Predicting a Multi-Parametric Probability Map of Active Tumor Extent Using Random Forests

Matthew Kelsey, PhD, Alicia Boyd, Fred Prior, PhD, Sarah J. Fouke, MD, Tammie Benzinger, MD, Michael Chicoine, MD, Sharath Cholleti, PhD,
Bart Keogh, PhD, Lauren Kim, MD, Mikhail Milchenko, PhD, David G. Politte, PhD, Stephen Tyree, Kilian Weinberger, PhD, Daniel Marcus, PhD

Introduction

- Glioblastoma multiforme (GBM) is the most common primary malignant brain tumor.
- GBM's infiltrative nature makes radiological tumor margin delineation challenging, which in turn affects the extent of surgical resection.
- Our aim is to develop a rule-based multi-parametric approach which incorporates multiple MRI markers in a concerted fashion as an improved method of characterizing the extent of viable tumor within a GBM lesion.
- Further, we propose a machine learning based multi-parametric approach, which uses radiologist-generated labels to train a classifier that is able to classify tissue on a voxel-wise basis and automatically generate a tumor segmentation.

Methods

- Preoperative MRI examinations of subjects with GBM were chosen from the COmprehensive Neuro-oncology Data Repository (CONDR) at Washington University in St. Louis and Swedish Neuroscience Institute (Seattle, WA).
- 8 MRI sequences, primary and derived, [T1 pre-contrast (Fig.1αA), T1 post-contrast (B), T2 (C), Fluid Attenuated Inversion Recovery (FLAIR) (D), Susceptibility Weighted Imaging (SWI) (E), Apparent Diffusion Coefficient (ADC) (F), relative Cerebral Blood Volume (rCBV) (G), and TraceW (H)] were co-registered and transformed to standard template space with 1 mm isotropic voxels. A board-certified radiologist manually segmented each MRI volume (Figure 1α) to produce a set of 6 total object maps (Figure 1ß).
- Voxel Labels were generated by combining manual segmentations (Table 1) based on the radiologist's rule set and estimate of the probability of active tumor (Fig. 1γ & Table 2).
- A Random Forests classifier was trained using a leave-one-out experimental paradigm. Linear regression analysis was also implemented for comparison.
- Tumor Infiltration (TI) was calculated at 20 locations across 7 subjects using needle biopsy pathology results.
- Receiver Operating Characteristic (ROC) analysis was used to compare the predictions of the Random Forests (RF) classifier and a linear regression-based classifier relative to the radiologist's manual segmentation.
- Further ROC analysis compared radiologist segmentations and RF predictions to ground truth values at the 20 locations with labeled pathology.

Acknowledgments

This work was partially supported by NIH grant 1R01NS066905-01. Data were provided by clinical trial NCT01124461.

References

1. Prior, F.W., et al., *Predicting a Multi-Parametric Probability Map of Active Tumor Extent Using Random Forests*. Engineering in Medicine and Biology Society (EMBC), 2013 Annual International Conference of the IEEE. *In review.*
2. McMillan, K.M., et al., *An objective method for combining multi-parametric MRI datasets to characterize malignant tumors*. Medical Physics, 2007. 34(3): p. 1053-1061.
3. Breiman L. Random Forests. Machine Learning. 2001;45(1):5-32.
4. Corso, Jason J., et al. *Efficient multilevel brain tumor segmentation with integrated bayesian model classification*. Medical Imaging, IEEE Transactions on 27.5 (2008): 629-640.

