Bioinformatics Lab Assignment 6: Image Processing of Cancer Cells in Python

By Noah Segal-Gould

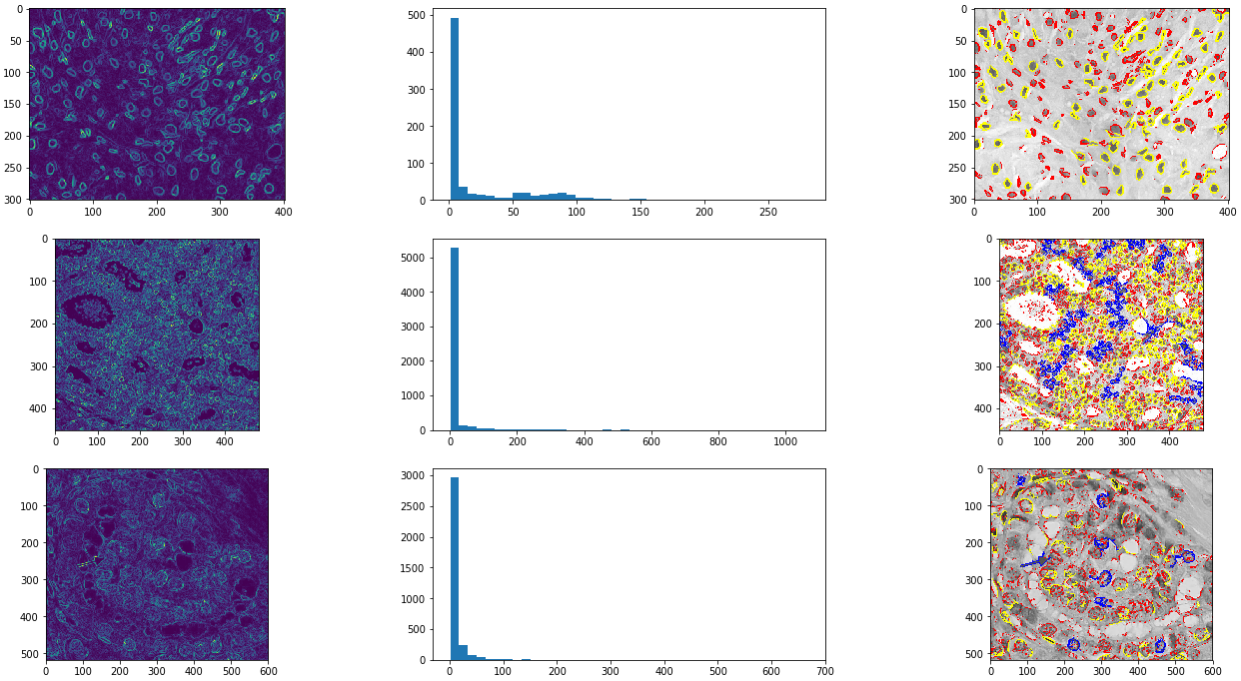
Background: In pursuit of cancer research, biologists can identify and count cells by looking at stained tissue samples. For images featuring more than hundreds of stained cells, a trained human eye can only work so fast. Instead of manually counting the cells, a microscopic picture can be taken and an image processing program can be utilized to do the work typically done by trained humans. The goal of this lab assignment is to apply algorithmic techniques for processing images of cells as well as for counting the cells present within those images using the Python programming language.

Methods: For this assignment, I created a function called `image_processing_algorithm` which takes as its arguments a list of dictionaries containing image filenames and my own personal estimates of the number of cells in each of those images, as well as a desired filter mode. The function utilizes the Stevens filtering method, the Stevens filtering method without any Gaussian blur filtering, and my own method I developed. It iterates through the three input images and applies the desired filtering method. My version takes advantage of the filters that are made available by default within the Python Image Library (PIL) in order to smooth the input image and apply an `UnsharpMask` to get better borders around the cells. Using `matplotlib`, I consolidated the 3 input images, 3 histograms, and 3 output images with included cell borders into one larger plot for easy viewing. The errors and mean errors are computed and printed along with the statistics on the number of small, medium, and big blobs as well as my personal cell count estimates. This function is run three times at the bottom of my program for steps 2, 3, and 4 of the assignment description.

