**Outline of Work Done on Project**

* Work done with clustering
  + K-means Clustering with k = 16, 32, 64, 128
    - K = 64 should probably be the final decision because the results look best (maybe we can compare this in the results)
  + Mean-shift clustering
    - While this seems to deliver decent results, the algorithm scales very poorly and will not be used further in the project
* Work done with neural nets
  + Image segmentation
    - Using a convolutional neural network, the image is segmented into nucleus and non-nucleus by classifying each pixel
    - Each image was scaled down to 256 x 256, RGB and grayscale images were tested
    - Training took about 2 days for each neural network
    - Provided decent results, but probably won’t show in report
    - The main problem with this method is the inability to correctly segment each individual nucleus
    - Tried to fix this by adding 3 classes for each pixel, non-nucleus, border of nucleus, interior of nucleus, but this didn’t improve the results enough to justify further investigation
  + Instance segmentation
    - Similar to instance segmentation, but first we find bounding boxes for each nucleus instance
    - Once we determine a bounding box contains a single nucleus, we segment the image into non-nucleus, nucleus and output the mask
    - The algorithm used for this was Mask-RCNN
    - This method shows potential but has not been investigated enough yet. Results seem good but training time for the neural network takes a very long time. Only RGB images have been tested so far.
* Future work
  + Further investigation with Mask-RCNN
  + We can try PCA with the same subimages that we run with clustering
  + We can improve the results across all categories by transforming our images into SIFT features, or at least a subset of SIFT features (these are the 11 dimensional vectors we talked about)