Universität des Saarlandes
FR 4.7 Sprach- und Signalverarbeitung

Prof. Dietrich Klakow
Pattern and Speech Recognition

Exercise Sheet 9

Convolutional Neural Networks

Deadline: 24.01.2017, 23:59

Gabor filters and Convolution

Exercise 9.1 *(2 points)*

Consider the signal

$$g(x) = e^{-(x-1)^2} \cos 3x$$

And the filter

$$e^{-x^2} \cos kx$$

Calculate the continuous convolution of the signal with the Gabor filter analytically. For which parameters of the Gabor filter k and at which displacement do you get the maximum output?

Exercise 9.2 *(2 points)*

The fourier transform of a signal $f(x)$ is given as

$$F(v) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x v} dx$$

For two signals f and g. Show that the Fourier transform of the convolution is equal to the product of the Fourier transforms. This is also known as the Convolution theorem. How does this help us in Convolution?

Exercise 9.3 *(2 points)*

Use the convolution theorem from **Exercise 9.2** and apply it to **Exercise 9.1** by transforming both signal and filter. What do you observe in frequency space when multiplying signal and filter?

Max and Average Pooling

Exercise 9.4 *(2 + 1 + 1 + 2 = 6 points)*

In this excercise you will implement a simple max pooling function. And see how it can detect a feature (brightest pixel) in a certain region. We will be using the same image **clock.png** from the previous assignment. Follow the same procedure from **Excercise 8.2** to load the image.

a) Implement a function **maxPool** that takes two parameters: **data** and **split**.
Your function should take a square 2^nx2^n matrix as **data** and return a square matrix of size **split**x**split** where each entry is the result of a max pool in that patch.
For e.g. for our image of size 256x256, and a **split** of value 2 , we have an output 2x2 matrix, where each entry will correspond to the result of a max pool on each quadrant of the image, with no overlaps.

b) Apply your function on the image with **split** as 8,4,2, then 1.
What is the maximum value? Where can you find it in each of the splits? Provide the position(s) of that maximum value in each output matrix >1x1 as a tuple (i,j).

c) Change your function to an average pooling function. Name it **meanPool** and use it on **clock.png** with **split** as 128. Save the output image as **clockMean.png** and submit it with your code. Do the same thing with **maxPool** and **split** as 128. Save the image as **clockMax.png**. What do you observe?

d) To do Backpropagation in the pooling layer we need to get the derivative of the pooling function. Derive the derivatives for the max pooling and mean pooling w.r.t. the input. Explain how the error would be backpropagated for each case, also known as the Upsampling operation.

Training a Convolutional neural network

Exercise 9.5 *(2 points)*

Now that you are familiar with the parts that build up a Convolutional Neural network, derive the backpropagation equations for the weights of the Convolutional layer. Provide a pseudocode for the implementation.

In practice, CNN are rarely trained from scratch. Explain how Transfer Learning can help us train CNNs faster.

Revision Questions

Exercise 9.6 *(2 + 2 + 2 = 6 points)*

a) Explain how you would go about implementing a ConvNet to be used in the CIFAR-10 task. Show the layers you would use and ellaborate on the function of each layer, its connections, and dimensions. You answer should not exceed 10-12 sentences.

b) Present the Backpropagation algorithm for Multi-Layer-Perceptrons and explain its steps. Assume you have a squared error function as the cost function and the logistic function as the activation function, derive the equation(s) for the derivative of the error w.r.t the weights in terms of the output of the next layer and the target output.

c) Given a fully connected network with 3 neurons in the input layer, 5 neurons in a hidden layer and 2 neurons in the output layer. If the activation function is linear, show that this network can be simplified into a one-layer network. Draw both networks.

The following instructions are mandatory. If you are not following them, tutors can decide to not correct your exercise.

Submission architecture

You have to generate a **single ZIP file** respecting the following architecture:

```
tutorial2_<matriculation_nb1>_<matriculation_nb2>_<matriculation_nb3>
|
+--- source
|       |
|       +------- file 1
|       +------- file 2
|       +------- ...
+--- report.pdf
+--- README.txt
```

where

- **source** contains the source code of your project,
- **report.pdf** is the report where you present your solution with **the explanations** and the plots.
- **README** which contains group member informations (name, matriculation numbers and emails) and a **clear** explanation about how to compile and run your source code

The ZIP filename has to be :

`tutorial2_<matriculation_nb1>_<matriculation_nb2>_<matriculation_nb3>.zip`

Some hints

We advice you to follow the following guidelines in order to avoid problems :

- Avoid building complex systems. The exercises are simple enough.
- Do not include any executables in your submission, as this will cause the e-mail server to reject it.

Grading

Send your assignment to the tutor who is responsible of your group:

- Merlin Köhler s9mnkoeh@stud.uni-saarland.de
- Yelluru Gopal Goutam goutamyg@lsv.uni-saarland.de
- Ahmad Taie ataie@lsv.uni-saarland.de

The email subject should start with `[PSR TUTORIAL 9]`