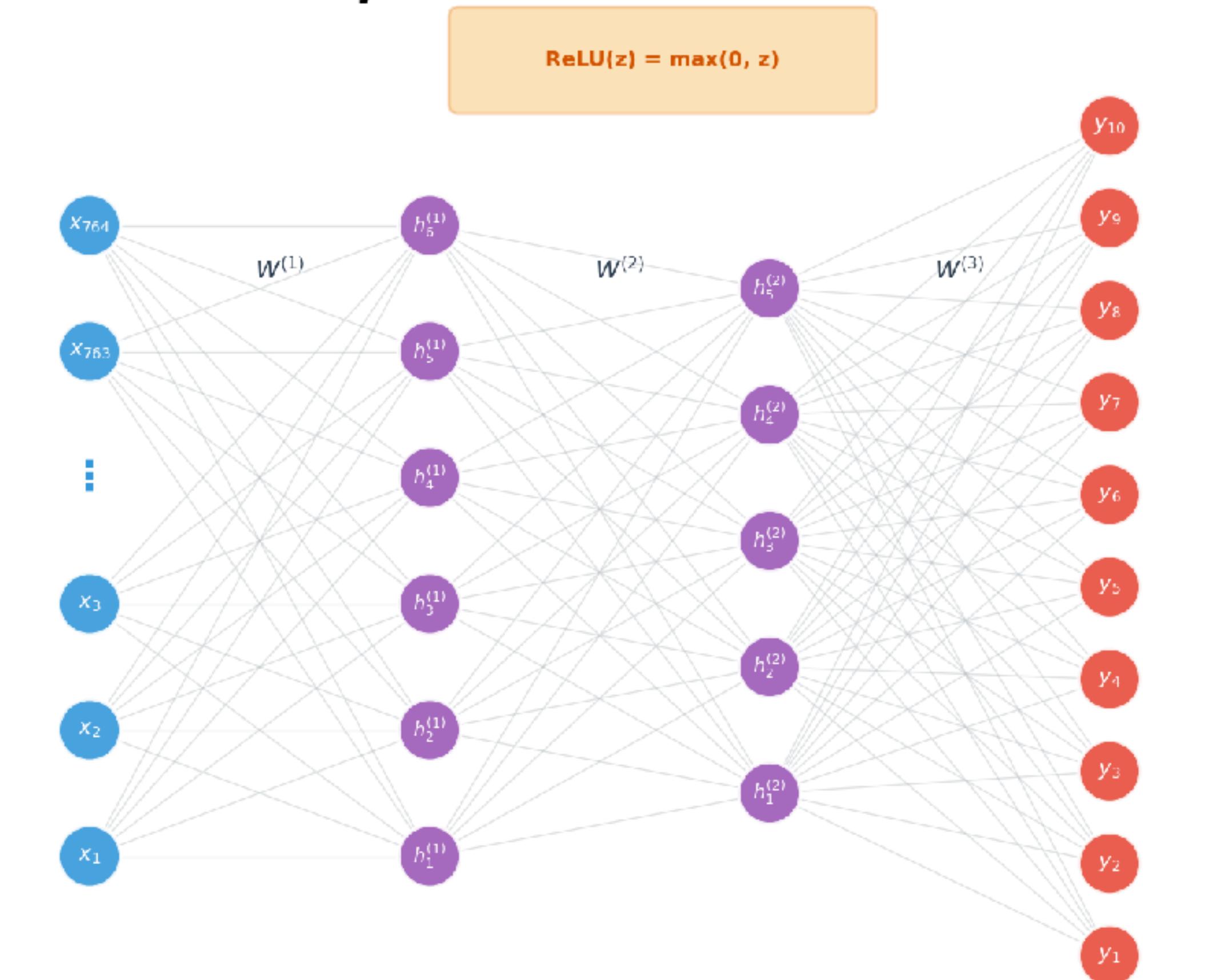
Multilayer Perceptrons

And why we need better

Extracting local features
at each layer is enough!



