One class classifier development for PAP smear image screening

CHALLENGES: Health, demographic change and wellbeing

PRODUCTIVE SECTOR: medical sector
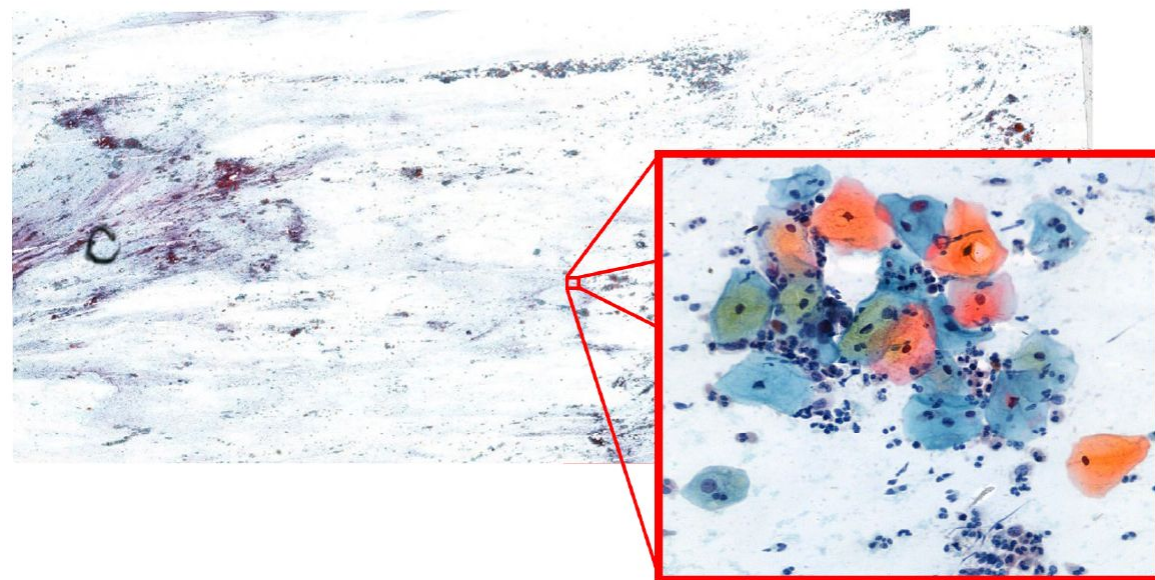
PROBLEM DESCRIPTION

The issue that we have addressed with our project is that much higher ammounts of healthy cells can be found in the training set containing manually classified/annotated cells extracted from digitalized PAP smears.

CHALLENGES

AND GOALS

The aim of our project was to investigate possibilities about how we can successfully improve the accuracy of CNN to find all the diseased cells in the digitalized smear test, in the case of an unbalanced dataset.

MATHEMATICAL AND COMPUTATIONAL METHODS

- We investigated the available CNN architectures and chose one appropriate for the task at hand.

- We examined the use of autoencoders as well as variational autoencoders for the generation of synthetic data

- We worked out a hibrid network using a combination of variational autoencoders and classification networks

- We developed a system that significantly improved the performace of CNN regarding diseased cells

Example of a digitalized smear

Deep learning to the aid of citology
One class classifier development for PAP smear image screening

Results and Benefits

The issue that we have addressed with our project is that much higher healthy cells can be found in the training set containing manually classified/annotated cells. The aim of our project was to investigate possibilities about how we can successfully improve the accuracy of CNN to find all the diseased cells in the digitalized smear test, in the case of an unbalanced dataset.

Thanks to the developed method for data balancing, the proportion of diseased cells found increased significantly. The developed algorithm is based on generative networks, which can be used as a hybrid network to generate synthetic images, that can be utilized to train a convolutional neural network with a modern architecture by means of data balancing and a combined cost function.

We developed a system that solves the unbalanced dataset problem and improves the performance of the CNN used for cell classification

SIGHTSPOT NETWORK

SightSpot Network Ltd.