The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

COURSEWORK 1

Name : Chai Wen Jye

Student ID: 20113681

Lecturer: Mr. Amr Ahmed

Module: Introduction to Image Processing (COMP2032)

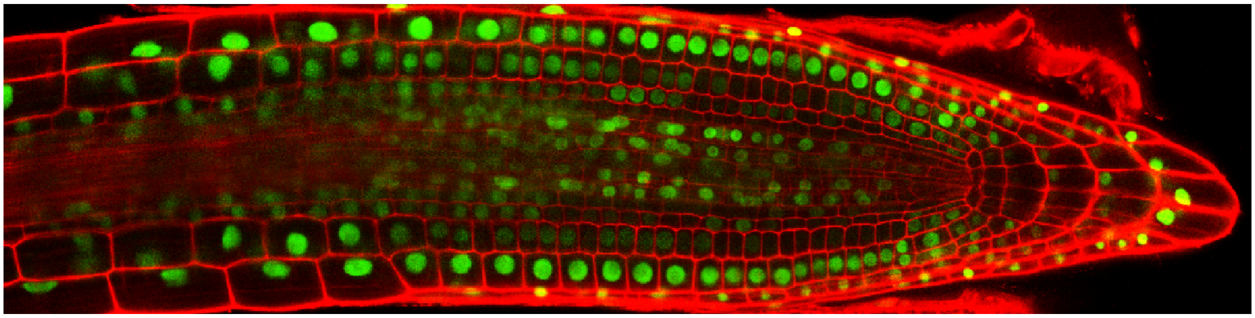


Fig 1. The sample image

1. Image Processing Technique Applied

1.1 Converting Color Space

The image provided had nucleus in green colour and cell wall in red colour. (Ajmal, Hollitt, Frean, & Al-Sahaf, 2018) compared the performance of the different color spaces for image processing. The studies showed that HSV color space gives more perceivable information for extracting interesting regions than the RGB color space. Hence, the image is first converted into HSV image.

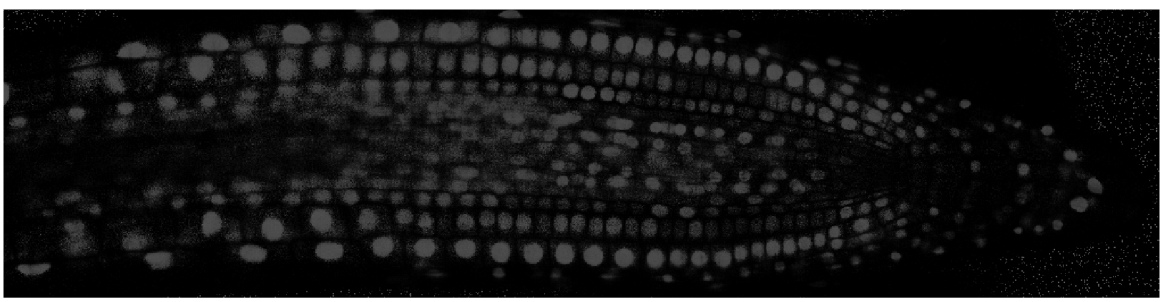


Fig 2. The hue representation of the sample image in HSV form.

1.2 Noise Reduction

We can observe that due to the nature of the sample image, the red colour region (cell wall) is suppressed as compared to the green colour. Hence, we can use this for the nucleus identification process. However, noise can be observed surrounding the image. So, the next step we carried out to eliminate those noise by performing element-wise multiplication on the image with a binary image that highlights the green region of the image. It is obtained by binarizing the green layer of the sample image in rgb form.

Fig 3. The binary image mask

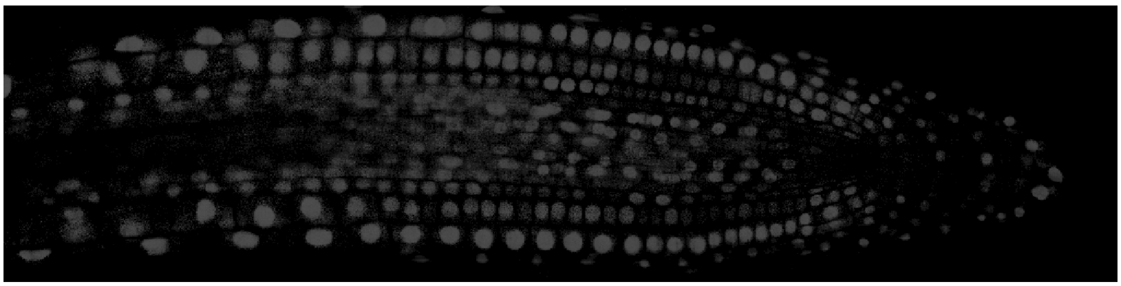


Fig 4. The resulting image after noise suppression.

Next, we want to eliminate the red region of the image. Although the value of the region is insignificant, it still serves as speckle noise in the image. We eliminate it by resetting the red region to 0. To achieve so, we perform otsu thresholding to the image to differentiate the foreground (green region) and the background(red region) of the image. After doing so, we reset the value smaller than the thresholding value to zero by using matrix operation. Next, we performed a further noise reduction process by implementing gaussian filtering.

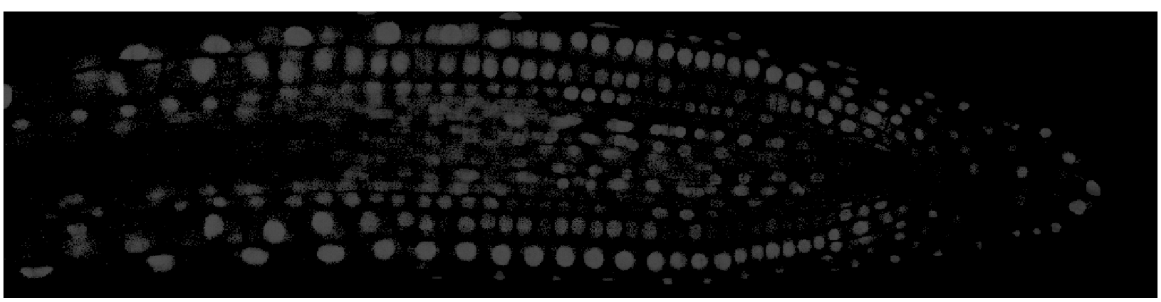


Fig 4. The resulting image after otsu thresholding.

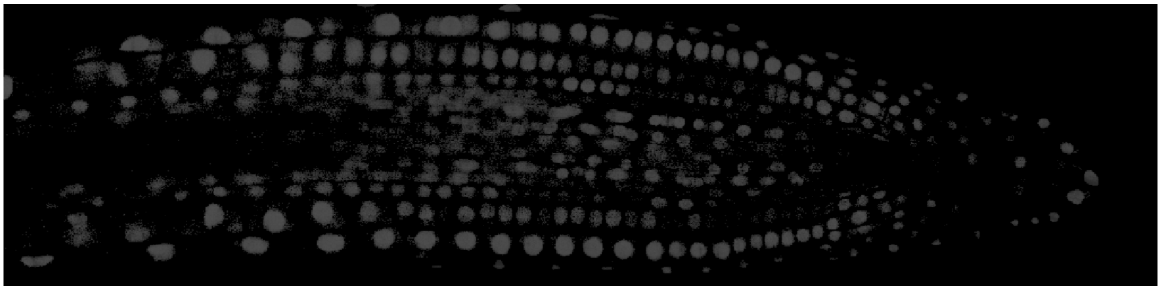


Fig 5. The resulting image after gaussian filtering.

The last step we performed for noise reduction is morphological opening. This process is to remove blur edging of the nucleus. However, in order to prevent images from being overfiltered, we use a matlab builtin function called ‘imreconstruct’ which helps to morphological reconstruction of the image marker under (morphological opening image) the image mask (gaussian filtering image).

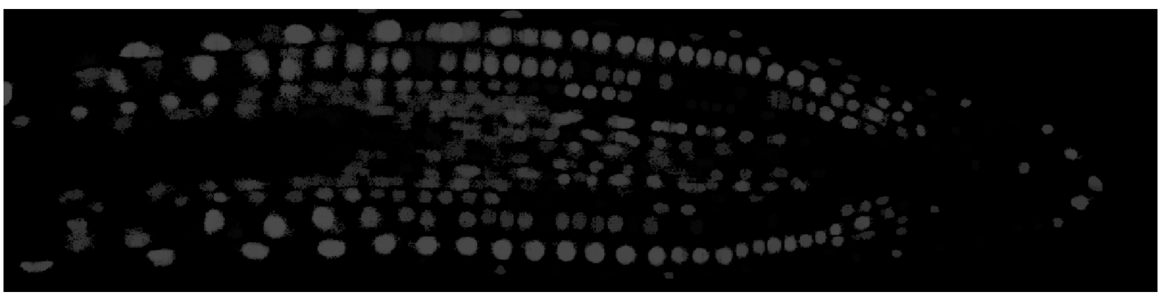


Fig 6. The resulting image after morphological opening and reconstruction.

1.3 Nucleus Identification(Segmentation)

With the result image above, we can identify the nucleus for performing a matlab builtin function called ‘imregionalmax’. It helps to identify the nucleus as we assume that each of the nucleus should be a region maxima.

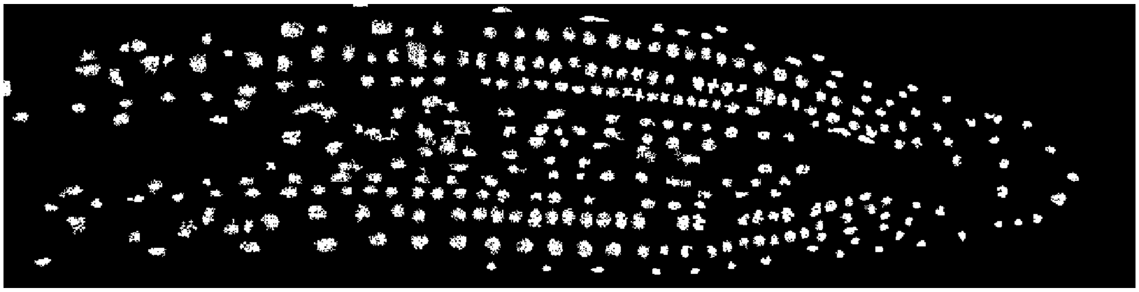


Fig 7. The binary image after performing ‘imregionalmax’ function.

After we obtain the image, we can then perform watershed function on the binary image to analyse the watershed ridgeline. As we can see, due to the nature of the ‘imregionalmax’ function, the identified nucleus seems to be eroded. In order to restore its shape, we dilate the binary image by using the same structuring element that we used earlier for the morphological opening process. To prevent that dilation process might merge the identified nucleus we subtract the result image with the watershed ridge line image that we obtained earlier.

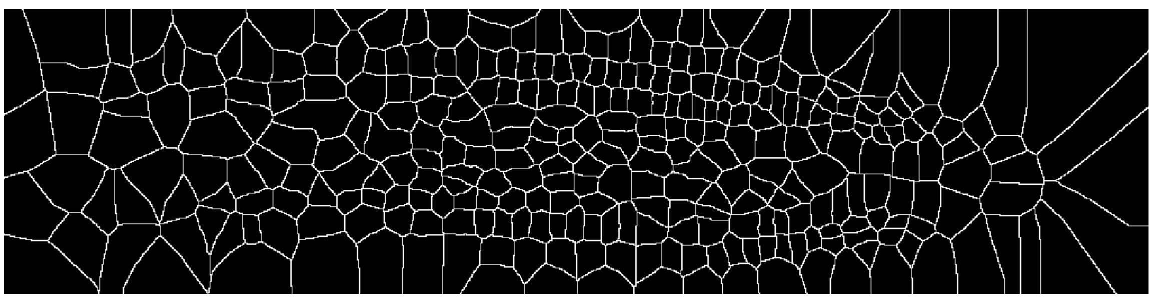


Fig 8. The resulting binary image of watershed ridgeline from figure 7

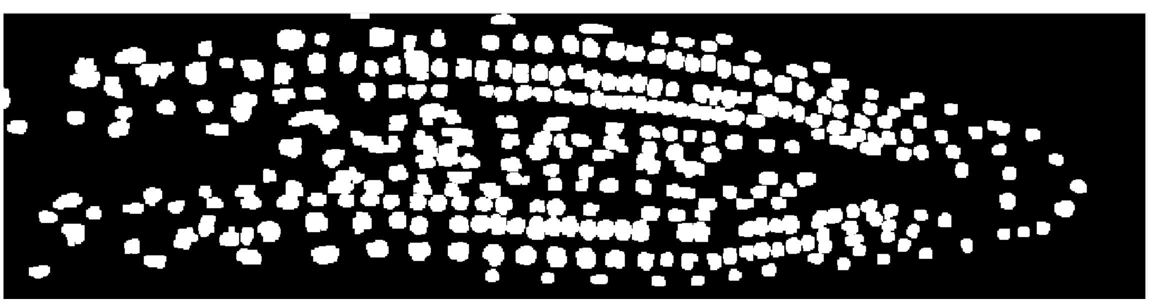


Fig 9. The binary image after dilation from figure 7

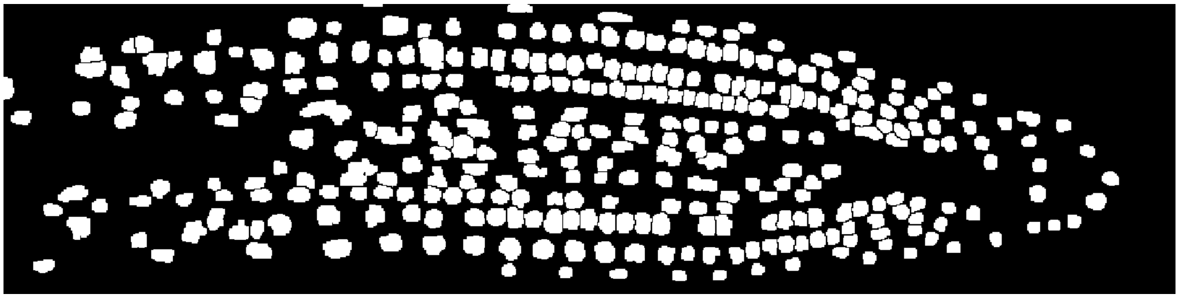


Fig 10. The binary image after subtracting figure 8 from figure 9

Now that we obtain the binary image of our segmented nucleus, we can perform the image segmentation. The function we used segmenting the nucleus in figure 10 is called 'bwlabel'. It is designed to label connected components in a 2-D binary image.