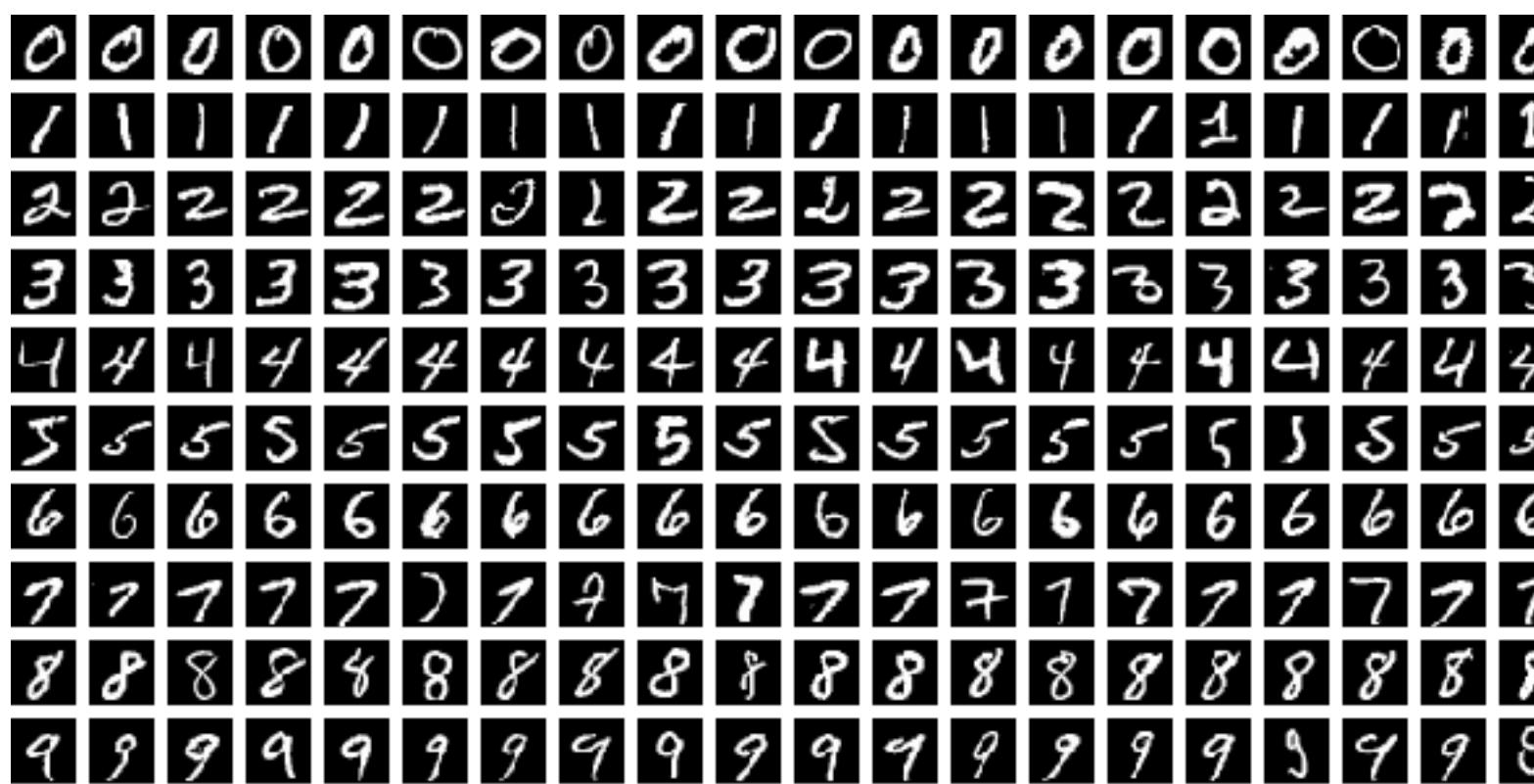
3-Layer MLP with ReLU Activations



Multilayer Perceptrons

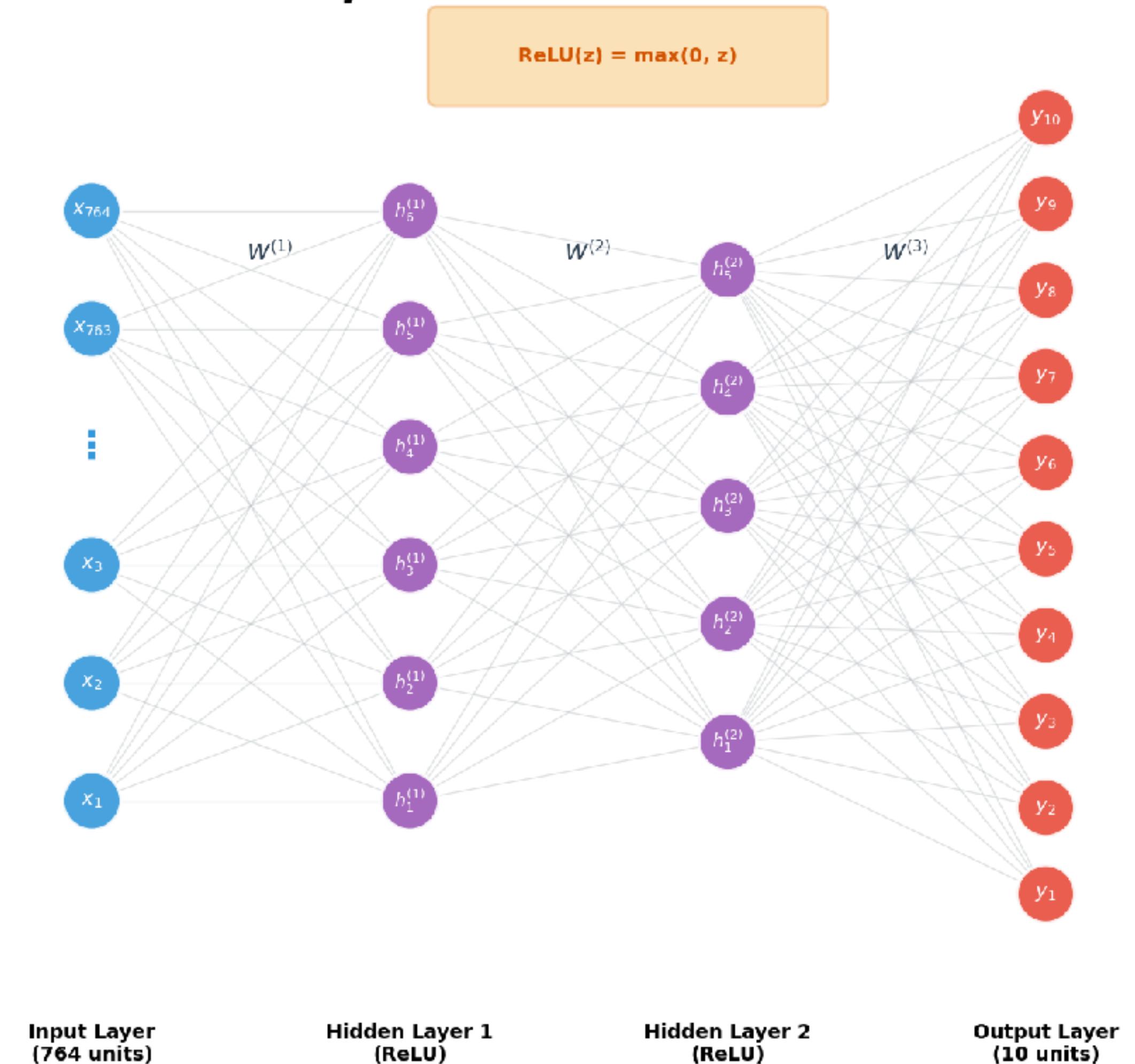
And why we need better

Extracting local features
