Machine Learning Using Tensorflow

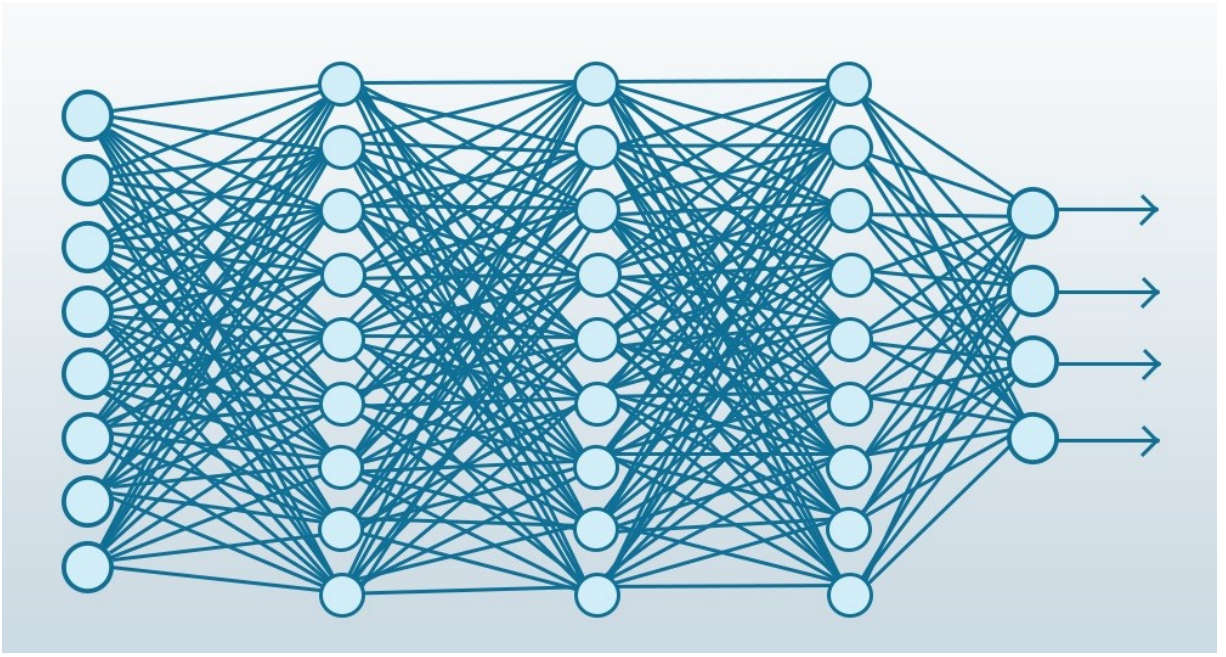
Week 6:
Convolution Neural Network

Shu-Ting Pi, PhD
UC Davis

TensorFlow ™

How many parameters?