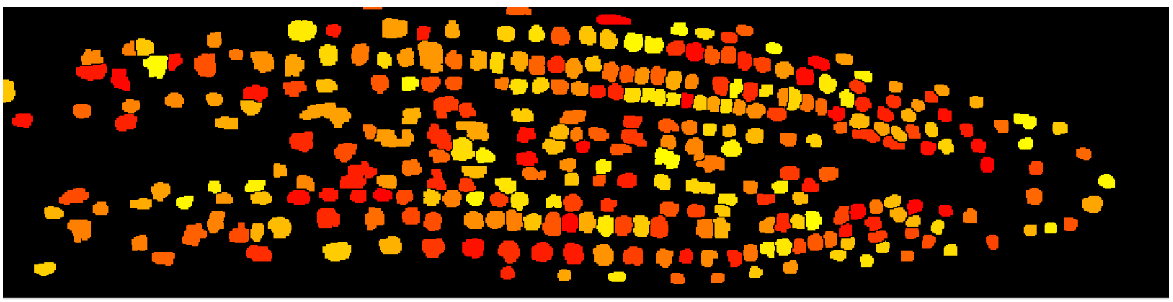


Fig 11. The resulting binary image after performing 'bwlabel' function. (Each label is coloured)

2. Nucleus Analysis

2.1 Total count of nuclei

The total count of regions in the binary image above simply reflects the total number of nuclei detected in the original image. For the sample image, a total of 309 nuclei were detected.

2.2 Size distribution of detected nuclei

In order to tabulate the size of each nucleus, we calculate the number of pixels of each region detected. Since the label image obtained, differentiate each region with an index, we can simply use a matlab builtin function called 'tabulate' to obtain the frequency table of the label image. After obtaining the frequency table, we are able to plot the histogram of the size distribution of the detected nuclei. In order to critically evaluate the size distribution, we can convert normalised data to evaluate the probability density of nucleus size. By calculating the standard deviation and mean value of the data, we can plot a normal distribution by using the probability distribution function.

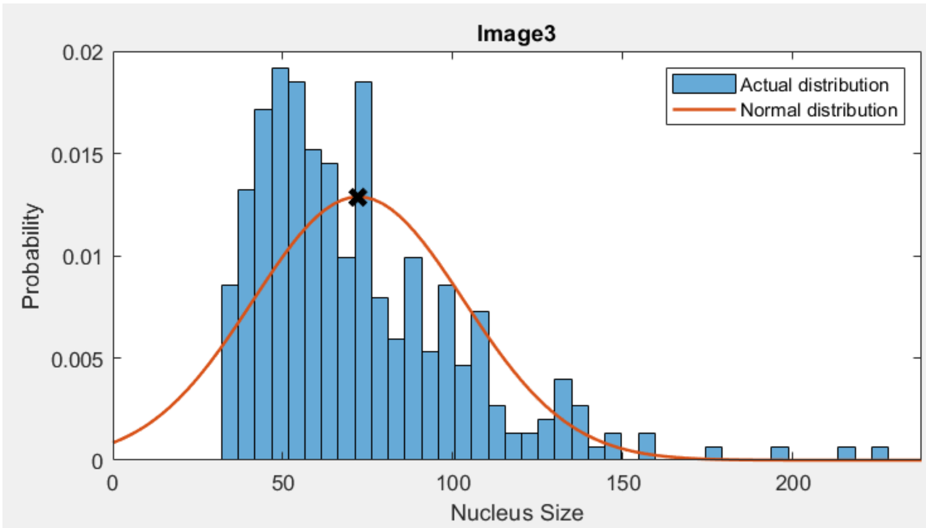


Fig 12. Distribution probability density graph of the nucleus size on the sample image

2.3 Shape distribution of the detected nuclei

In order to tabulate the shape of the nuclei, we calculate the eccentricity and the roundness of the detected nucleus. This is because the shape of a nucleus is expected to be oval. The eccentricity of a nucleus reflects how much a nucleus deviates from being circular. If the value is 0, it means that the nucleus is circular; and if the value is 1, it means that the nucleus is being a line segment. We also calculate the compactness of the nucleus. The IPQ (Osserman 1978) has become one of the most widely accepted compactness measures. It is calculated as:

$$C_{IPQ} = \frac{4\pi A}{P^2}$$

where:
A, denotes the area of the shape.
P, denotes the perimeter of the shape.

According to Li, Goodchild and Church (2013), a shape with a high value of C_{IPQ} is considered to be more compact than a shape with a lower C_{IPQ}. We perform scatter plots to study the shape distribution of the nucleus. A circle is the most compact shape and by definition above it will have a compactness value of 1. As we proposed above, we also plot a multivariate normal distribution to critically analyse the shape of the nucleus.

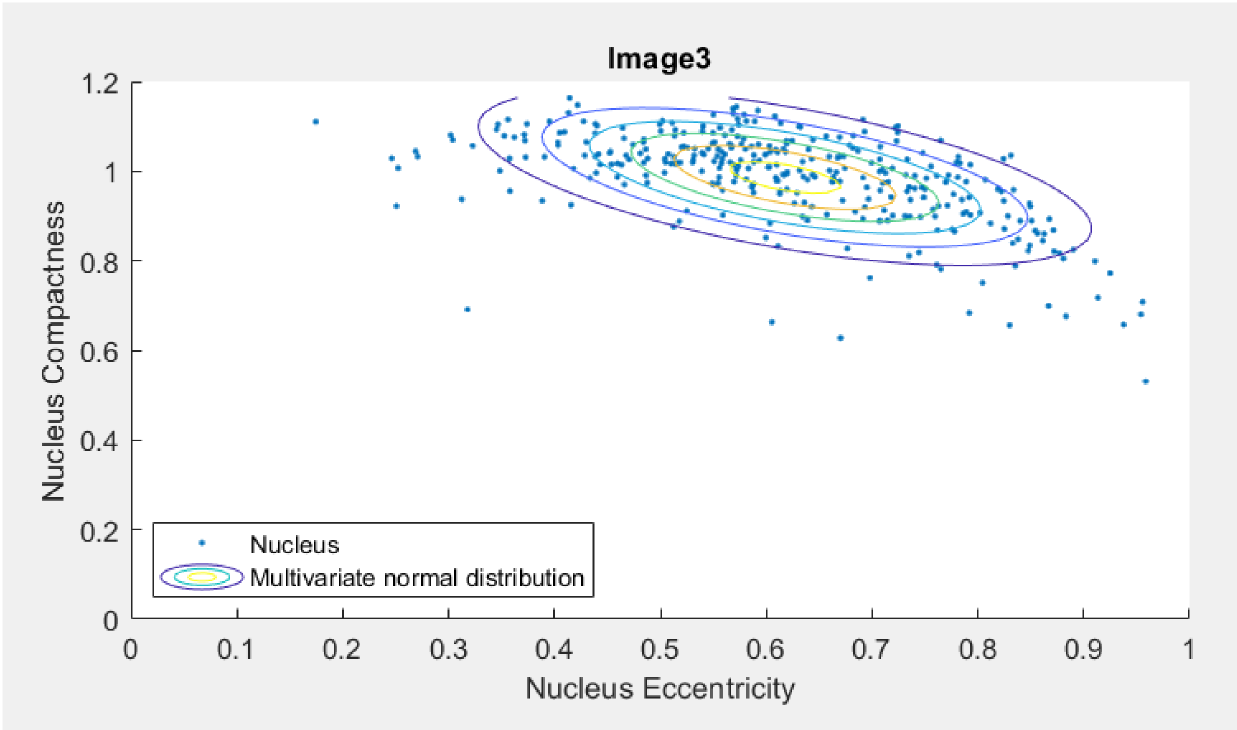


Fig 12. Scatter plot of the nucleus shape on the sample image

2.4 Brightness distribution of the detected nuclei

In order to tabulate the brightness of the nucleus, we used a matlab image processing tools function called ‘regionprop’ where it returns the mean intensity of the region. The input that we passed to this function is the binary image in figure 10 mask over the greyscale of the sample image. It can be noticed there is bimodal in the probability distribution graph. Hence, we can use a matlab function called ‘fitgmdist’ to fit the gaussian model into the bimodal distribution.

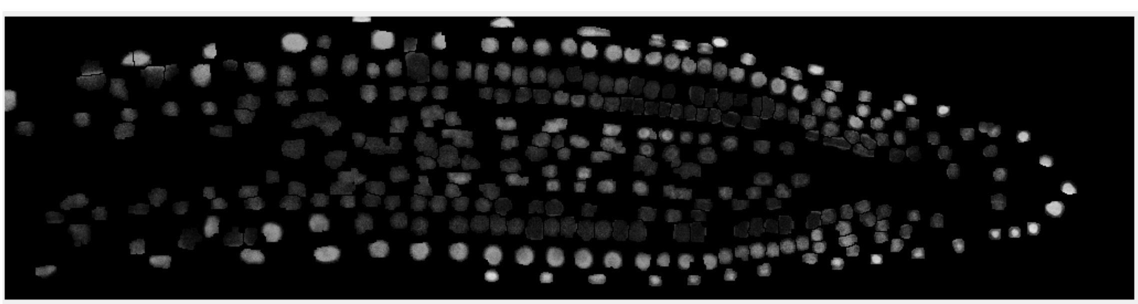


Fig 13. Input image passed to the function

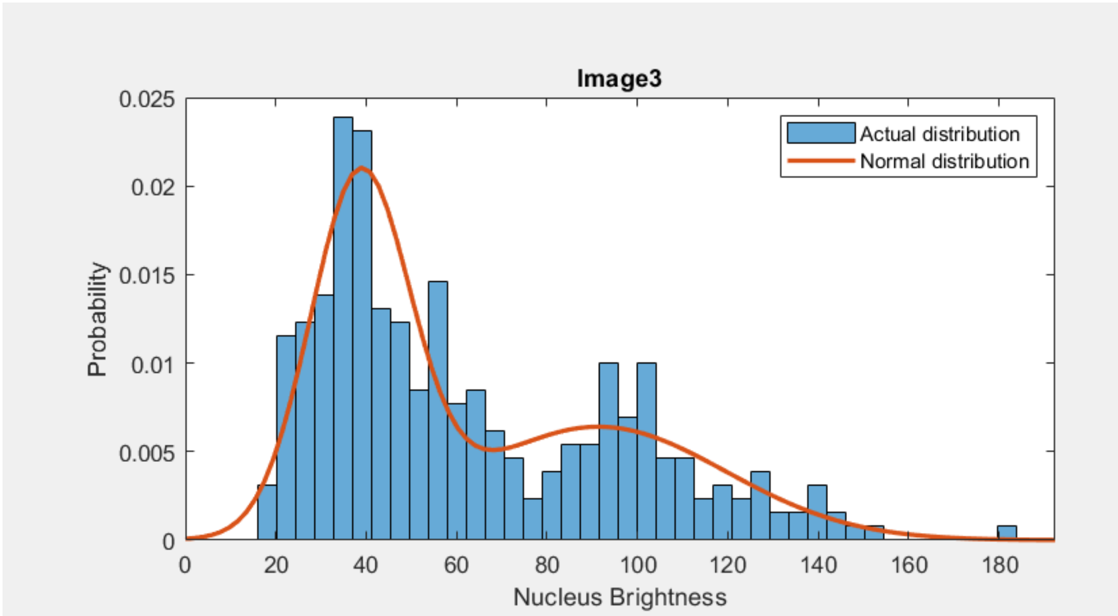
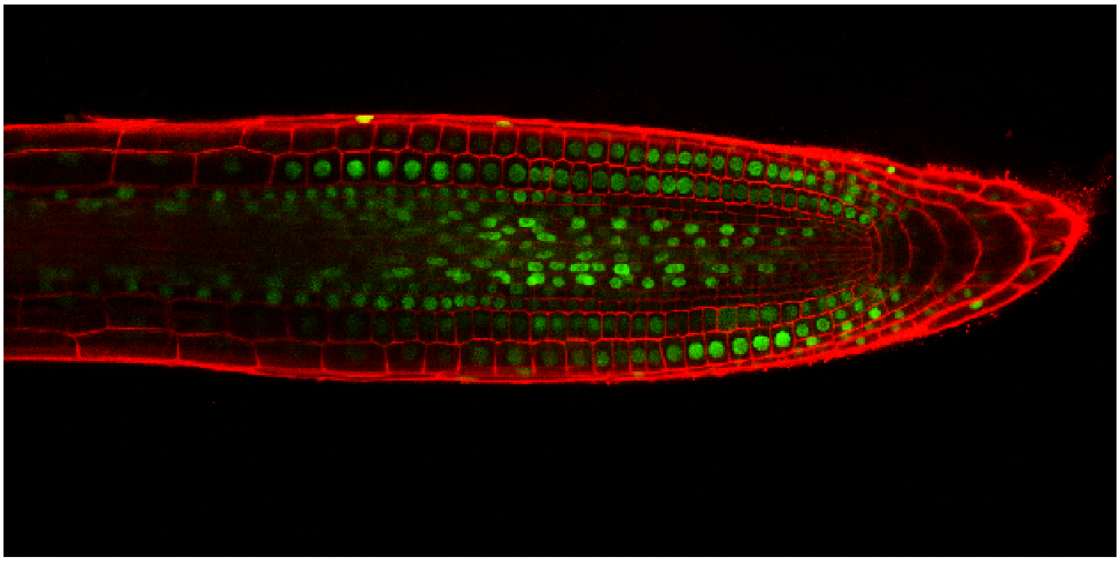


