

Project Presentation

Biomedical Image Analysis

Image Analysis Methods for Automated Cell Counting

13th April 2020

Introduction

- The number of cells in a cell culture can give useful information about cell density to understand the culture condition.
- Usually, cell counting can be achieved manually, but the reliability highly depends on the experience and concentration of the examiners.
- So automated cell counting based on image analysis and computer vision methods is an essential tool to improve the accuracy of the process.
- The general goal of this project is the development of image analysis methods to enable the automated counting of human promyelocytic leukemia (HL60) cancer cells using images obtained with a light microscope and a hemocytometer (piece of glass on which lines are drawn in a certain size to help manual counting).
- This project is inspired in the following publication:

Akin Ozkan *et al.*, “Computer Vision based automated cell counting pipeline: a case study for HL60 cancer cell on hemocytometer”, Biomedical Research 2018; 29 (14): 2956-2962.

Full set of images is available at http://datasets.akinuzkan.net/HL60_HEM40X_CCData

Objectives

The project has two main goals:

1. Image Region of Interest (ROI) delineation (Task 1)

The first goal of the project is the automatic delineation of the rectangular region of the image where cells will be counted, usually called Region of Interest (ROI). The ROI delineation data includes the original image, paired with the manual delimitation (ground truth-GT) in the form of a binary mask.

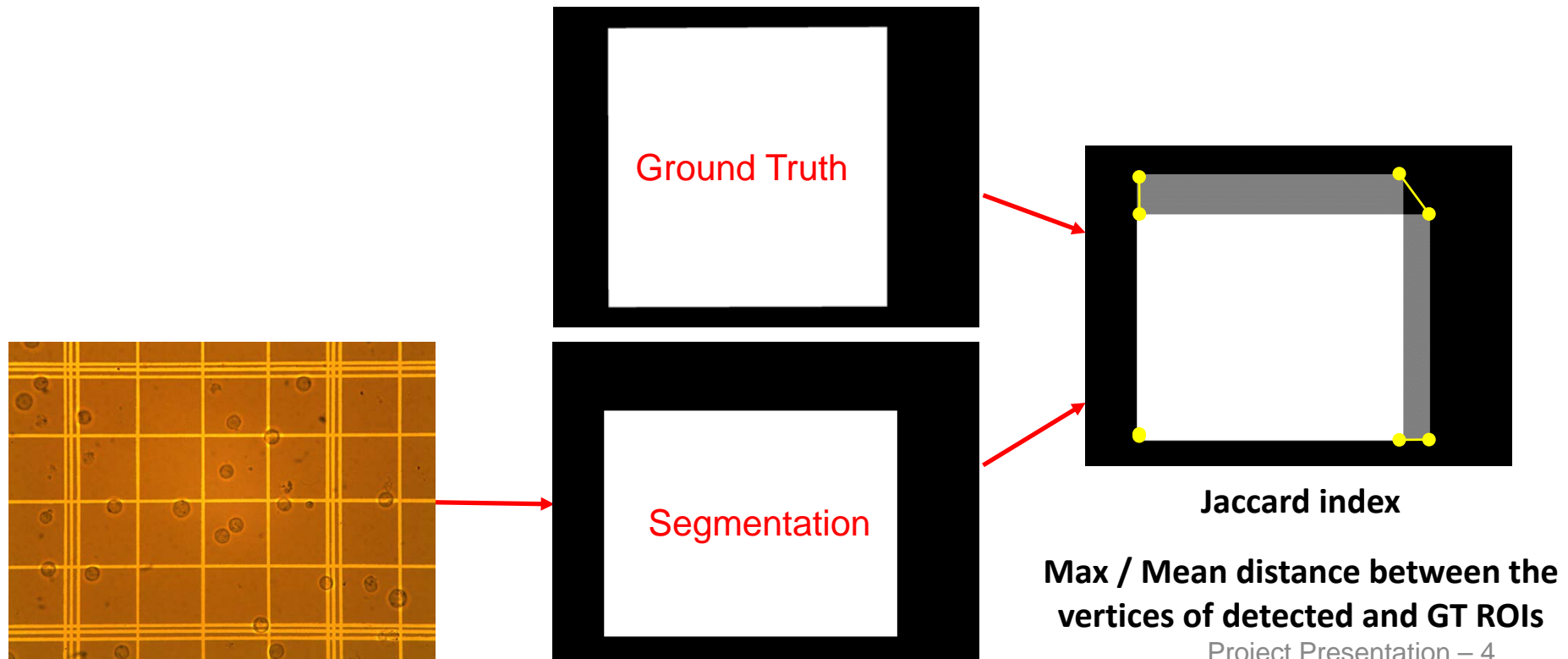
2. Cell segmentation and counting (Task 2)

