Detecting and measuring cervical cancer cells.

Erick Calderon-Morales

5/11/2021

Introduction

Cancer is a disease that affects XXXX humans each year in low income countries. Early detection and treatment of this diseases can potentially improve the probability of fully recover. Despite of this, early detection of cancer cells can be challenging. In this report I show the results of a work flow that demonstrates the potential use of several computer vision techniques for detecting and measuring cancer cells from microscopic RGB images.

The dataset consist of 499 cancerous smear images coming from Wroclaw Academic Hospital in Poland. Images come from patients suffering from various forms of cervical cancer. For this report I used 7 types of cancerous cells. Columnar epithelial cells (type 1) are characterized by being rectangular, taller than it is wide, and are most commonly found in a single-layer arrangement. Parabasal squamous epithelial cells (type 2) are typically round or nearly round and have a high nuclear to cytoplasmic ratio. Intermediate squamous epithelial cells (type 3) are polygonal-shaped with a thin and transparent cytoplasm. Superficial squamous epithelial cells (type 4) are characterized by having a polygonal shaped. Mild (type 5), moderate (type 6) and severe nonkeratinizing dysplastic cells (type 7) are characterized by oval to spindled nuclei, high nuclear to cytoplasmic ratios, indistinct cell borders and little to no surface maturation. (Win et al, 2020).

Workflow

For the analysis of the cancer cell images I followed the steps proposed by Win et al (2020).

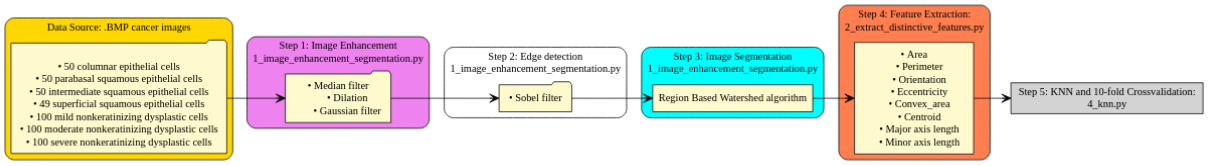

Figure 1: Workflow followed. Each box represent each step followed. The name of the code for carrying the analysis has the extention .py rigth after the step name

