**SC 4001 CE/CZ 4042 Neural Network and Deep Learning**

*Last update: 10 October 2023 9:10AM*

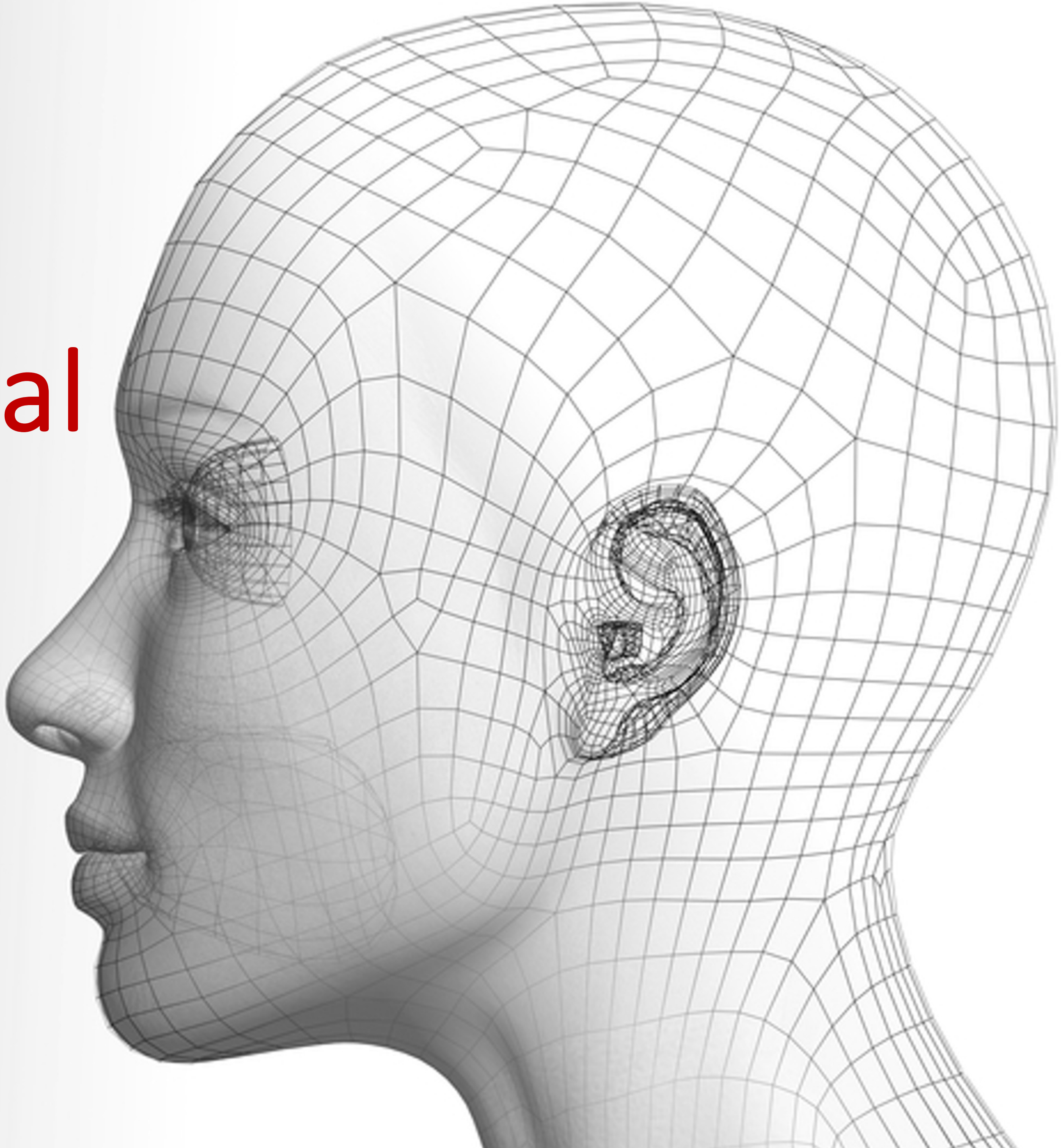
# Tutorial 6
# Convolutional Neural Networks I

Chen-Change Loy

吕健勤

https://www.mmlab-ntu.com/person/ccloy/

https://twitter.com/ccloy



# Question 1

1. The first hidden layer of a convolution neural network (CNN) has a convolution layer consisting of two feature maps with filters $\boldsymbol{w_1}$ and $\boldsymbol{w_2}$ and biases = 0.1, and neurons having sigmoid activation functions, and a pooling layer with a pooling windows of size 2x2:

$$\boldsymbol{w}_1 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \text{ and } \boldsymbol{w}_2 = \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

The input layer is of 6x6 size and receives an input image $\boldsymbol{I}$:

Filters=weights

$$\boldsymbol{I} = \begin{pmatrix} 0.7 & 0.1 & 0.2 & 0.3 & 0.3 & 05 \\ 0.8 & 0.1 & 0.3 & 0.5 & 0.1 & 0.0 \\ 1.0 & 0.2 & 0.0 & 0.3 & 0.2 & 0.7 \\ 0.8 & 0.1 & 0.5 & 0.6 & 0.3 & 0.4 \\ 0.1 & 0.0 & 0.9 & 0.3 & 0.3 & 0.2 \\ 1.0 & 0.1 & 0.4 & 0.5 & 0.2 & 0.8 \end{pmatrix}$$

a. Find the outputs at the first convolution layer if
   i. padding = 0 and strides = (1,1)
   ii. padding = 1 and strides = (2,2)

b. Find the outputs at the first pooling layer for Part (a), assuming strides of $(2, 2)$ and pooling is
   i. max pooling
   ii. mean pooling

t6q1.ipynb

# Question 1

$$\boldsymbol{w}_1 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \boldsymbol{w}_2 = \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \text{ and } b = 0.1.$$

$$\boldsymbol{I} = \begin{pmatrix} 0.7 & 0.1 & 0.2 & 0.3 & 0.3 & 05 \\ 0.8 & 0.1 & 0.3 & 0.5 & 0.1 & 0.0 \\ 1.0 & 0.2 & 0.0 & 0.3 & 0.2 & 0.7 \\ 0.8 & 0.1 & 0.5 & 0.6 & 0.3 & 0.4 \\ 0.1 & 0.0 & 0.9 & 0.3 & 0.3 & 0.2 \\ 1.0 & 0.1 & 0.4 & 0.5 & 0.2 & 0.8 \end{pmatrix}$$

Synaptic inputs to the feature map with filter $\boldsymbol{w}_1$:
$\boldsymbol{u}_1 = Conv(\boldsymbol{I}, \boldsymbol{w}_1) + b_1$

Padding = 0 and strides = (1,1)

$$\boldsymbol{u}_1(1,1) = 0.7{\times}0 + 0.1{\times}1 + 0.2{\times}1 + 0.8{\times}1 + 0.1{\times}0 + 0.3{\times}1 + 1.0{\times}1 + 0.2{\times}1 + 0.0{\times}0 + 0.1 = 2.7$$
$$\boldsymbol{u}_1(1,2) = 0.1{\times}0 + 0.2{\times}1 + 0.3{\times}1 + 0.1{\times}1 + 0.3{\times}0 + 0.5{\times}1 + 0.2{\times}1 + 0.0{\times}1 + 0.3{\times}0 + 0.1 = 1.4$$
$$\boldsymbol{u}_1(2,1) = 0.8{\times}0 + 0.1{\times}1 + 0.3{\times}1 + 1.0{\times}1 + 0.2{\times}0 + 0.0{\times}1 + 0.8{\times}1 + 0.1{\times}1 + 0.5{\times}0 + 0.1 = 2.4$$

