DATA.ML.300 Computer Vision Exercise 2

Uyen Phan

01 2024

1

a) The output of the hidden unit:
$net_h = w_1.i + b_1 = 1. - 2 + 2 = 0$
$y = \sigma(net_h) = \frac{1}{1+exp(-0)} = 0.5$

The output unit:
$out_o = w_2.out_h + b_2 = 0.5.4 + 0 = 2$

b) Loss result: $E = \frac{1}{2}(1-2)^2 = 0.5$

c) The derivative of the loss with respect to w_2:

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

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

$\frac{\partial y}{\partial w_2} = y = 0.5$

Therefore, $\frac{\partial E}{\partial w_2} = 1 * 0.5 = 0.5$

c) The derivative of the loss with respect to w_1:

$$\frac{\partial E}{\partial w_1} = \frac{\partial E}{\partial out_h} * \frac{\partial out_h}{\partial net_h} * \frac{\partial net_h}{\partial w_1}$$

$\frac{\partial E}{\partial out_h} = \frac{\partial E}{\partial y} * \frac{\partial y}{\partial out_h} = 1 * w_2 = w_2 = 4$

$\frac{\partial out_h}{\partial net_h} = out_h(1 - out_h) = 0.5(1 - 0.5) = 0.25$

$\frac{\partial net_h}{\partial w_1} = i = 1$

Therefore, $\frac{\partial E}{\partial w_1} = 4 * 0.25 * 1 = 1$

2

a) Euclidean distance

$\mathbf{Q}$ and $\mathbf{A} = \begin{bmatrix}1 & 1 & 9 & 0 & 1\end{bmatrix}^T = \sqrt{12} = 3.46$

$\mathbf{Q}$ and $\mathbf{B} = \begin{bmatrix}1 & 0 & 4 & 9 & 9\end{bmatrix}^T = \sqrt{23} = 4.79$

Cosine similarity

$\mathbf{Q}$ and $\mathbf{A}$

$\mathbf{Q} \cdot \mathbf{A} = \begin{bmatrix}2 & 2 & 18 & 16 & 2\end{bmatrix}^T = 40$

$||\mathbf{Q}||\ .\ ||\mathbf{A}|| = \sqrt{61}\sqrt{31} = 43.49$

$\cos\theta = \frac{40}{43.49} = 0.92$

$\mathbf{Q}$ and $\mathbf{B}$

$\mathbf{Q} \cdot \mathbf{B} = \begin{bmatrix}6 & 1 & 24 & 4 & 10\end{bmatrix}^T = 45$

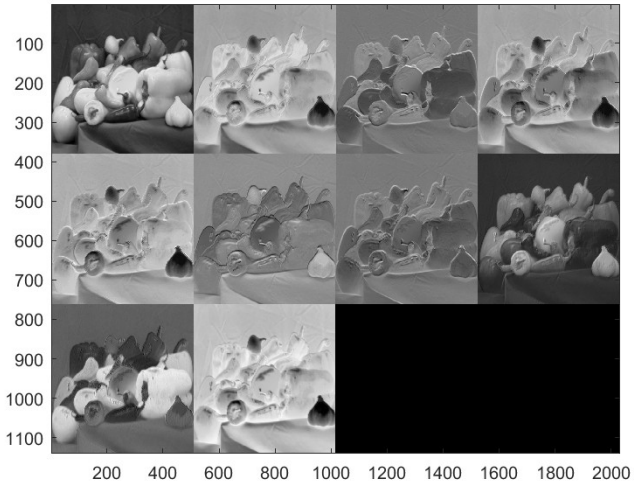
$||\mathbf{Q}||\ .\ ||\mathbf{B}|| = \sqrt{61}\sqrt{52} = 56.32$

