**EEL 4930/ 5934**

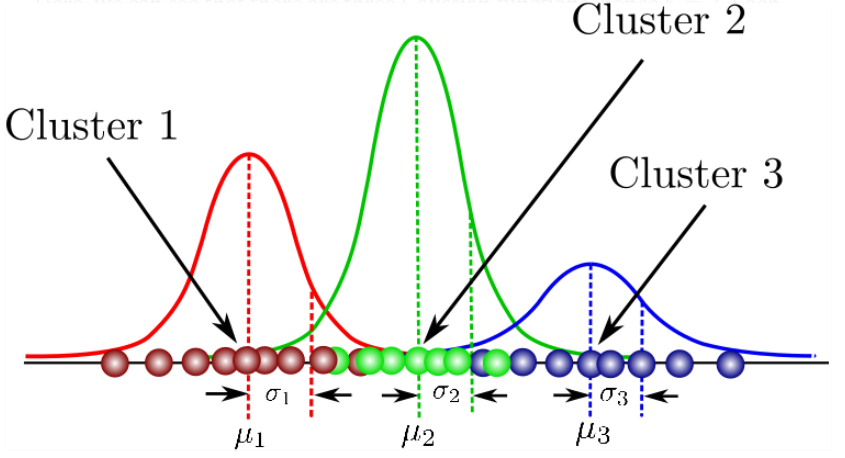
**Introduction to Biomedical Image Analysis**

**Assignment – 11**

**Due: 04/18/2024, Noon**

**Image Segmentation via Clustering Algorithms:**

During the process of image segmentation, we look for characteristic features of the structures we are interested in extracting from the image. These features can be related to color, size, shape, and texture. Separating the objects in the image that we are interested in from objects in the background can also be accomplished using clustering algorithms. In class we went through an example using K-Means on the RGB channels of an image. For this assignment we will be looking at the Mixture of Gaussians algorithm which uses an iterative technique to estimate the parameters (mean and standard deviation) of a certain number of Gaussian distributions that are combined to mimic the distribution of the provided data. An example of this is shown below.



Taken from: <https://towardsdatascience.com/gaussian-mixture-models-explained-6986aaf5a95>

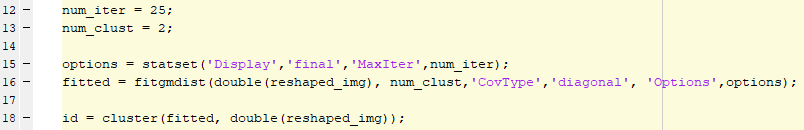
**Part 1: Preparing Image data for clustering.**

1. Read in the image into your workspace (‘*HE\_img.PNG’)*.
2. Reshape the image matrix from (m x n x 3) to (L x 3). This puts the different color channels into columns.
3. Generate a 3-D scatterplot of RGB values. Label each axis according to the color channel those values are derived from.

**(10 pts.)**

**Part 2: Clustering the data.**

1. The command for fitting a mixture of Gaussians in Matlab is *fitgmdist()*. Look at the documentation page by typing *fitgmdist()* in the search bar in the upper-right corner. What is the name of the algorithm used to optimize Gaussian parameter values? What is the default method used to initialize distributions?
2. Run the following lines of code. (Replace ‘reshaped\_img’ with the name of your 3 column matrix from Part 1).



1. Reshape the ‘id’ matrix to the (m x n) height and width of the original image and display using *imagesc()*. Include the resulting image and original image below.
2. Try out a handful of different values (5) for the ‘num\_clust’ variable. How do the cluster-labeled images change as the number of clusters increases? What are the log-likelihood values for each number of clusters?

**(20 pts.)**

**Part 3: Clustering in different color spaces.**

1. Use the commands *rgb2hsv()*, *rgb2lab()*, and *rgb2ycbr()* to transform your original RGB image into the HSV, LAB, and YCbCr color spaces. Run the mixture of Gaussians clustering again using ‘num\_clust’ equal to 4. Include the results for each color space below. How do the different color spaces compare? Which one seems to best isolate individual cell types within the image?

**(20 pts.)**