# Question 1

$$\boldsymbol{u}_1 = \begin{pmatrix} 2.7 & 1.4 & 1.4 & 1.9 \\ 2.4 & 2.0 & 2.0 & 2.1 \\ 1.7 & 2.0 & 2.6 & 2.6 \\ 2.8 & 2.0 & 3.1 & 2.0 \end{pmatrix}$$

Similarly, $\boldsymbol{u}_2 = Conv(\boldsymbol{I}, \boldsymbol{w}_2) + b_2 = \begin{pmatrix} 0.3 & 0.0 & -0.2 & 2.0 \\ 0.3 & 0.4 & 0.6 & 0.8 \\ -0.1 & 0.8 & 1.1 & -0.3 \\ 0.2 & -0.1 & -0.2 & 0.3 \end{pmatrix}$

Feature maps at the convolutional layer

$$\boldsymbol{y}_1 = f(\boldsymbol{u}_1) = \frac{1}{1+e^{-\boldsymbol{u}_1}} = \begin{pmatrix} 0.94 & 0.8 & 0.8 & 0.87 \\ 0.92 & 0.88 & 0.88 & 0.89 \\ 0.85 & 0.88 & 0.93 & 0.93 \\ 0.94 & 0.88 & 0.96 & 0.88 \end{pmatrix}$$

$$\boldsymbol{y}_2 = f(\boldsymbol{u}_2) = \frac{1}{1+e^{-\boldsymbol{u}_2}} = \begin{pmatrix} 0.57 & 0.5 & 0.45 & 0.55 \\ 0.57 & 0.60 & 0.65 & 0.69 \\ 0.48 & 0.69 & 0.75 & 0.43 \\ 0.55 & 0.48 & 0.45 & 0.57 \end{pmatrix}$$

Output spatial size
= (Input size – Filter size +2(Padding))/Stride + 1

= (6 – 3)/1 + 1 = 4

# Question 1

Feature maps at the convolutional layer:

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} \boxed{\begin{matrix} 0.94 & 0.8 \\ 0.92 & 0.88 \end{matrix}} & \begin{matrix} 0.8 & 0.87 \\ 0.88 & 0.89 \end{matrix} \\ \begin{matrix} 0.85 & 0.88 \\ 0.94 & 0.88 \end{matrix} & \begin{matrix} 0.93 & 0.93 \\ 0.96 & 0.88 \end{matrix} \end{pmatrix} \\ \begin{pmatrix} 0.57 & 0.5 & 0.45 & 0.55 \\ 0.57 & 0.60 & 0.65 & 0.69 \\ 0.48 & 0.69 & 0.75 & 0.43 \\ 0.55 & 0.48 & 0.45 & 0.57 \end{pmatrix} \end{pmatrix}$$

Pooling 2×2 and strides = 2

Max-pooling:

$$o = \begin{pmatrix} \begin{pmatrix} 0.94 & 0.89 \\ 0.94 & 0.96 \end{pmatrix} \\ \begin{pmatrix} 0.60 & 0.69 \\ 0.69 & 0.75 \end{pmatrix} \end{pmatrix}$$

Mean-pooling:

$$p_{ave} = \begin{pmatrix} \begin{pmatrix} 0.88 & 0.86 \\ 0.89 & 0.92 \end{pmatrix} \\ \begin{pmatrix} 0.56 & 0.58 \\ 0.55 & 0.55 \end{pmatrix} \end{pmatrix}$$

# Question 1

$$\boldsymbol{w}_1 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \boldsymbol{w}_2 = \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \text{ and } b = 0.1.$$

$$\begin{array}{ccc}
 & & \begin{matrix} 0 & 0 & 0 \end{matrix} \\
\boldsymbol{I} = & \begin{matrix} 0 \\ 0 \\ \\ \\ \\ \\ \end{matrix} & \begin{pmatrix} 0.7 & 0.1 & 0.2 & 0.3 & 0.3 & 05 \\ 0.8 & 0.1 & 0.3 & 0.5 & 0.1 & 0.0 \\ 1.0 & 0.2 & 0.0 & 0.3 & 0.2 & 0.7 \\ 0.8 & 0.1 & 0.5 & 0.6 & 0.3 & 0.4 \\ 0.1 & 0.0 & 0.9 & 0.3 & 0.3 & 0.2 \\ 1.0 & 0.1 & 0.4 & 0.5 & 0.2 & 0.8 \end{pmatrix}
\end{array}$$

Synaptic inputs to the feature map with filter $\boldsymbol{w}_1$:
$\boldsymbol{u}_1 = Conv(\boldsymbol{I}, \boldsymbol{w}_1) + b_1$

Padding = 1 and strides = (2,2)

$$\boldsymbol{u}_1(1,1) = 0.0\times0 + 0.0\times1 + 0.0\times1 + 0.0\times1 + 0.7\times0 + 0.1\times1 + 0.0\times1 + 0.8\times1 + 0.1\times0 + 0.1 = 1.0$$

# Question 1

$$\boldsymbol{w}_1 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \boldsymbol{w}_2 = \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \text{ and } b = 0.1.$$

$$\boldsymbol{I} = \begin{array}{c} \begin{matrix} \phantom{0.7} & 0 & 0 & 0 & \phantom{0.3} & \phantom{05} \end{matrix} \\ \begin{pmatrix} 0.7 & 0.1 & 0.2 & 0.3 & 0.3 & 05 \\ 0.8 & 0.1 & 0.3 & 0.5 & 0.1 & 0.0 \\ 1.0 & 0.2 & 0.0 & 0.3 & 0.2 & 0.7 \\ 0.8 & 0.1 & 0.5 & 0.6 & 0.3 & 0.4 \\ 0.1 & 0.0 & 0.9 & 0.3 & 0.3 & 0.2 \\ 1.0 & 0.1 & 0.4 & 0.5 & 0.2 & 0.8 \end{pmatrix} \end{array}$$

Synaptic inputs to the feature map with filter $\boldsymbol{w}_1$:
$\boldsymbol{u}_1 = Conv(\boldsymbol{I}, \boldsymbol{w}_1) + b_1$

Padding = 1 and strides = (2,2)

$$\boldsymbol{u}_1(1,1) = 0.0\times 0 + 0.0\times 1 + 0.0\times 1 + 0.0\times 1 + 0.7\times 0 + 0.1\times 1 + 0.0\times 1 + 0.8\times 1 + 0.1\times 0 + 0.1 = 1.0$$

$$\boldsymbol{u}_1(1,2) = 0.0\times 0 + 0.0\times 1 + 0.0\times 1 + 0.1\times 1 + 0.2\times 0 + 0.3\times 1 + 0.1\times 1 + 0.3\times 1 + 0.5\times 0 + 0.1 = 0.9$$

# Question 1

$$\boldsymbol{w}_1 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \boldsymbol{w}_2 = \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \text{ and } b = 0.1.$$

$$\begin{array}{c} \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad 0 \quad\;\; 0 \quad\;\; 0 \\ \boldsymbol{I} = \begin{pmatrix} 0.7 & 0.1 & 0.2 & 0.3 & 0.3 & 05 \\ 0.8 & 0.1 & 0.3 & 0.5 & 0.1 & 0.0 \\ 1.0 & 0.2 & 0.0 & 0.3 & 0.2 & 0.7 \\ 0.8 & 0.1 & 0.5 & 0.6 & 0.3 & 0.4 \\ 0.1 & 0.0 & 0.9 & 0.3 & 0.3 & 0.2 \\ 1.0 & 0.1 & 0.4 & 0.5 & 0.2 & 0.8 \end{pmatrix} \end{array}$$

