Cell Detection based on Gaussian Process Regression

Wei Wang Niklas Hansson Reidar Luid

Semester Project for Advanced Machine Learning , Technical University of Denmark

Problem

Microscopy images usually contain a huge number of cells. Detecting cells through visual analysis the images can be tedious and time-consuming work because of drawbacks like

- Inter-observer variability
- Intra-observer variability

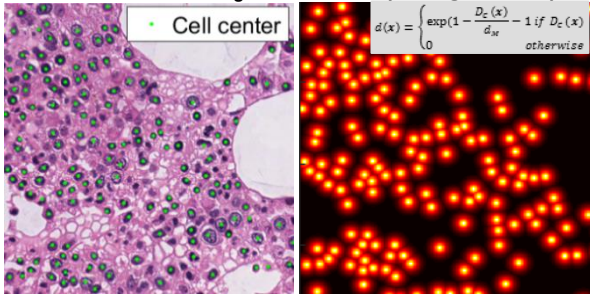
Therefore, results from visual inspection are often inconsistent.

Solution? Machine Learning!

Data

- 11 600x600 bone marrow cell images
- Corresponding ground truth cell locations

Patch of 600x600 image **Corresponding score-map**



Learning goal

Learn a machine to detect if a given pixel in an input image is a cell center or not. We do this by:

- Using GP regression to train on cell images and corresponding score map
- Subsequently predict a score map of a given input image

Key equations

Kernel function (also called the covariance function)

$$k(x_q, x_p) = \theta \exp(-\lambda |x_p - x_q|^2)$$

Predictive mean

$$\bar{f}_* = \boldsymbol{k}_*^T (K + \sigma^2 I)^{-1} \boldsymbol{y}$$

Predictive variance

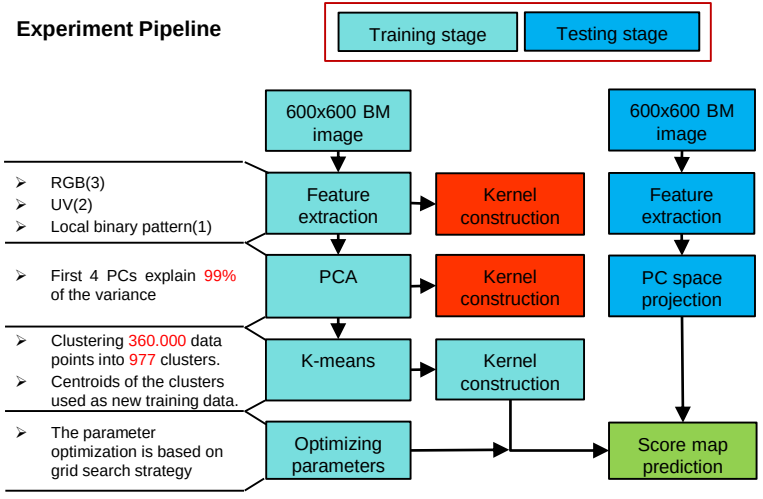
$$V[f_*] = k(\boldsymbol{x}_*, \boldsymbol{x}_*) - \boldsymbol{k}_*^T (K + \sigma^2 I)^{-1} \boldsymbol{k}_*$$

Potential issues

360.000 data points for each image:

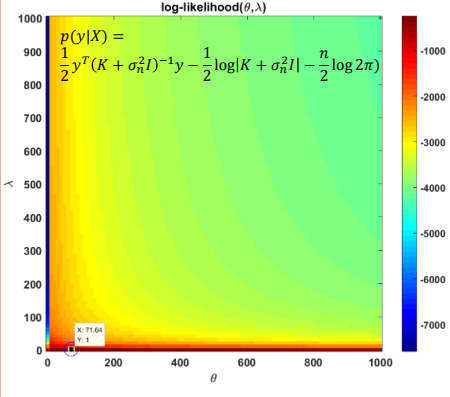
- Memory issues with a too large K matrix
- Invertability of K matrix questionable

Experiment Pipeline



Hyper Parameter Optimization

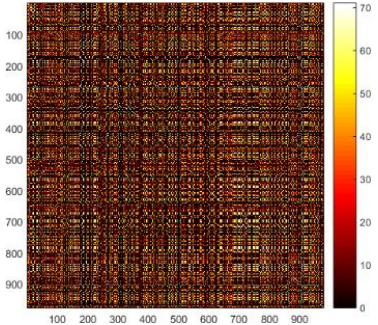
Coordinate search strategy is applied, in order to find the local maximum of the log-likelihood function



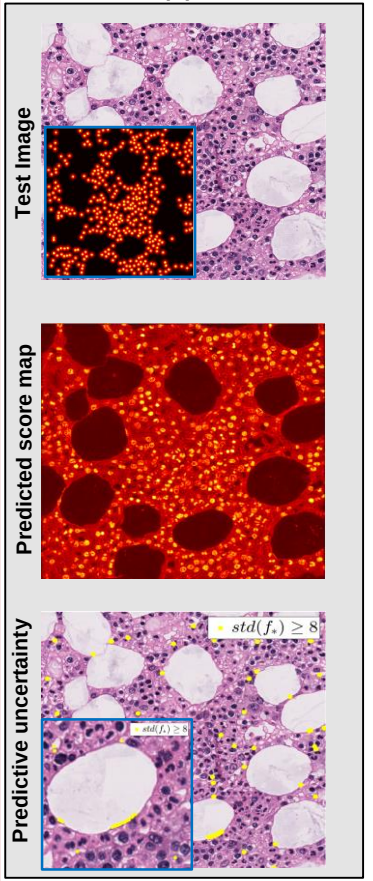
Kernel matrix construction

The dimension input data. In our case, number of training data after Kmean is 977

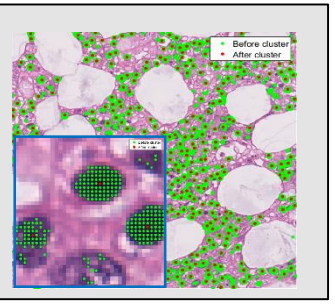
We conduct **PCA** to reduce the **dimension** of **the feature**, and Kmeans to reduce the needed **number of data** in order to compute *K* matrix



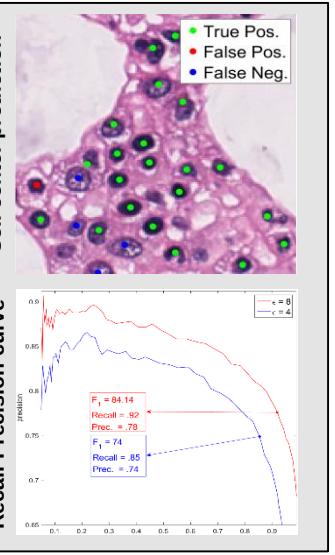
Score map prediction



Post-processing predictions



Model evaluation



$$Precision = \frac{TP}{(TP + FP)}$$

$$Recall = \frac{TP}{(TP + FN)}$$

$$F_1 = 2 * \frac{Precision * Recall}{(Precision + Recall)}$$

F_1-measure	
$F_{1_{MICCAI}}$	0.8717
$F_{1_{Ours}}$	0.8414

Training time time ≈ $12h$

Testing time ≈ $1min.$