Counting stems cells with Machine Learning

Lionel Chiron

Report

Abstract

Counting cells in an image by hand when datasets are composed of thousands of acquired images is very time consuming, it is prone also to errors due to the lack of attention when the operator is tired. Automating such a task even semi automatically gives accelerates and makes the accomplishment of the tasks more robust. Currently stem cells images are acquired at a very low resolution (10x) this does not permit to use Machine Learning directly since the models at this resolution cannot be precise enough to avoid false detections. We use hence Machine Learning also for determining the region of interest with two different methods: A statistical one and another one using the patterns linked locally to the presence of the cells.

Keywords: Machine Learning, Stem cells

Counting stems cells with Machine Learning

# Introduction

Counting cells by hand is long, fastidious, prone to errors. With recent techniques like Deep Learning using convolutional neural networks, this task can be automated robustly enough without spending an excessive amount of work for conceiving the algorithms. If the patterns targeted contain sufficiently details for the algorithms to be able to recognize them with a high level of fidelity, the difficulty rely more on the technical problem of selecting the good training sets. Here the patterns we want to recognize are not precise enough for the Machine Learning to be able to distinguish between a true and a false positive.

We need to add supplementary filtering steps so as to remove the false positive detections.

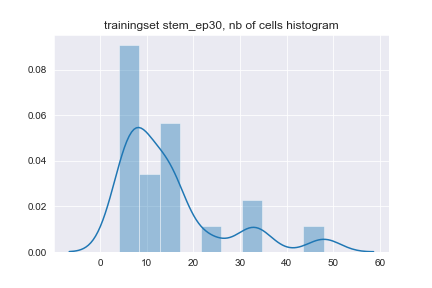
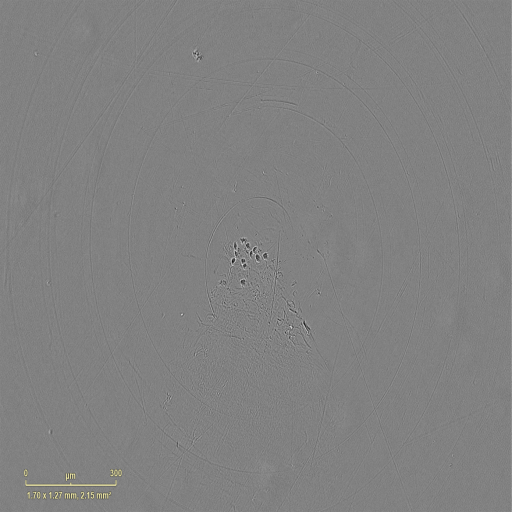
## Detecting cells patterns

A method of choice for detecting the stem cells is to use U-Net algorithm trained on a set of images with a large variety of shapes and size (rounded cells, elongated ones etc...).

### Testing various trainingset with different characteristics.

#### Rounded shapes varying the number of cells on clean images

A first flavor of models was divised for rounded stem cells at focus with black pattern inside and grey white edges for the contours. Those models were produced from trainingset (training\_hemato\_stem\_cells) with images with an average of 14 cells on 20 pictures.



#### Cells with dirty background

Here the idea is to train on images not perfectly clean so as to handle both dirty and clean experiments to target a larger range of images. Those models contain the prefix “training\_hemato\_stem\_cells1”.

C:\Users\Lionel\AppData\Local\Microsoft\Windows\INetCache\Content.Word\3.tiff

#### Larger variety of cells, grey cells, out of focus

#### Larger trainingsets

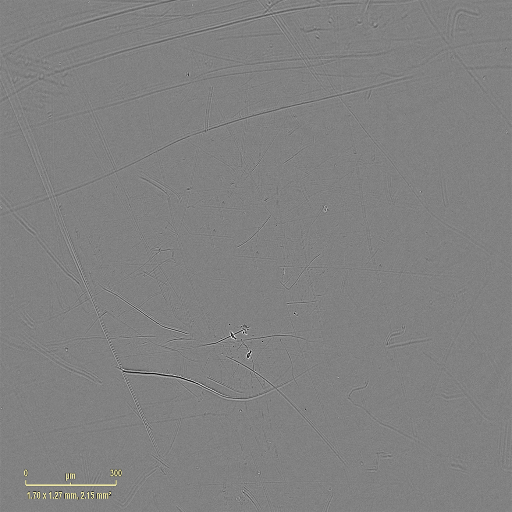
#### Few cells with dirt trainingsets

C:\Users\Lionel\AppData\Local\Microsoft\Windows\INetCache\Content.Word\21.tiff

#### Few cells clean trainingsets

C:\Users\Lionel\AppData\Local\Microsoft\Windows\INetCache\Content.Word\132.tiff

#### Only flat cells trainingsets



### Conclusion

#### 

## Combining multiple models

### Principle

Each model (few cells, cells with dirt, large span in number, flat cells) has its limits so the idea is to run many models and combine what they see in one prediction image.

### Experiment

### Conclusion

## Filtering the false positives

The idea is to perform an extensive detection of all the possible cells including false positives and to use afterwards to filters out the false positives.

### Different ways to filter

Various way to peter out the false positives can be envisaged.

* Proximity between the detections
* Not close to forbidden shapes
* Shapes move more than the mean
* Typical area of displacement

### Statistical method

### Associated patterns method

### Experiment

### Conclusion

References

Last Name, F. M. (Year). Article Title. *Journal Title*, Pages From - To.

Last Name, F. M. (Year). *Book Title.* City Name: Publisher Name.

Footnotes