Computer Vision – Week 2

Anh Huy Bui – 293257

1. Let $x_1 = 1$ represents the input
   a) The output of the hidden unit for this training case:

$$x_2 = \frac{1}{1 + \exp(-w_1 x_1 - b_1)} = \frac{1}{2}$$

   The output of the output unit for this training case:

$$y = w_2 x_2 + b_2 = 2$$

   b) The loss of this training case:

$$E = \frac{1}{2}(t - y)^2 = \frac{1}{2}$$

   c) The derivative of the loss with respect to w2

$$\frac{\partial E}{\partial w_2} = \frac{\partial E}{\partial y}\frac{\partial y}{\partial w_2}$$

$$= \big((-1) * (t - y)\big)x_2 = (w_2 x_2 + b_2 - t)x_2$$

$$= \frac{w_2}{(1+\exp(-w_1 x_1 - b_1))^2} - \frac{t}{1+\exp(-w_1 x_1 - b_1)}$$

$$= \frac{4}{(2)^2} - \frac{1}{2} = \frac{1}{2}$$

   d) The derivative of the loss with respect to w1

$$\frac{\partial E}{\partial w_1} = \frac{\partial E}{\partial y}\frac{\partial y}{\partial \sigma}\frac{\partial \sigma}{\partial w_1}$$

$$= \big((-1) * (t - y)\big) * w_2 * \frac{\exp(-w_1 x_1 - b_1)}{(1 + \exp(-w_1 x_1 - b_1))^2} * x_1$$

$$= 2 * 4 * \frac{1}{2^2} * 1 = 2$$

2.

a) Euclidean distance:

$$d_{QA} = \sqrt{\sum_{i=1}^{5}((Q_i - A_i)^2)} = 3.46$$

$$d_{QB} = \sqrt{\sum_{i=1}^{5}((Q_i - B_i)^2)} = 4.80$$

Cosine similarity:

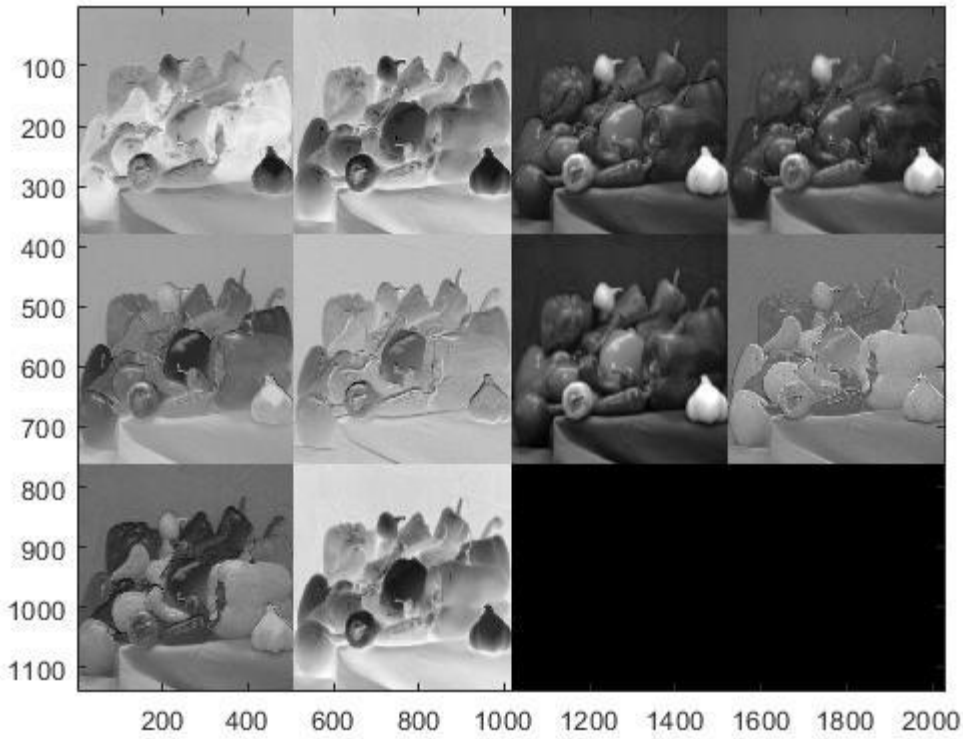
$$similarity(Q, A) = \frac{Q.A}{\|Q\|\|A\|} = 0.919$$