Image enhancenment and segmentation

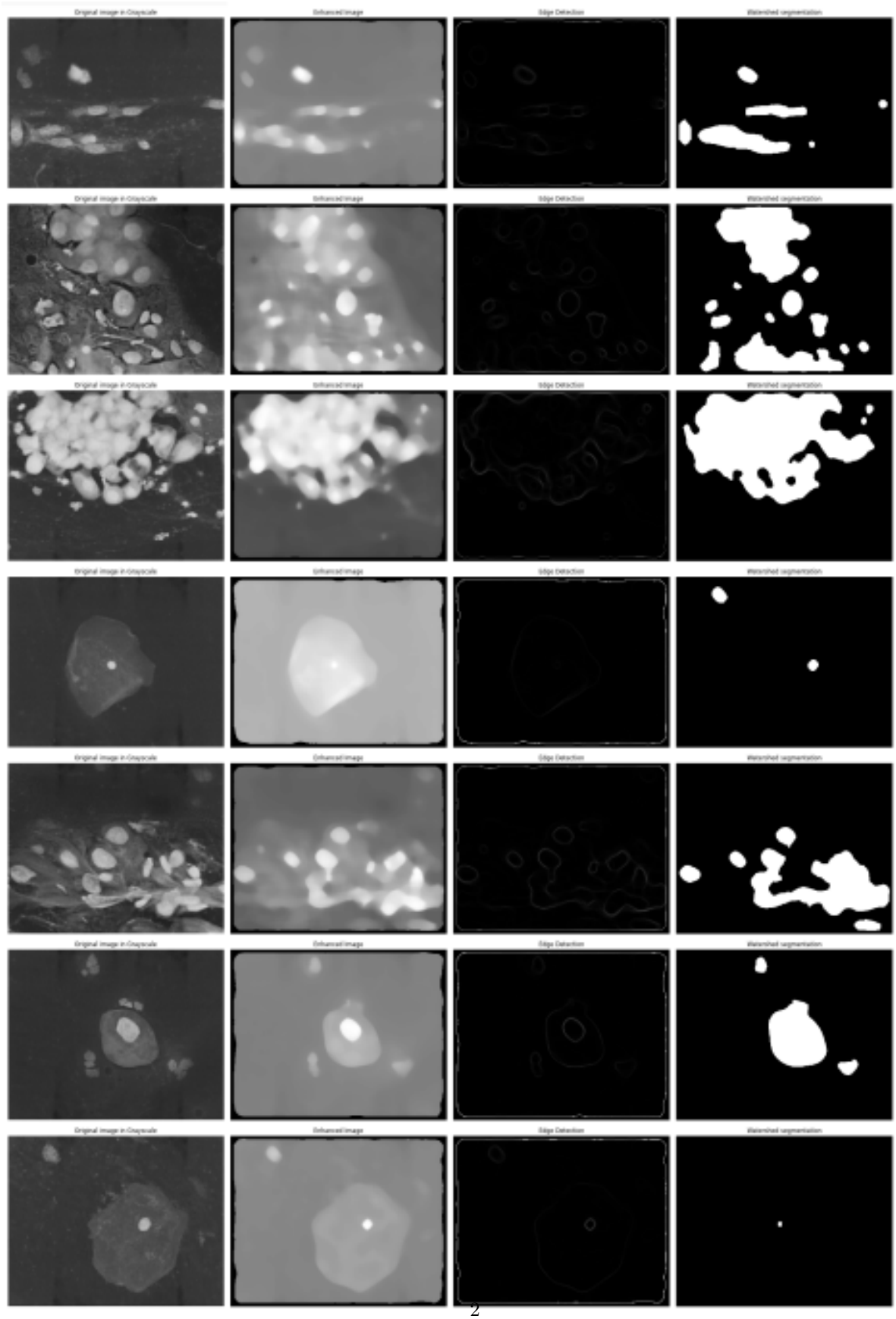

Figure 2 (above) shows the results of the Steps 1, 2 and 3 obtained with the code called 1_image_enhancement_segmentation.py.The first column in figure 2, from left to right, shows the original original image, the second column shows the image after the median, dilation and gaussian filters were applied. The third shows the result after the Sobel algorithm was applied. Finally, the last column shows the result after the watershed algorithm was applied.

KNN and 10-Cross Validation

K = 3

- Scores: [34.51327433628318, 32.743362831858406, 43.36283185840708, 45.13274336283185, 30.08849557522124, 28.31858407079646, 27.43362831858407, 38.05309734513274, 32.743362831858406, 35.39823008849557]
- Maximum Accuracy: 45.132743%
- Mean Accuracy: 34.779%

K = 5

- Scores: [36.283185840707965, 40.707964601769916, 30.973451327433626, 39.823008849557525, 27.43362831858407, 40.707964601769916, 31.858407079646017, 24.778761061946902, 40.707964601769916, 40.707964601769916]
- Maximum Accuracy: 40.707965%
- Mean Accuracy: 35.398%

K = 7

- Scores: [31.858407079646017, 30.08849557522124, 34.51327433628318, 37.16814159292036, 22.123893805309734, 29.20353982300885, 27.43362831858407, 32.743362831858406, 36.283185840707965, 27.43362831858407]
- Maximum Accuracy: 37.168142%
- Mean Accuracy: 30.885%

K = 9

- Scores: [30.973451327433626, 34.51327433628318, 30.08849557522124, 26.548672566371685, 26.548672566371685, 28.31858407079646, 24.778761061946902, 29.20353982300885, 31.858407079646017, 32.743362831858406]
- Maximum Accuracy: 34.513274%
- Mean Accuracy: 29.558%

K = 11

- Scores: [28.31858407079646, 24.778761061946902, 23.893805309734514, 33.6283185840708, 25.663716814159294, 26.548672566371685, 26.548672566371685, 27.43362831858407, 31.858407079646017, 29.20353982300885]
- Maximum Accuracy: 33.628319%
- Mean Accuracy: 27.788%

Reference

Win, K. P., Kitjaidure, Y., Hamamoto, K., & Myo Aung, T. (2020). Computer-Assisted Screening for Cervical Cancer Using Digital Image Processing of Pap Smear Images. Applied Sciences, 10(5), 1800. rt_3/project_report_part_3/diagram.png)