**LIDC Weekly Meeting** **9/8/2016**

PCA on Dicom Images:

For the code visit <https://github.com/ysriram1/Lung-Nodule-Identification>

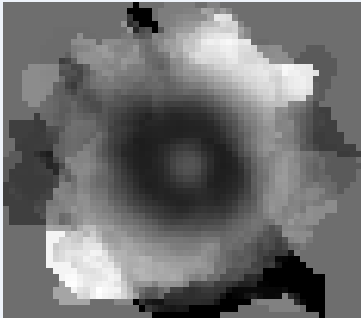
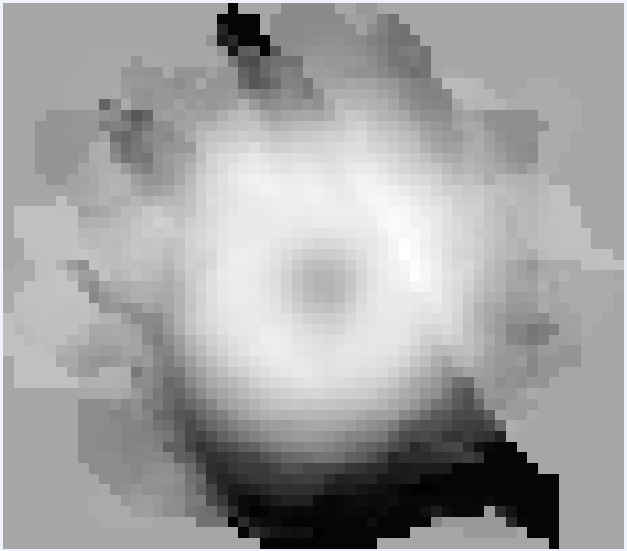
* Obtained the raw DICOM image files from \\Medixsrv\lidc\LIDC\All Extracted Images\nodules
* Extracted the 809 DICOM images that correspond to the 809 instances in the REU file
* There is a lot of variation in the sizes on the DICOM images. The maximum number of rows is 51 and columns is 58.
* All the image matrices have been symmetrically padded to be of size 51 x 58.
* All the matrices are unrolled into a single arrays and stacked up to form 809x2958 matrix.
* This matrix is standardized by centering and scaling.
* PCA is performed on this matrix
* We extracted the top 10 Eigen vectors. Here is the % of variation each of the Eigen Vectors (PCs) account for:

**0.26953955, 0.10198192, 0.08870162, 0.06421865**, 0.05518049, 0.04118215, 0.03424778, 0.03166701, 0.02642079, 0.02131871

* The Eigen vectors that correspond to the first 4 highest
* eigen values have been extracted and scaled to 0 to 255. And recreated as images.

Here are the eigen nodules:

PC1: 26.9% PC2: 10.2%

PC3: 8.9% PC4: 6.4%