Fig 14. Distribution probability density graph of the nucleus intensity on the sample image

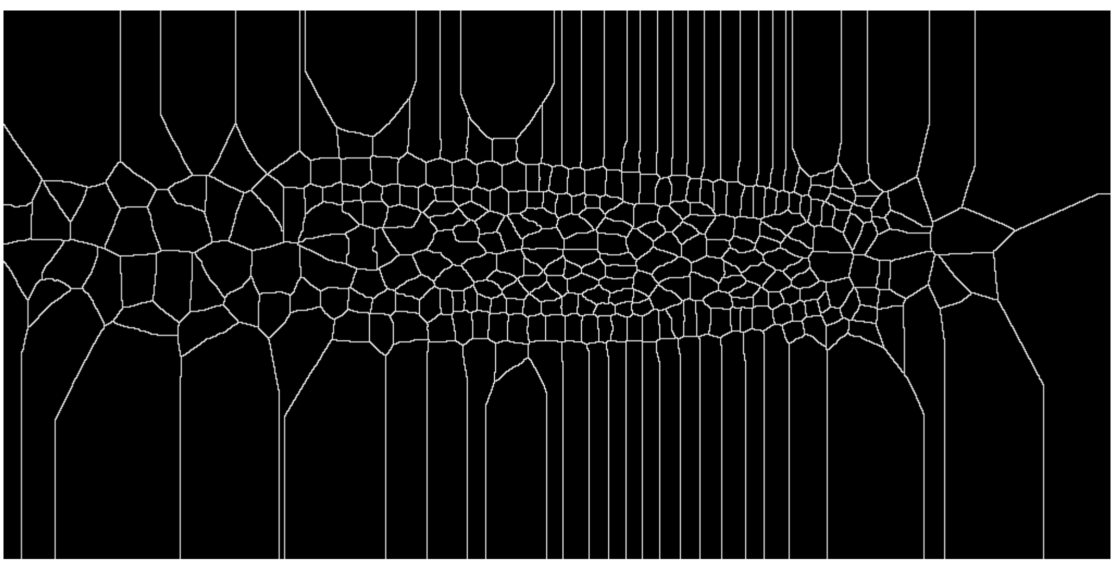
3. Result Presentation

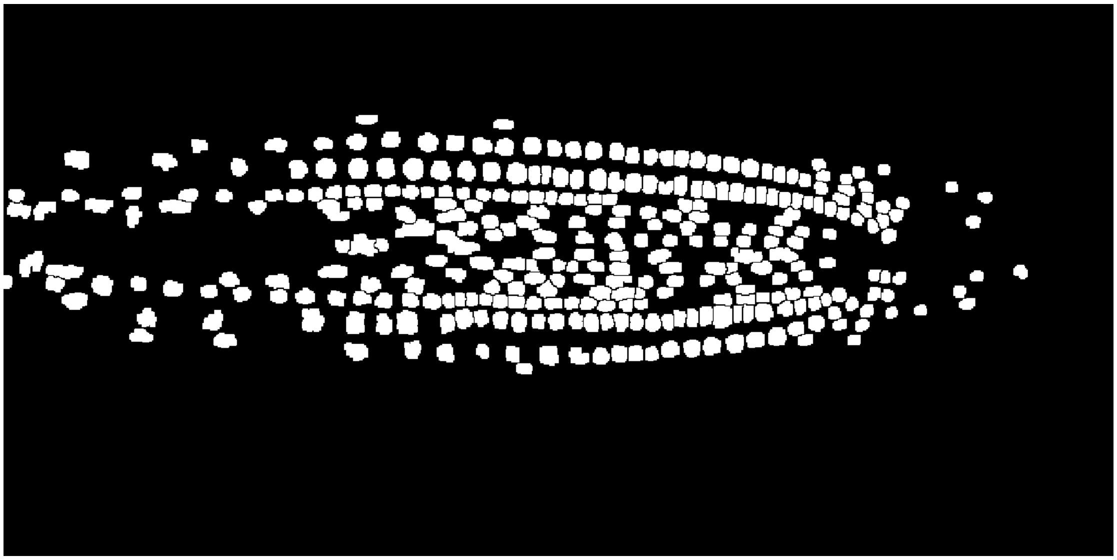
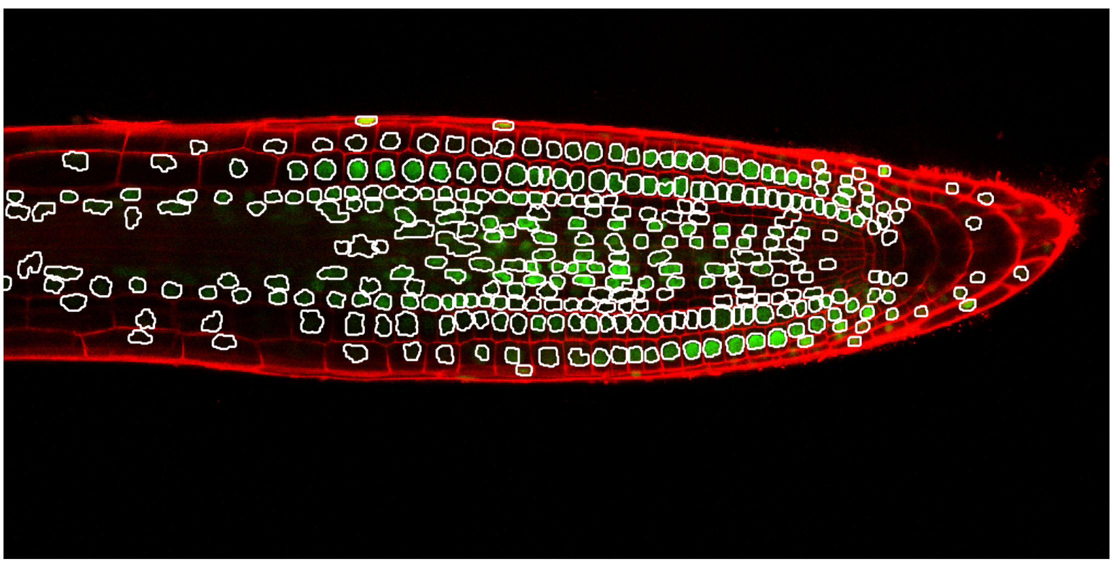
3.1 Supplied Image 1 [320 nuclei detected]

Original image

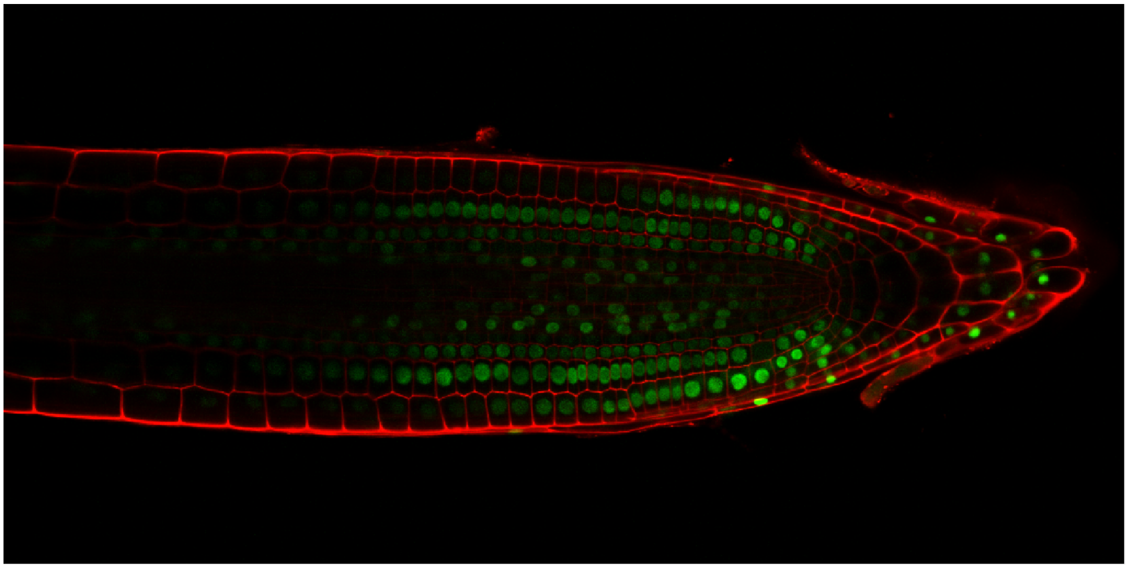
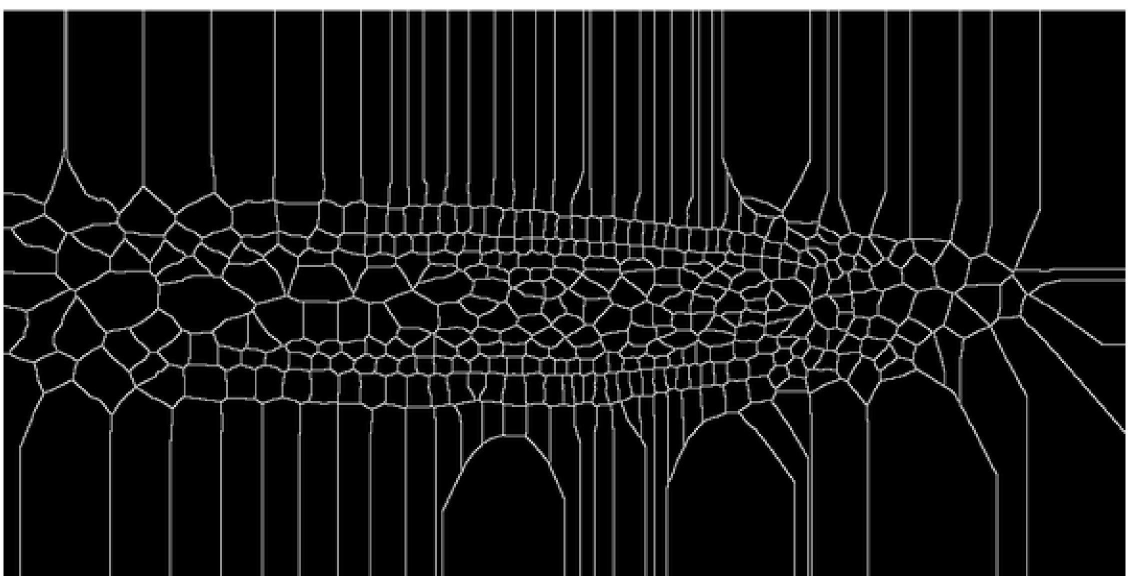