at each layer is enough!

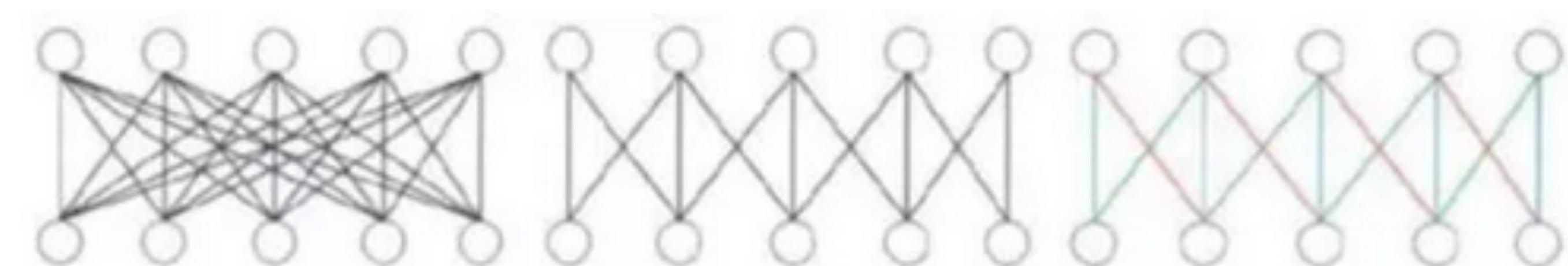
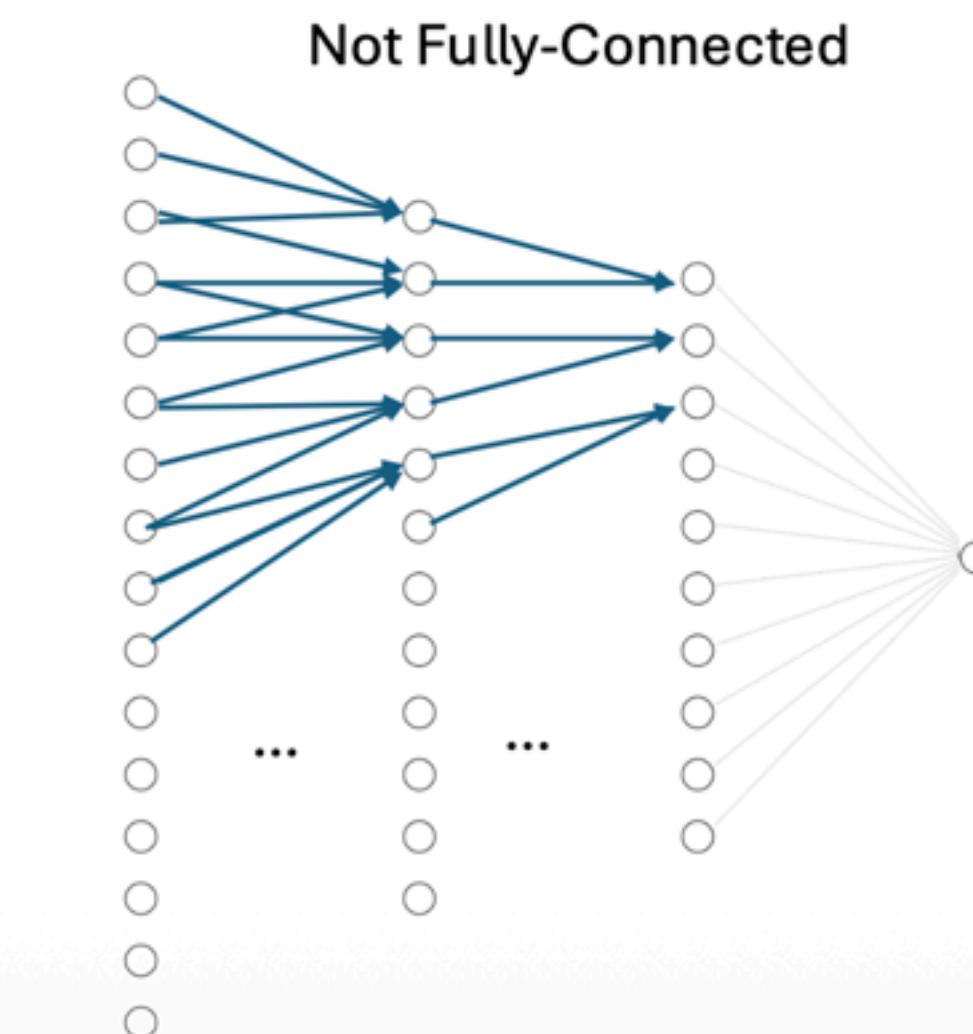
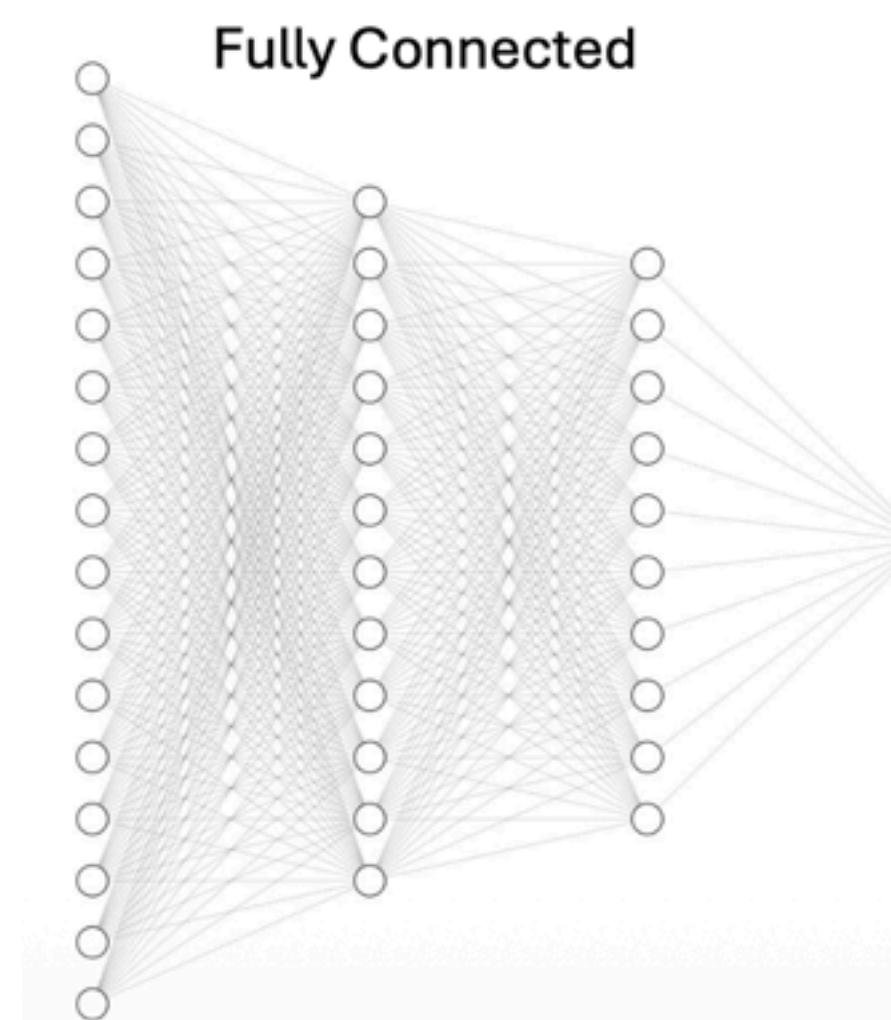
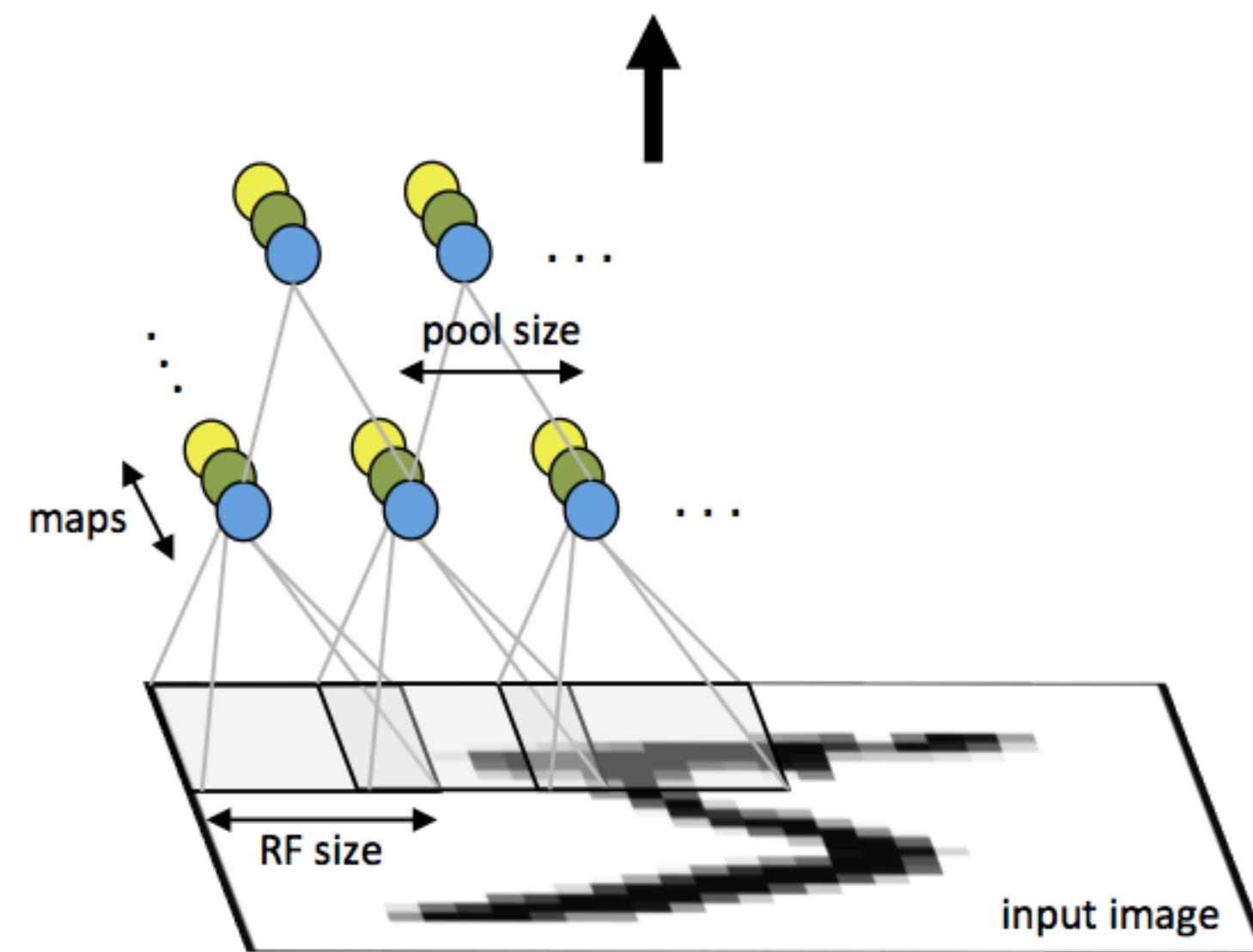