In this task, students are asked to segment the cells inside the ROI and count them. The cell segmentation data includes, for each image, the same data provided for ROI delineation (image + ROI), and, as GT, the position of each of the counted cells.

Project development

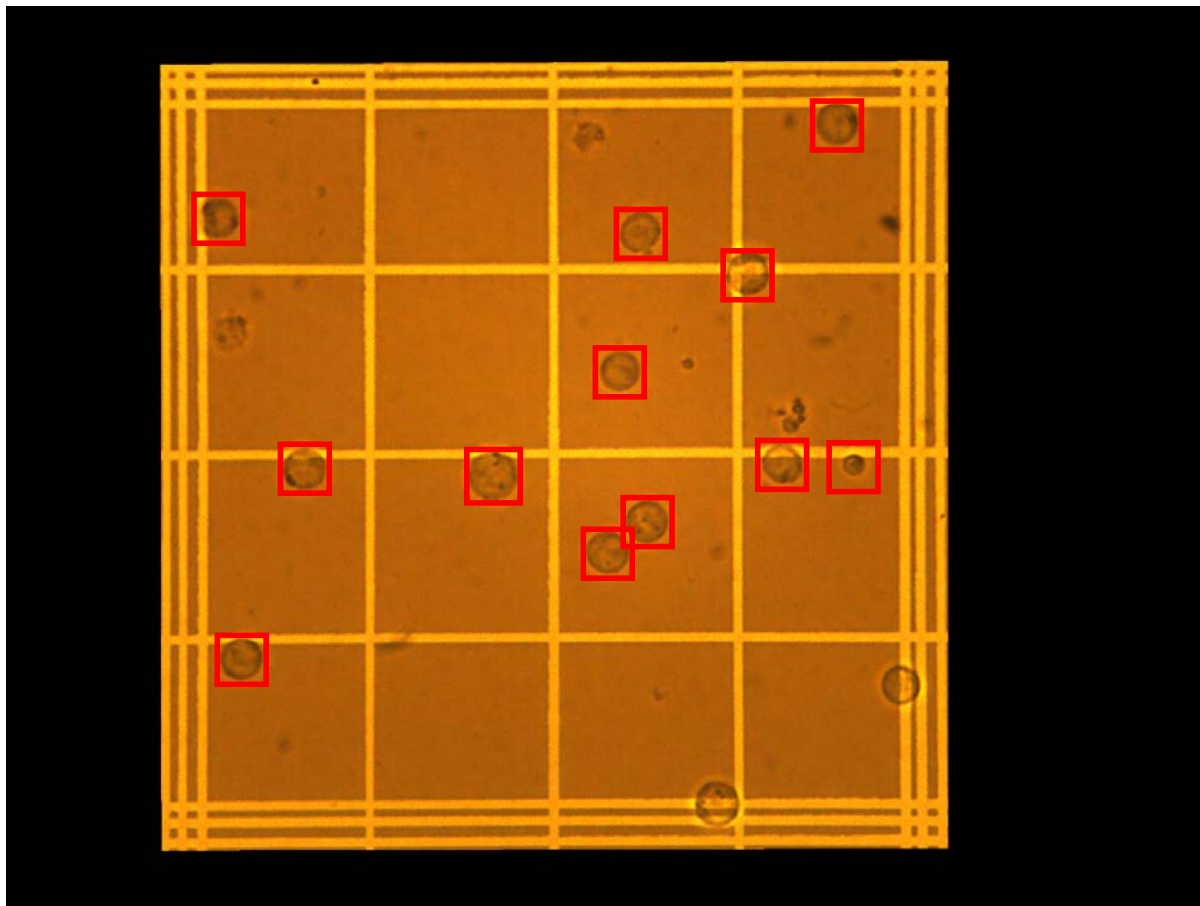
Task 1 - ROI delineation: The algorithm for ROI Delineation should have as input an original image and as output a binary image with the ROI area represented as a white rectangle in a black background.