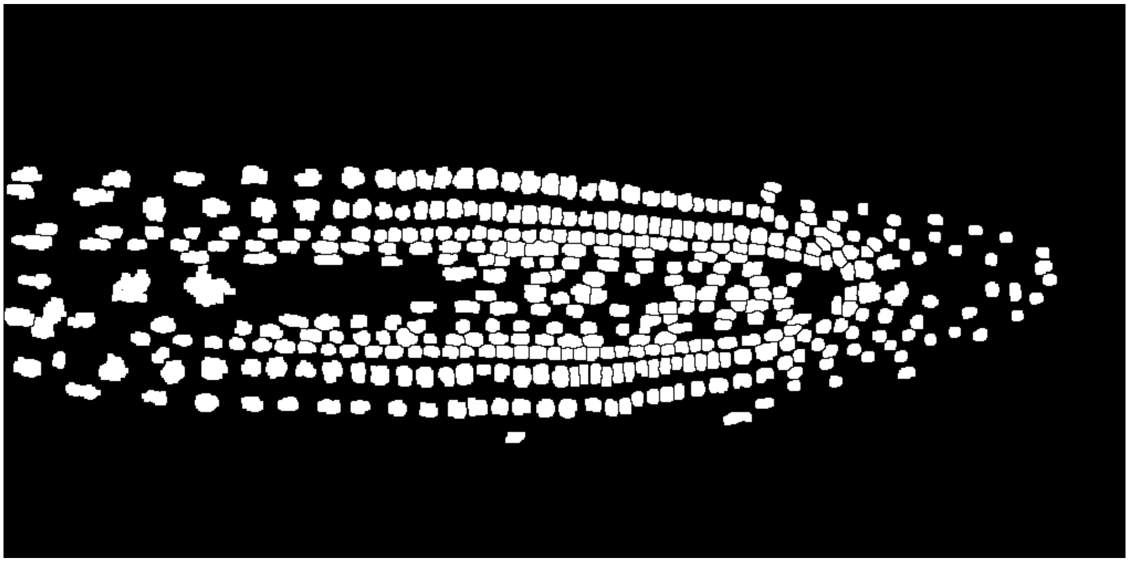
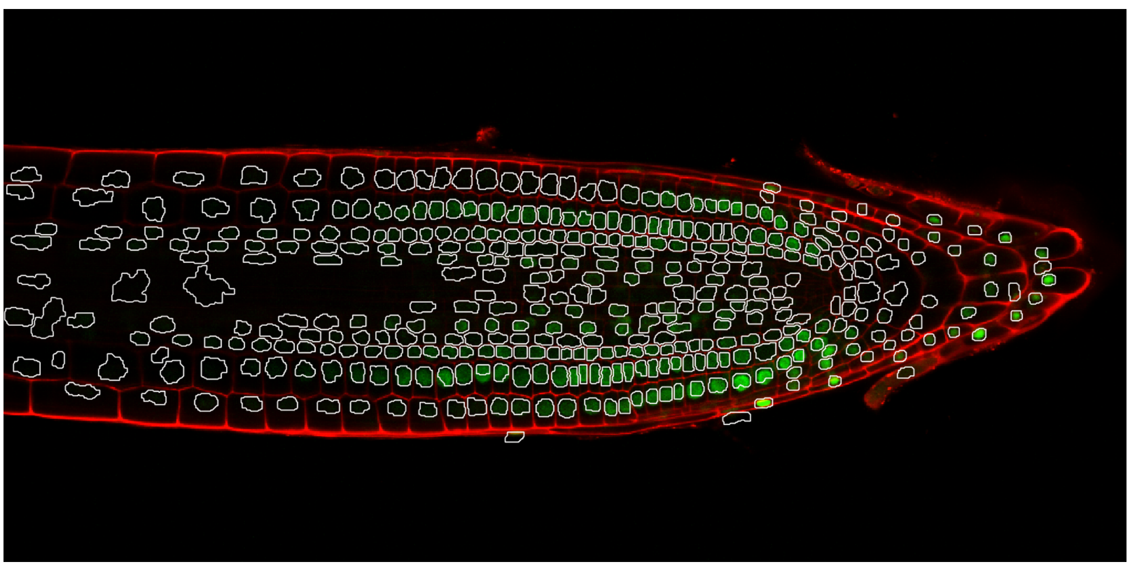
Watershed ridgeline of nuclei region



Binary Image masking region corresponding to nuclei

Labeled nuclei over original image


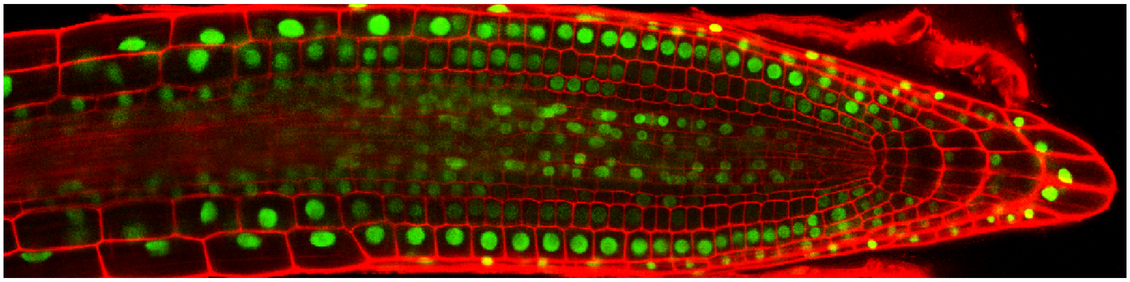
3.2 Supplied Image 2 [386 nuclei detected]

Original image

Watershed ridgeline of nuclei region


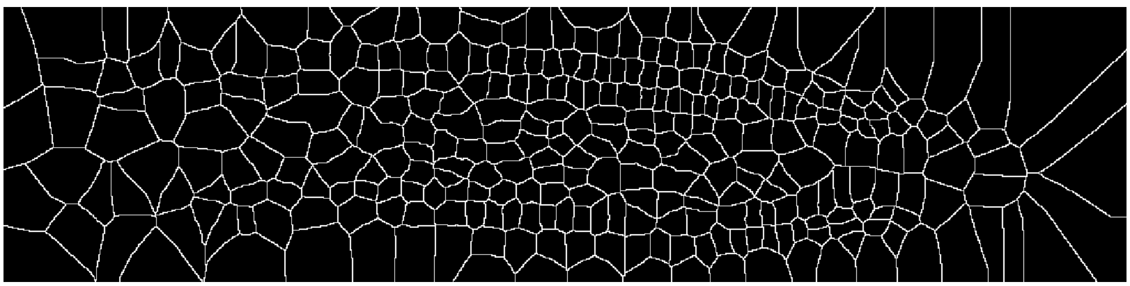
Binary Image masking region corresponding to nuclei

Labeled nuclei over original image


3.3 Supplied Image 3 [338 nuclei detected]

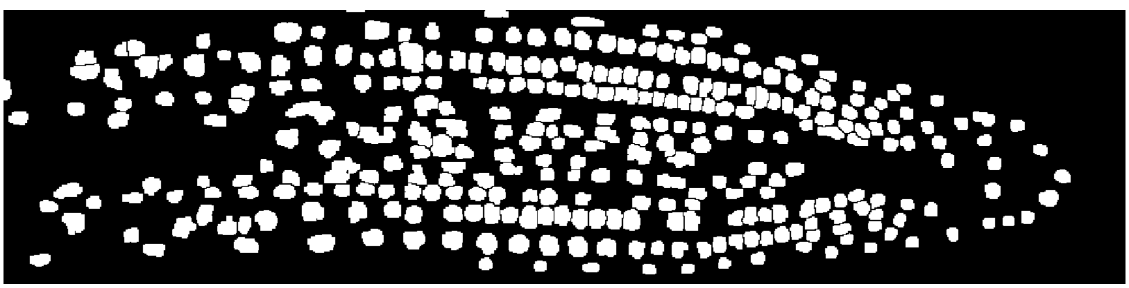
Original image



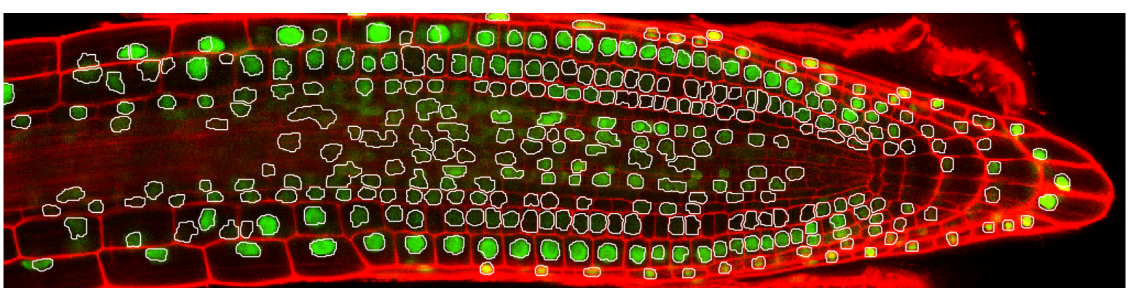
Watershed ridgeline of nuclei region



Binary Image masking region corresponding to nuclei



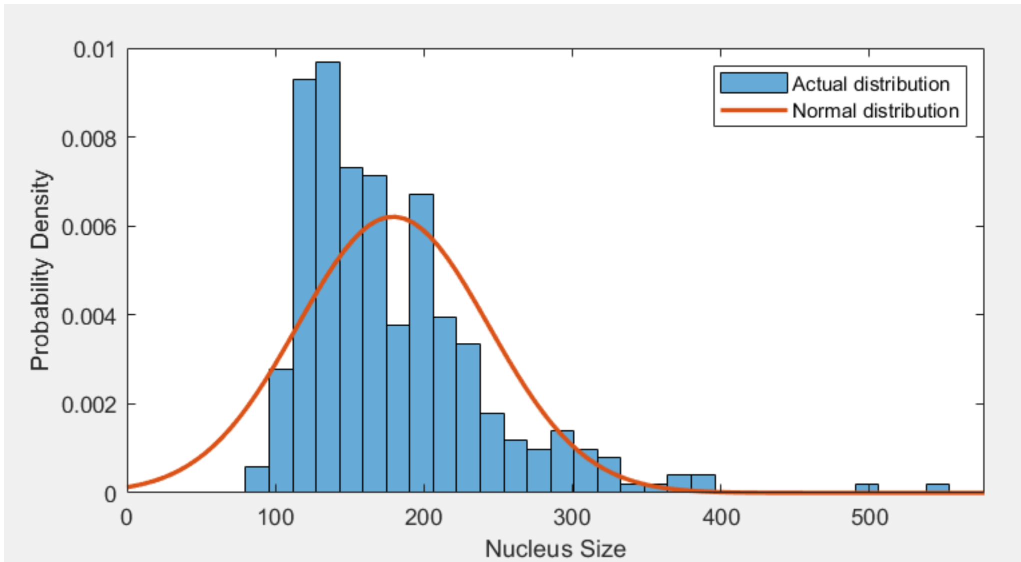
Labeled nuclei over original image



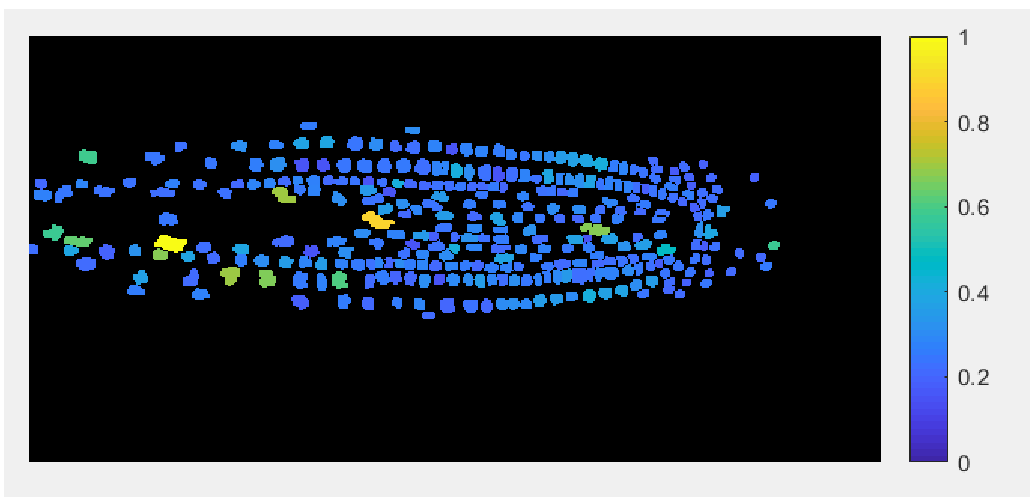
3.4 Size Distribution of Detected Nuclei

3.4.1 Supplied Image 1

Nuclei size distribution analysis



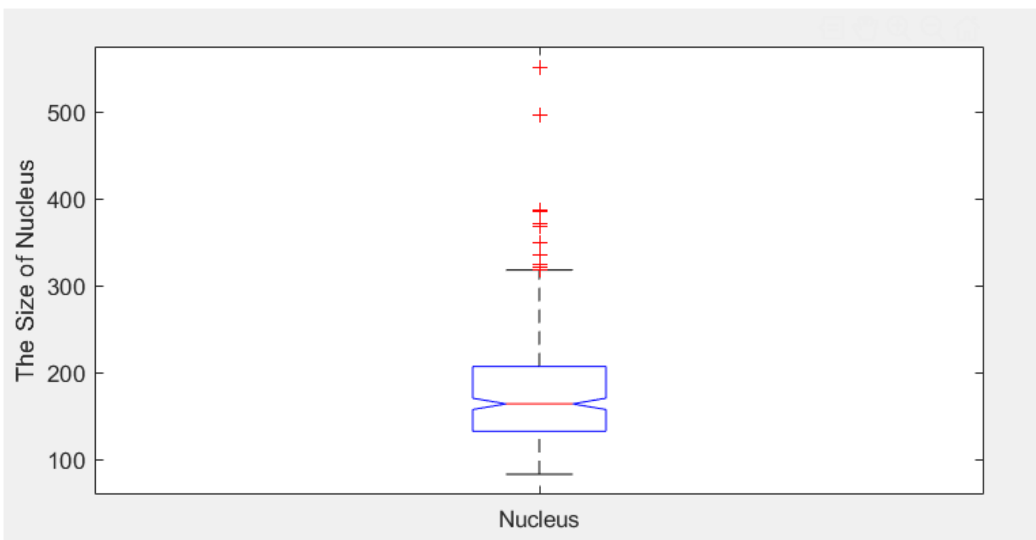
Labeled nuclei when color based on their size (0 indicating smallest; 1 indicating largest)