Not only local features... but
translation invariant features!

3-Layer MLP with ReLU Activations



How to Diagram



fully connected layer

All different weights

locally connected layer

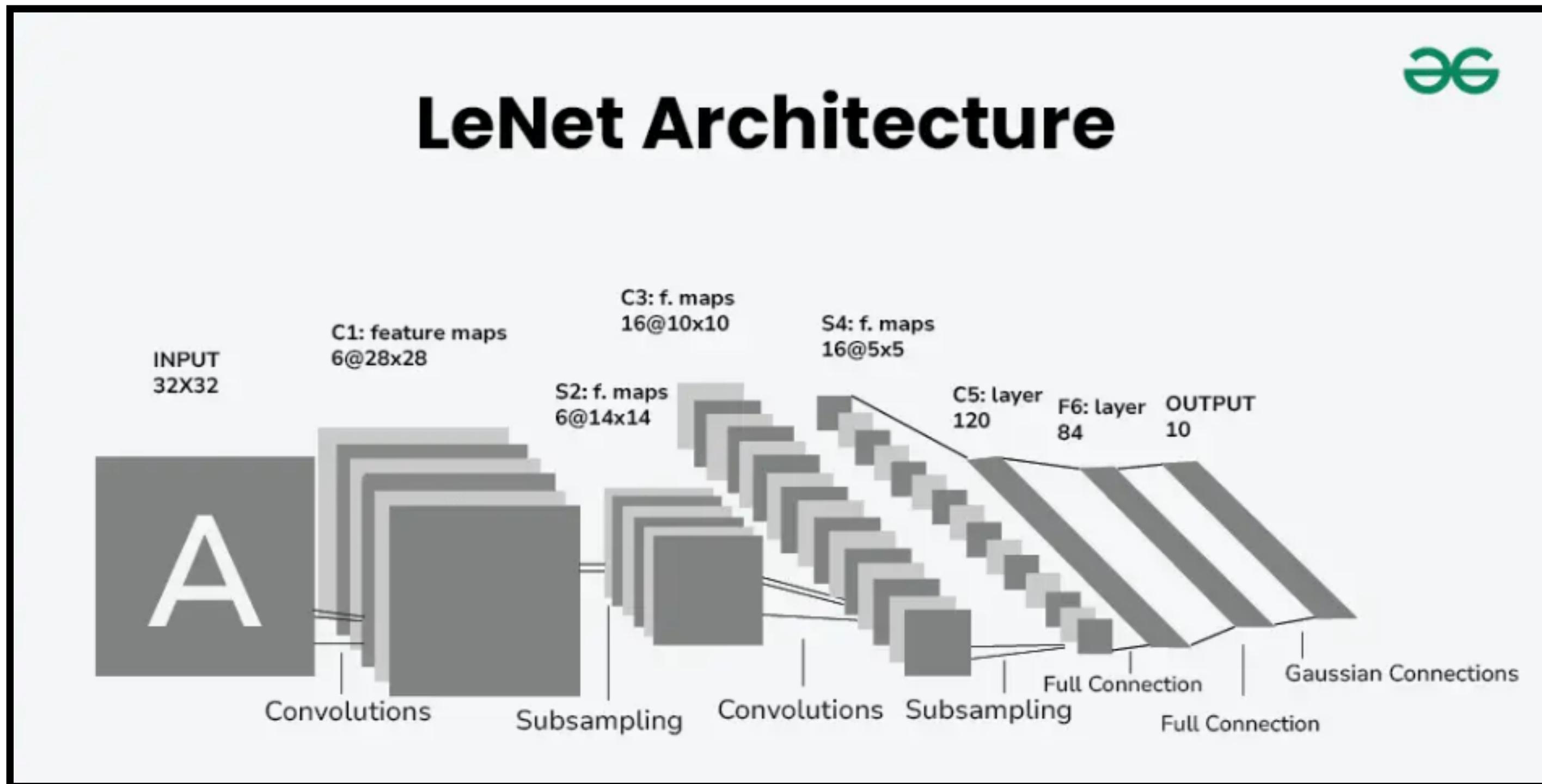
All different weights

convolution layer

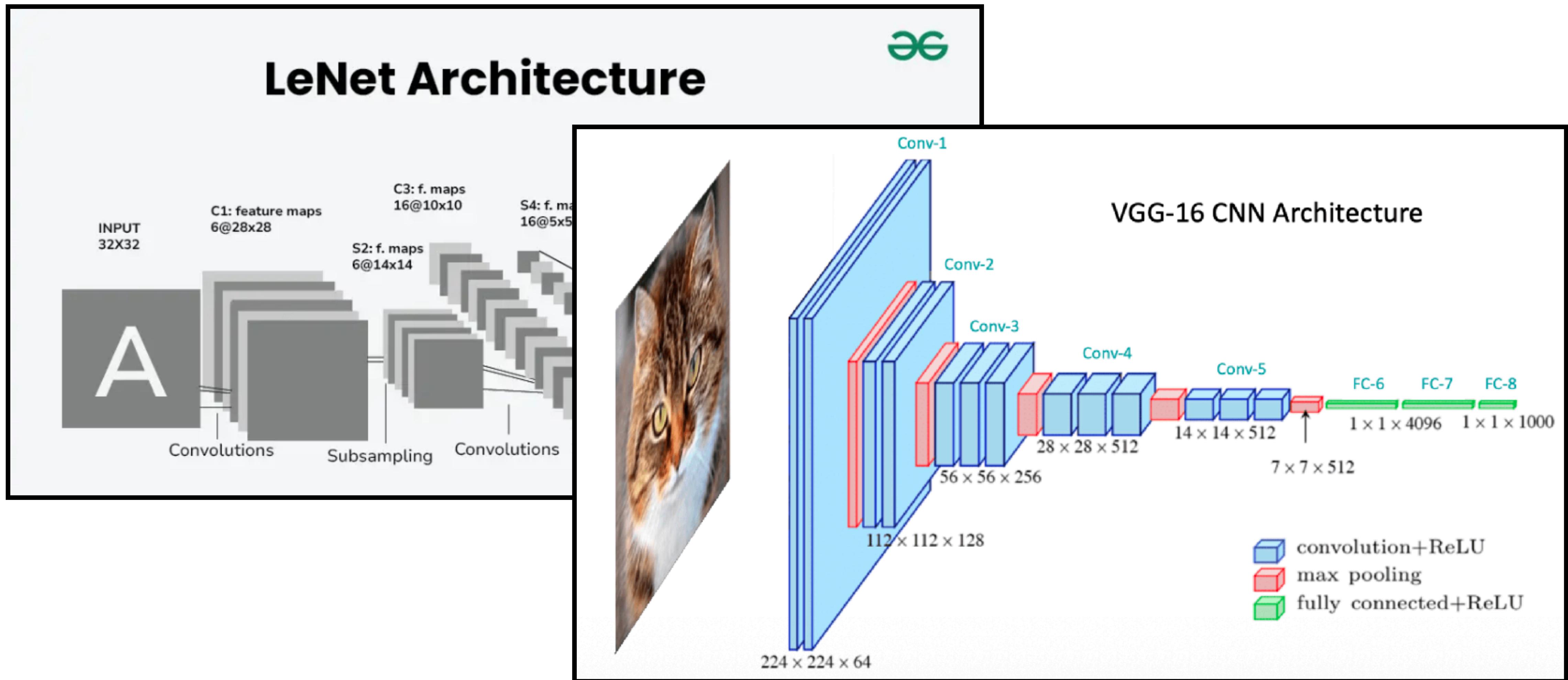
Shared weights

Deep Convolutional Networks

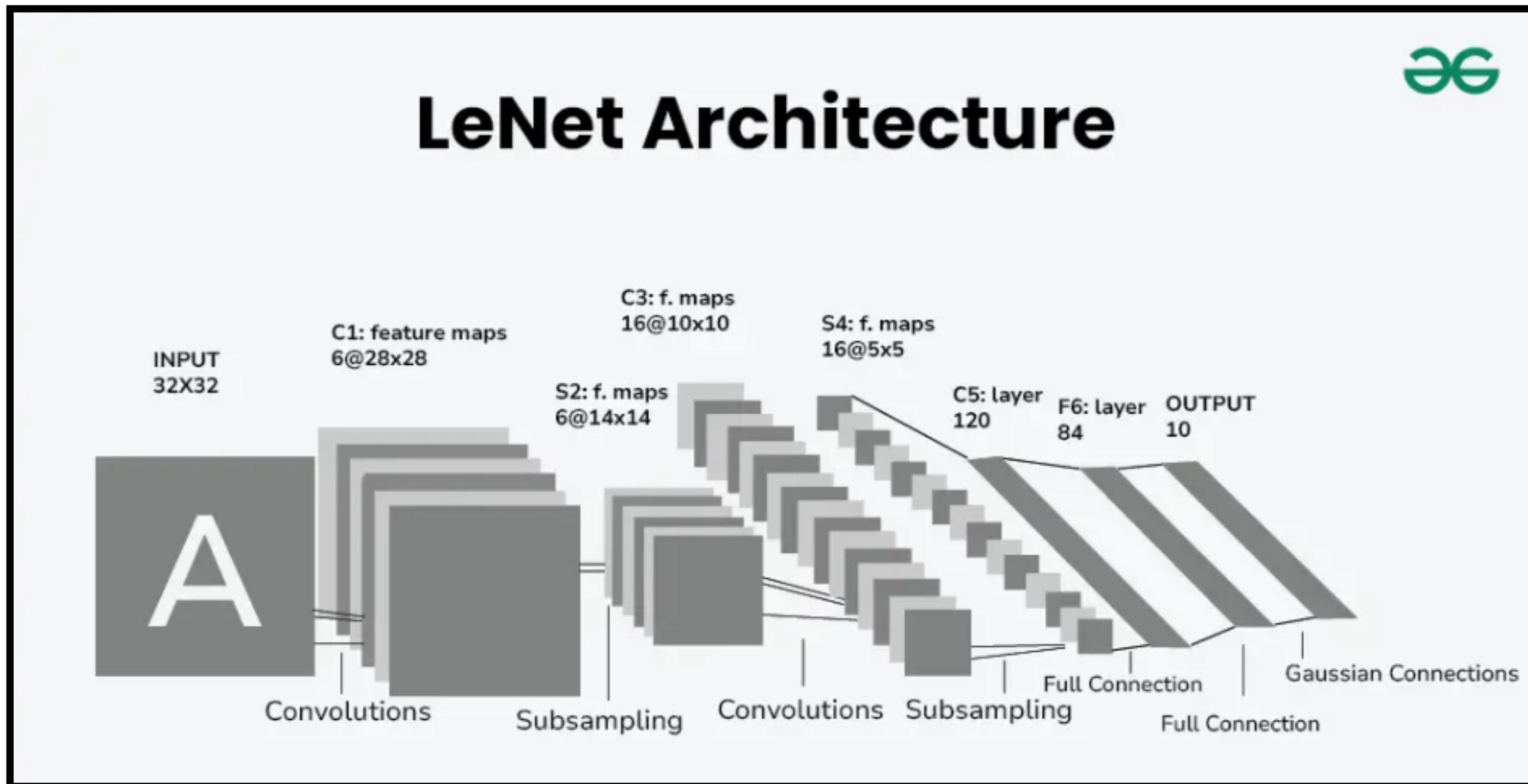
Deep Convolutional Networks



Deep Convolutional Networks

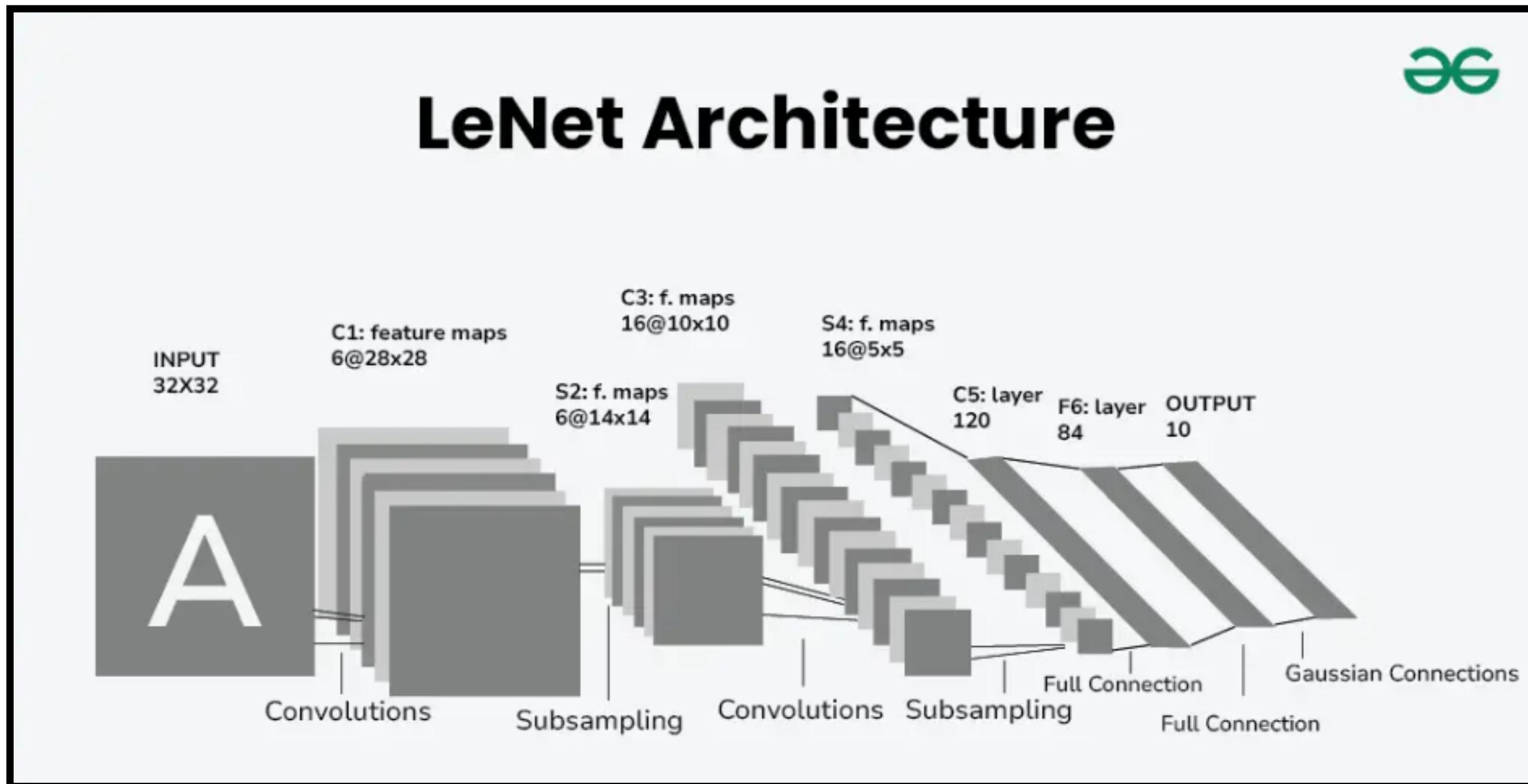


Deep Convolutional Networks



Can you guess how many parameters it has?

