

Machine Learning

Course Work 5

to complete by: **01.12.2016**

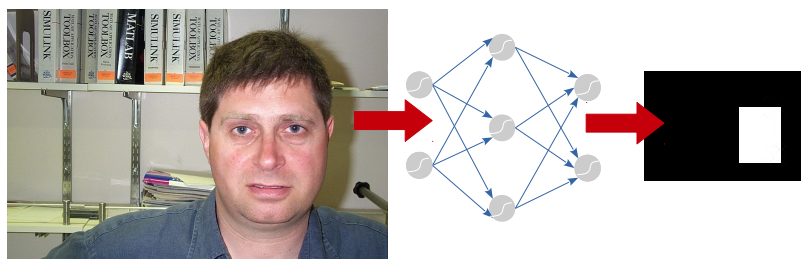
Please execute this course work (as well as any following) in teams of two.

In this exercise, you will build, train and evaluate your own Convolutional Neural Network (CNN). Again, we will use Google's Tensorflow library. Tensorflow gives you the necessary building blocks for your CNN (`tf.nn.conv2d()`, `tf.nn.max_pool()`, etc.). Before working on the exercise, you may want to consult the Tensorflow tutorial again.

Exercise 5.1 (Face Detection)

Build a CNN which – given an input image – outputs a segmentation mask marking each pixel as “face” (target: 1) or “background” (target: 0).

- Use the “Caltech Faces 1999” Dataset. On the course homepage, you will also find the ground truth in csv format (same content as the official `ImageData.mat`). From this ground truth, create a rectangular bounding box for each image modeling the face region.
- Your CNN's output segmentation mask should be scaled down by a factor of 16 compared to the input image (i.e., your segmentation does not have to be pixel-accurate).



- Use the AdamOptimizer for training. Optimize the cross-entropy as a cost function (summed over all pixels in the output mask).
- Evaluate your network, once with a random split, once splitting by person. In both cases, use about 100 test images and the rest for training.
- Generate the final segmentation mask by thresholding the output mask's pixels at 0.5. Measure segmentation quality using the Jaccard Index. Also, prepare a few visualizations of your output masks. Do they match the desired output?

- Conduct at least three different "runs", each with a different setting (varying the number, sizes and types of layers, or other parameters). For each run, keep track of all relevant parameters and results.

Exercise 5.2 (Report)

Put together a presentation of 2-3 slides, showing your quantitative results, some representative visualizations segmentation masks, and any open questions.