Automatic Prostate Tumor Detection from T2-Weighted MRIs

Shalini Dangi (304465145)
Kelvin Zhang (304569586)
Department of Computer Science
University of California, Los Angeles

Background

- Prostate cancer (PCa) is one of the most common types of cancer in men, and manifests in the form of malignant growths on the prostate gland.
- The current standard for PCa diagnosis is transrectal ultrasound biopsy, however manual physician diagnoses is subjective and difficult to standardize and test for errors.
- MRIs represent a faster, non-invasive alternative and existing research has been directed at automated, software based solutions to prostate cancer detection that are faster, more scalable, and less prone to error than human processing methods.
- This project will focus on applying machine learning techniques to T2-Weighted MRIs, one of the main parameters in multiparametric MRI analysis, in order to detect the presence and location of PCa growths.

Methods

- A data set of 62 patients with varying levels of cancerous tissue were used, and each T2 MRI was paired with a corresponding "mask" showing cancerous, prostate, and non-prostate regions (*Figure 2*).
- For each pixel in each image, we extracted a feature vector consisting of its immediate neighbors within an N-pixel radius (*Figure 1)*.
- The labeled feature vectors were used to train supervised learning models provided by SciKit-Learn, and tested using k-fold cross validation.

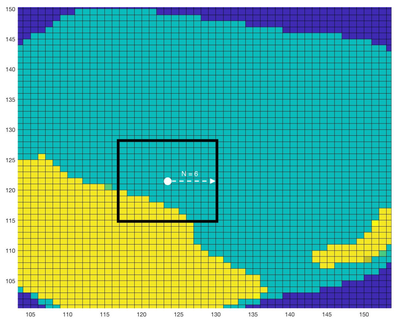


Figure 1. Illustration of feature vector consisting of all of the pixels within the black box with the center pixel designated by a white circle. The size of the patch is determined by N. (In this example, N = 6.) Feature vectors are discarded if the center pixel corresponds to a 0 in the mask (marked by a dark blue color) as this indicates a region outside of the prostate.

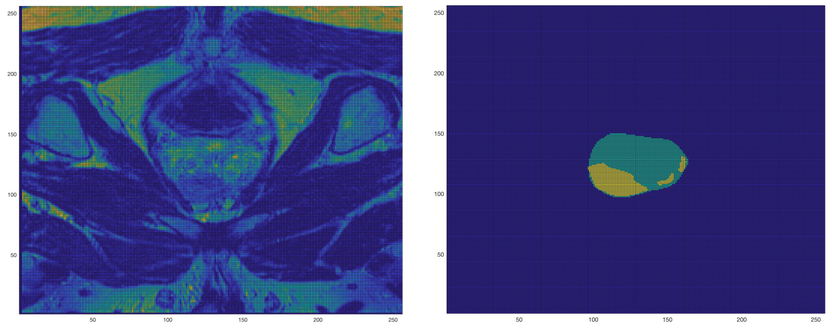


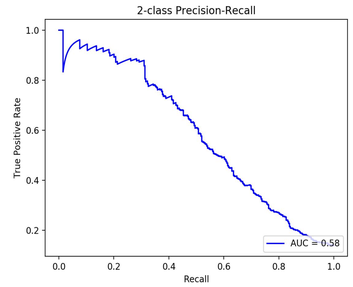
Figure 2. Example T2-weighted image (left) and corresponding mask (right). Mask colors represent 0 - blue, 1 - green, 2 - yellow for non-prostate, prostate, and cancerous prostate regions, respectively. Note that the quality of scans and masks varies greatly among patients.

Results

- We tested the following models: random forest classification, AdaBoost, gradient boosting, k-nearest neighbors, decision tree classification, and Multilayer Perceptron classification.
- Experimental adjustments that we explored in pursuit of optimal performance included feature vector size (N) variation, feature vector pruning, k-fold variation for cross validation, selective pruning of patients, prediction methods after model fitting (binary vs. probabilistic), threshold variation, and utilization of multiple evaluation methods (ROC-AUC, PR-AUC, Accuracy, Sensitivity, Specificity, etc.).
- ROC-AUC varied dramatically across models, with some underperforming near 0.5 and GBC with N=6 achieving the highest score of 0.8404.
- PR-AUC tended to fluctuate between specific images as expected since it is affected by class balance. As a result of the large imbalance in data and our focus on positive labels, it may be the more useful metric for this application.
- We considered precision recall performance by taking the difference between PR-AUC and baseline precision of positive labels/total labels; this was maximized by GBC at n=4 with a difference of 0.411.

Conclusion

- Ultimately, we observed the best ROC-AUC and PR-AUC performance from the SciKit-Learn Gradient Boosting Classifier (*Figure 3)*.
- Although we were unable to generalize strong results through all of the models we tested, based on the data we gathered we hypothesize that with tuning of the individual estimators and careful curation of the data (for example, normalizing the pixel values or more intelligently pruning outliers), machine learning models offer a very powerful approach to medical imaging problems.



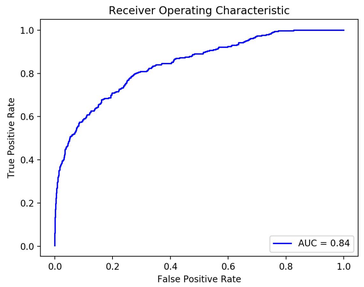


Figure 3. Precision Recall curve for GBC with n=4, Receiver Operating Characteristic curve for GBC with n=6. These were the best performing models for their respective diagnostic statistic.

References

- L. Hussain, A. Ahmed, S. Saeed, S. Rathore, I. A. Awan, S. A. Shah, A. Majid, A. Idris, and A. A. Awan, "Prostate cancer detection using machine learning techniques by employing combination of features extracting strategies," Cancer Biomarkers, vol. 21, no. 2, pp. 393–413, 2018.
- X. Wang, W. Yang, J. Weinreb, J. Han, Q. Li, X. Kong, Y. Yan, Z. Ke, B. Luo, T. Liu, and L. Wang, "Searching for prostate cancer by fully automated magnetic resonance imaging classification: deep learning versus non-deep learning," Scientific Reports, vol. 7, no. 1, 2017.