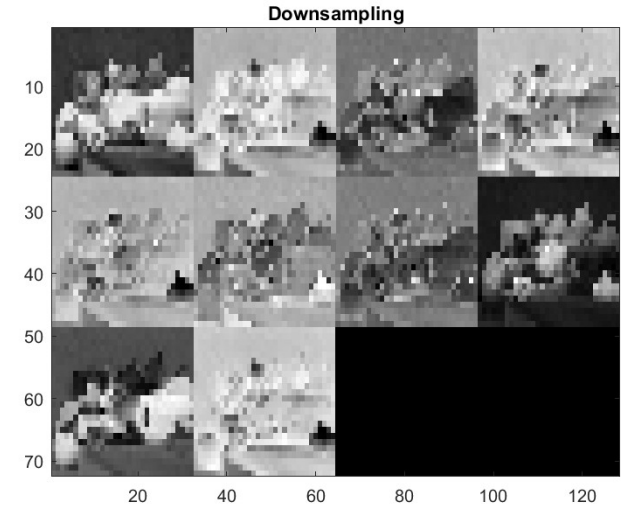
$\cos\theta = \frac{45}{56.32} = 0.79$

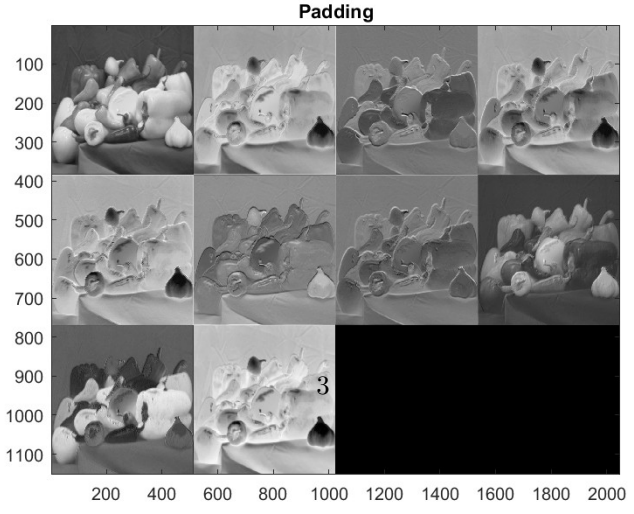
b) The image $\mathbf{A}$ is more similar to the query image $\mathbf{Q}$ because it has a shorter Euclidean distance and a higher value of cosine similarity.