Synaptic inputs to the feature map with filter $\boldsymbol{w}_1$:
$\boldsymbol{u}_1 = Conv(\boldsymbol{I}, \boldsymbol{w}_1) + b_1$

Padding = 1 and strides = (2,2)

$$\boldsymbol{u}_1(1,1) = 0.0{\times}0 + 0.0{\times}1 + 0.0{\times}1 + 0.0{\times}1 + 0.7{\times}0 + 0.1{\times}1 + 0.0{\times}1 + 0.8{\times}1 + 0.1{\times}0 + 0.1 = 1.0$$
$$\boldsymbol{u}_1(1,2) = 0.0{\times}0 + 0.0{\times}1 + 0.0{\times}1 + 0.1{\times}1 + 0.2{\times}0 + 0.3{\times}1 + 0.1{\times}1 + 0.3{\times}1 + 0.5{\times}0 + 0.1 = 0.9$$
$$\boldsymbol{u}_1(1,3) = 0.0{\times}0 + 0.0{\times}1 + 0.0{\times}1 + 0.3{\times}1 + 0.3{\times}0 + 0.5{\times}1 + 0.5{\times}1 + 0.1{\times}1 + 0.0{\times}0 + 0.1 = 1.5$$

# Question 1

$$\boldsymbol{w}_1 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \boldsymbol{w}_2 = \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \text{ and } b = 0.1.$$

$$\boldsymbol{I} = \begin{array}{c} \\ 0 \\ 0 \\ 0 \\ \\ \\ \end{array} \begin{pmatrix} 0.7 & 0.1 & 0.2 & 0.3 & 0.3 & 05 \\ 0.8 & 0.1 & 0.3 & 0.5 & 0.1 & 0.0 \\ 1.0 & 0.2 & 0.0 & 0.3 & 0.2 & 0.7 \\ 0.8 & 0.1 & 0.5 & 0.6 & 0.3 & 0.4 \\ 0.1 & 0.0 & 0.9 & 0.3 & 0.3 & 0.2 \\ 1.0 & 0.1 & 0.4 & 0.5 & 0.2 & 0.8 \end{pmatrix}$$

Synaptic inputs to the feature map with filter $\boldsymbol{w}_1$:
$\boldsymbol{u}_1 = Conv(\boldsymbol{I}, \boldsymbol{w}_1) + b_1$

Padding = 1 and strides = (2,2)

$$\boldsymbol{u}_1(1,1) = 0.0\times0 + 0.0\times1 + 0.0\times1 + 0.0\times1 + 0.7\times0 + 0.1\times1 + 0.0\times1 + 0.8\times1 + 0.1\times0 + 0.1 = 1.0$$

$$\boldsymbol{u}_1(1,2) = 0.0\times0 + 0.0\times1 + 0.0\times1 + 0.1\times1 + 0.2\times0 + 0.3\times1 + 0.1\times1 + 0.3\times1 + 0.5\times0 + 0.1 = 0.9$$

$$\boldsymbol{u}_1(1,3) = 0.0\times0 + 0.0\times1 + 0.0\times1 + 0.3\times1 + 0.3\times0 + 0.5\times1 + 0.5\times1 + 0.1\times1 + 0.0\times0 + 0.1 = 1.5$$

$$\boldsymbol{u}_1(2,1) = 0.0\times0 + 0.8\times1 + 0.1\times1 + 0.0\times1 + 1.0\times0 + 0.2\times1 + 0.0\times1 + 0.8\times1 + 0.1\times0 + 0.1 = 2.0$$

# Question 1

$$\boldsymbol{u}_1 = \begin{pmatrix} 1.0 & 0.9 & 1.5 \\ 2.0 & 2.0 & 2.1 \\ 2.0 & 2.0 & 2.0 \end{pmatrix}$$

Similarly, $\boldsymbol{u}_2 = Conv(\boldsymbol{I}, \boldsymbol{w}_2) + b_2 = \begin{pmatrix} 1.0 & 1.0 & 0.7 \\ 0.1 & 0.4 & 0.8 \\ 0.3 & -0.1 & 0.3 \end{pmatrix}$

Feature maps at the convolutional layer

$$\boldsymbol{y}_1 = f(\boldsymbol{u}_1) = \frac{1}{1 + e^{-\boldsymbol{u}_1}} = \begin{pmatrix} 0.73 & 0.71 & 0.82 \\ 0.88 & 0.88 & 0.89 \\ 0.88 & 0.88 & 0.88 \end{pmatrix}$$

$$\boldsymbol{y}_2 = f(\boldsymbol{u}_2) = \frac{1}{1 + e^{-\boldsymbol{u}_2}} = \begin{pmatrix} 0.73 & 0.73 & 0.67 \\ 0.52 & 0.60 & 0.69 \\ 0.57 & 0.48 & 0.57 \end{pmatrix}$$

# Question 1

Feature maps at the convolutional layer:

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} \boxed{\begin{matrix} 0.73 & 0.71 \\ 0.88 & 0.88 \end{matrix}} & \begin{matrix} 0.82 \\ 0.89 \end{matrix} \\ \begin{matrix} 0.88 & 0.88 \end{matrix} & 0.88 \end{pmatrix} \\ \begin{pmatrix} \boxed{\begin{matrix} 0.73 & 0.73 \\ 0.52 & 0.60 \end{matrix}} & \begin{matrix} 0.67 \\ 0.69 \end{matrix} \\ \begin{matrix} 0.57 & 0.48 \end{matrix} & 0.57 \end{pmatrix} \end{pmatrix}$$

Pooling VALID 2x2 strides = 2:

Max-pooling:

$$o = \begin{pmatrix} (0.88) \\ (0.73) \end{pmatrix}$$

Mean-pooling:

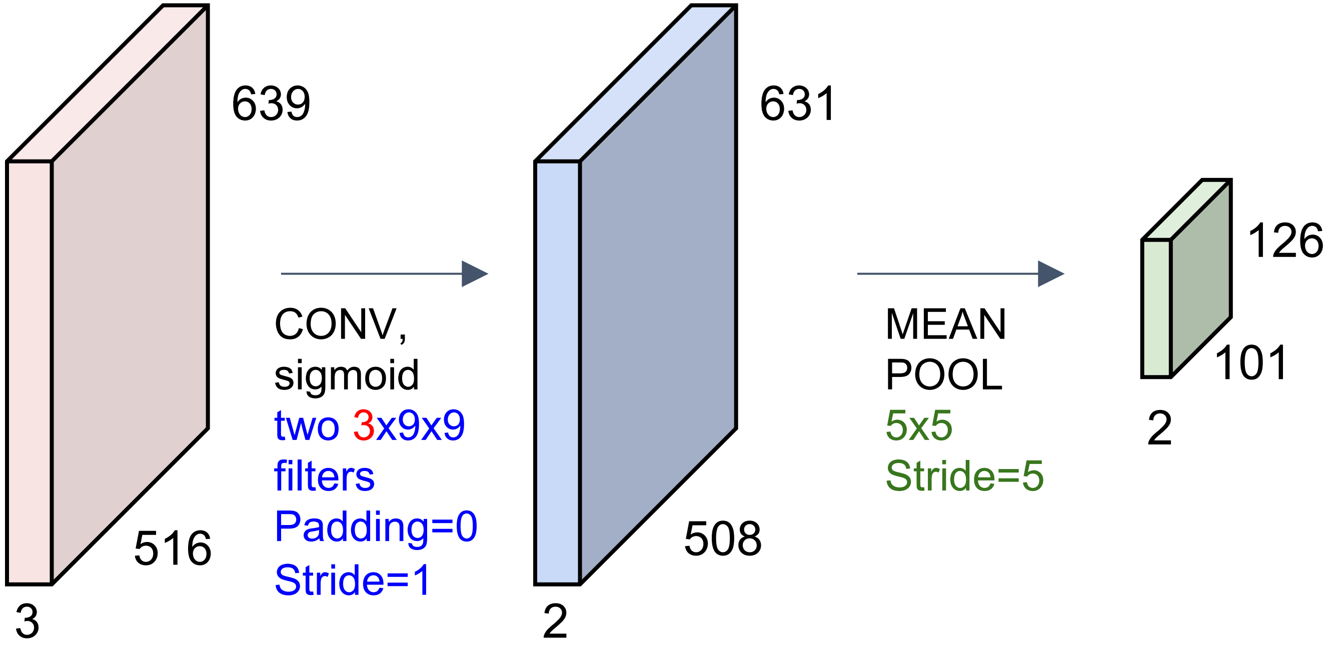
$$p_{ave} = \begin{pmatrix} (0.80) \\ (0.65) \end{pmatrix}$$