Deep Convolutional Networks



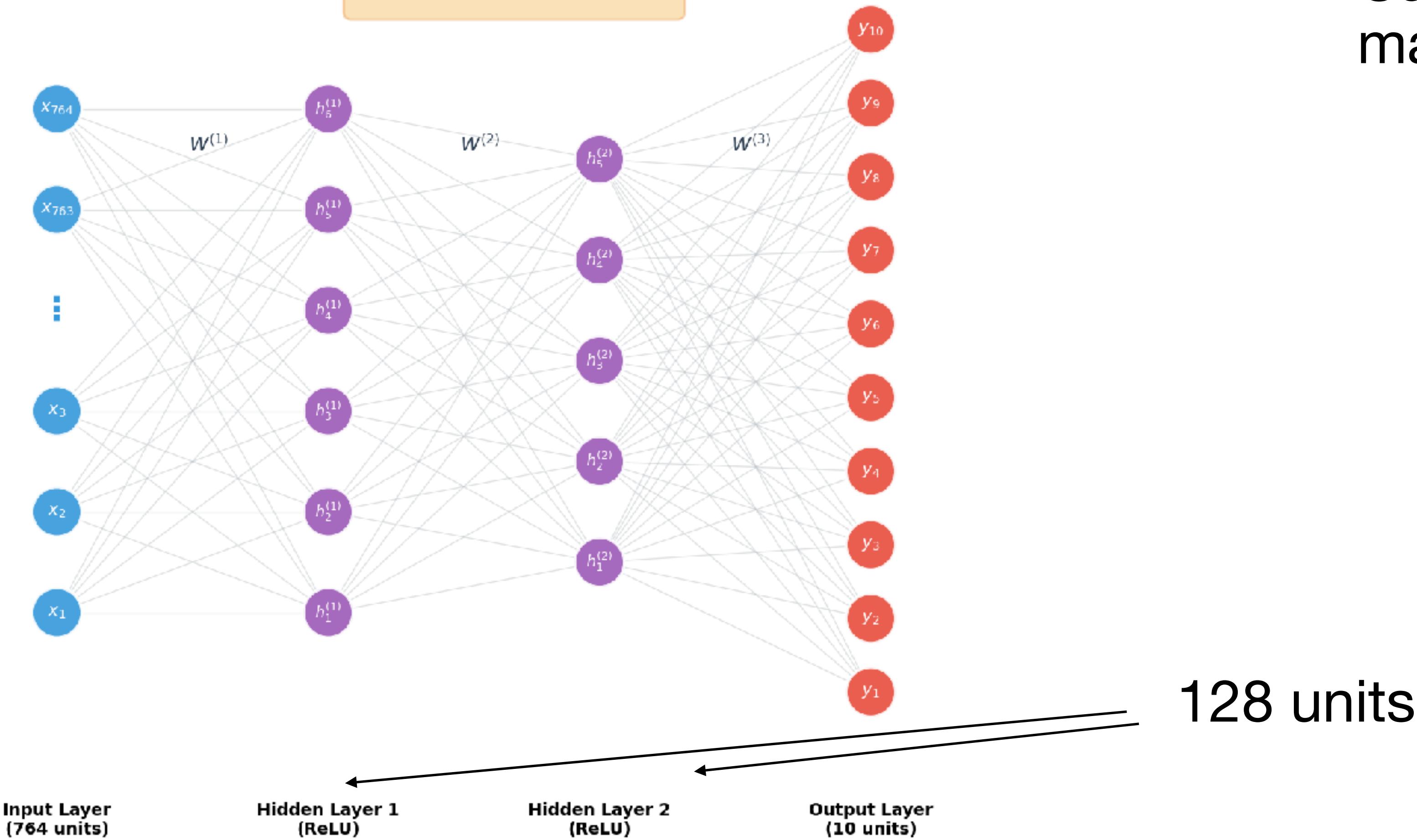
Can you guess how many parameters it has?

~60K!

Deep Convolutional Networks

3-Layer MLP with ReLU Activations

$$\text{ReLU}(z) = \max(0, z)$$

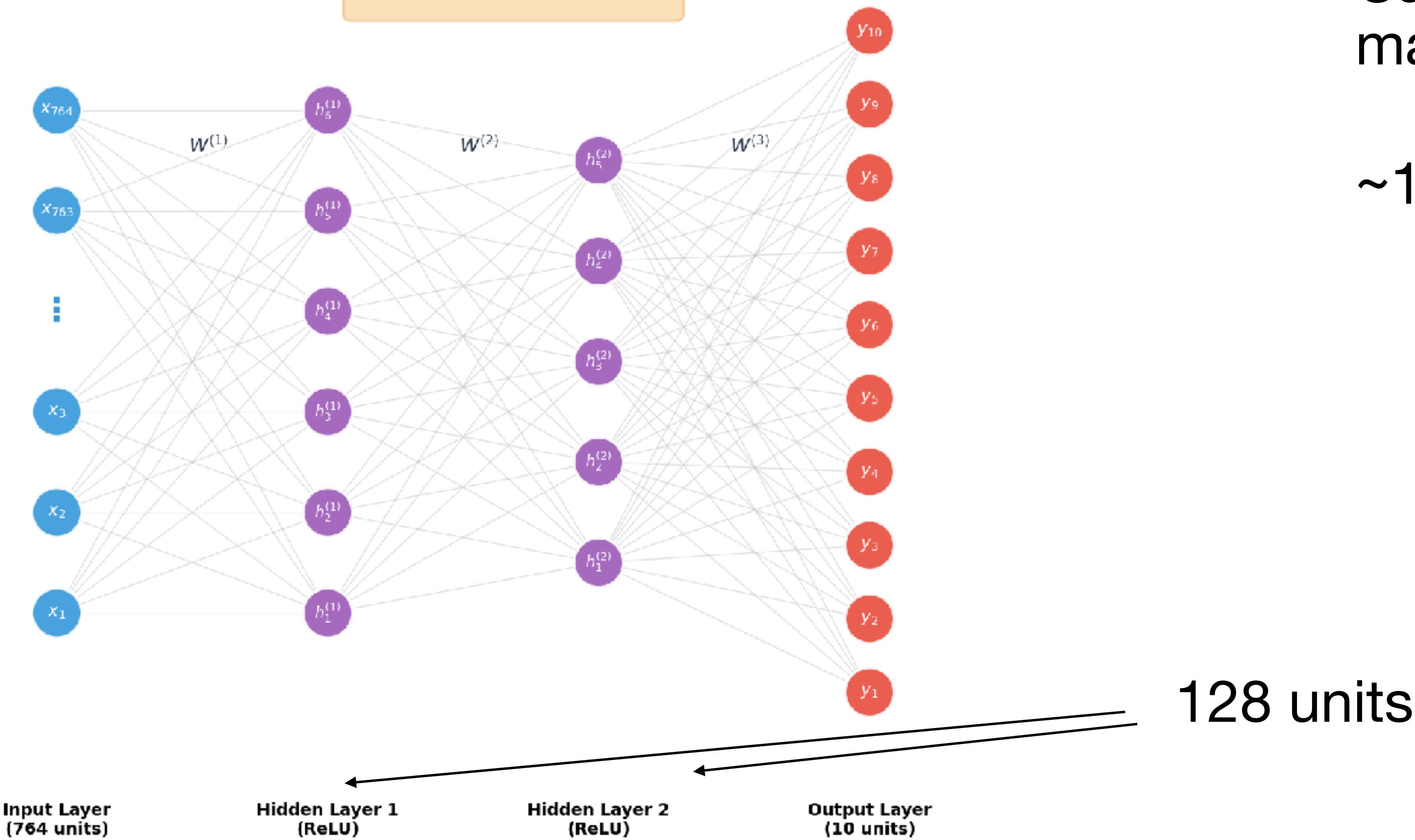


Can you guess how many parameters it has?

Deep Convolutional Networks

3-Layer MLP with ReLU Activations

$$\text{ReLU}(z) = \max(0, z)$$



Can you guess how many parameters it has?

~120K!

