Cervical cancer is the second most common cancer affecting woman worldwide but at the same time it is one of the most preventable and treatable cancers. Since the most common form of cervical cancer starts with pre-cancerous changes and develops very slowly, up to 90% of cervical cancers may be prevented if cell changes are detected and treated early. Early detection is undertaken using a Pap smear which is evaluated by investigating a microscope image. An automatic detection can help to process more Pap smears at a time and decrease number of mistakes caused by the human factor.

In order to identify cancerous cells in Pap smear images, several steps has to be accomplished. These are image preprocessing, image segmentation, cell classification (type of a cell) and cervical cell classification (normal or abnormal). This problem is rather difficult as a Pap smear images can contain several types of cells. There are two types of cervical cells, Squamous cells (Exocervix) and Glandulas cells (Endocervix). Cervical cancer is usually caused by the first type and is called Squamous cell carcinoma. Additionally, white blood cells (Neutrophil), metaplastic cells, yeast strands, cell debris or bacteria can appear in samples. They all can be clustered in an arbitrary way which makes even segmentation very difficult as we are interested mostly just in separation of squamous cell cytoplasm and nuclei. After segmentation, the most important features have to be extracted from squamous cell cytoplasm and nuclei and further used for classification. In this project, I will focus only on the classification part.

Classification  The classification can either determine a cell to be normal or abnormal, or assign it one of various levels of dysplasia. As it is crucial to determine whether a patient has to be treated or not, this project will consider only normal or abnormal cervical cell classification. Published methods for cervical cancer classification works with cervical features that are either extracted manually by a human expert [2] or automatically [1]. The most important features for classification are nucleus and cytoplasm area, nucleus and cytoplasm brightness/minima/maxima and nucleus roundness. A lot of methods make use of Artificial Neural Networks (ANN) [2], Mat-Isa et al. [3] developed Hybrid Multilayer Perceptron which obtains better results than classical ANN.

Goal  The goal of this project is to classify cervical cells as normal or abnormal from provided cervical cell features. Depending on the method used for the feature extraction, cell type classification might be investigated too. The used method will be ANN and might compared with some other methods if the results of ANN are not satisfactory enough.

References