256 x256 = 65,536 pixels



input
65,536

hidden1
x 32738

hidden2
x 16384

hidden3
x 8192

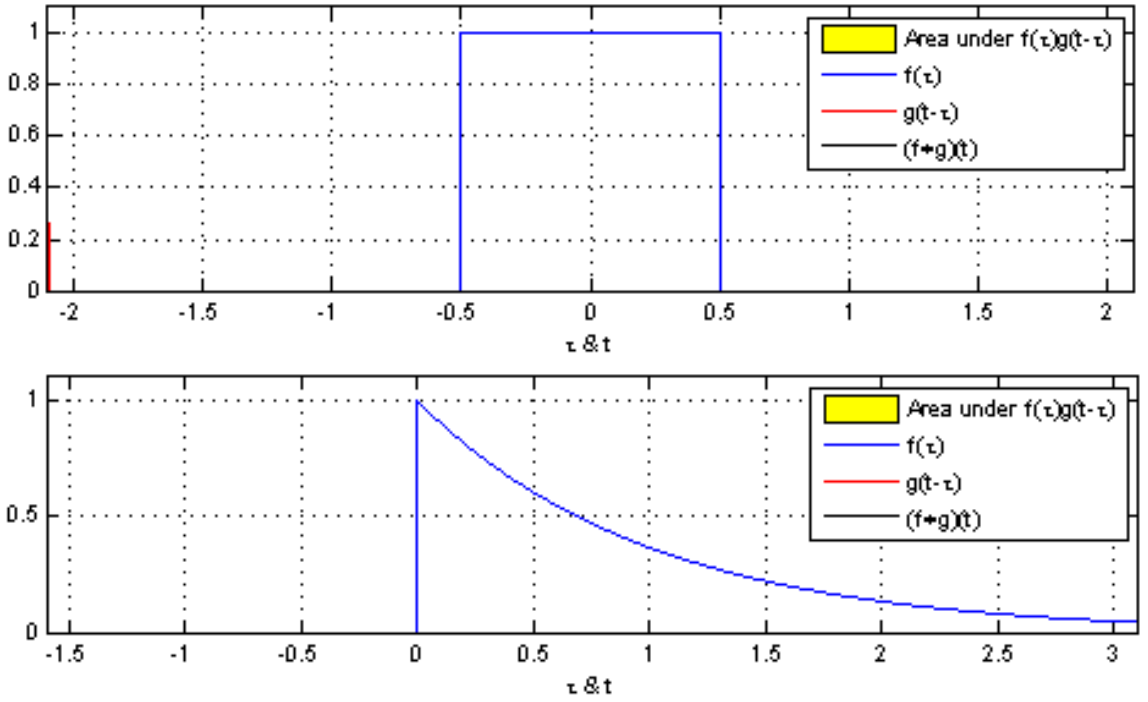
Output
x 10

Trainable Parameters ~ 3 x 10^18 ! Ridiculous !

What is Convolution?

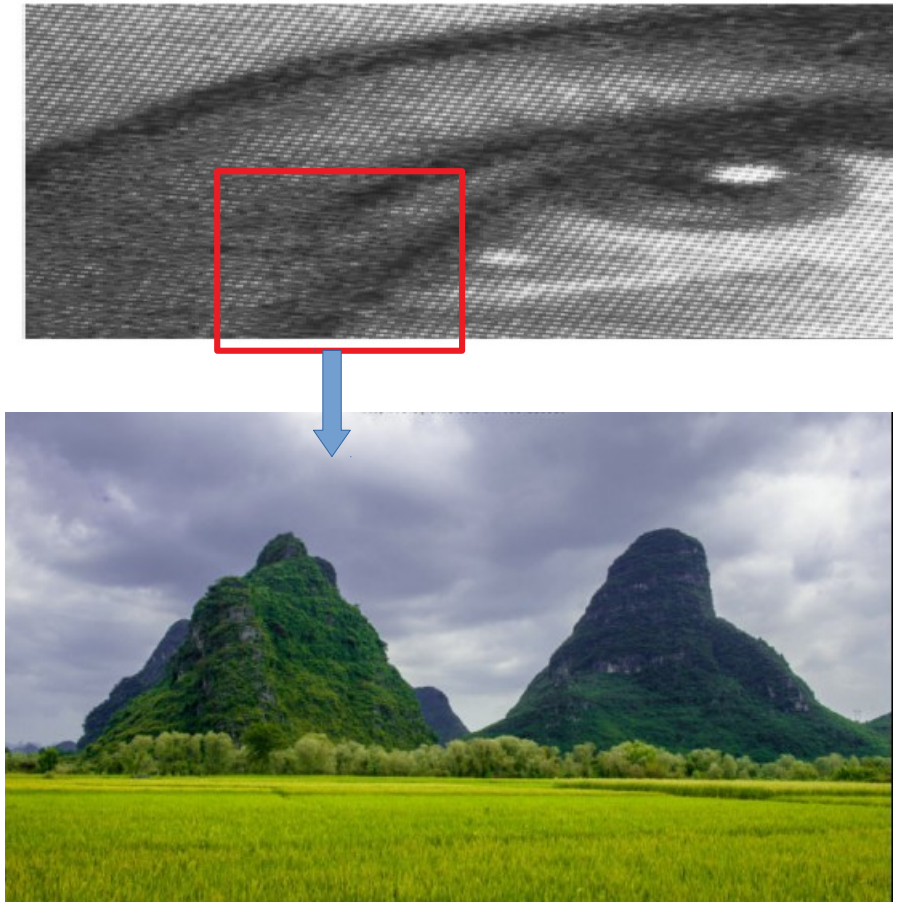
$$(f * g)(n) = \int_{-\infty}^{\infty} f(\tau)g(n-\tau)d\tau$$

$$(f * g)(n) = \sum_{\tau=-\infty}^{\infty} f(\tau)g(n-\tau)$$



Weighted sum of a function !

