Deep Learning
Prof. S. Harmeling

Wednesday, 20. May 2020
DUE 23:55 Wednesday, 27. May 2020

Exercises #5

Please upload a scan of your solution or use LaTeX to typeset your solution. Use the same naming scheme as usual and submit a single zip-file to the sciebo dropoff folder for the exercises. The link will be available on the ILIAS page.

1. **Spatial Dimensions.**
   Please hand in your solutions of this exercise to the mailbox.

   In the following we denote the shape of the input tensor of a convolutional layer with $W_1 \times H_1 \times D_1$. Furthermore, let $F \times F \times D_1$ be the shape of the filter. We denote the number of filters with K, the stride with S and the amount of zero padding with P.

   A convolutional layer outputs a volume of size $W_2 \times H_2 \times D_2$, where

$$
\begin{aligned}
W_2 &= \frac{(W_1 - F + 2P)}{S} + 1 \\
H_2 &= \frac{(H_1 - F + 2P)}{S} + 1 \\
D_2 &= K.
\end{aligned}
$$

   (a) Consider an input tensor with shape $55 \times 43 \times 3$. What shape has the output tensor of a convolutional layer with six $3 \times 3 \times 3$ filters, a stride of two and a zero padding of one?

   (b) Consider an input tensor with shape $122 \times 122 \times 5$, an output tensor with shape $42 \times 42 \times 10$, a stride of three and zero padding of two. What is the number of filters and which shape do they have?

   (c) Consider an input tensor with shape $73 \times 73 \times 3$, an output tensor with shape $11 \times 11 \times 6$, zero padding of two and filters of shape $7 \times 7 \times X$. What is the correct value for X? Which stride has been used?

   (d) Consider an input tensor with shape $256 \times 256 \times 3$, an output tensor with shape $86 \times 86 \times 10$, a stride of three and multiple $9 \times 9 \times 3$ filters. What amount of zero padding has been used and what is the number of filters?

   Please provide a detailed solution, not just the results. If you describe a filter shape please provide all of its dimensions.

   50 points

2. **Higher Level Representations: Image Features.**
   We have seen that we can achieve reasonable performance on an image classification task by training a linear classifier on the pixels of the input image. In this exercise we will show that the classification performance can be improved by training linear classifiers on features that are computed from the raw pixels.

   For detailed instructions please we refer to the Jupyter notebook *features.ipynb*.

   50 points