

Project Plan

The starting point of this work is a previous thesis that dealt with the classification of printers using a SVM and the extraction of features using the wavelet transformation.

The question still arises whether it is possible to extend the features on the basis of expert knowledge such as droplet size (size of printed matrix dots if you zoom in), distances between the droplets, solutions dpi, etc...

Furthermore, since the current model cannot yet provide sufficient information with real data, another interesting observation would be whether factors such as different substrates (normal paper/photo paper), print settings (normal/high), resolution, etc. can influence the classification.

The goal of the thesis is therefore to investigate how far further informative features can be extracted and whether a classification model can be learned taking into account confounding factors such as different substrates, print settings, resolution.

The first milestone of the thesis is the extraction of the features based on the expert knowledge. The image processing library OpenCV shall be used for edge detection. The newly extracted features shall be compared and evaluated with the model in the previous thesis (baseline). Ideally, a satisfactory result should be achieved here.

The next milestone deals with the question whether a classification model can be learned considering confounding factors. Ideally, the goal is to achieve a sufficient result in the classification of printers, despite disturbing factors.

It's interesting to investigate how the classification model perform on different paper (photo, normal).

For instance, there should be at least 2 prints for 1 printer: 1 photo and 1 paper and the goal is to train a model which can identify the scans based on those differences.

As for now we only considered Photo paper.

For this process, different machine learning (SVM, XGBoost) and possibly, if time permits, deep learning methods (CNN, ANN) are to be used.

Accordingly, the performance of the learned model on confounding factors is to be investigated.

The results will be evaluated and discussed and will serve as a basis for the later production system.