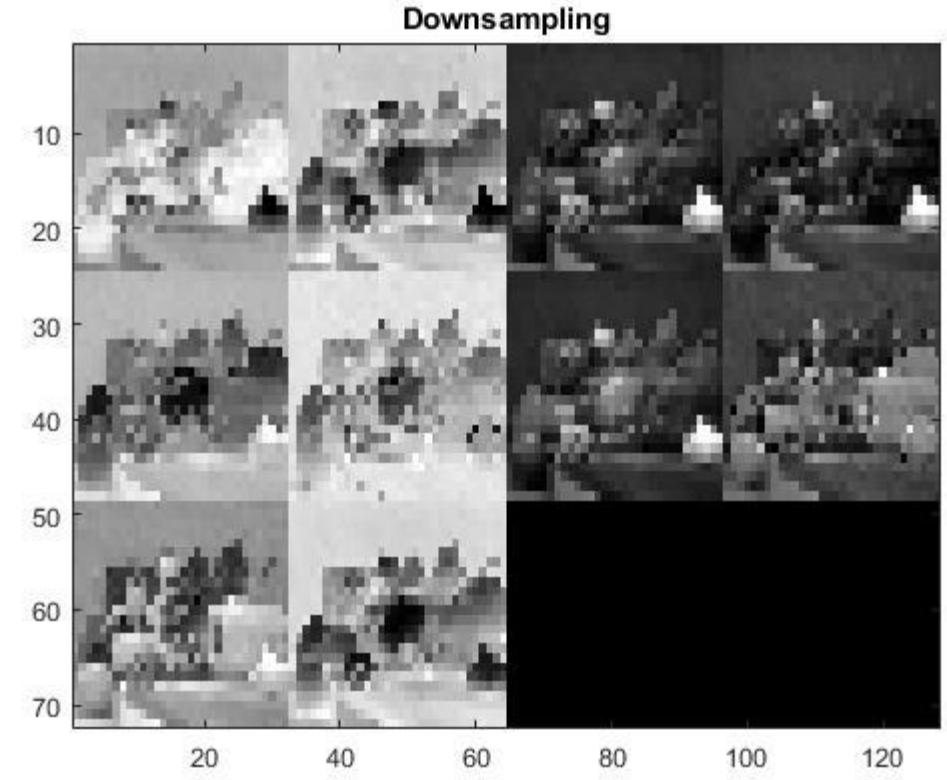
$$similarity(Q, B) = \frac{Q.B}{\|Q\|\|B\|} = 0.799$$

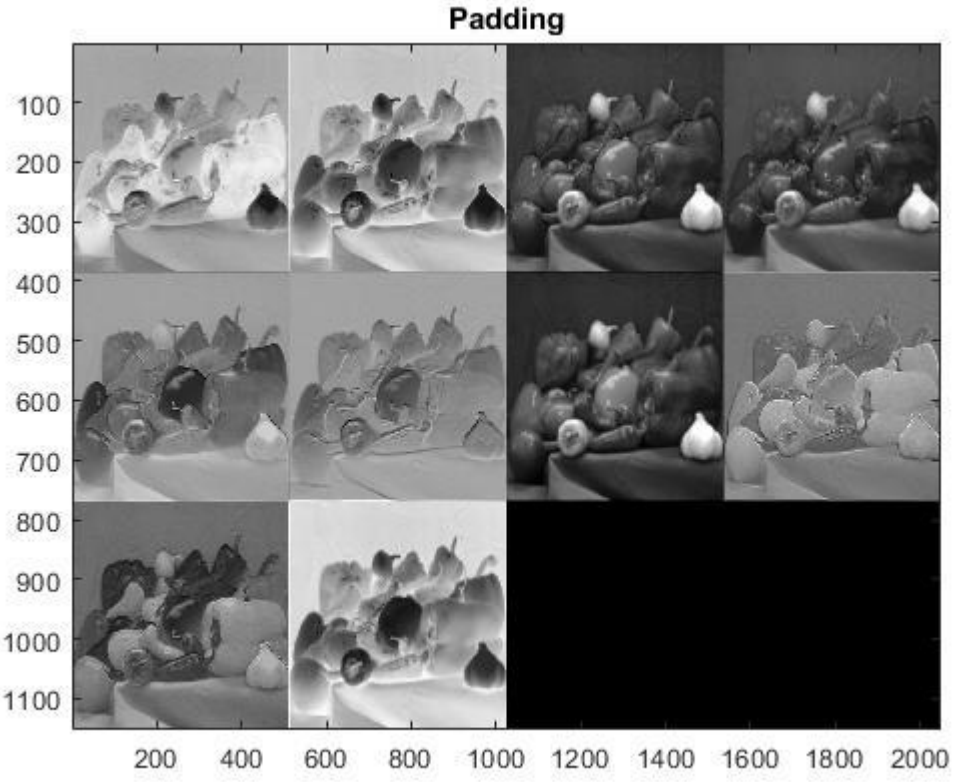
b) Because distance from A to Q is shorter than B, and similarity of A and Q is also higher than Q and B so A is more similar to Q than B

