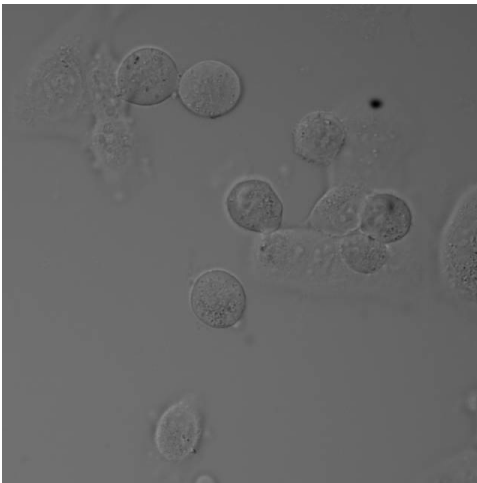


EEL 4930/5934
Introduction to Biomedical Image Analysis
Assignment – 8
Due: 04/02/2024, Noon

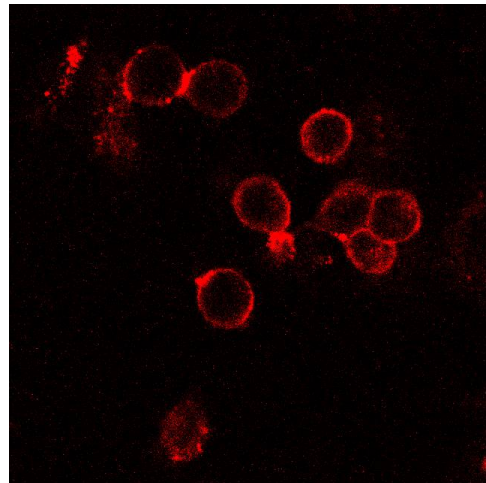
Segmentation Performance Quantification

In last week's lecture, we talked about different metrics used to quantify segmentation performance compared to ground-truth annotations. For this assignment, you will try to segment the cells from two different imaging modalities, DIC (Differential Interference Contrast) and fluorescence.

DIC

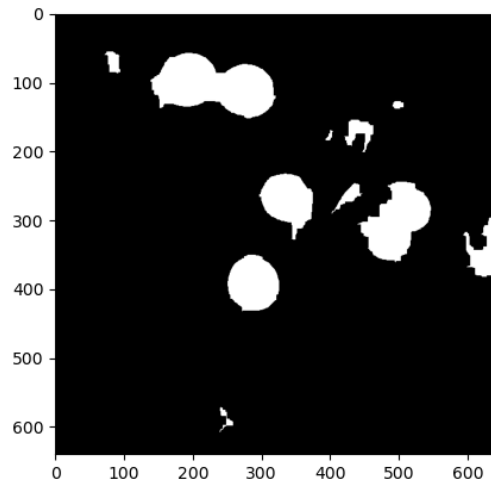


Fluorescence

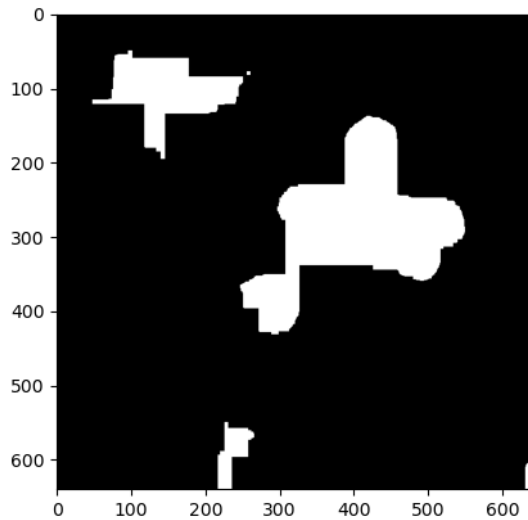


- 1) Segment cells from both of these images. Output should be two binary masks (one for each imaging modality). **(10 pts X 2 = 20 pts)**

DIC



FLUOR:



- 2) Now we will compare the segmentations to the provided ground truth (*SS-ANT_GT.jpg*). For each of the two binary masks generated in question 1, generate the F1 score, precision, recall, and specificity. (2 pts X 2 images X 4 metrics = 16 pts)

a.
$$F1 \text{ score} = \frac{True \text{ Positives}}{True \text{ Positives} + \frac{1}{2}(False \text{ Positives} + False \text{ Negatives})}$$

$$\text{b. Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

$$\text{c. Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

$$\text{d. Specificity} = \frac{\text{True Negatives}}{\text{True Negatives} + \text{False Positives}}$$

DIC first, then fluor

```
{'Precision': 0.8486555425656287, 'Recall': 0.7127767198286122, 'Specificity': 0.9883787989249301, 'F1 Score': 0.7748039136411418}
{'Precision': 0.5679579632398206, 'Recall': 0.7686562723161152, 'Specificity': 0.9465430191183991, 'F1 Score': 0.6532392656653012}
```

- 3) Describe the segmentation performance for each one of these images. What do the metrics generated in question 2 tell you about your segmentation quality? Which imaging modality resulted in a better/easier segmentation? (4 pts)

DIC:

The precision is relatively high, meaning that most pixel classified as parts of a cell are indeed parts of a cell. The recall is not as high but still high, meaning that most of the cells are not fully segmented out but at least partially not detected. The specificity is very high – very few cells are classified as a false positive. The F1-score not terrible indicating that despite the false negatives present, the overall segmentation is still passable

Fluor:

The precision is low: about half of pixels classified as cells are not parts of a cell. The recall is higher than DIC, indicating that more of the cells are segmented out, but of course it is still not perfect. The specificity is very high, but lower than DIC – very few pixels are classified as a false positive. The F1-score is lower than DIC significantly; the presence of many false positives as well as false negatives mean that this segmentation is rather poor.

Overall, the DIC resulted in a better segmentation. This may be because it is better for segmentation, or because of the crude segmentation methods.

