

MLPR - LAB 1 REPORT

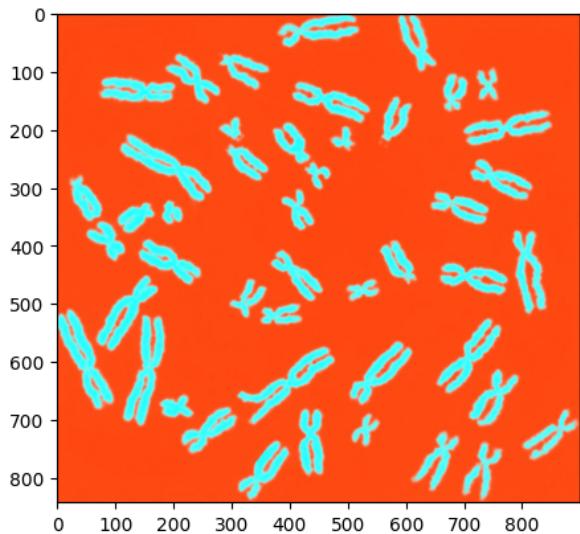


Image loaded and plotted in BGR format (openCv convention)

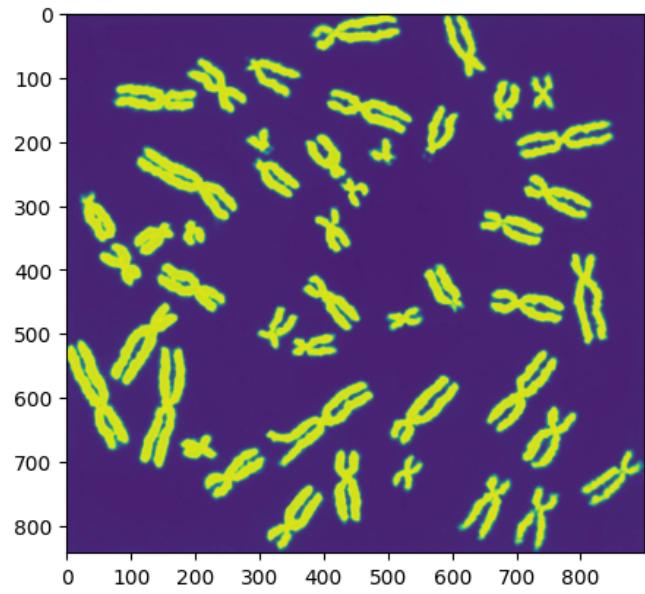
If we wanted to display the original image , we may use matplotlib convention (RGB) using :

```
image_rgb = cv2.cvtColor(image_reading , cv2.COLOR_BGR2RGB)
```

Converting the original “chromosomes.jpg” into grey-scale image

converting the image into grey color reduces the amount of data the computer needs to process (Ex: R-G-B) . But with grey color , the process depends only on intensity

The grey-scale image which we have converted is used in thresholding process further as it works on a single channel intensity image.



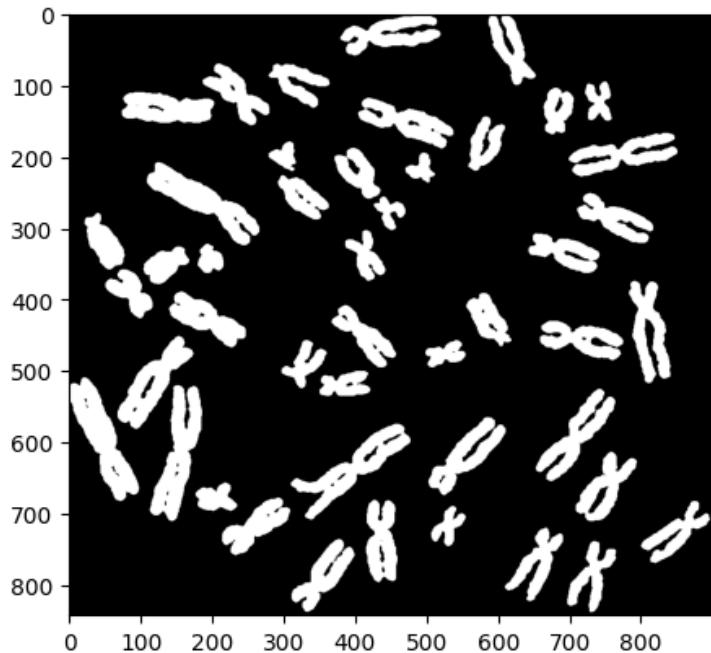
Morphological Opening , Thresholding

Opening -

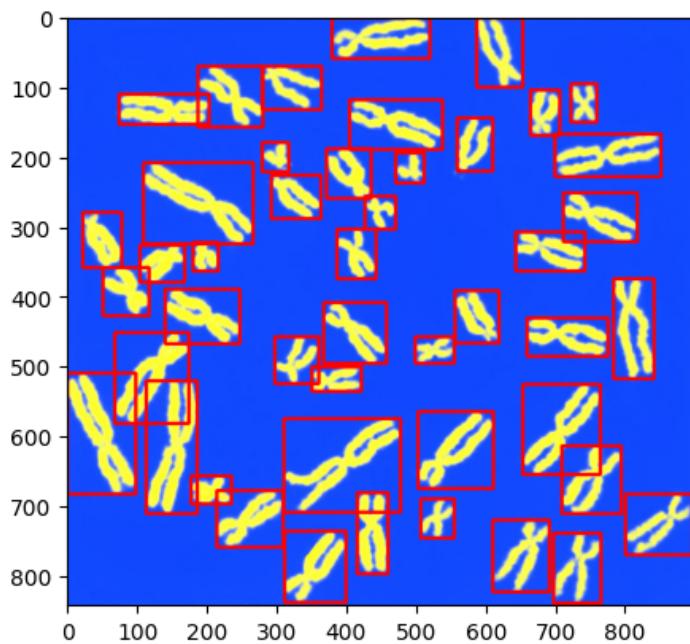
Erosion followed by Dilation which is used to remove noise , smoothens objects's boundary in the image . In this case , we used a kernel of size 5X5 (Ellipse) which scans the image and operates .

**We used Binary Thresholding method to convert a Grey-scale image into a Black and White image by setting a threshold value of 127
Yellow Pixels intensity > 127 (They turned white) —> Chromosomes
Purple background —> Intensity < 127 (They turned black) .**

Below is the image after Thresholding :



Bounding Boxes :



Q1: How can contour detection be used to identify objects in an image?

Contour detection identifies continuous boundaries of connected pixels having similar intensity. Each contour corresponds to a distinct object, allowing us to analyze objects such as chromosomes based on their shape and size in this scenario .

Q2: What is the importance of standardization of data? What difference did you observe before and after standardization?

Standardization removes scaling issues by centering features around zero with unit variance. Before standardization, features like area and perimeter dominated the analysis , whereas after standardization, all features contributed equally and this enabled us fair comparison.

Q3: Let's consider one of the values in the width column is missing. How to handle this missing value?

The missing width value can be handled by imputing it with a statistical measure such as the mean or median of the respective column. This prevents distortion during analysis or model training.

Q4: What is the importance of data normalization? What difference did you observe before and after normalization?

Normalization scales all features to a fixed range, typically [0,1], making them directly comparable. After normalization, features become easier to visualize and rank, and no feature numerically dominates others due to scale .

Q5: How might you adapt the bounding box construction process to handle overlapping or touching chromosomes?

Overlapping or touching chromosomes can be handled by merging intersecting bounding boxes or applying morphological operations before contour detection. This helps group related contours making sure a single bounding box represents each chromosome.