Learning cancer grade in an automated manner

Grading differs to the differentiability of cancers-that is, how different cancer tissue looks from the surrounding normal tissue. Automated cancer grading would mean that we no longer need multiple radiologists to painstakingly manually annotate cancer aggressiveness. Grading leading to staging would mean that we learn the characteristics of poorly differentiated tissue and can assign stage to that particular cancer. This would have far-reaching implications in the way that we are used to dealing with the cancer staging problem, which can be a very tricky problem to solve, especially for cancers that are non muscle invasive and non-metastatic, however growing at a fast pace (that is, highly aggressive).

The project is structured as follows:

- Train a CNN to classify cancer versus normal tissue ROIs from labeled segments of bladder cancer MRI scans.
- For each image, fit a bimodal Gaussian mixture model on just the cancer ROI.
- Replicate cancer region from one embedding in the bright region and one embedding in the dark region (closest to the means of the modes obtained from the trained GMM model), using a genetic algorithm.
- Compute Manhattan distance between the closest embedding achieved to the cancer ROI and the cancer ROI embedding itself.
- Compute a correlation between the closest distances and the labeled cancer grades.

This project has to be completed by the middle of February, owing to administrative constraints.