# Question 2

2. Given ‘3wolfmoon.jpg’ color image of size $639 \times 516$.
   a. Initialize weights and biases of a convolutional layer with two kernels of size $9 \times 9$. Note that the input image is in color and has three channels.
   b. Display the feature maps at the convolution layer, assuming sigmoid activation functions. Use VALID padding and strides = 1.
   c. Display the outputs of a mean pooling layer with a pooling window size $5 \times 5$ and strides = 5.

t6q2.ipynb

# Question 2


639
516
3
CONV,
sigmoid
two 3x9x9
filters
Padding=0
Stride=1
631
508
2
MEAN
POOL
5x5
Stride=5
126
101
2

# Question 2

```
# open image and normalize
img = Image.open('3wolfmoon.jpg')
img = np.asarray(img, dtype='float32') / 256.
print(img.shape)
```

```
---------------------------------------------------------------------------
FileNotFoundError                         Traceback (most recent call last)
<ipython-input-2-becdc6409bf5> in <module>()
      1 # open image and normalize
----> 2 img = Image.open('3wolfmoon.jpg')
      3 img = np.asarray(img, dtype='float32') / 256.
      4 print(img.shape)

/usr/local/lib/python3.7/dist-packages/PIL/Image.py in open(fp, mode)
   2841
   2842     if filename:
-> 2843         fp = builtins.open(filename, "rb")
   2844         exclusive_fp = True
   2845

FileNotFoundError: [Errno 2] No such file or directory: '3wolfmoon.jpg'
```

You need to mount Google Drive locally and use the correct path

https://colab.research.google.com/notebooks/io.ipynb#scrollTo=u22w3BFiOveA

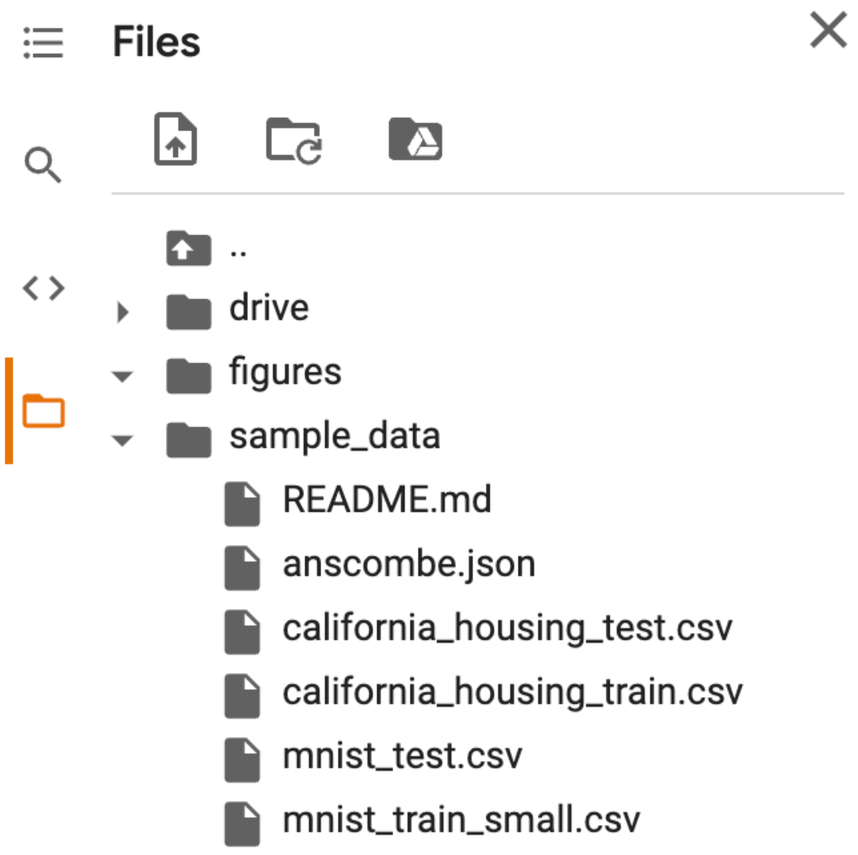
# Question 2

```python
# mount google drive
from google.colab import drive
drive.mount('/content/drive')
```

```
Mounted at /content/drive
```

```python
# open image and normalize
img = Image.open('/content/drive/MyDrive/3wolfmoon.jpg')
img = np.asarray(img, dtype='float32') / 256.
print(img.shape)
```