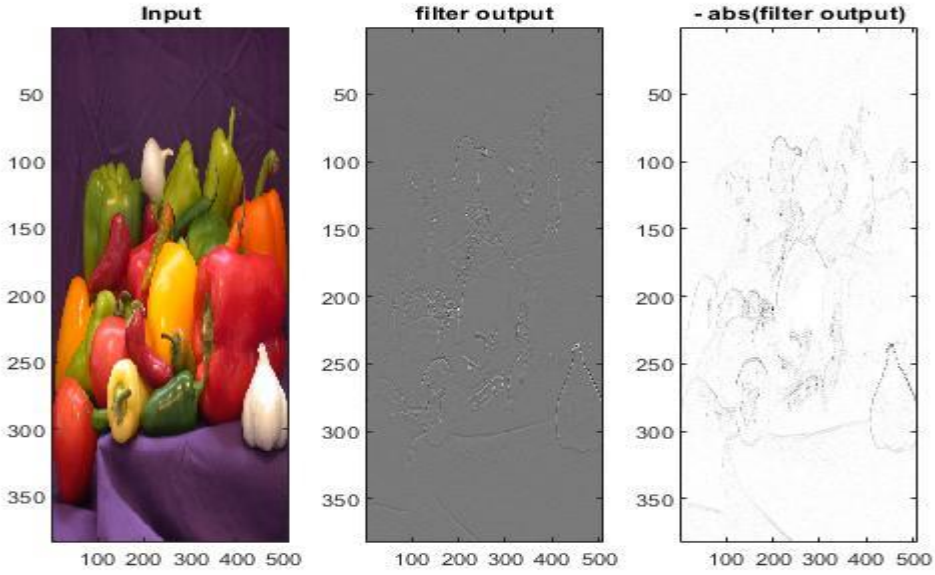
3.