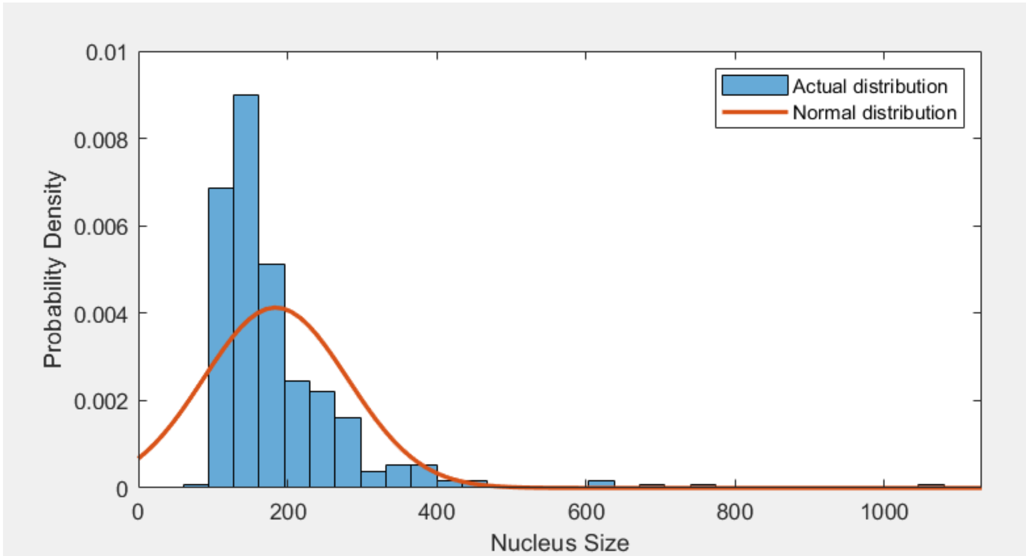
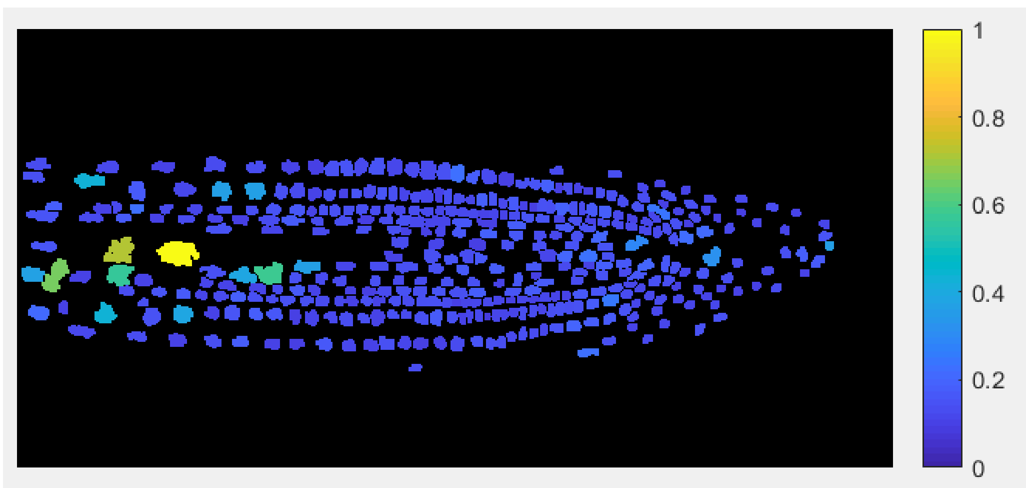
Statistic analysis regarding the size of nuclei

	1	2
1	Mean Size	179.3188
2	Maximum Size	552
3	Minimum Size	83
4	Standart Deviation	64.3581
5	Most Frequent Size	123

Box plot of the size distribution of nuclei



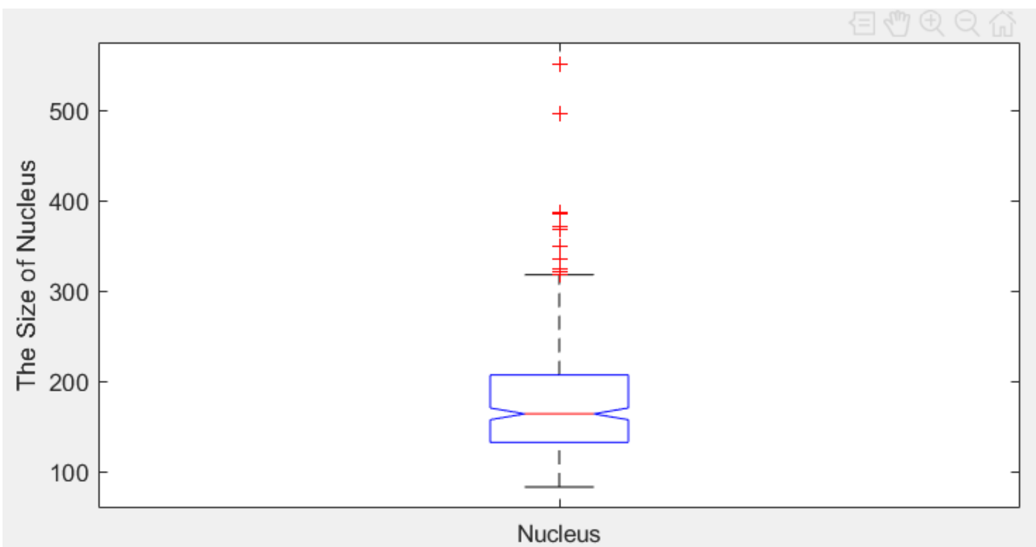
3.4.2 Supplied Image 2

Nuclei size distribution analysis

Labeled nuclei when color based on their size (0 indicating smallest; 1 indicating largest)


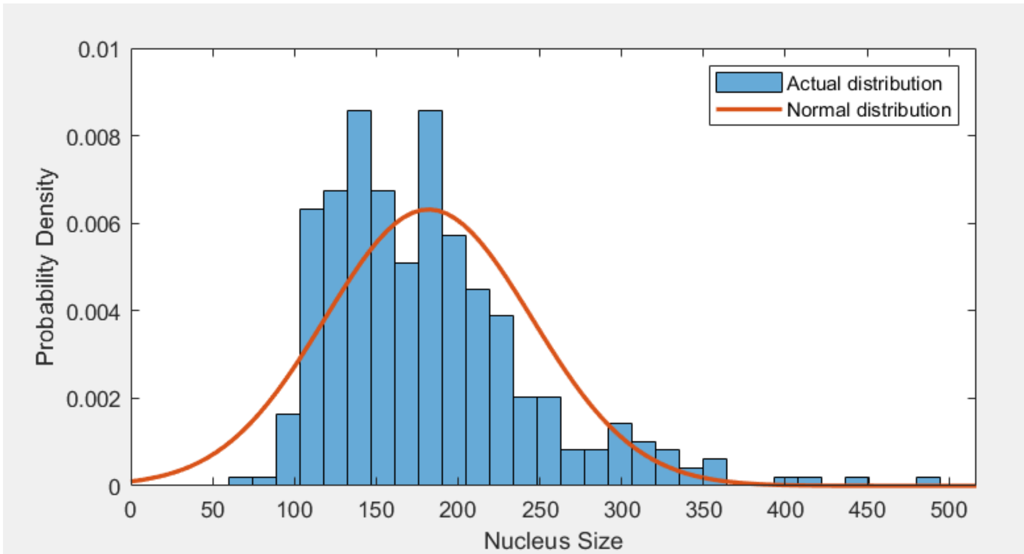
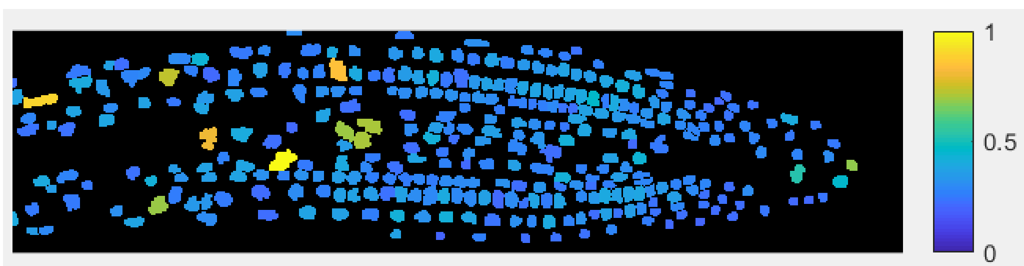
Statistic analysis regarding the size of nuclei

	1	2
1	Mean Size	184.0181
2	Maximum Size	1068
3	Minimum Size	83
4	Standart Deviation	96.7399
5	Most Frequent Size	132

Box plot of the size distribution of nuclei



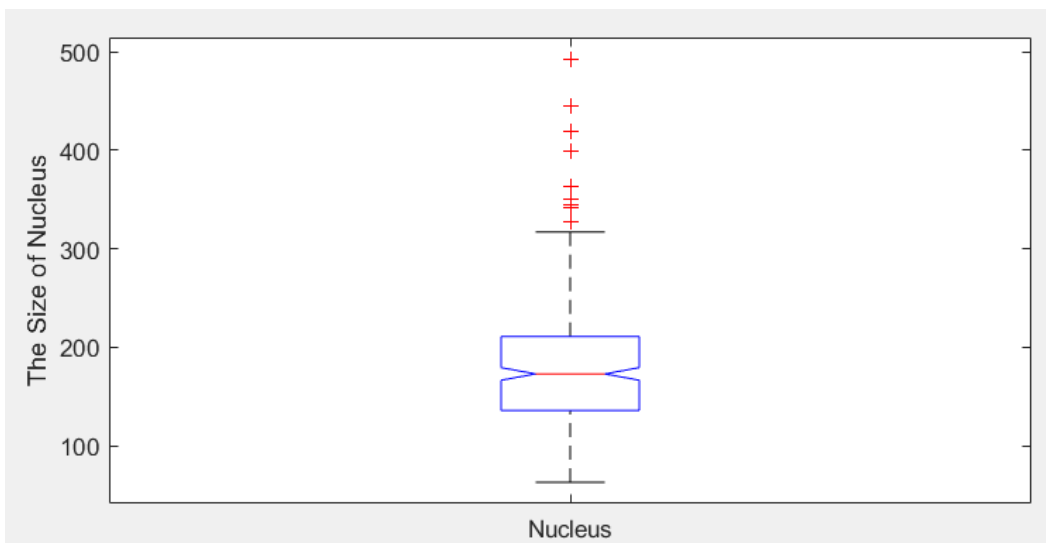
3.4.3 Supplied Image 3

Nuclei size distribution analysis

Labeled nuclei when color based on their size (0 indicating smallest; 1 indicating largest)


Statistic analysis regarding the size of nuclei

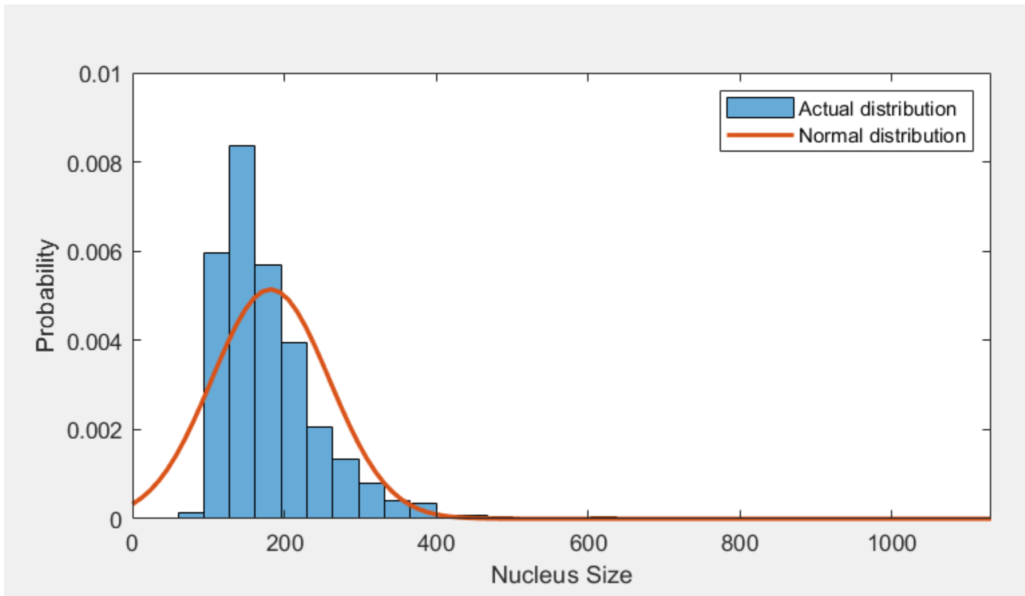
	1	2
1	Mean Size	181.9882
2	Maximum Size	493
3	Minimum Size	63
4	Standart Deviation	63.2691
5	Most Frequent Size	114