This binary image should be compared with the corresponding ground truth (GT) image in order to evaluate the performance of the developed algorithm using: a) the Jaccard index; b) the maximum and mean values of the Euclidean distance between the vertices of the detected and GT ROIs.



Project development (cont.)

Task 2 - Cell segmentation and counting: The algorithm for cell segmentation and counting should use the GT for the ROI mask provided with the dataset, and segment the areas of the ROI that can be considered as cells. Then, the number of detected individual areas will be considered as the total number of counted cells.



Number of cells = 12

Project development (cont.)

Task 2 - Cell segmentation and counting

For evaluating the performance of the algorithm, the GT for task 2 is given as a “.mat” file for each original image (with the same name of the original image but with a “.mat” extension).

Each file contains a matrix that has as many rows as the number of counted cells, and 4 columns with the position of the cells. Relevant data for each cell are the coordinates of the top left corner of square region delineating the cell (in the two first columns of the corresponding matrix line), and the size of the square region that contains the cell (in the third and forth columns).

As some cells can be located over the border of the ROI, the protocol for counting cells should include all cells that do not touch the border of the ROI and all the cells that overlap the top and left borders of the ROI. Cell positioned over the bottom and right sides of the ROI are not included in the GT, so should not be considered for the final counting result.

Project development (cont.)

Task 2 - Cell segmentation and counting

The results to obtain for each image of the are:

- . The **number of counted cells** measured by the number of regions obtained by the developed segmentation method;
- . The **number of true positives (TP)**:
A segmented area is considered a true positive if the corresponding Jaccard index is equal or larger than 0.5; the number of true positives is identical to the number of correctly detected cells;
- . The **number of false positives (FP)**:
A segmented area is considered a false positive if the corresponding Jaccard index is lower than 0.5;
- . The **number of false negatives (FN)**:
A false negative is a cell from the GT that does not correspond to any segmented region;
- . The **Recall, Precision and F_1 -measure (R, P, F_1)**

Average values for the whole dataset should also be reported.

Datasets

The dataset to be used for algorithm development (training set) is composed of 50 images. The corresponding GT data, already described previously, is formed by the binary image with the ROI GT and the “.mat” file with the number and position of the counted cells.

A second, independent, test set of 50 images will be provided in the day before submission. Results for the training and test sets should be reported separately.

Major guidelines

- Students should form groups of **2-3 students**; however, upon request, a group can be formed by just one student. Groups with just one student are only required to develop Task 2.
- The delivery deadline for the report, code and contributions is May 22nd, 2020.
- The delivery deadline for the material of the presentation is May 24th, 2020.
- These dates can be changed depending on decisions taken by FEUP regarding the academic calendar.

Deliverables

Each group is required to deliver:

- A four A4 page report in IEEE Transactions article format describing the methodology and obtained results, both qualitatively and quantitatively;
- The MATLAB code developed during the project;
- The material prepared for the public presentation of the project (15 minutes maximum for groups with 2/3 students; 10 minutes maximum for individual groups);
- A brief description of the contribution of each element of the group to the project.