Use it to smooth an image



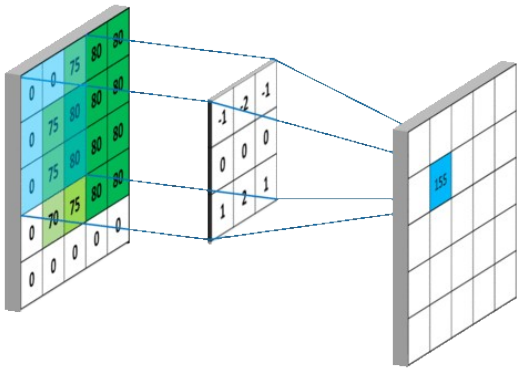
$$f = \begin{bmatrix} a_{0,0} & a_{0,1} & a_{0,2} \\ a_{1,0} & a_{1,1} & a_{1,2} \\ a_{2,0} & a_{2,1} & a_{2,2} \end{bmatrix}$$

$$\begin{bmatrix} a_{0,0} & a_{0,1} & a_{0,2} & \cdots & a_{0,n} \\ a_{1,0} & a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,0} & a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m,0} & a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \Longrightarrow \begin{bmatrix} c_{0,0} & c_{0,1} & c_{0,2} & \cdots & c_{1,n} \\ c_{1,0} & c_{1,1} & c_{1,2} & \cdots & c_{1,n} \\ c_{2,0} & c_{2,1} & c_{2,2} & \cdots & c_{2,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{m,0} & c_{m,1} & c_{m,2} & \cdots & c_{m,n} \end{bmatrix}$$

$$c_{1,1} = f * g$$

$$g = \begin{bmatrix} \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \end{bmatrix}$$

Continous Convolution



Input Volume (+pad 1) (7x7x3)

x[:,:,0]

0	0	0	0	0	0	0
0	0	1	1	0	2	0
0	2	2	2	2	1	0
0	1	0	0	2	0	0
0	0	1	1	0	0	0
0	1	2	0	0	2	0
0	0	0	0	0	0	0

x[:,:,1]

0	0	0	0	0	0	0
0	1	0	2	2	0	0
0	0	0	0	2	0	0
0	1	2	1	2	1	0
0	1	0	0	0	0	0
0	1	2	1	1	1	0
0	0	0	0	0	0	0

x[:,:,2]

0	0	0	0	0	0	0
0	2	1	2	0	0	0
0	1	0	0	1	0	0
0	0	2	1	0	1	0
0	0	1	2	2	2	0
0	2	1	0	0	1	0
0	0	0	0	0	0	0

Filter W0 (3x3x3)

w0[:,:,0]

-1	1	0
0	1	0
0	1	1

w0[:,:,1]

-1	-1	0
0	0	0
0	-1	0

w0[:,:,2]

0	0	-1
0	1	0
1	-1	-1

Bias b0 (1x1x1)

b0[:,:,0]

1

Filter W1 (3x3x3)

w1[:,:,0]

1	1	-1
-1	-1	1
0	-1	1

w1[:,:,1]

0	1	0
-1	0	-1
-1	1	0

w1[:,:,2]

-1	0	0
-1	0	1
-1	0	0

Bias b1 (1x1x1)

b1[:,:,0]

0

Output Volume (3x3x2)

o[:,:,0]