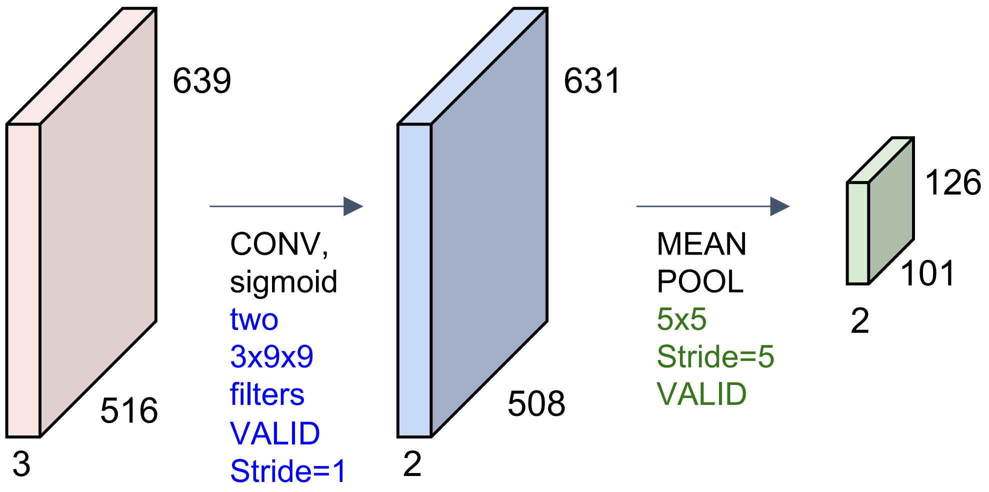
```
(639, 516, 3)
```

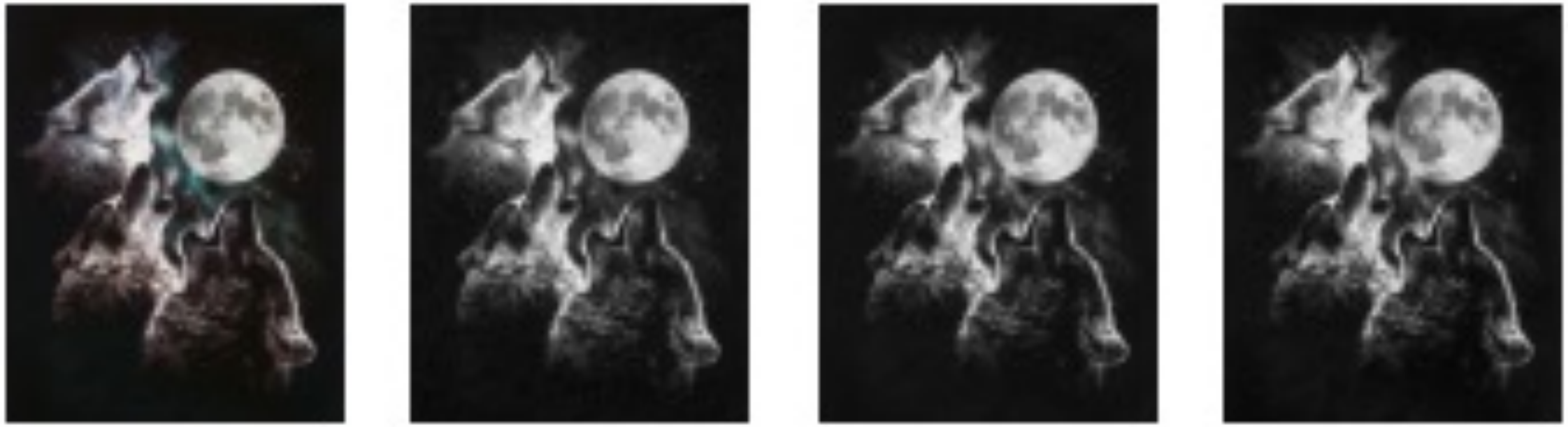



You can copy the path from the browser

# Question 2

Plot the three channels of the input image



Original image

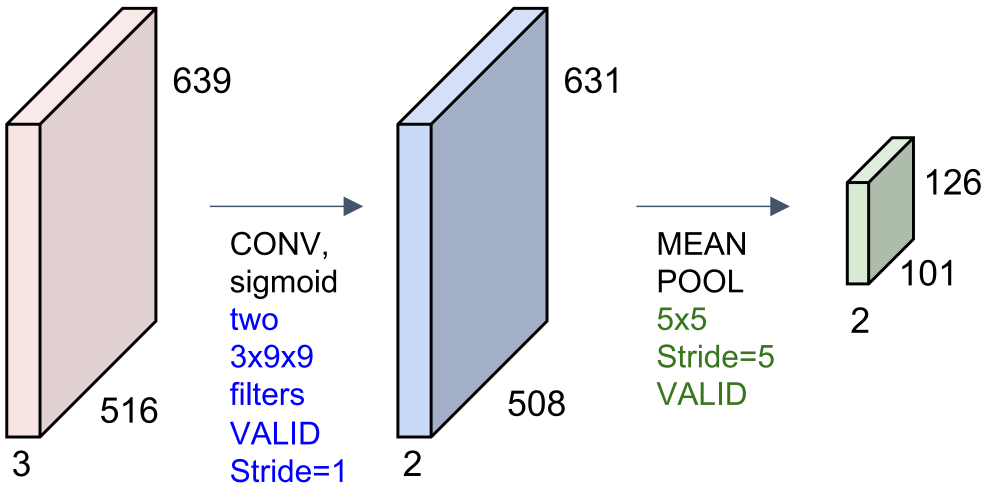
Red

Green

Blue

# Question 2

Plot the original image and the first and second components of conv output





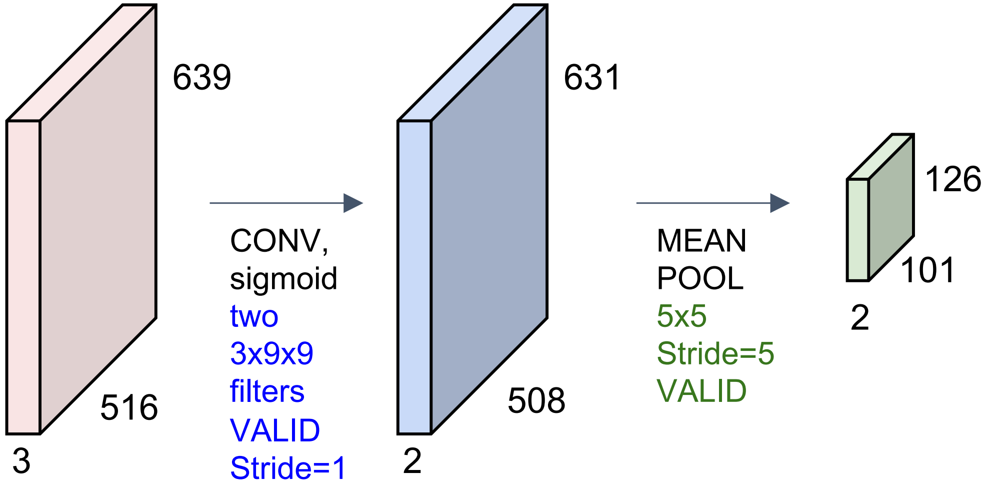
Original image


Feature map 1

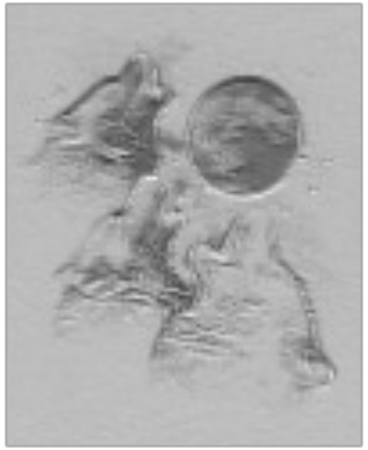

Feature map 2

# Question 2

Plot the original image and the first and second components of pooling output




Original image



Pooled feature map 1

Pooled feature map 2

# Question 3

3. Design a CNN with one hidden layer to recognize digit images in MNIST database:

   http://yann.lecun.com/exdb/mnist/

   The convolution layer consists of 25 filters of dimensions 9x9 and the pooling layer has a pooling window size 4x4. Assume VALID padding and default strides for both convolution and pooling layer. Train the network with mini batch gradient decent learning with learning factor $\alpha = 10^{-3}$ and batch size = 128.

