PA3 Individual Report

Zhenrui Yue
Computer Science & Engineering
UC San Diego
La Jolla, CA 92093
yuezrhb@gmail.com

1 Understanding convolutional network basics

Filtered Inputs:

x	x	x	x	x
x	4	4	3	x
x	-4	-1	0	x
x	-2	-1	0	x
x	x	x	x	x

x	x	x	x	x
x	-2	-2	0	x
x	2	3	2	x
x	2	2	3	x
x	x	x	x	x

Output Feature Map:

x	x	x	x	x
x	2	2	3	x
x	-2	2	2	x
x	0	1	3	x
x	x	x	x	x

2 Maximumly activating patch

Activating Patches:

-1	0	0
x	x	x
1	1	1

0	x	1
1	x	x
1	1	1

3   Spatial pooling

Output Feature Map:

2	3
2	3

4   Number of learnable parameters

(i)   The number of input channels to conv1: 1

(ii)   The number of input channels to conv2: 12

(iii)   The number of input channels to conv3: 10

(iv)   The number of the incoming dimensions to fc1 will be $491 \times 491 \times 8 = 1{,}928{,}646$ in total. First, the grayscale image with 512×512 resolution will be processed by a 8×8 kernel with 12 channels of output, resulting in a $505 \times 505 \times 12$ output of the layer. Then, it will be processed by another 8×8 kernel with 10 channels of output, resulting in a $498 \times 498 \times 10$ output size. Another convolutional layer with a 6×6 kernel and 8 output channels will further shrink the input size to $493 \times 493 \times 8$. Finally, this will be max-pooled by a 3×3 kernel, leaving the next output before the fully connected layer in the size of $491 \times 491 \times 8$.