Machine Learning 2015: Project 3 - Feature Engineering

Ivaylo Toskov
itoskov@student.ethz.ch

Maximilian Wurm
mwurm@student.ethz.ch

Martina Rivizzigno
rimartin@student.ethz.ch

December 19, 2015

Experimental Protocol

This is the protocol of the group "I don't care".

1   Tools

In this project we splitted the work into a feature extraction and a classfication part. On the one hand we used Malab for the image processing and feature extraction. On the other hand, python together with numpy and the machine learning library "scikit-learn" was used for the subsequent classification.

2   Approach

Like mentioned above, the workflow for this project can be subdevided into two separate areas. Firstly it is important to extract as many interesting features as possible from the underlying images.
Only after that, one has enough information to classify the cells.

2.1   Feature Engineering

The area of computer vision offers several powerful methods and algorithms to extract different kinds of informations out of an image. The following ones have been taken into account within our reflections.

2.1.1   PHOG

2.1.2   SIFT: Scale Invariant Feature Transform

This method describes local features in an image.

3 Classification

For the classification we used the basically same algorithm like in the last project. An ExtraTreeClassifier Scikit-learn GridSearchCV function runs a cross validation based, fitting-performance test over all combinations of possible parameters, which are given in a separate list. The cross validation made it possible to immediately estimate the precision of a certain configuration.

4 Lessons Learned

In this project it became obvious that it is very important to extract proper features. Having proper features means not only to extract as many reasonable dimensions from the image as possible, but also to select the most interesting ones.

Concerning classification, the same aspects as for the last project were valid: Tweaking a single classifier, at some points reached its limit, so that no further progress was achievable by tweaking the parameters and the ensemble ExtraTreeClassifier has been chosen again.