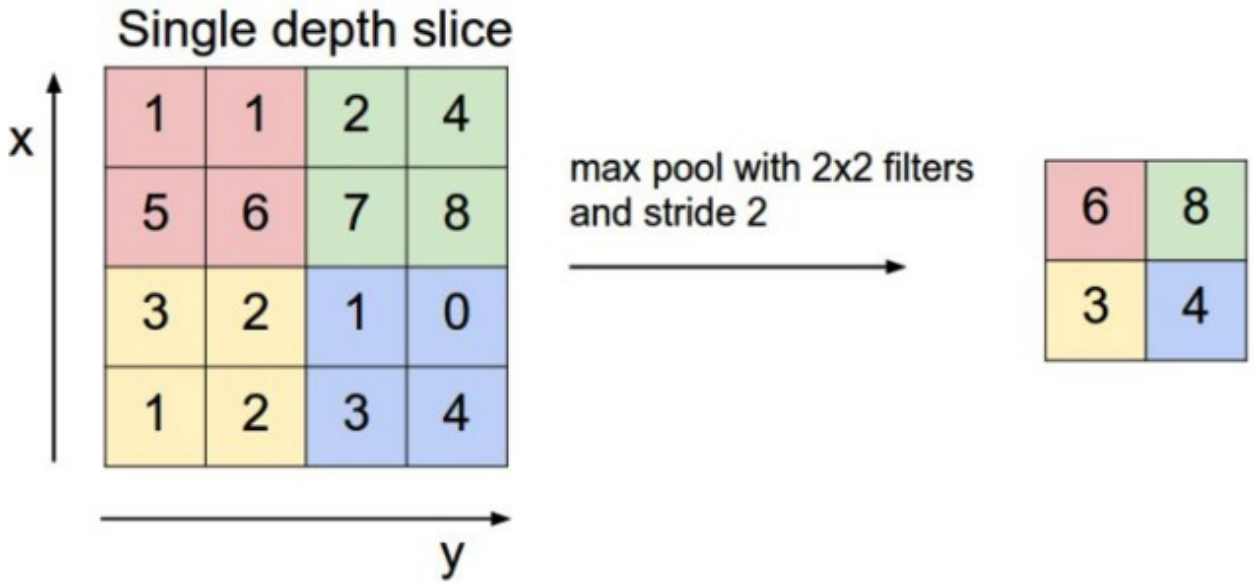
6	7	5
3	-1	-1
2	-1	4

o[:,:,1]

2	-5	-8
1	-4	-4
0	-5	-5

toggle movement

Max Pooling

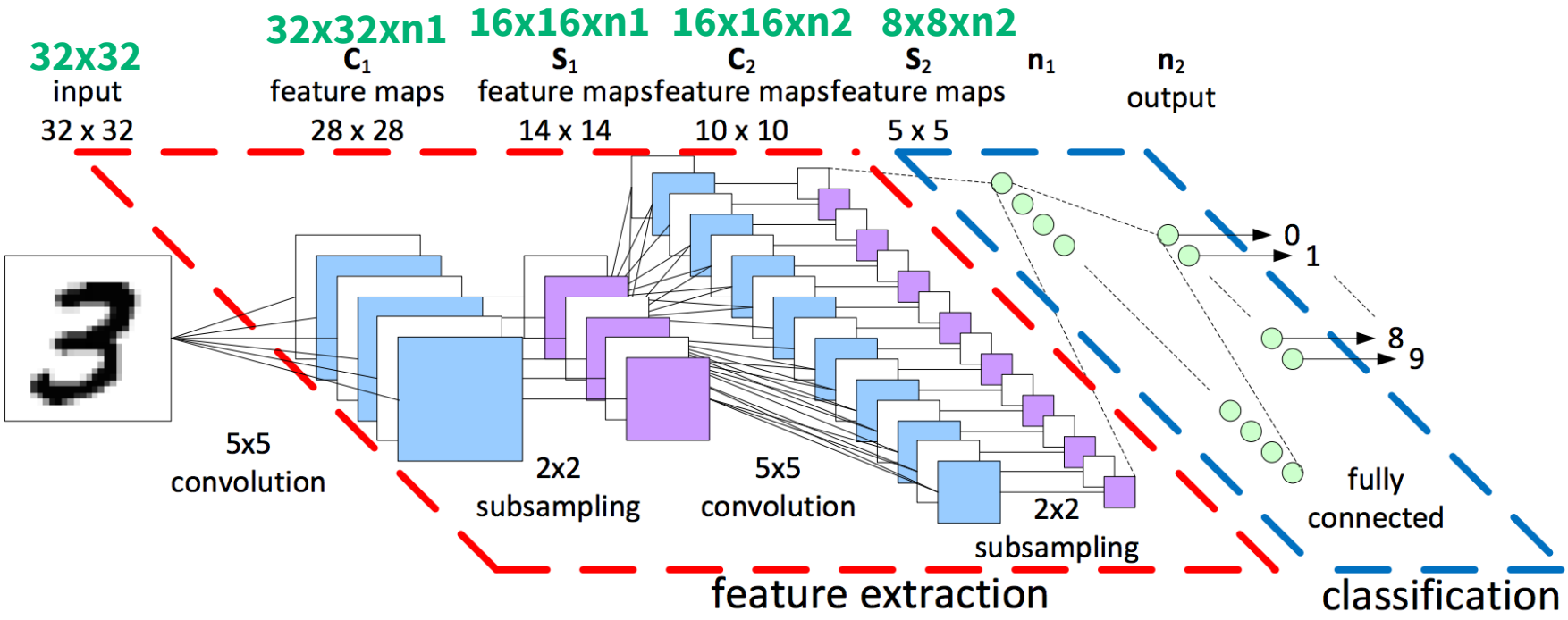


Use the maximal value to represent the main feature of a block

Procedure of CNN

Check this video released by Google:

- **https://www.youtube.com/watch?v=jajksuQW4mc**



All image block shares the same "filter lens", parameters are highly reduced!