   Plot

   a. The training and test errors against learning epochs.
   b. Final filter weights
   c. Feature maps at the convolution and pooling layers for a representative test pattern
   d. Repeat training by introducing decay parameter $\beta = 10^{-6}$ and momentum term with $\gamma = 0.5$, and compare the learning curves

t6q3a.ipynb

t6q3b.ipynb

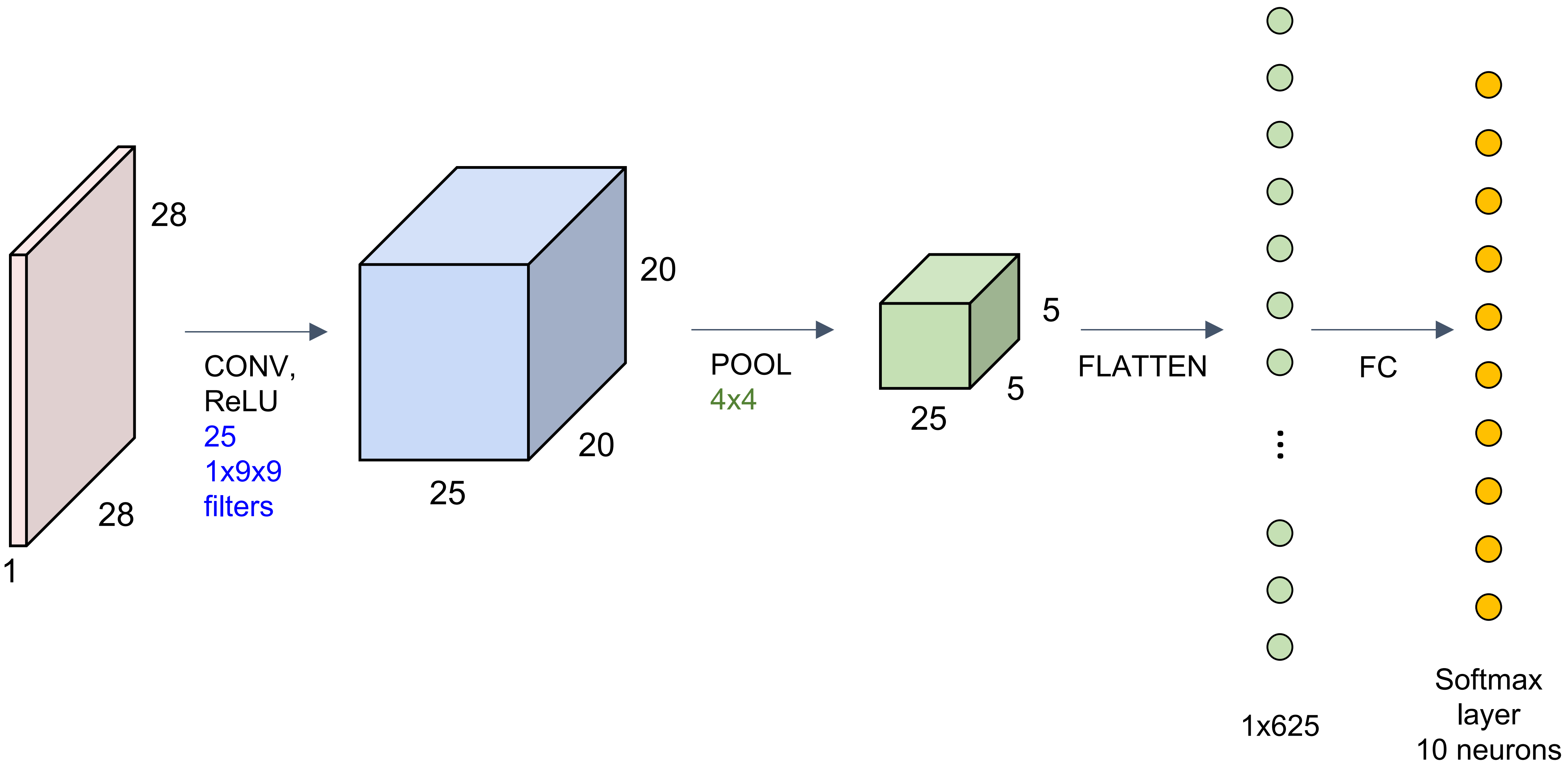
# Question 3

**MNIST**

- Size-normalized and centred 1x28x28 =784 inputs
- Training set = 60,000 images
- Testing set = 10,000 images