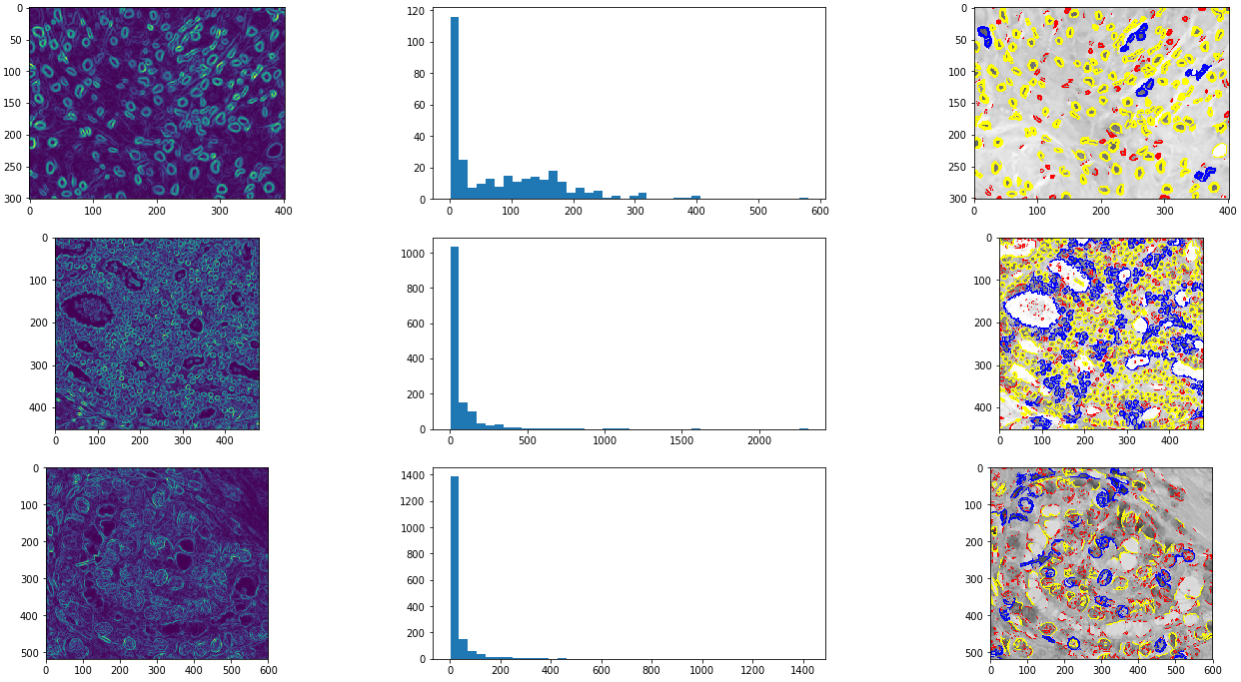
Results: As illustrated below, the mean errors between the Stevens method with and without Gaussian filtering are much closer to one another than either is to the mean error when my method is applied:

	CellsHE1.jpg		CellsHE2.jpg		CellsHE3.jpg		
	Estimate	Actual	Estimate	Actual	Estimate	Actual	Mean Error
Stevens	80	100	278	450	98	150	31.0%
Stevens without filtering	138		449		210		26.1%
My method	93		431		158		5.5%

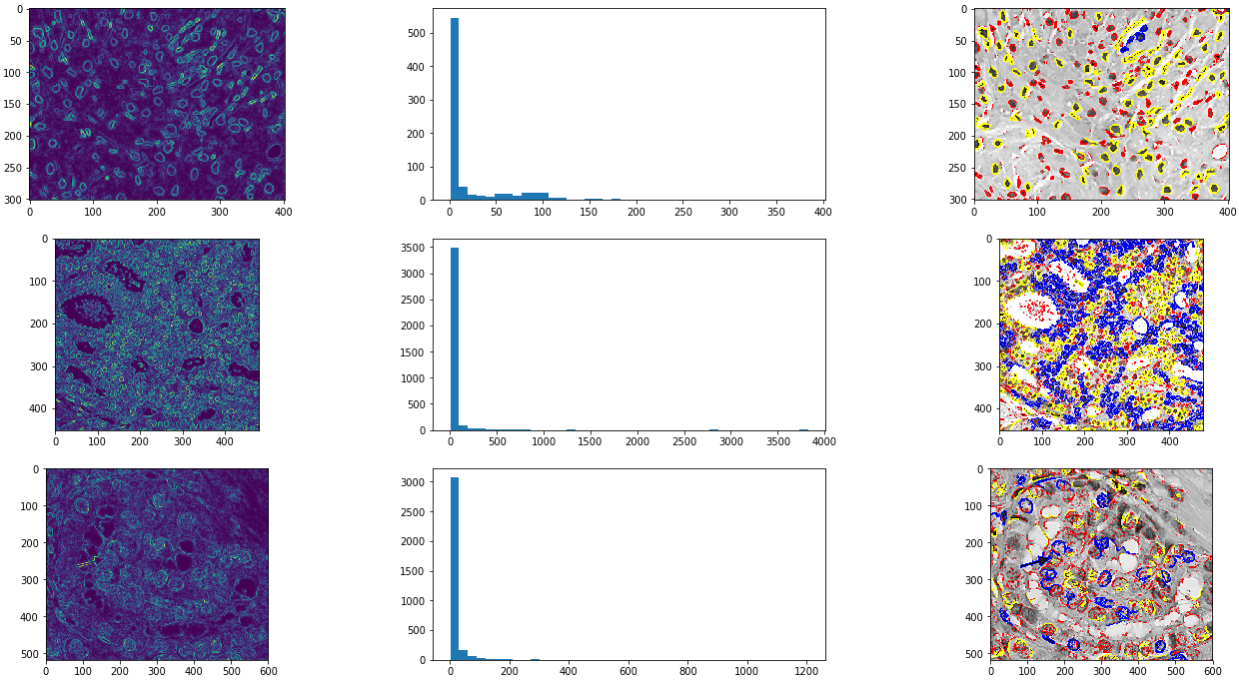
Stevens:



Stevens without filtering:



My method:



Conclusion: I successfully completed all the steps assigned in the lab assignment description. My program outputs errors and images appropriately.

Acknowledgements: I worked alone on this assignment. I used code from the course Moodle 2, the textbook, and the Python Image Library documentation.