Box plot of the size distribution of nuclei



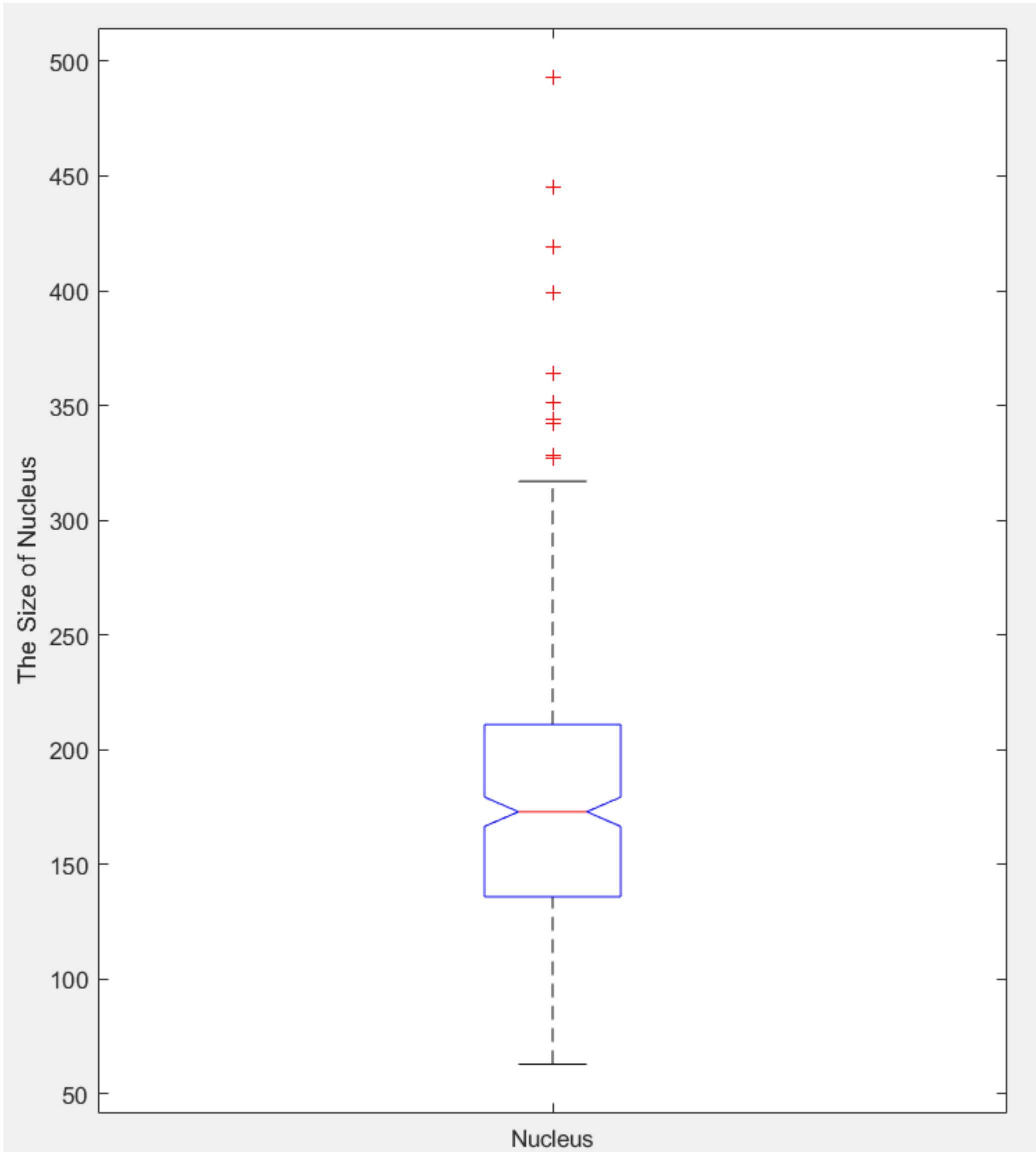
3.4.4 Cumulative Analysis

Nuclei size distribution in all supplied image



Statistic analysis regarding the size of nuclei in all supplied image

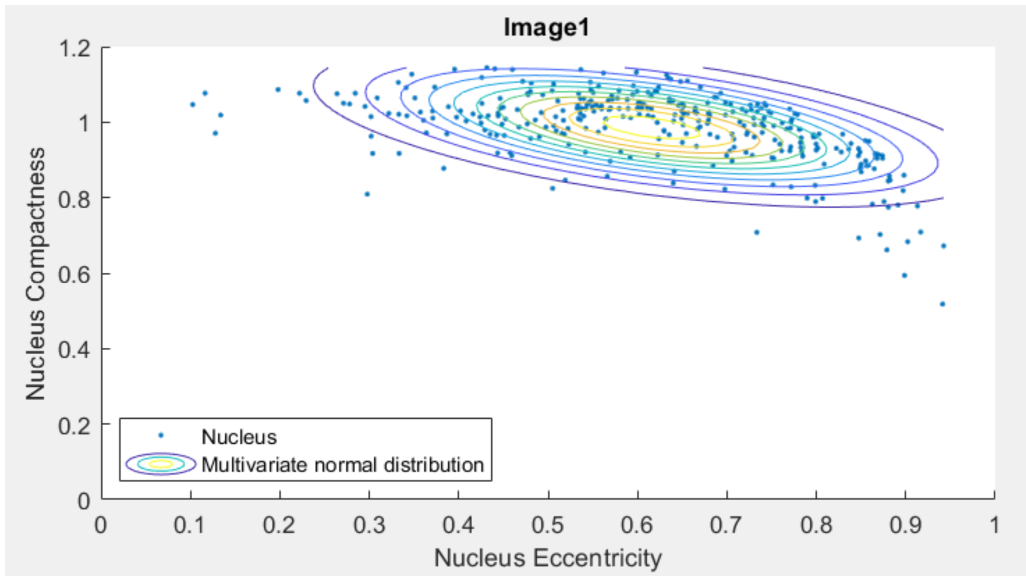
	1	2
1	Mean Size	181.9205
2	Maximum Size	1068
3	Minimum Size	63
4	Standart Deviation	77.5785
5	Most Frequent Size	127

Box plot of the size distribution of nuclei in all supplied image


3.5 Shape Distribution of Detected Nuclei

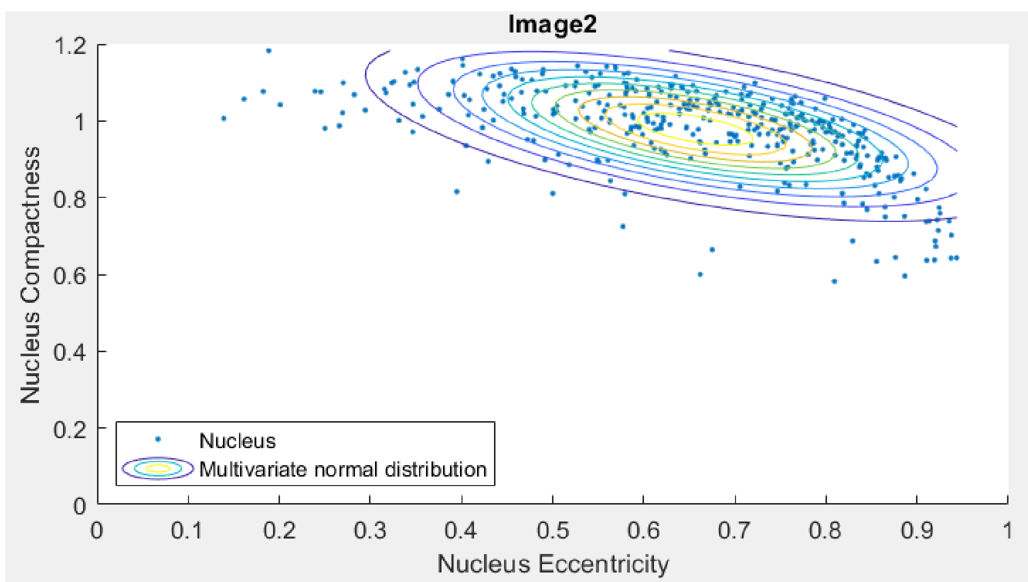
3.5.1 Supplied Image 1

Scatter plot of the shape properties of nuclei

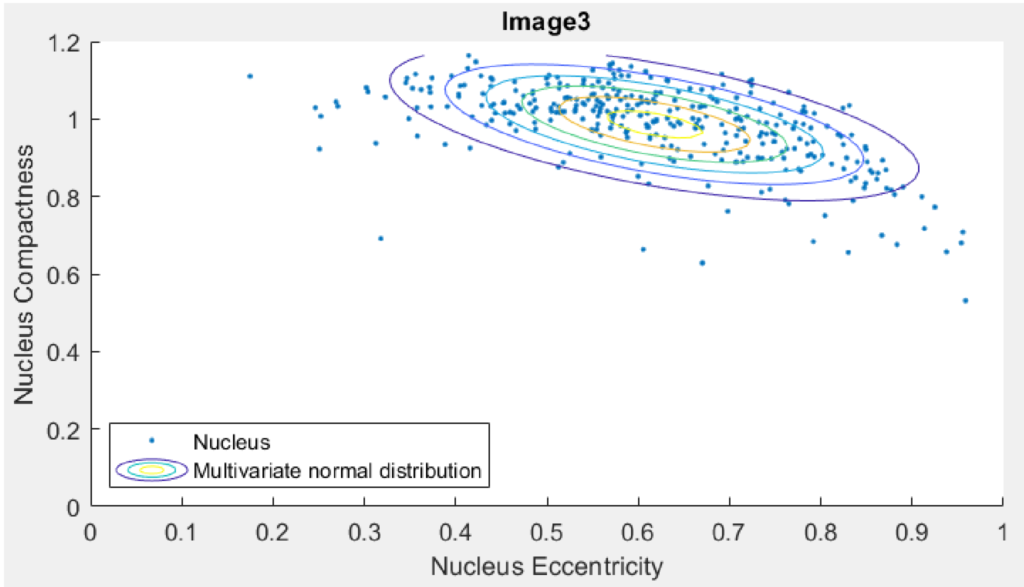


3.5.2 Supplied Image 2

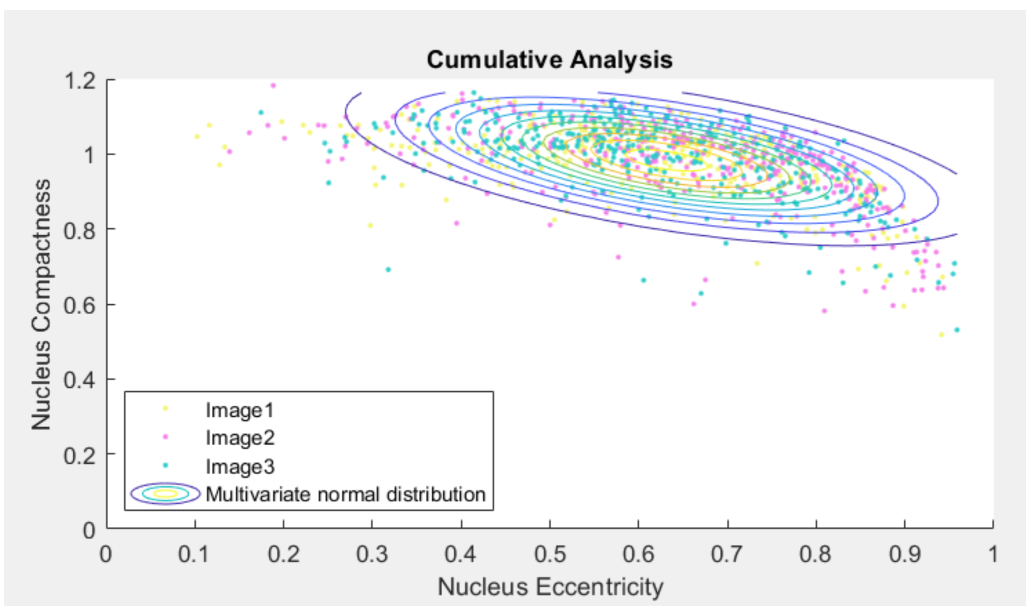
Scatter plot of the shape properties of nuclei