Figure1. Rule-Based Radiologist Analysis Establishes "Truth"

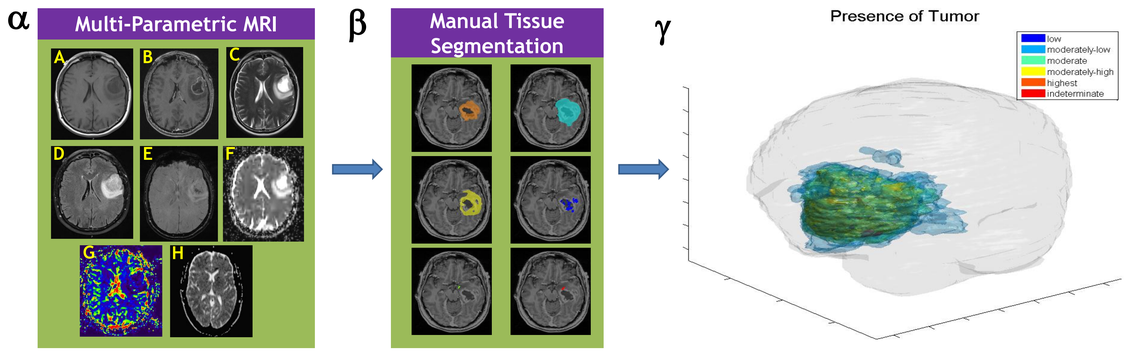


Table 1. Summary of Segmentations

MRI Parameter	Criteria	Classification for Viable Malignancy
SWI	Discontinuous areas of signal void on SWI	Indeterminate: hemorrhage
Necrosis	T2 hyperintensity suppressed on FLAIR	Negative: liquefactive necrosis
FLAIR	Hyperintensity on FLAIR	Positive: possible micro-invasion of tumor
Diffusion Restriction (DR)	Hyperintensity on TraceW; hypointensity on ADC	Positive: viable tumor
rCBV	Areas demonstrated 1.75 times the cerebral blood volume compared to normal brain tissue	Positive: viable tumor
Enhancement	Hyperintensity on T1 post-contrast not present on T1 pre-contrast	Positive: viable tumor

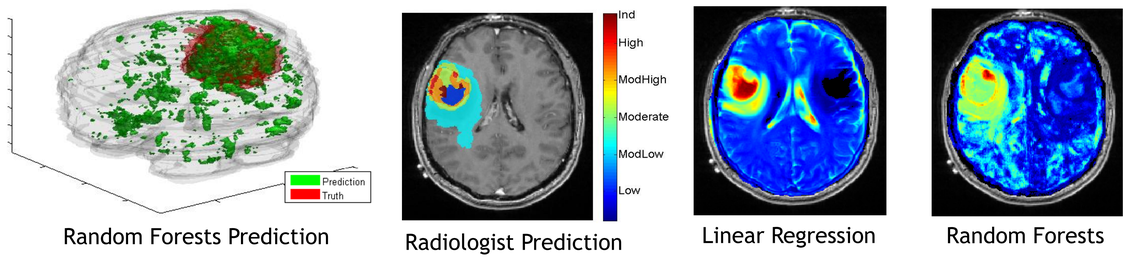
Table 2. Assigning Probability of Viable Tumor

Probability of Viable Tumor	Criterion
Normal Brain	Voxel not included in any object map
Indeterminate	Any voxel containing susceptibility artifact
Low	Any voxel containing necrosis
Moderately Low	Any voxel containing FLAIR hyperintensity in the absence of other positive indicators (enhancement, diffusion restriction, or elevated CBV)
Moderate	Any voxel containing FLAIR hyperintensity and enhancement without additional positive indicators (diffusion restriction or elevated CBV)
Moderately High	Any voxel containing FLAIR hyperintensity and enhancement with one additional positive indicators (diffusion restriction or elevated CBV)
Highest	Any voxel containing all the positive indicators for viable tumor (FLAIR hyperintensity, enhancement, diffusion restriction, and elevated CBV)

