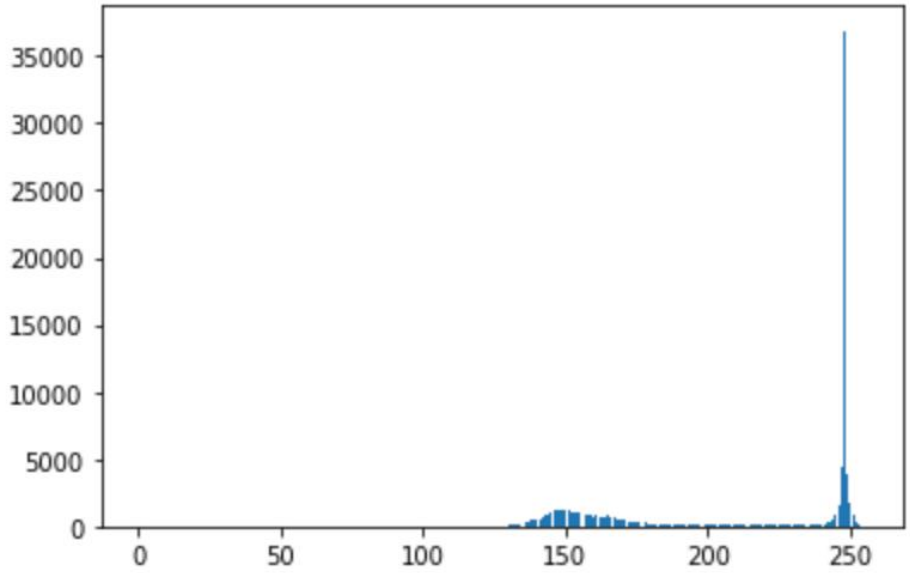
Assignment 1 Report

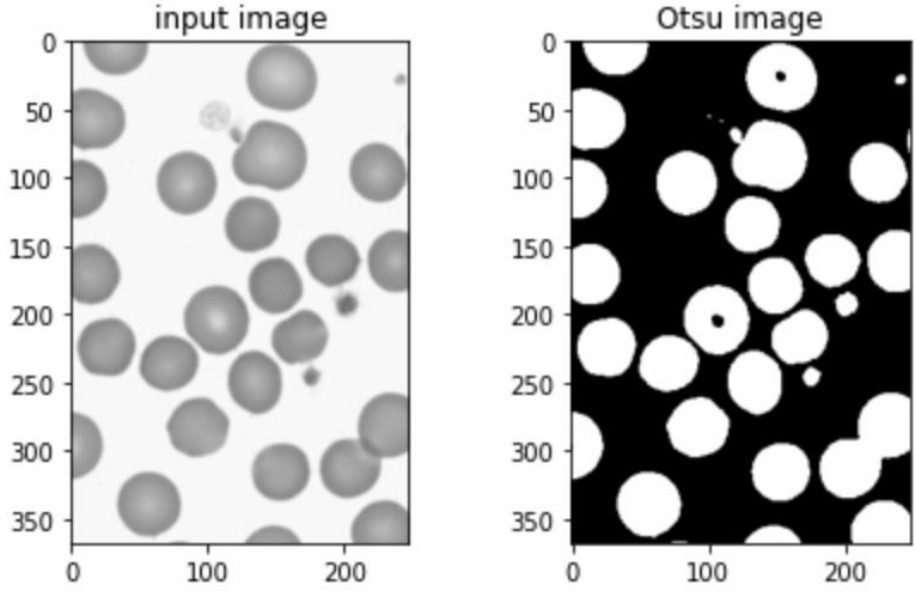
Written by XiaoHu z5223731
The version of Python is 3.8.2
The version of OpenCV is 4.5.1
“c3.jpg” is using for explanation in this report.

Task1: Otsu method

Read the gray scale of image of c3, then using the Otsu method to find the appropriate threshold of the image. Helping functions such as “otsu_threshold()” which can be found in Ass1.ipynb. Below is the image of gray scale, output of Otsu method, histogram of each pixels and threshold:

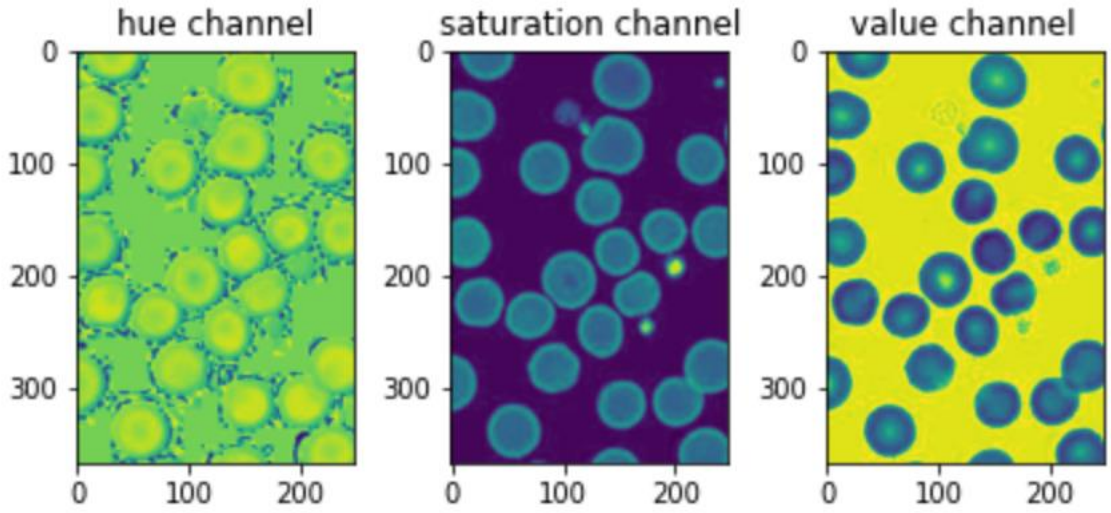


The threshold of image is 201

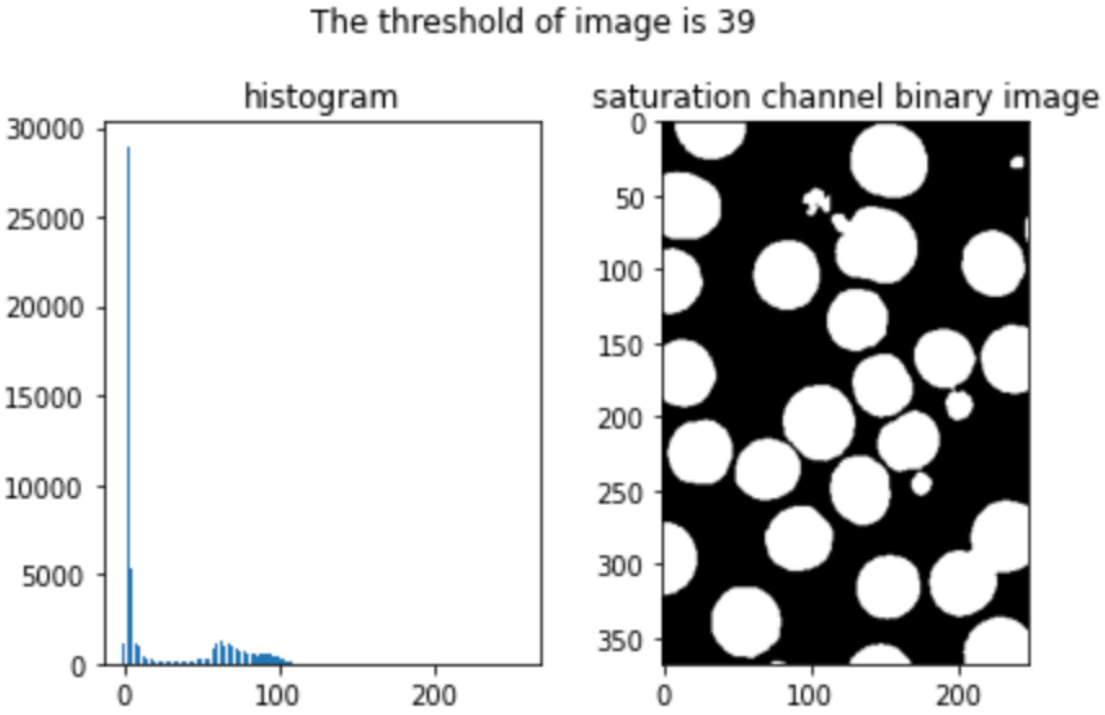


Task1: HSV

Read normal image like “img = cv2.imread('c3.jpg')”, then splitting each channel of HSV image by using cv2.split() function which is permitted.



Using the saturation or value channel to distinguish the cells and background is much way easier, but the value channel is similar to the RGB format image whose threshold have determined before. Therefore, taking saturation channel and apply Otsu method to determine the threshold. The output image and threshold are the following:



Task2: Counting the number of cells

Firstly, applying median filter to the image produced by task1 in order to reduce the noisy. In the second, using the connected component labeling method to distinguish the cell and background. Since the disjoint cell will have different label, using the 'dictionary' which is data type to store and compute the area of the label.

Determine minimum and overlapping area

To achieve these, copying the area in the dictionary contains area of each label into a list data type, and sort it in an ascending order. using functions which are 'find_threshold_growth()', 'find_threshold_var()' and 'check_over_lapping()' to determine the minimum and overlapping area. The function 'find_threshold_growth()' is to check the growth rate between two area in the list, for example, a1 = list[1] which is 100 and a2 = list[2] which is 200, then growth rate is (a1-a2)/a1 = 1, assuming the growth rate between area of cell and noisy is 1.8. Since the list that contains area is sorted, if there is one growth rate of area is greater or equal to 1.8, then all the areas after is considered as cells. Calculating the mean area of cells, using the mean and function 'check_over_lapping()' to determine the area is overlapping or not.

Assumption: if the area is greater or equal to the $(mean\ area) * 2$, then the area can be considered as overlapping.

However, If the growth rate doesn't fluctuate too much, for example, the growth rate between two area are 1, 1.1, 1.2, 1.2, 1, 1.1, which is hard to distinguish the noisy and cell (the growth rate between two area are 1, 1.1, 1.2, 1.9, 1, 1.1, there is a "big jump" between 1.2 and 1.9, it is easier to distinguish the cell and noisy), in this case, the function 'find_threshold_var()' is using to compute the threshold area by maximising the inter-class variance, and take 200 as minimum area, and using the threshold area and function 'check_over_lapping()' to determine the area is overlapping or not.

Assumption: the overlapping error is 0.15, which means if the area of cell is greater or equal to $(threshold\ area) * 2 * 0.85$, then the area can be considered as overlapping.

The following image is the number of cells of image 'c3.jpg'