3.5.3 Supplied Image 3

Scatter plot of the shape properties of nuclei


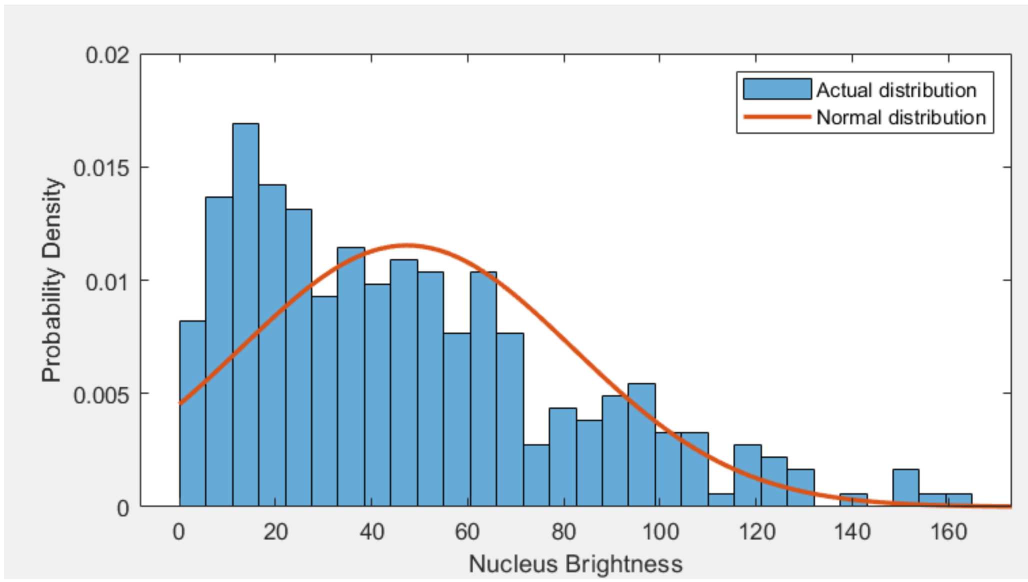
3.5.4 Cumulative Analysis

Scatter plot of the shape properties of nuclei in all supplied image


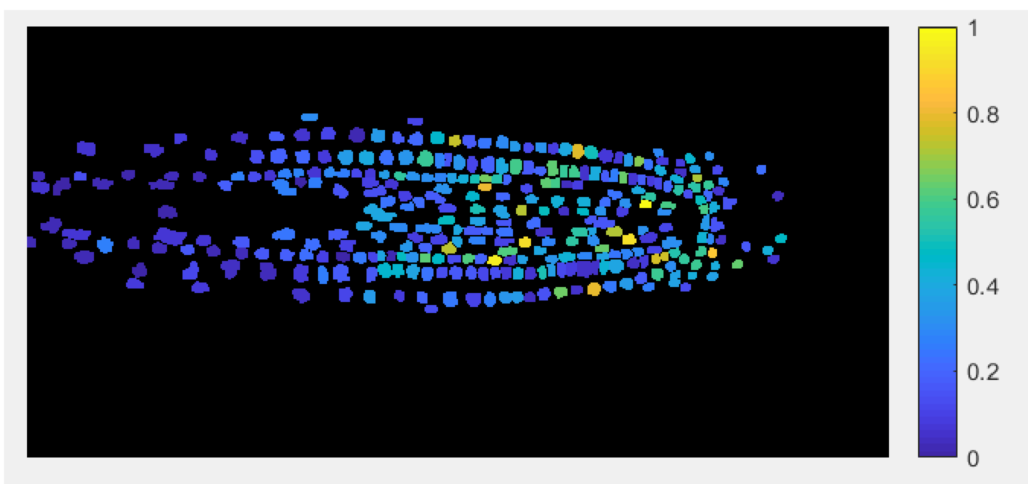
3.6 Intensity Distribution of Detected nuclei

3.6.1 Supplied Image 1

Nuclei intensity distribution analysis



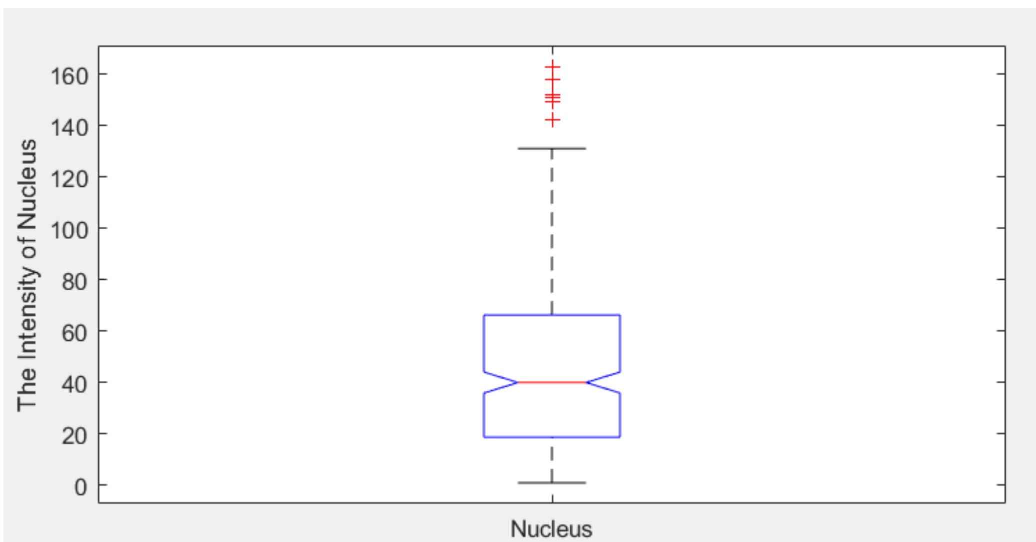
Labeled nuclei when color based on their intensity (0 indicating dimmest; 1 indicating brightest)



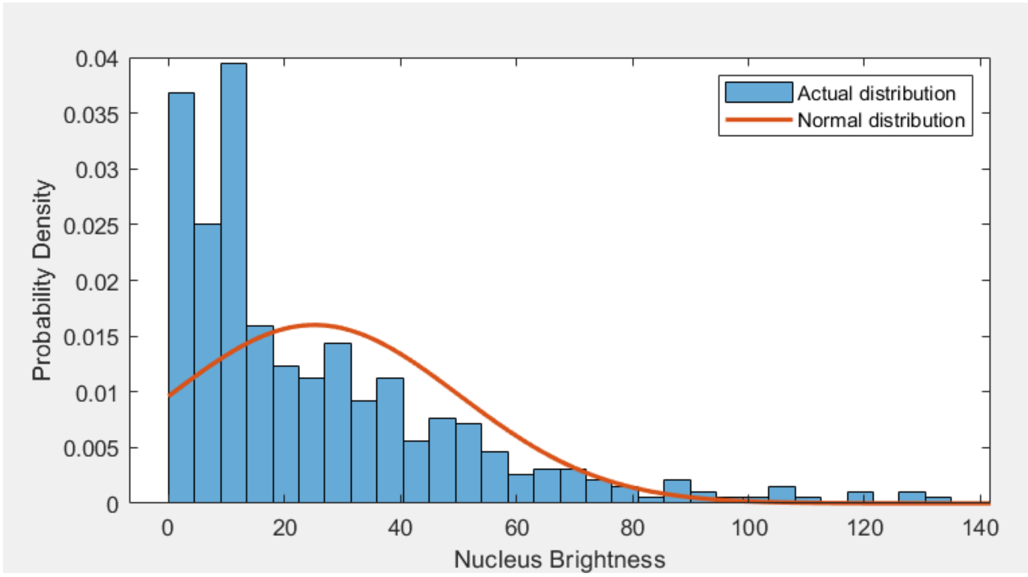
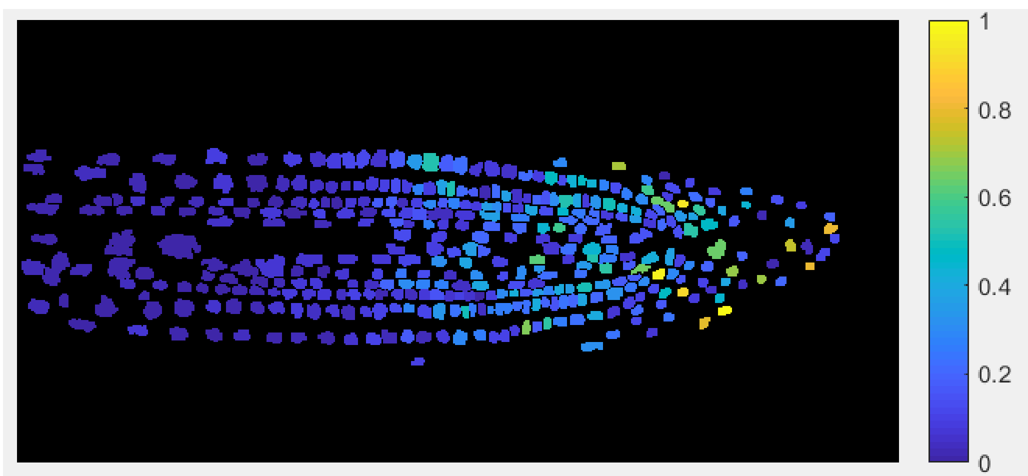
Statistic analysis regarding the intensity of nuclei

	1	2
1	Mean Intensity	47.3213
2	Highest Intensity	163
3	Lowest Intensity	1
4	Standart Deviation	34.6385
5	Most Frequent Intensity	18

Box plot of the intensity distribution of nuclei



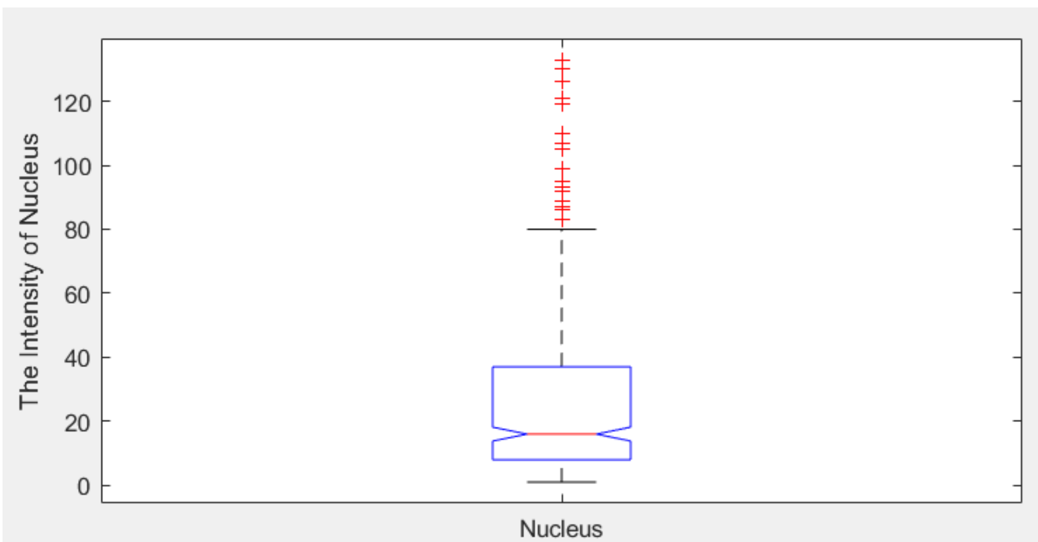
3.6.2 Supplied Image 2

Nuclei intensity distribution analysis

Labeled nuclei when color based on their intensity (0 indicating dimmest; 1 indicating brightest)


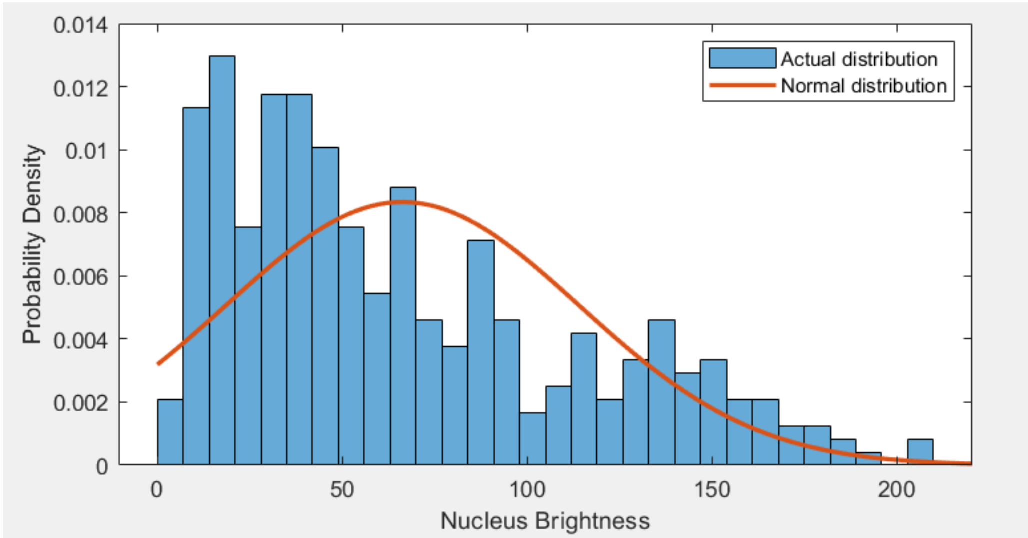
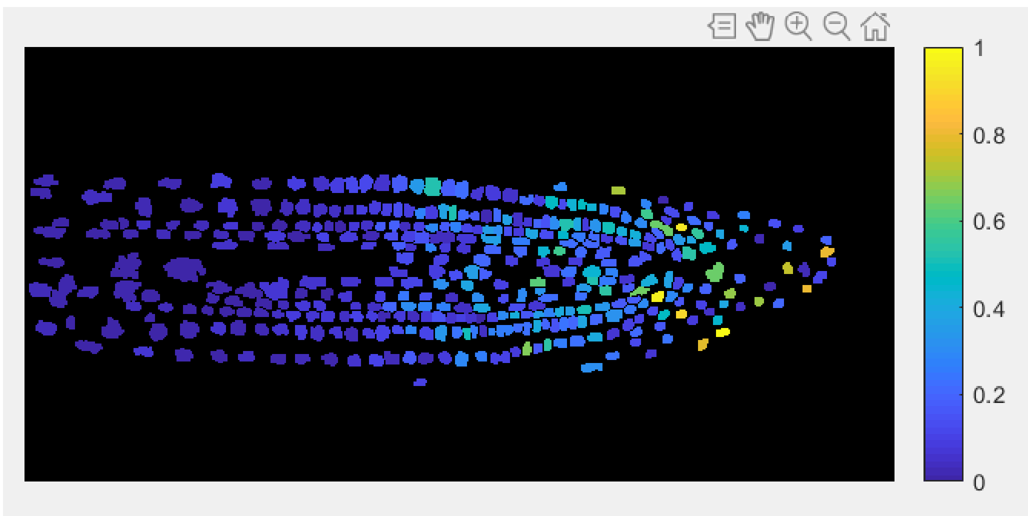
Statistic analysis regarding the intensity of nuclei