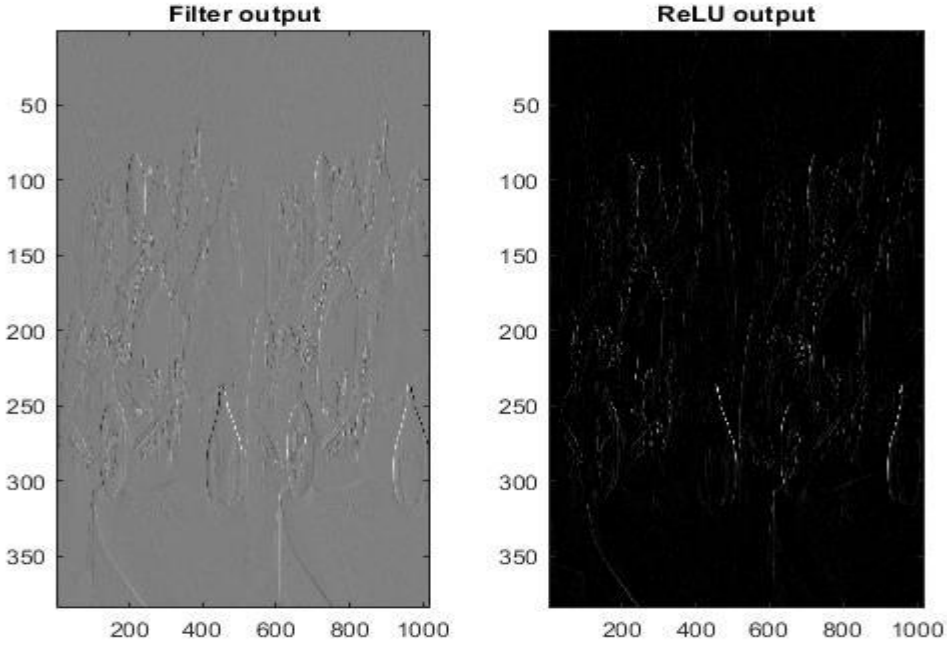

50
100
150
200
250
300
350
50 100 150 200 250 300 350 400 450 500

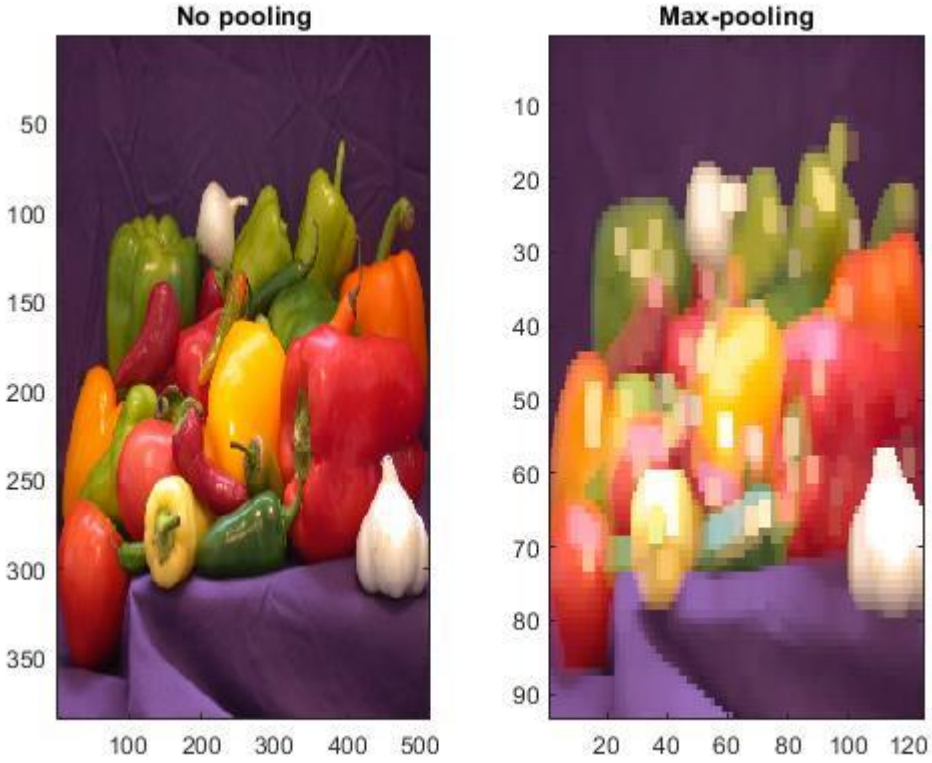

100
200
300
400
500
600
700
800
900
1000
1100
200 400 600 800 1000 1200 1400 1600 1800 2000

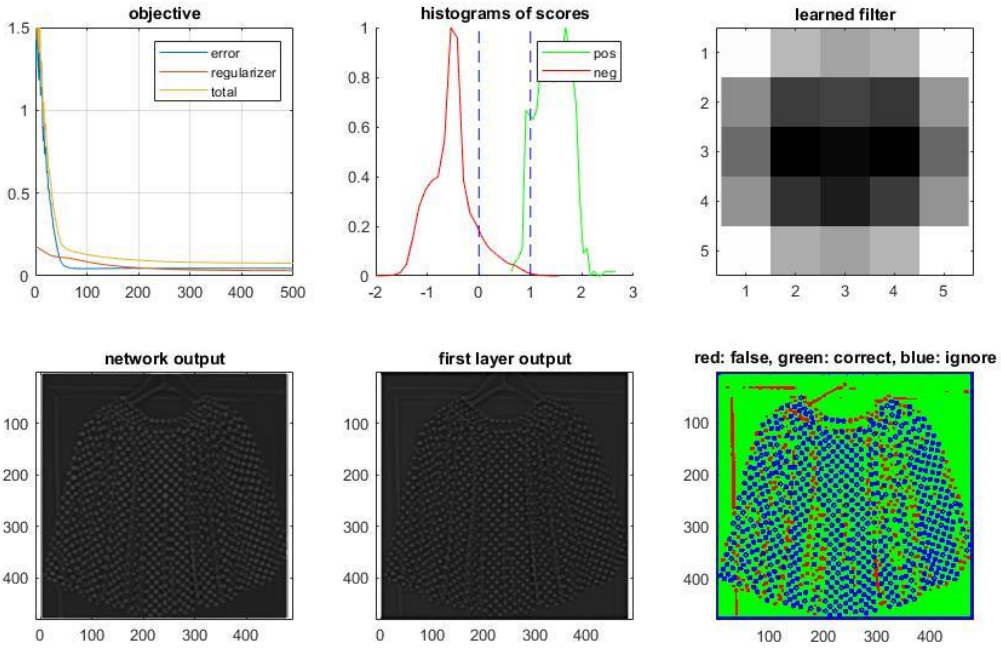

Downsampling
10
20
30
40
50
60
70
20
40
60
80
100
120


Padding
100
200
300
400
500
600
700
800
900
1000
1100
200
400
600
800
1000
1200
1400
1600
1800
2000


Input
filter output
- abs(filter output)
50
100
150
200
250
300
350
100 200 300 400 500


Filter output
ReLU output
50
100
150
200
250
300
350
200 400 600 800 1000


No pooling
Max-pooling
50
100
150
200
250
300
350
100
200
300
400
500
10
20
30
40
50
60
70
80
90
20
40
60
80
100
120


objective
error
regularizer
total
histograms of scores
pos
neg
learned filter
network output
first layer output
red: false, green: correct, blue: ignore