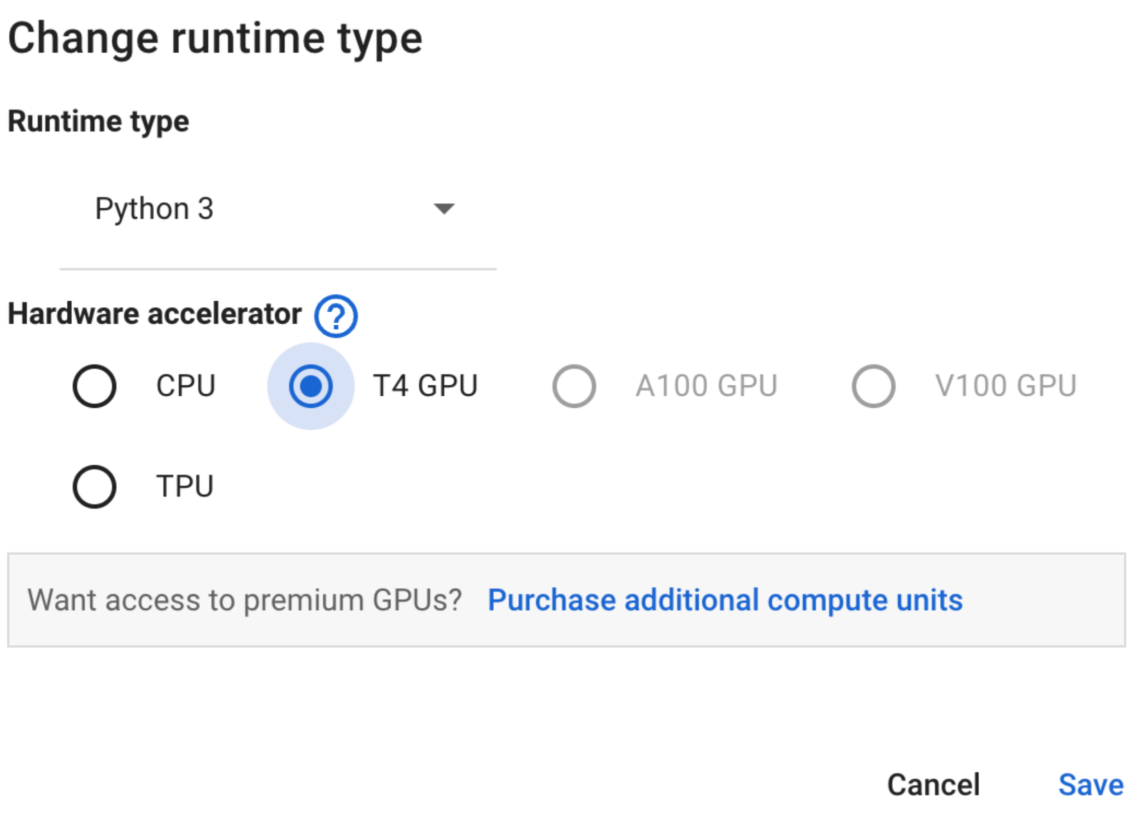
http://yann.lecun.com/exdb/mnist/

# Question 3



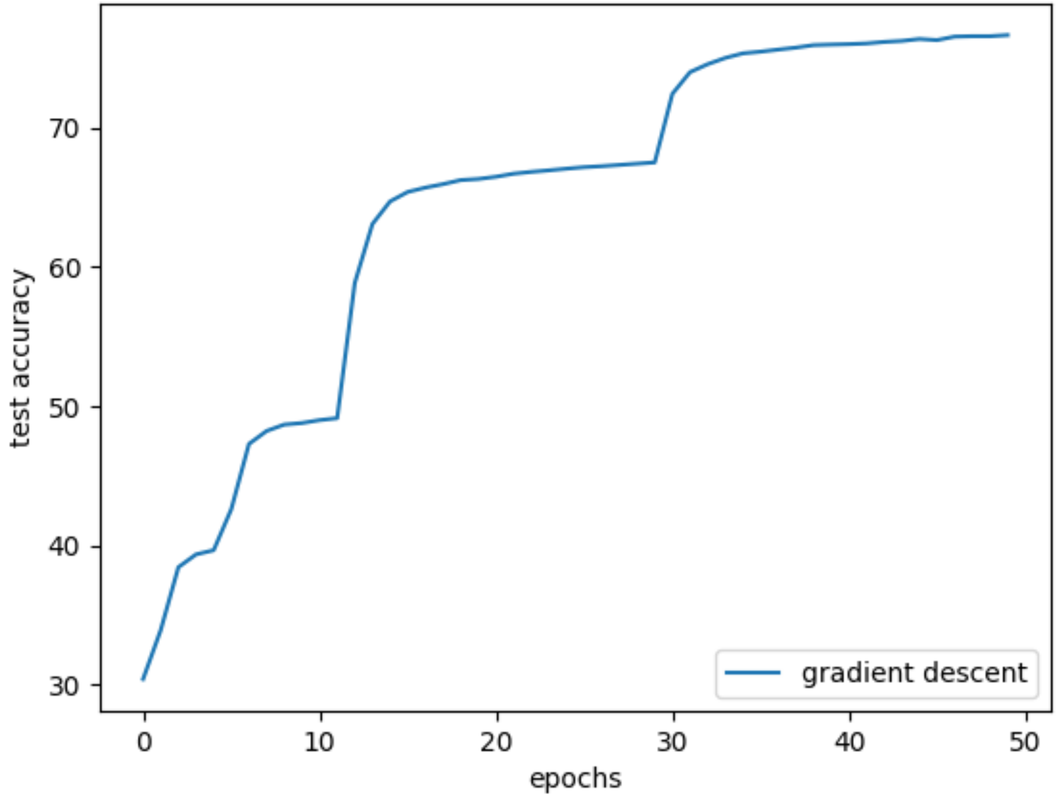
28
28
1
CONV,
ReLU
25
1x9x9
filters
20
20
25
POOL
4x4
5
5
25
FLATTEN
1x625
FC
Softmax
layer
10 neurons

# Question 3


Change runtime type
Runtime type
Python 3
Hardware accelerator
CPU
T4 GPU
A100 GPU
V100 GPU
TPU
Want access to premium GPUs? Purchase additional compute units
Cancel
Save

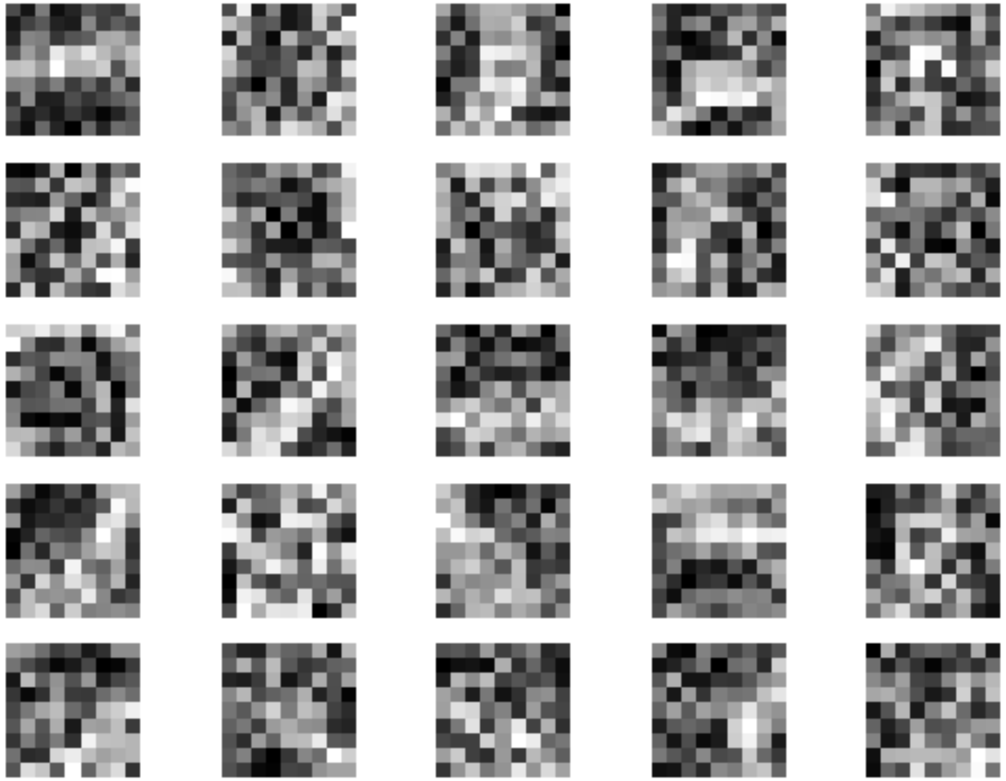
# Question 3a

Plot the learning curve (I only show the test curve) - you should be able get better learning curve if you run for more epochs