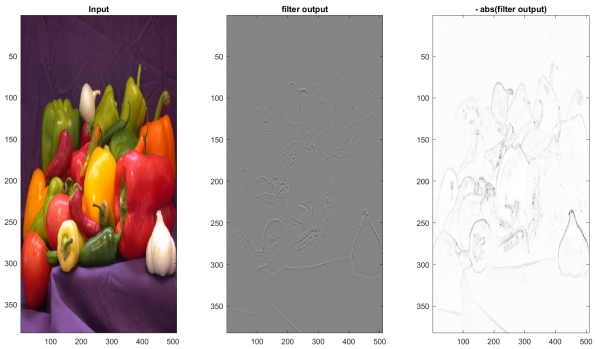
3

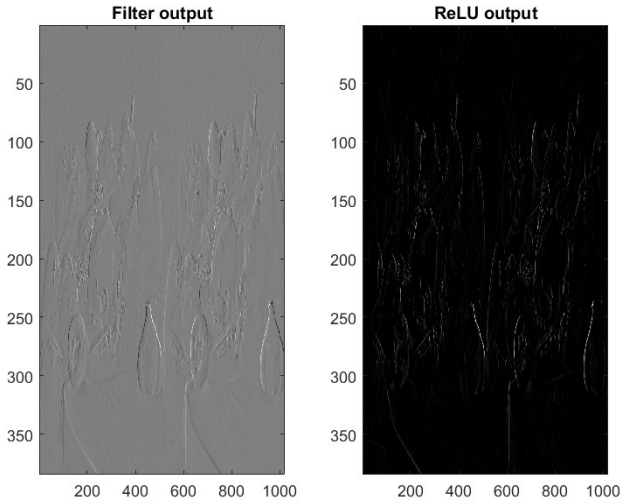
MATLAB answers

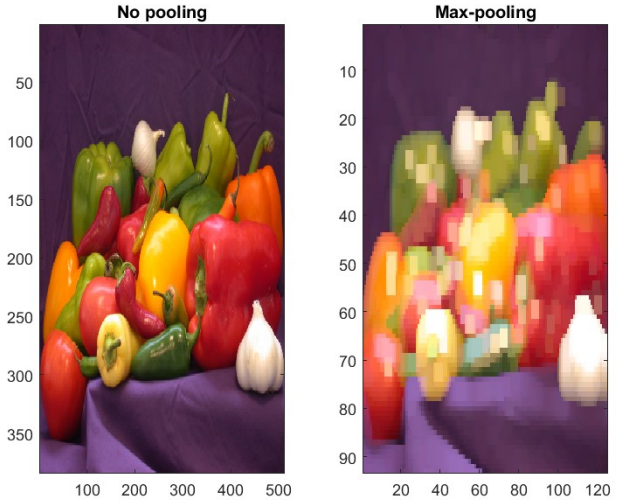

100
200
300
400
500
600
700
800
900
1000
1100
200
400
600
800
1000
1200
1400
1600
1800
2000


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)


Filter output
ReLU output


No pooling
Max-pooling

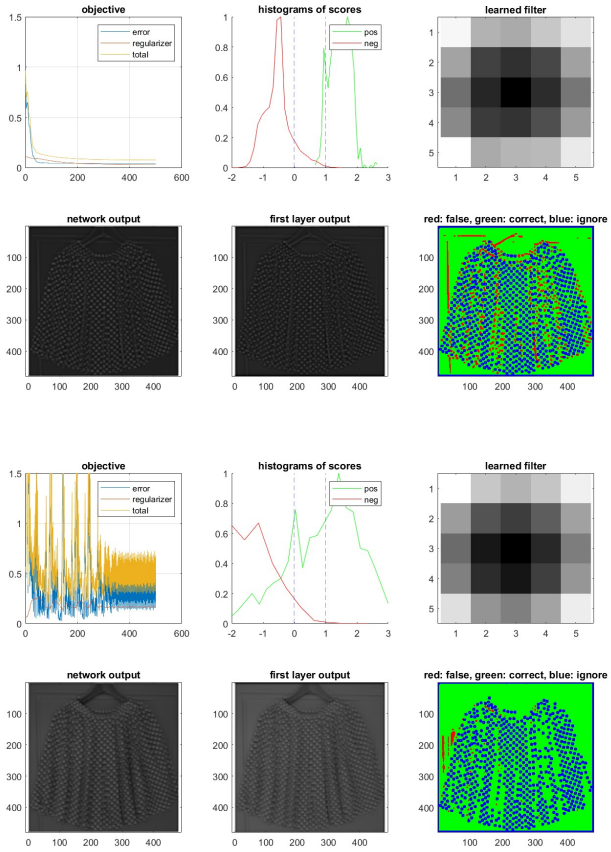


Figure 1: Doubled learning rate

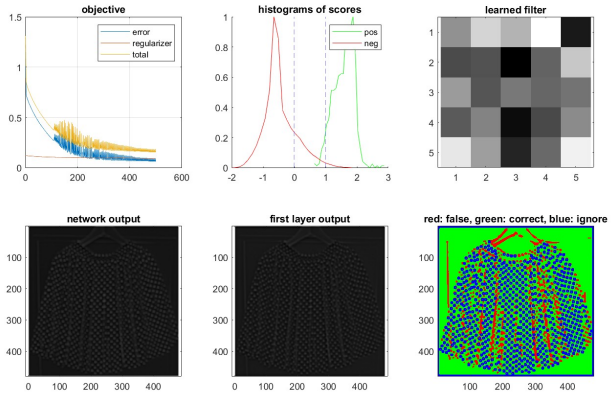


Figure 2: no momentum