128 units

Output Size of a Convolution Layer

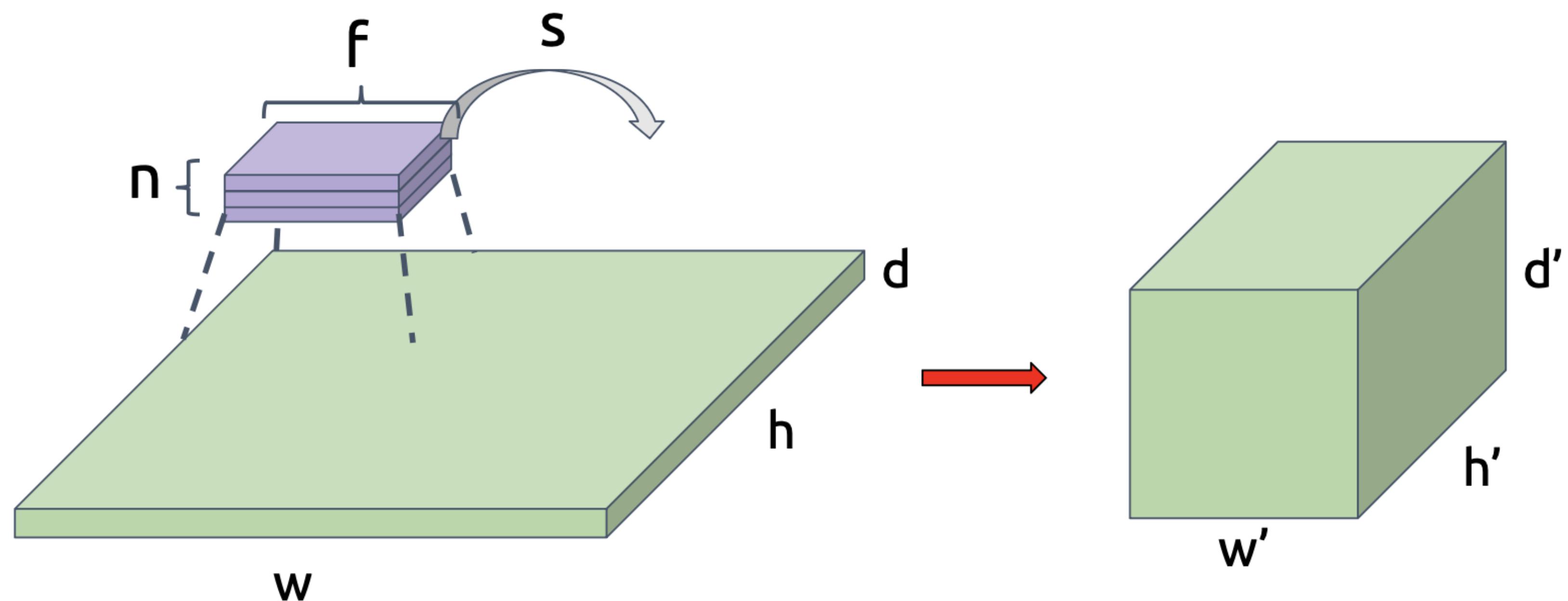
Suppose we know the number of filters, their size, the stride, and padding (n, f, s, p).

Then for a convolution layer with input dimension $w \times h \times d$, the output dimensions $w' \times h' \times d'$ are:

$$w' = \frac{w - f + 2p}{s} + 1$$

$$h' = \frac{h - f + 2p}{s} + 1$$

$$d' = n$$



Output Size for “VALID” Padding

$$w' = \frac{w - f + 2p}{s} + 1$$

Let $w = 4$

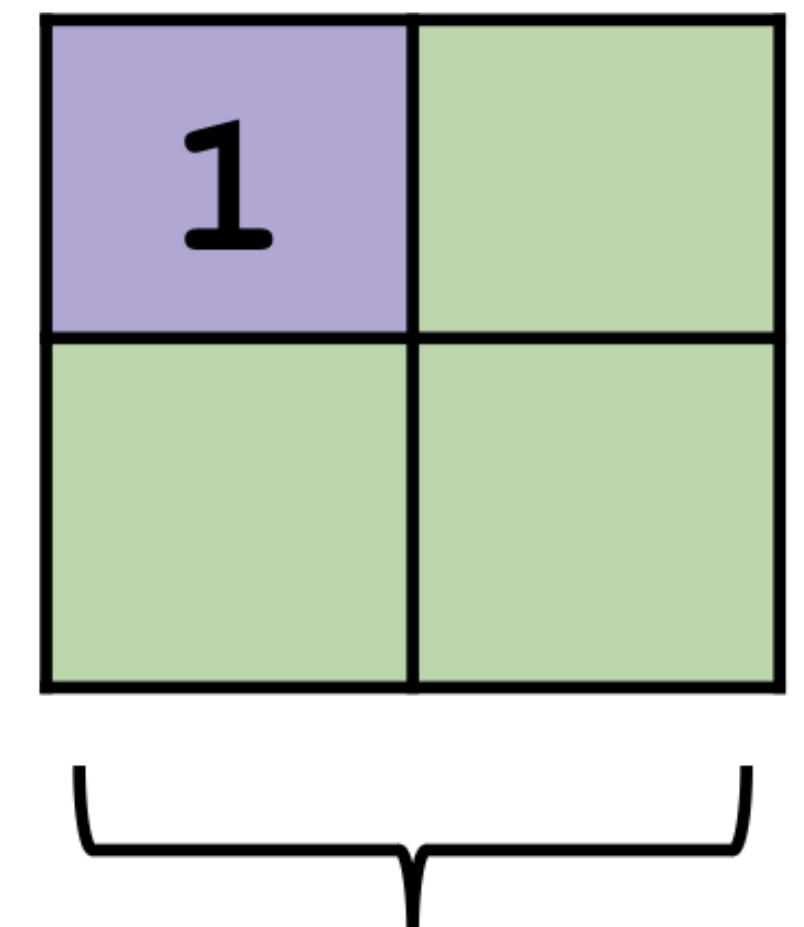
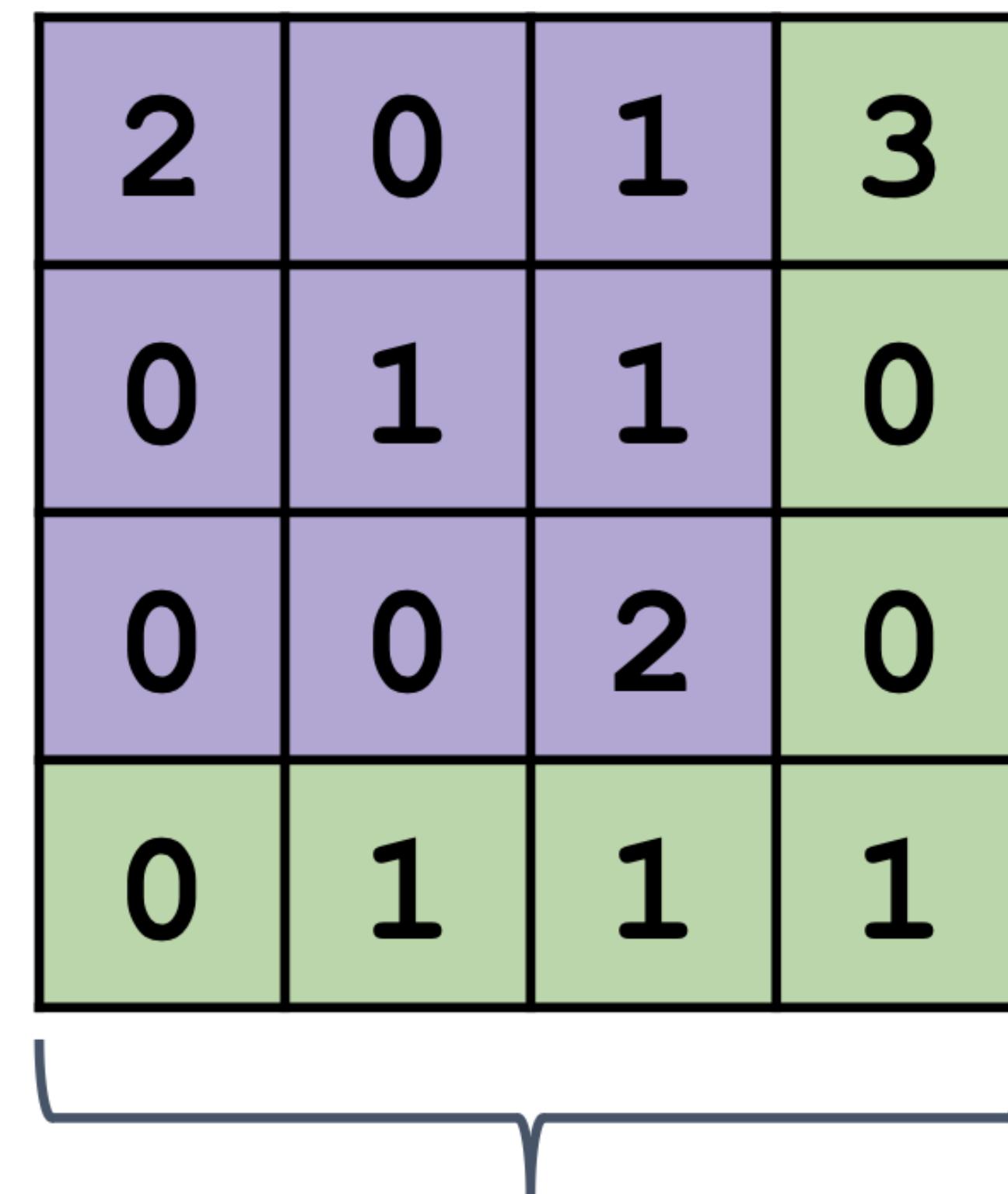
num filters $n = 1$
filter size $f = 3$
stride $s = 1$
padding $p = 0$