	1	2
1	Mean Intensity	25.1935
2	Highest Intensity	133
3	Lowest Intensity	1
4	Standart Deviation	24.9613
5	Most Frequent Intensity	1

Box plot of the intensity distribution of nuclei



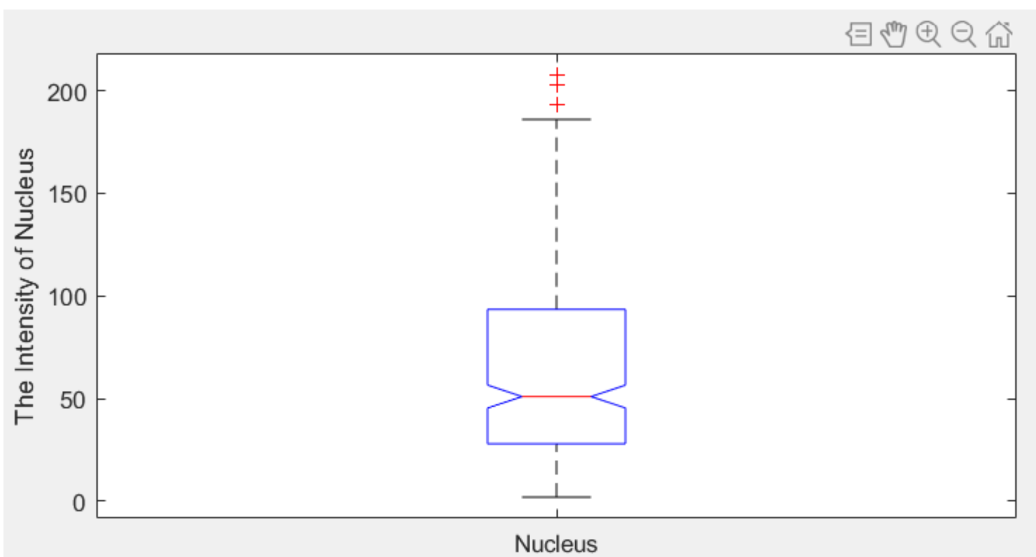
3.6.3 Supplied Image 3

Nuclei intensity distribution analysis

Labeled nuclei when color based on their intensity (0 indicating dimmest; 1 indicating brightest)


Statistic analysis regarding the intensity of nuclei

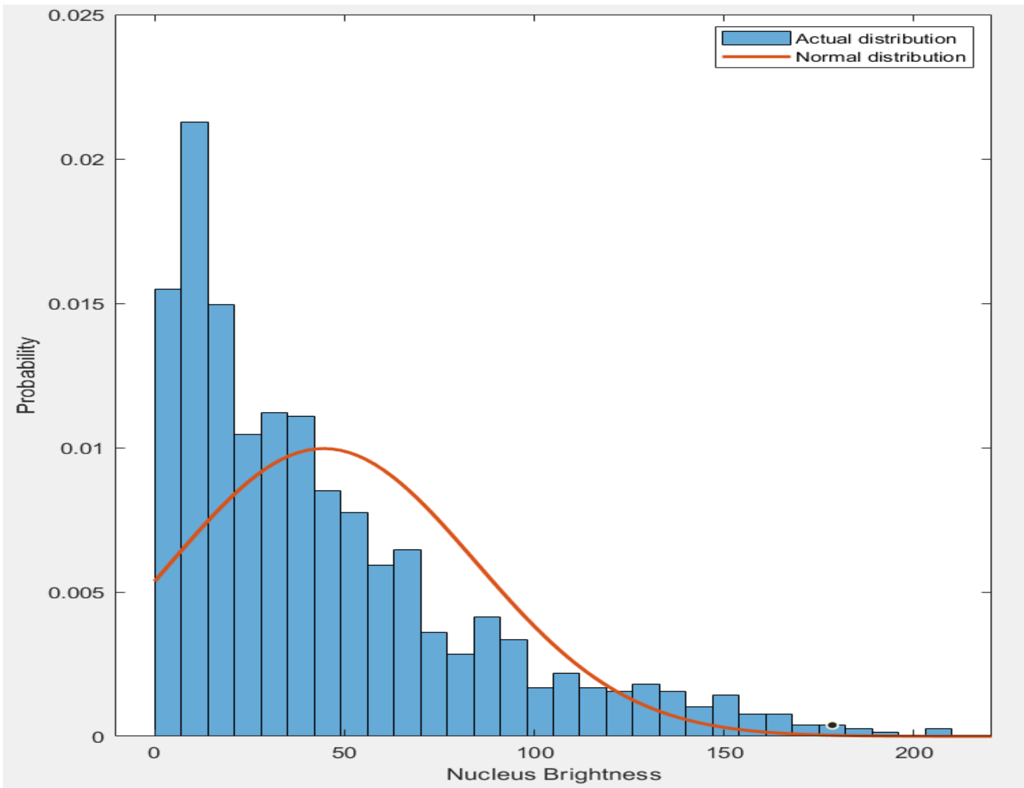
	1	2
1	Mean Intensity	66.3314
2	Highest Intensity	208
3	Lowest Intensity	2
4	Standart Deviation	47.8931
5	Most Frequent Intensity	40

Box plot of the intensity distribution of nuclei



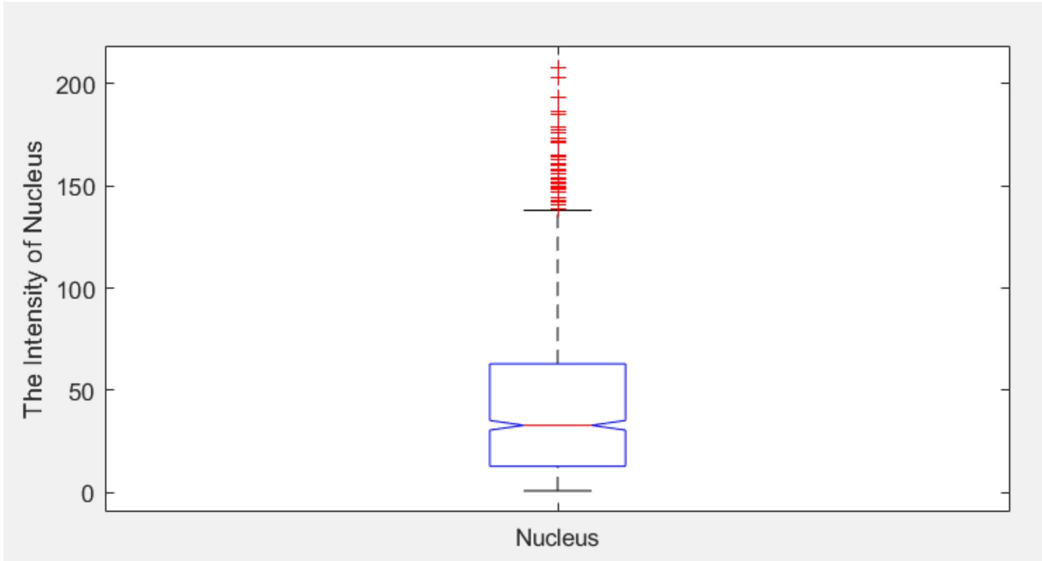
3.6.4 Cumulative Analysis

Nuclei intensity distribution in all supplied image



Statistic analysis regarding the intensity of nuclei in all supplied image

	1	2
1	Mean Intensity	44.5045
2	Highest Intensity	208
3	Lowest Intensity	1
4	Standart Deviation	40.0423
5	Most Frequent Intensity	1

Box plot of the intensity distribution of nuclei in all supplied image


4. Results Discussion

4.1 Detected nuclei on the supplied image

4.1.1 Strength

By implementing the image processing approach we discussed above, we are able to extract the nucleus out of most of the cell even if it is underlying below the red region (cell wall). This can be seen in the result that shown in the section 5.1 to 5.3, especially on supplied image 2. The original image has low intensity and the nucleus can hardly be identified using the bare eye. By converting the color space, we managed to capture the hue of the colour in the image and perform the following detection process.

Another strength in this technique is that we manage to eliminate the noise that surrounds the supplied image by highlighting the foreground of the image. This is important because imregionalmax function is highly sensitive to noise, especially salt and pepper noise due to the fact that salt noise will be detected as a region when passing into the function.

The last strength is that we managed to display several ways of analysis,including a color map that corresponds to the properties of nucleus and normal distribution and multivariate normal distribution of detected nuclei based on the calculated mean and standard deviation. This helps to visualise the distribution of nucleus and also examine the quality of the image processing technique.

4.1.2 Weakness

One of the main weaknesses is that the approach implemented fails to detect some of the nuclei, especially when nuclei are overlapping each other. This is because the segmentation technique we implement is imregionalmax function. Although it helps to detect the underlying nucleus, it will only mark one regional maxima even if there are multiple nuclei present in the same evaluating region. As we mentioned earlier, the method is sensitive to noise, it might return poor results even if there are slight differences among the background pixels that cannot be identified by bare eyes.

As we mentioned earlier, the color space we chose to process our image is hsv color space. Although we managed to minimise the red region of the image. This has caused a ridgeline formed for the nucleus that are underlying in between the cell wall and being segmented into 2 regions during the segmentation process.

4.1.3 Ways of improvement

One way to improve our current result is by introducing a laplacian filter to our process to capture fine details and also thinning the edge of the nucleus. We can also try to implement a different approach to each of the supplied images based on the properties of each image.

4.2 Analysis on the properties of detected nuclei

4.2.1 Size distribution of the detected nuclei

As shown in the section above, the size distribution of the detected nuclei mostly followed the normal distribution with calculated mean and standard deviation which indicates the nuclei detection is promising. From the analysis, we can observe that there exists some outlier and there is no nucleus that falls under the first 10 % percentile of the normal distribution. This is because a nucleus with that size is too small to be detected and being filtered away.

4.2.2 Intensity distribution of the detected nuclei

Unlike the size distribution, intensity distribution of the detected nuclei does not follow the normal distribution graph. We can observe that it shows as a bimodal distribution which indicates that the nuclei can be categorised into two groups, ‘dim nuclei’ and ‘bright nuclei’. We can hence make a hypothesis that this is because this is a 2D image and the nuclei that are on different depths are overlapping each other and the bright nuclei are land over the other. Hence the amount of light that they are reflecting shows a bimodal distribution.

4.2.3 Shape distribution of the detected nuclei

Since we are analysing the shape of the nuclei in terms of their eccentricity and compactness, we can observe that their distributions follow the bivariate normal distribution where the variance of the nucleus eccentricity is larger than the other and thus formed a shape of ellipsoid. This indicates that the shape of detected nuclei is overall oval with slight compactness compared to a circle.

5. Reference

A. Ajmal, C. Hollitt, M. Frean and H. Al-Sahaf, "A Comparison of RGB and HSV Colour Spaces for Visual Attention Models," 2018 International Conference on Image and Vision Computing New Zealand (IVCNZ), 2018, pp. 1-6, doi: 10.1109/IVCNZ.2018.8634752.

Li, W., Goodchild, M. F., & Church, R. (2013). An efficient measure of compactness for TWO-DIMENSIONAL shapes and its application In regionalization problems. *International Journal of Geographical Information Science, 27*(6), 1227-1250. doi:10.1080/13658816.2012.752093

Osserman, R., 1978. Isoperimetric inequality. Bulletin of the American Mathematical Society, 84 (6), 1182-1238.

Santiago, R.S. and Bribiesca, E., 2009. State of the art of compactness and circularity measures. International Mathematical Forum, 4 (27), 1305-1335.

6. Video Link

Video Link: https://youtu.be/ewUgt0EaQYg

7. Word Count

The word count of this report is ~2100 excluding the references, table of content and labeling of images.