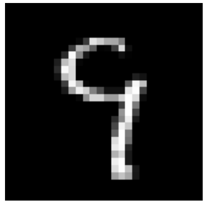


# Question 3b

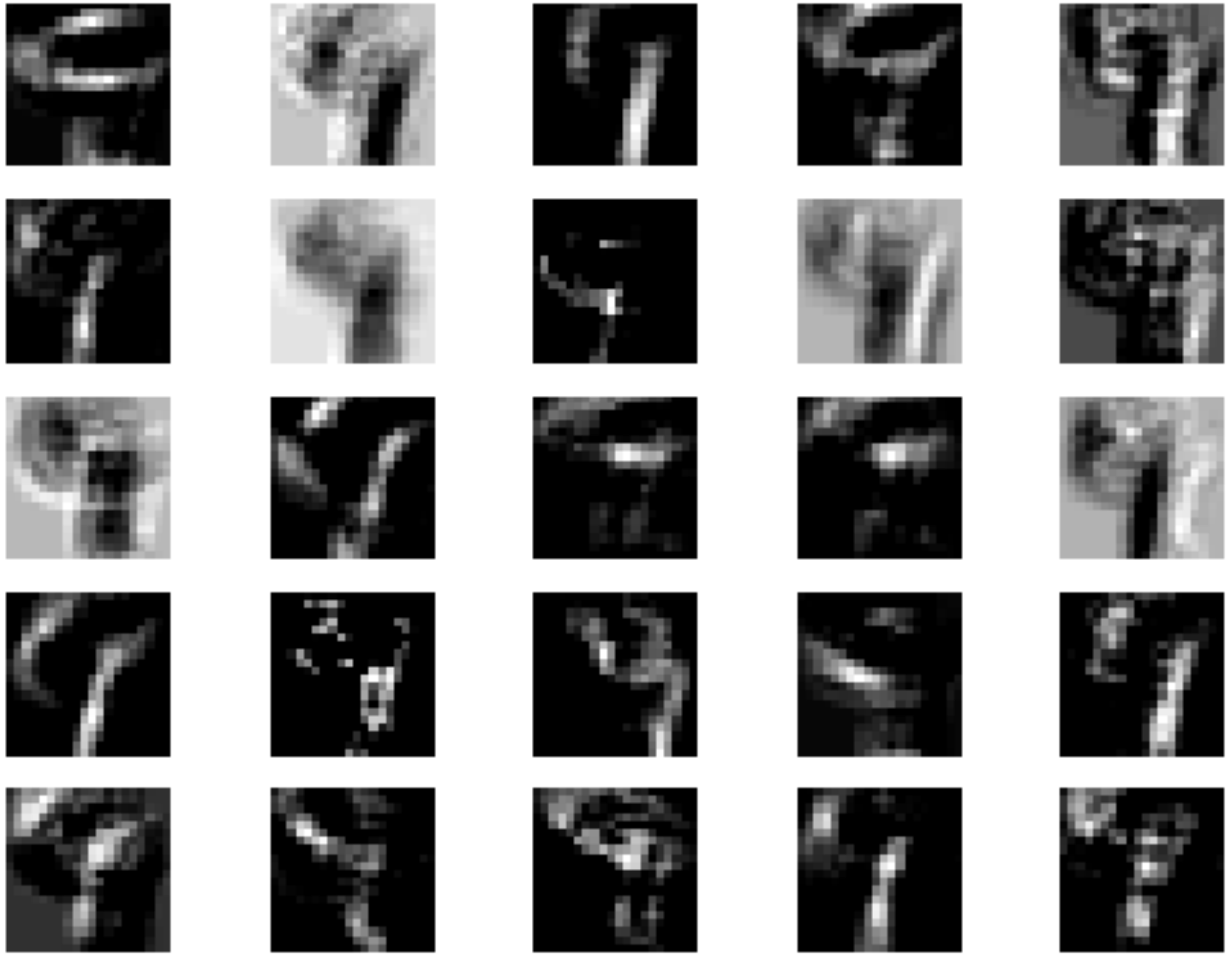
Plot filters learned in the conv layer



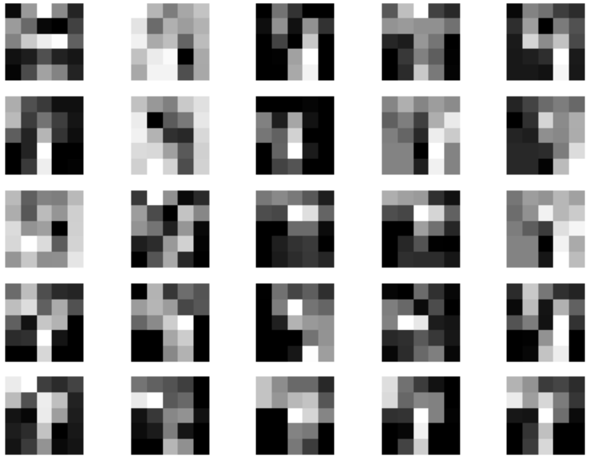
# Question 3c



Input 28x28

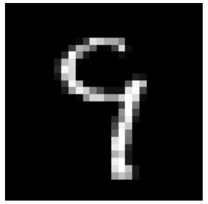


Feature maps of the convolution layer 25x20x20



Feature maps of pooling layer 25x5x5

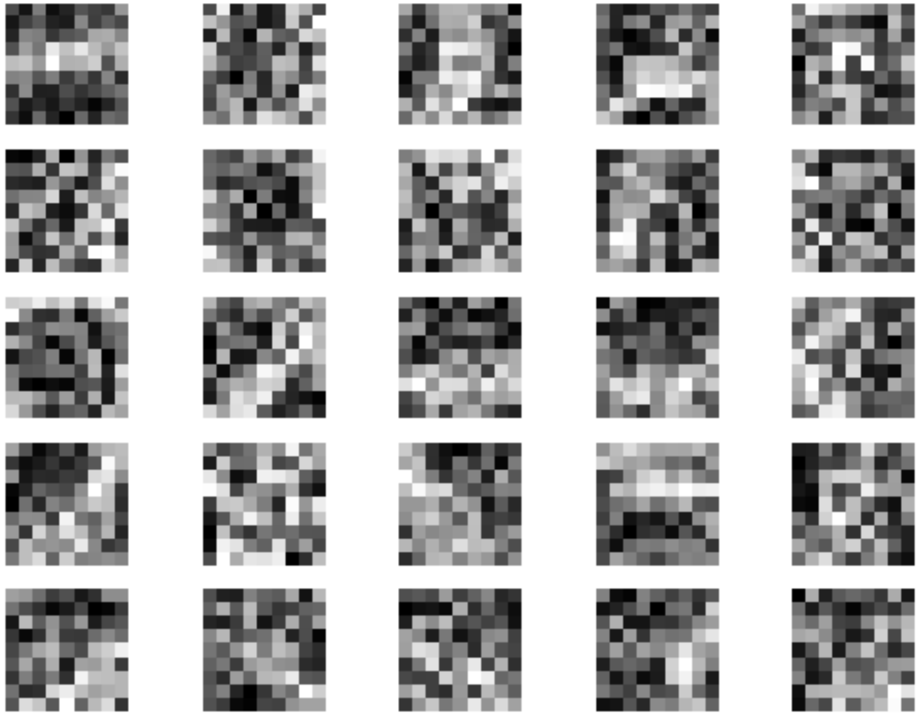
# Question 3c



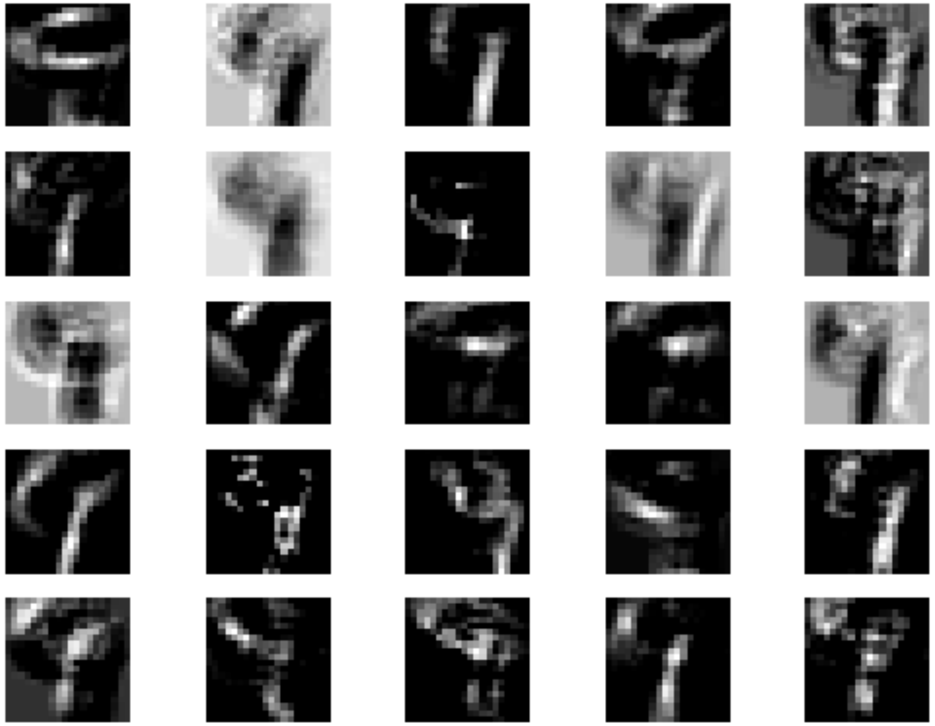
Input 28x28

Compare the filters learned and the corresponding feature maps.

What do you see?



Filters learned in the conv layer



Feature maps of the convolution layer 25x20x20

# Question 3d

Gradient descent:

$$W = W - \alpha \nabla_W J$$

Weight decay (regularization):

$$J_1(W, b) = J(W, b) + \beta_2 \sum_{ij} (w_{ij})^2$$

Weight decay:

$$W = W - \alpha(\nabla_W J + \beta W)$$

Sgd with momentum:

$$V = \gamma V - \alpha \nabla_W J$$
$$W = W + V$$

Sgd with decay and momentum:

$$V = \gamma V - \alpha(\nabla_W J + \beta W)$$
$$W = W + V$$