Image size: 32x32 → 8x8xn2 (width → deepth)

Why it works?

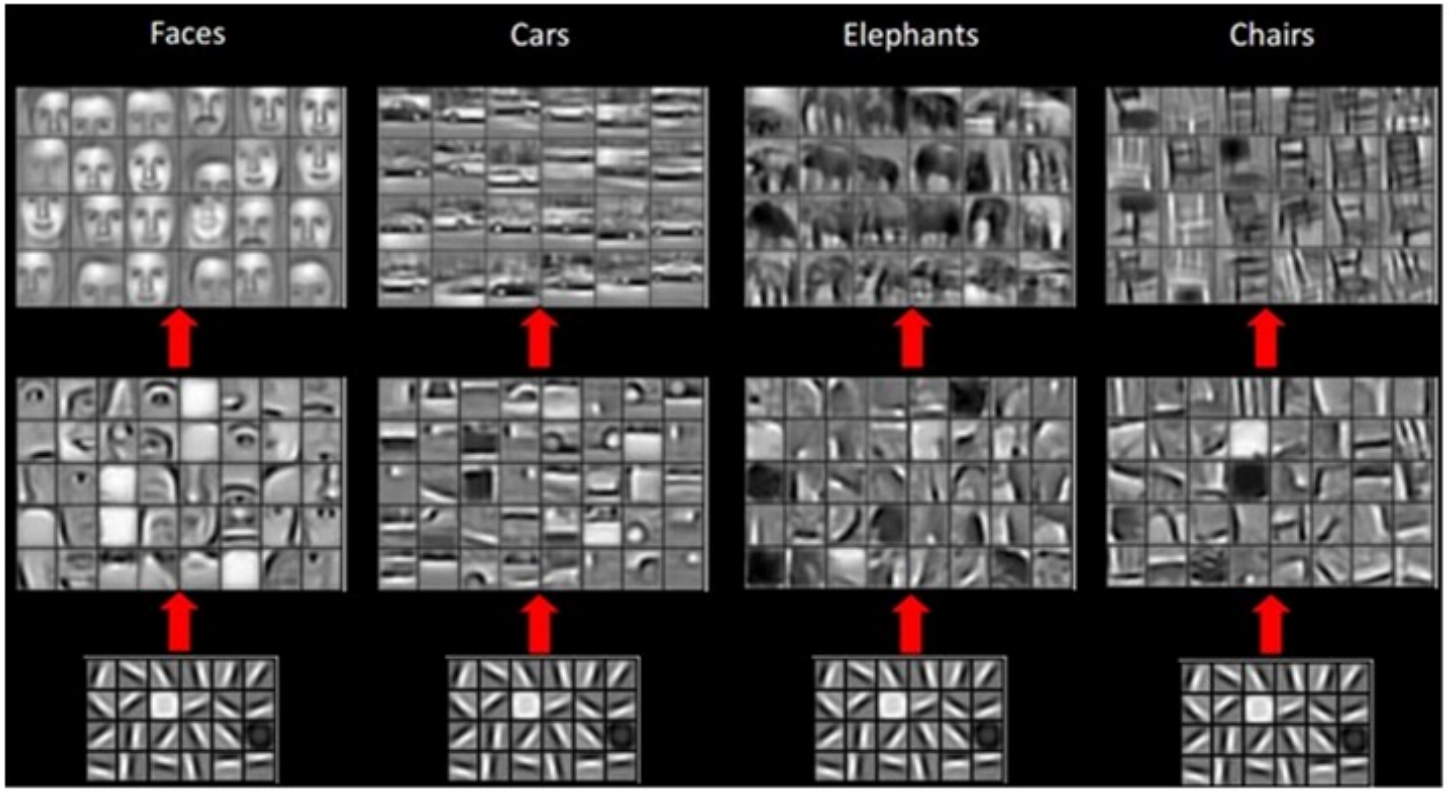
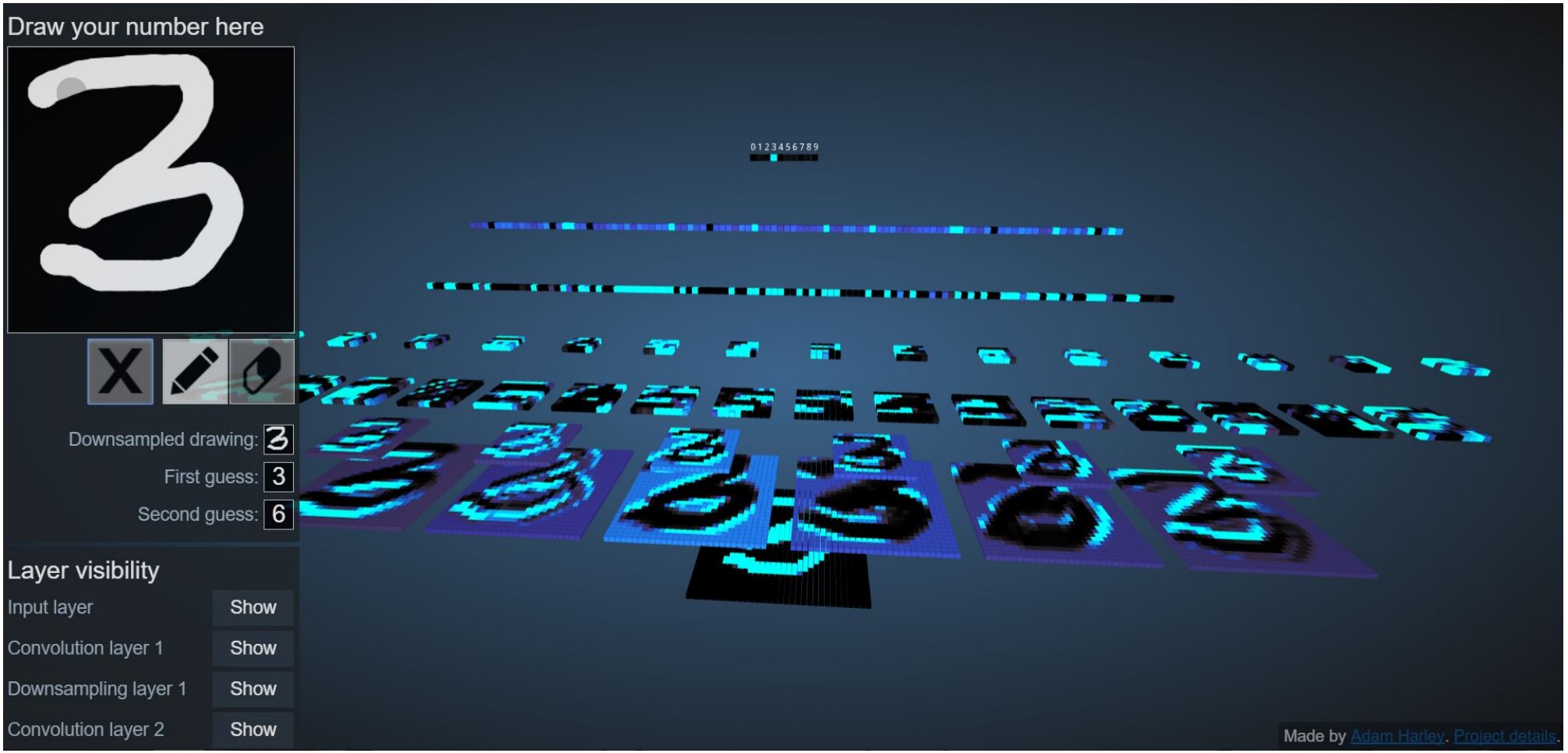


Image = local features merge to global feature !
(actually not only image but almost all kinds of information)

Convolution as a “filter”

http://scs.ryerson.ca/~aharley/vis/conv/



Exercise

code06_ex1_v1:

Using TensorFlow to perform CNN

code06_ex1_v2:

Using TensorFlow to perform CNN

Using t-SNE to cluster the output of flatten layer

code06_ex2_v2:

Using Keras to perform CNN