

Assignment 4

COS 470/570: Image Processing and Computer Vision

Objective:

Understand the internal workings of a convolutional neural network (CNN) by manually calculating the output dimensions and the number of parameters for each layer in the CNN.

Tasks:

Task 1: Calculate the output feature map of this convolutional layer. The input feature map is $5 \times 5 \times 4$, two kernels ($3 \times 3 \times 4$), stride=1, zero padding=1. Can you calculate the output feature map?

You have the option to calculate the output feature map either manually or by writing code. However, I **recommend** employing both methods to enhance your understanding of the convolution process. For the assignment, providing the code file is not necessary; just submit the calculated results.

Input, $5 \times 5 \times 4$

```
[0.40, 0.13, 0.53, 0.35, 0.04],
[0.50, 0.87, 0.80, 0.24, 0.88],
[0.69, 0.45, 0.02, 0.75, 0.52],
[0.24, 0.76, 0.16, 0.70, 0.08],
[0.09, 0.90, 0.41, 0.27, 0.08]

[0.86, 0.98, 0.46, 0.80, 0.48],
[0.43, 0.54, 0.66, 0.75, 0.90],
[0.43, 0.47, 0.56, 0.96, 0.06],
[0.95, 0.03, 0.99, 0.64, 0.82],
[0.95, 0.61, 0.14, 0.03, 0.75]

[0.74, 0.32, 0.89, 0.33, 0.40],
[0.63, 0.61, 0.10, 0.47, 0.10],
[0.27, 0.55, 0.99, 0.51, 0.23],
[0.07, 0.51, 0.91, 0.32, 0.30],
[0.52, 0.36, 0.25, 0.91, 0.94]

[0.71, 0.71, 0.70, 0.62, 0.32],
[0.88, 0.06, 0.17, 0.56, 0.04],
[0.42, 0.18, 0.78, 0.43, 0.77],
[0.94, 0.14, 0.25, 0.13, 0.61],
[0.92, 0.99, 0.23, 0.25, 0.92]
```

Kernel 1: $3 \times 3 \times 4$

```
[0.32, 0.05, 0.43],
[0.91, 0.37, 0.60],
[0.23, 0.52, 0.40]

[0.31, 0.91, 0.47],
[0.49, 0.42, 0.92],
[0.64, 0.49, 0.62]

[0.58, 0.41, 0.17],
[0.39, 0.45, 0.77],
[0.11, 0.58, 0.79]

[0.92, 0.43, 0.93],
[0.38, 0.15, 0.52],
[0.74, 0.45, 0.53]
```

Kernel 2: $3 \times 3 \times 4$

```
[0.07, 0.01, 0.59],
[0.02, 0.94, 0.68],
[0.78, 0.64, 0.73]

[0.14, 0.55, 0.42],
[0.90, 0.46, 0.80],
[0.65, 0.96, 0.81]

[0.32, 0.27, 0.45],
[0.76, 0.52, 0.13],
[0.52, 0.95, 0.88]

[0.34, 0.69, 0.81],
[0.10, 0.44, 0.03],
[0.77, 0.06, 0.67]
```

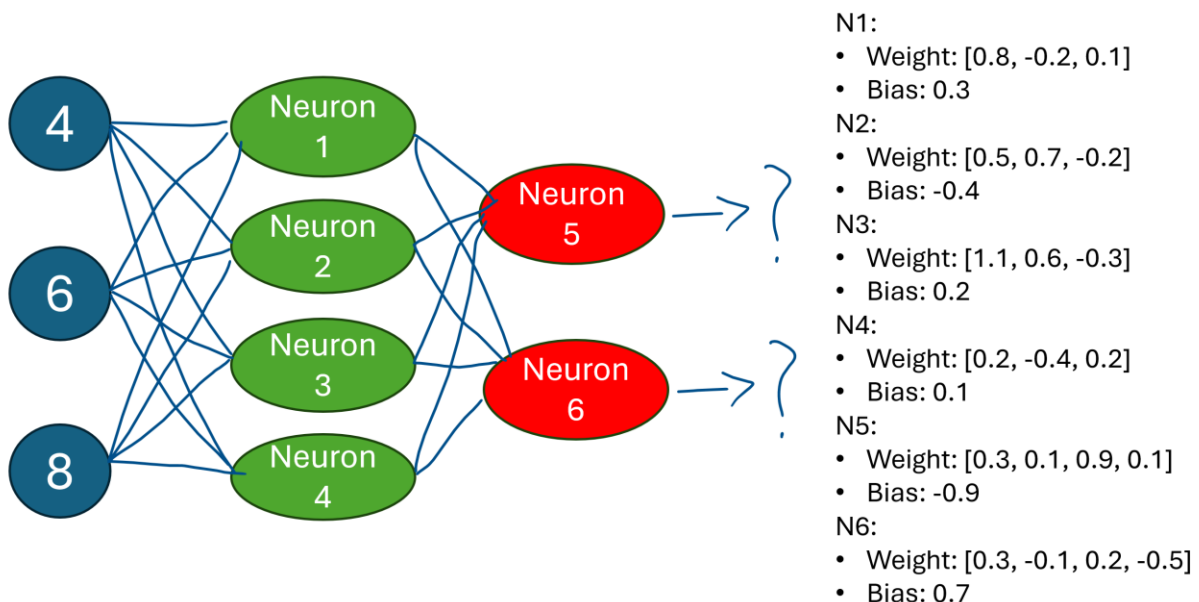
Task 2: Calculate the output of the pooling layers separately for max pooling and average pooling. Use a filter size of 3 and a stride of 3.

You have the option to calculate the output feature map either manually or by writing code. However, I **recommend** employing both methods to enhance your understanding of the convolution process. For the assignment, providing the code file is not necessary; just submit the calculated results.

```
[0.63, 0.79, 0.13, 0.46, 0.56, 0.20, 0.05, 0.83, 0.83],
[0.76, 0.86, 0.31, 0.57, 0.22, 0.70, 0.95, 0.42, 0.82],
[0.34, 0.62, 0.13, 0.07, 0.99, 0.83, 0.22, 0.19, 0.99],
[0.44, 0.91, 0.23, 0.68, 0.37, 0.05, 0.14, 0.19, 0.62],
[0.78, 0.97, 0.06, 0.07, 0.88, 0.60, 0.74, 0.60, 0.57],
[0.02, 0.28, 0.06, 0.02, 0.20, 0.77, 0.98, 0.86, 0.95],
[0.37, 0.89, 0.70, 0.35, 0.06, 0.22, 0.20, 0.35, 0.89],
[0.42, 0.76, 0.56, 0.20, 0.75, 0.78, 1.00, 0.62, 0.95],
[0.34, 0.34, 0.70, 0.69, 0.44, 0.22, 0.66, 0.64, 0.80]
```

Task 3: Calculate the output of the fully connected layers. To simplify the calculation, do not consider the activation function in this task.

You have the option to calculate the output feature map either manually or by writing code. However, I **recommend** employing both methods to enhance your understanding of the convolution process. For the assignment, providing the code file is not necessary; just submit the calculated results.



Task 4: Calculate the output dimensions and the number of parameters (considering the weights and biases).

Task 4a): For a convolution layer, the input is $64 \times 64 \times 10$, this layer has 100 kernels, each kernel's size is 7×7 . Stride=1 and zero padding=1. What is the size of the output feature map? And can you calculate the number of parameters in this layer?

Task 4b): For a fully connected layer, the input is 1×1024 , this layer has 512 neurons. What is the size of the output? And can you calculate the number of parameters in this layer?