$$w' = \frac{4 - 3 + 2 \cdot 0}{1} + 1$$
$$= 1 + 1 = 2$$

Output Size for “VALID” Padding

$$w' = \frac{w - f + 2p}{s} + 1$$

num filters $n = 1$
filter size $f = 3$
stride $s = 1$
padding $p = 0$



$w' = 2$

Output Size for “SAME” Padding

$$w' = \frac{w - f + 2p}{s} + 1$$

num filters $n = 1$

filter size $f = 3$

stride $s = 1$

padding $p = 1^*$

Let $w = 4$

$$w' = \frac{4 - 3 + 2 \cdot 1}{1} + 1$$
$$= 3 + 1 = 4$$

Padding size needs to be determined

Output Size for “SAME” Padding

$$w' = \frac{w - f + 2p}{s} + 1$$

num filters $n = 1$
filter size $f = 3$
stride $s = 1$
padding $p = 1^*$

Padding size needs to be determined

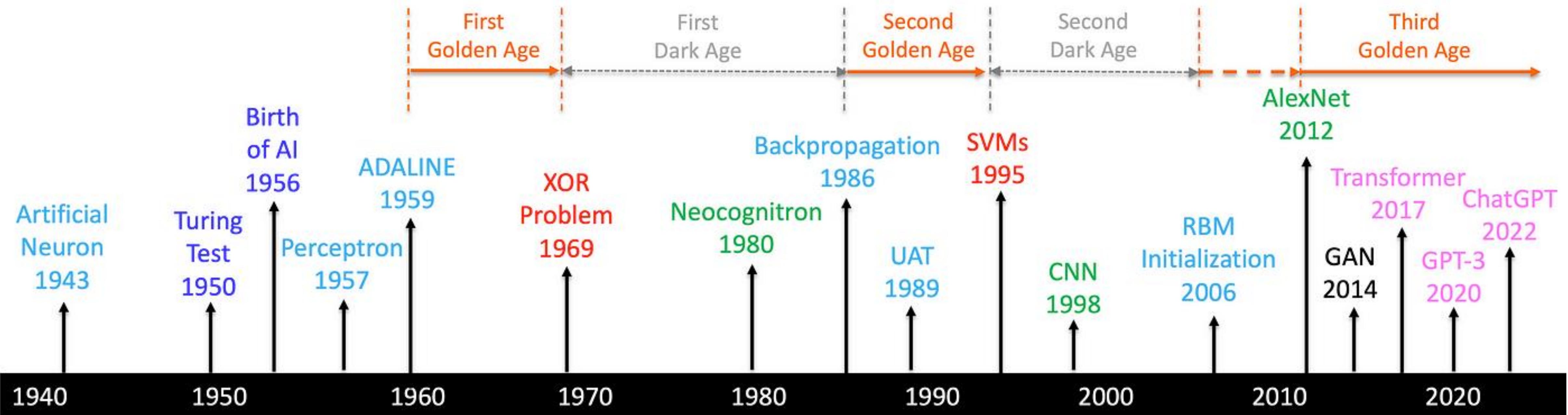
0	0	0	0	0	0
0	2	0	1	3	0
0	1	1	2	3	0
0	4	3	2	1	0
0	8	3	1	3	0
0	0	0	0	0	0

$w = 4$

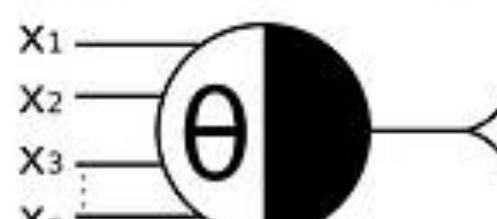
1			

$w' = 4$

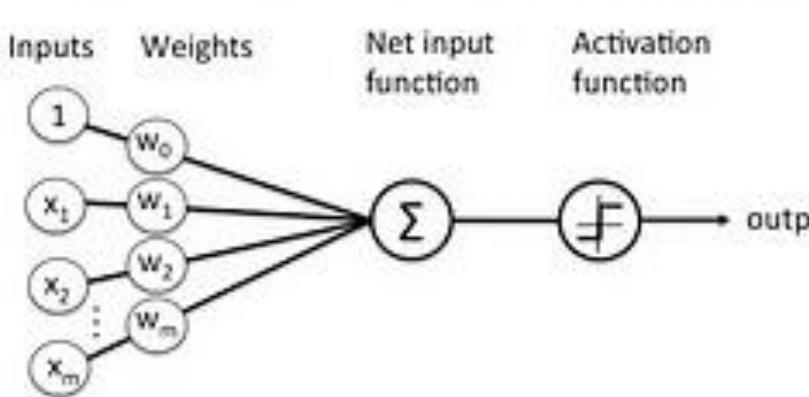
A Brief History of AI with Deep Learning



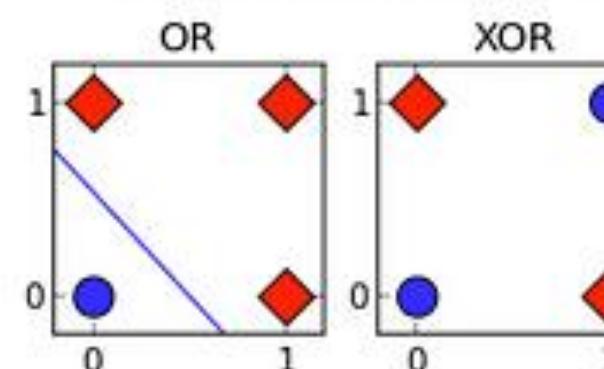
McCulloch-Pitts



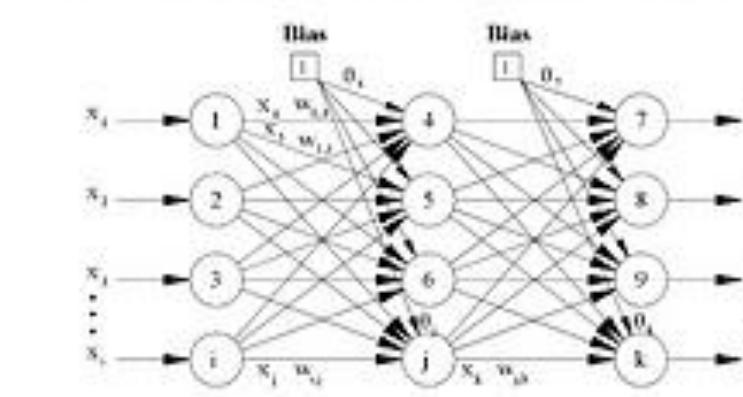
Rosenblatt



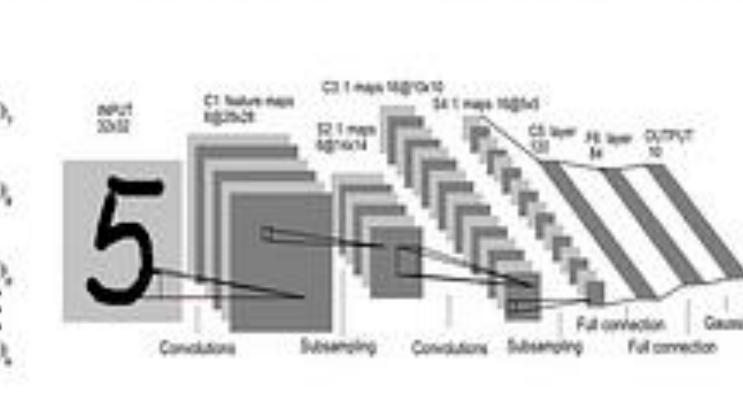
Widrow-Hoff



Minsky-Papert



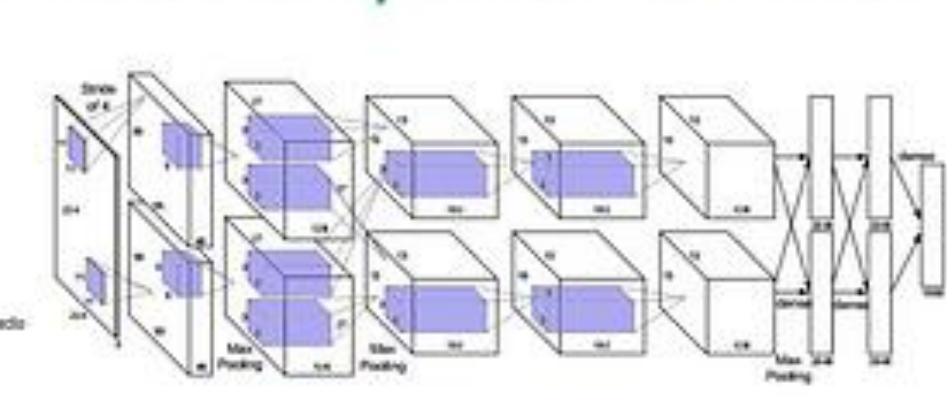
Rumelhart, Hinton et al.



LeCun

Hinton-Ruslan Krizhevsky et al.