Machine Learning Based Tumor Multi-Parametric Probability Maps

- The Random Forests (RF) algorithm constructs an ensemble of decision trees. Each decision tree is constructed by selecting a random subset of features and training examples, creating a variety of *experts*. The leaves of the trees are associated with constant predictions. RF combines the votes of all the trees for overall prediction.
- To construct a tumor probability estimate and tissue segmentations, feature vectors are constructed for each voxel in the brain, using the value of that voxel in each of the 8 MR data types.
- Once trained, the classifier is applied to every voxel (every feature vector) in the test set and classifies the tissue as normal or malignant.
- A leave-one-out experiment with N labeled data sets uses N-1 data sets to train the classifier and then predicts the labels of the N^{th}. This process is repeated until all data sets have been predicted by the Classifier. Figure 2 illustrates the radiologist's multi-parametric tumor map and the map predicted by the Random Forests Classifier for one case.

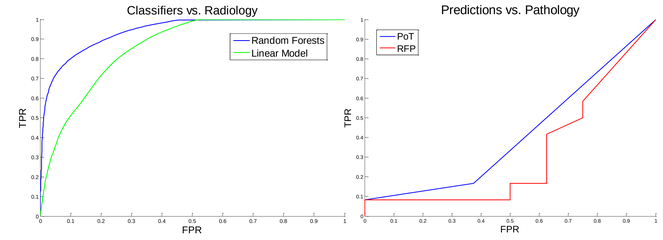
Figure 2. Tumor Segmentations



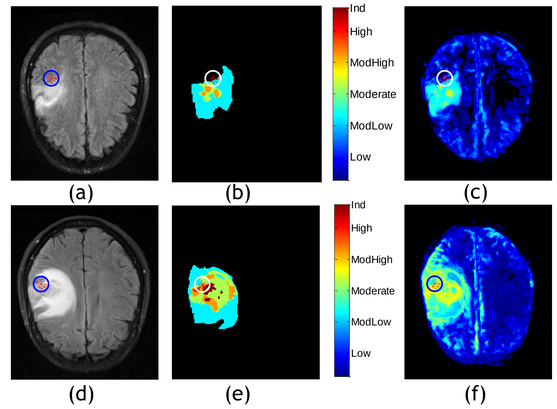
Random Forests Prediction | Radiologist Prediction | Linear Regression | Random Forests

The Random Forests algorithm classifies other brain tissue as having a high probability of being cancer. Separating false positives from true positives in these areas is an area of active research.

Results



- ROC analysis (**Classifiers vs. Radiology**) shows that both linear and Random Forests (RF) classifiers are able to generate reasonable multi-parametric probability maps predicting radiologist-generated segmentations and tumor extent. The RF classifier results in an Area Under Curve (AUC) of 0.92 compared to an AUC of 0.77 for the linear model.
- Using pathology as truth, ROC analysis (**Predictions vs. Pathology**) shows poor correlation with radiologist's designation of Presence of Tumor (PoT) and Random Forests Prediction (RFP). ROC AUC were 0.41 and 0.32, respectively.



- Pathology indicates 100% tumor infiltration at the point specified by * in (a) while the same coordinates place the marker outside the margins of a radiologist's Presence of Tumor (PoT) segmentation (b) and Random Forests Prediction (RFP) bounds (c).
- Similarly, pathology reports Low Probability of tumor infiltration (35%) at the point shown in (d). Corresponding coordinates in (e) and (f) show High and Moderately High probability of tumor predictions by PoT segmentation and RFP.
- Pathology localization appears to be compromised by brain shift and post hoc estimation of biopsy location.

Conclusions

- The infiltrative nature of gliomas makes assessment of tumor burden a challenge. Multi-parametric imaging markers may offer a method to improve our measures of tumor invasion and, ultimately, extent of resection.
- By enhancing our multi-parametric approach with Machine Learning we eliminate manual segmentation and generate a probability map that incorporates contrast enhancement with additional MRI markers to accurately predict radiologist defined tumor boundaries and tissue type.
- Validation of radiologist based PoT and machine learning based RFP tumor extent using pathology results were unsuccessful due to sample localization error and inadequate sample size.