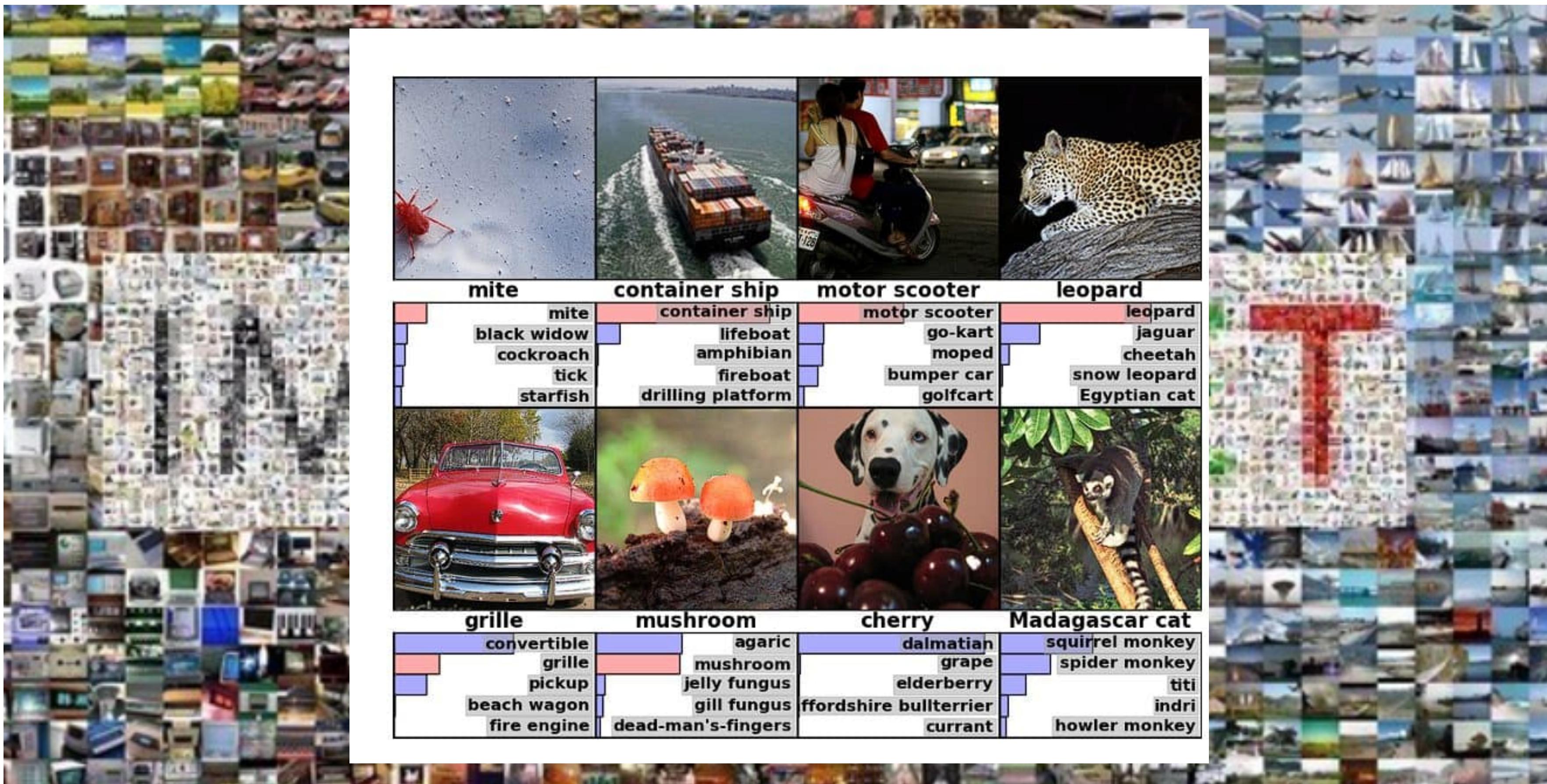
Vaswani



The AlexNet moment

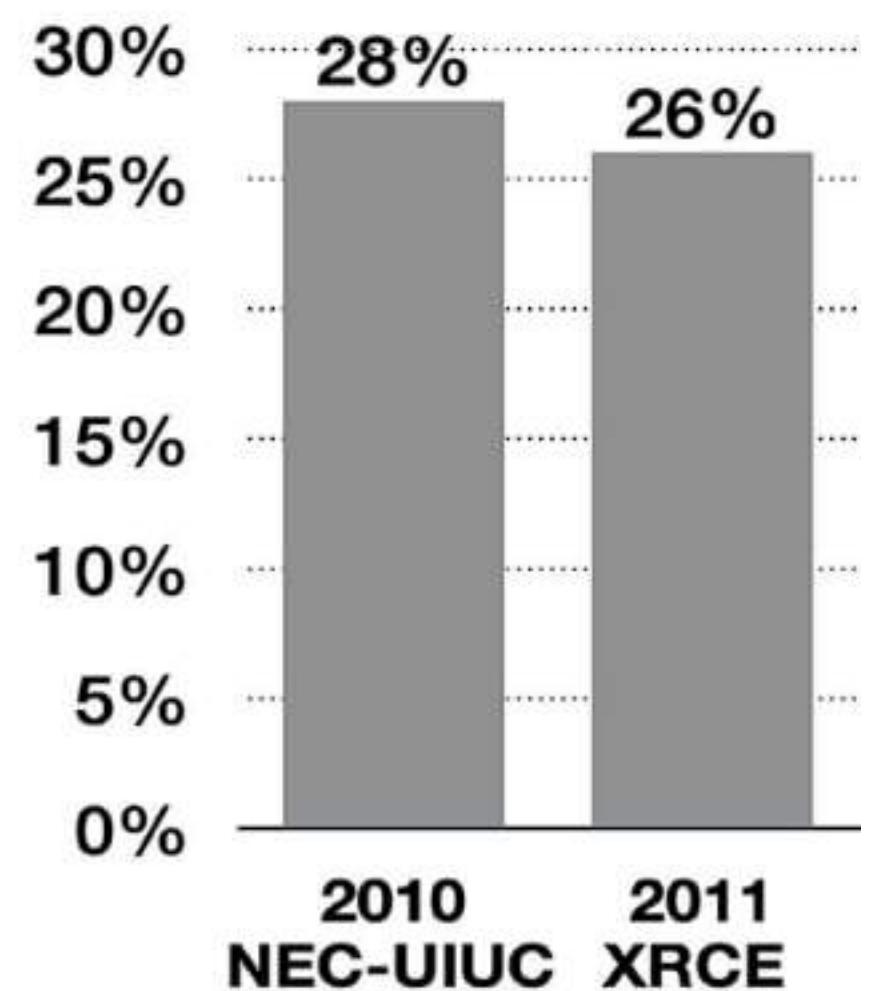


The AlexNet moment

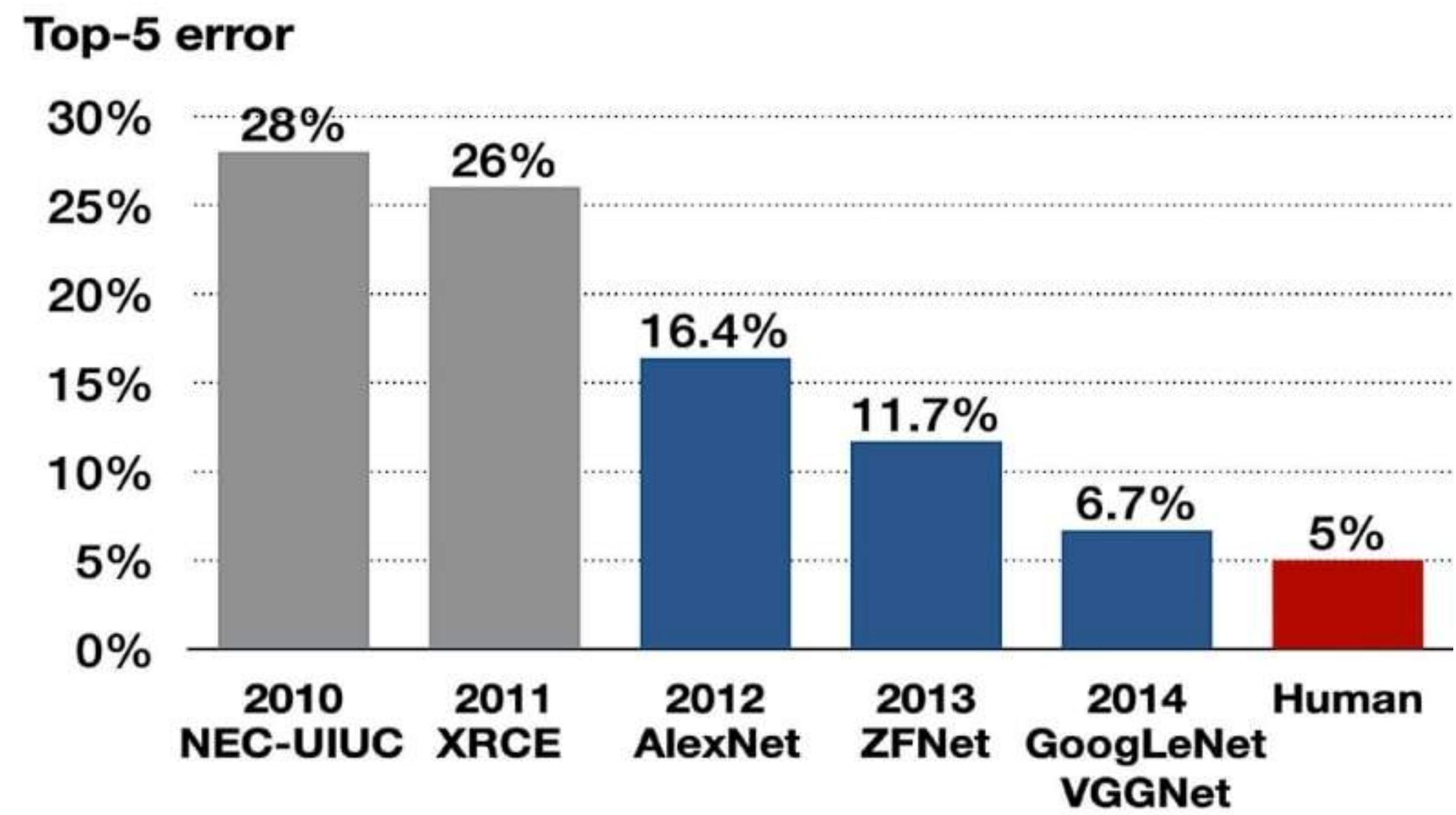


The AlexNet moment

Top-5 error



The AlexNet moment



The AlexNet moment

