35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

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

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

Abstract— **Glioblastoma Mulitforme is highly infiltrative, making precise delineation of tumor margin difficult. Multi-modality or multi-parametric MR imaging sequences promise an advantage over anatomic sequences such as post contrast enhancement as methods for determining the spatial extent of tumor involvement. In considering multi-parametric imaging sequences however, manual image segmentation and classification is time-consuming and prone to error. As a preliminary step toward integration of multi-parametric imaging into clinical assessments of primary brain tumors, we propose a machine-learning based multi-parametric approach that 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. A random forests classifier was trained using a leave-one-out experimental paradigm. A simple linear classifier was also trained for comparison. The random forests classifier accurately predicted radiologist generated segmentations and tumor extent.**

*Research supported in part by NIH grant 1R01NS066905-01. Data was provided by clinical trial NCT01124461.

Fred Prior is with the Mallinckrodt Institute of Radiology, Washington University School of Medicine, St. Louis, MO 63110 USA (phone: 314-747-0331; fax: 314-362-6971; e-mail: priorf@mir.wustl.edu).

Sarah J. Fouke is with the Department of Neurosurgery, Swedish Medical Center, Seattle, WA 98122, USA (e-mail: Sarah.Fouke@swedish.org).

Tammie Benzinger is with the Mallinckrodt Institute of Radiology, Washington University School of Medicine, St. Louis, MO 63110 USA (e-mail: benzingert@mir.wustl.edu).

Alicia Boyd is with the Mallinckrodt Institute of Radiology, Washington University School of Medicine, St. Louis, MO 63110 USA (e-mail: boyda@mir.wustl.edu).

Michael Chicoine is with the Department of Neurosurgery, Washington University School of Medicine, St. Louis, MO 63110 USA (email chicoinem@wudosis.wust.edu)

Sharath Cholleti is with the Center for Comprehensive Informatics, Emory University, Atlanta, GA 30322 USA (e-mail: sharath.cholleti@emory.edu).

Matthew Kelsey is with the Mallinckrodt Institute of Radiology, Washington University School of Medicine, St. Louis, MO 63110 USA (e-mail: kelseym@mir.wustl.edu).

Bart Keogh is with the Department of Radiology, Swedish Medical Center, Seattle, WA 98122, USA (e-mail: Bart.Keogh@swedish.org).

Lauren Kim is with the Department of Radiology, Massachusetts General Hospital, Boston, MA 02114, USA (e-mail: LKIM6@partners.org).

Mikhail Milchenko is with the Mallinckrodt Institute of Radiology, Washington University School of Medicine, St. Louis, MO 63110 USA (e-mail: milchenkom@mir.wustl.edu).

David G. Politte is with the Mallinckrodt Institute of Radiology, Washington University School of Medicine, St. Louis, MO 63110 USA (e-mail: politted@mir.wustl.edu).

Stephen Tyree is with the Department of Computer Science & Engineering, Washington University, St. Louis, MO 63130 USA (e-mail: swtyree@wustl.edu).

Kilian Weinberger is with the Department of Computer Science & Engineering, Washington University, St. Louis, MO 63130 USA (e-mail: kilian@seas.wustl.edu).

Daniel Marcus is with the Mallinckrodt Institute of Radiology, Washington University School of Medicine, St. Louis, MO 63110 USA (e-mail: dmarcus@wustl.edu).

I. Introduction

Glioblastoma Muliforme (GBM) is the most common primary malignant brain tumor. Outcomes for GBM patients remain poor even after multimodal therapy (including surgical resection, radiotherapy, and chemotherapy) with a median survival of approximately one year and a 2-year survival rate of less than 20%[1, 2]. GBM is highly infiltrative, making precise delineation of tumor margin difficult. Contrast enhancement on MRI has traditionally been used to plot the geographic extent of tumor involvement[3]. However, current consensus acknowledges the role of additional MRI markers in the characterization of tumor margins beyond the boundaries of contrast enhancement [4-7].

A significant predictor of patient outcome is the extent of surgical resection. Despite this, complete resection of enhancing tissue alone does not result in a resection of all tumor burden, and it is recognized that tumor remains beyond the limits visible on anatomic imaging sequences such as T1 + Gadolinium (contrast) MRI. One might suggest a more extensive surgical margin to improve clinical outcomes; however, extension of surgical margins can increase the risk of surgical morbidity as often tumor margins may approach functionally important regions of the brain. Thus accurate pre-operative definition of the geographic extent of tumor is highly desirable.

It is, however, very difficult for a neurosurgeon to integrate the multiple available MR sequences in a single operative session. Even considering multi-parametric sequences is technically challenging. Manual image segmentation and classification are time consuming and prone to error. Ultimately, a single map combining all the relevant MR parameters in the characterization of affected brain would have significant clinical utility [8].

Our aim is to develop a rule-based multi-parametric approach that 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.

II. MULTI-PARAMETRIC MRI

A. Imaging Protocol

Preoperative MRI examinations of subjects with GBM were chosen from the Comprehensive Neuro-oncology Data Repository (CONDR) at Washington University in St. Louis. All data was collected at Washington University and Swedish Neuroscience Institute (Seattle, WA) using a standardized imaging protocol implemented on all clinical scanners at each of the two participating institutions. This standardized brain tumor MR imaging protocol is based on the ACRIN 6686 (RTOG 0825) [9] protocol and includes MP-RAGE/SPGR (SPoiled Gradient Recalled) Gd-enhanced volumetric acquisition and T2-weighted and FLAIR sequences in addition to advanced imaging sequences including diffusion (acquired as diffusion tensor imaging (DTI, 12 direction, bmax 1400) [10]), dynamic susceptibility contrast (DSC) [11] to generate relative cerebral blood flow (rCBF) and cerebral blood volume (rCBV) parameter maps. The protocol includes 7 pulse sequences and sets of acquisition parameters that yield 11 derived image sets.

B. Image Processing

Eight MRI image sets, primary and derived [T1 pre- and post- contrast, T2, Fluid Attenuated Inversion Recovery (FLAIR), Susceptibility Weighted Imaging (SWI), Apparent Diffusion Coefficient (ADC), TraceW, and rCBV], were interpolated to 1 mm^3 isotropic voxels and co-registered using nordicICE (NordicNeuroLab AS, Bergen, Norway) perfusion and diffusion processing.

III. RULE-BASED EXPERT ANALYSIS

The literature has shown that the inclusion of multi-modality MR sequences can improve the performance of tumor tissue classification [8, 12]. Systematic methods to establish an accurate ground truth labeling, however, remain elusive. At present, most radiologists and neurological surgeons rely on qualitative assessment of independent image sequences when considering brain tumor diagnoses. A more quantitative and rigorous method focusing on the integration of imaging sequences to, collectively, answer questions of tissue diagnosis, tumor margin, and even therapeutic response relies upon a more quantitative and specific methodology for image analysis. In this work, we propose such a method, utilizing expert radiologist evaluation along with techniques developed to combine MR parameter maps into a unified probability map suggesting a more sensitive and specific measure of tumor margin.

A. Approach

A board-certified radiologist manually segmented regions of signal abnormality on each MRI parameter map (Table 1) to produce a set of 6 object maps (Figure 1). All images were registered to a common, 1 mm isotropic atlas space prior to segmentation. These segmentations were reviewed and approved by a CAQ-certified neuroradiolgist. The various types of images contained information about the presence or absence of tumor within each voxel in the image volume based on current neuroradiology conventions. Although each type of image contains valuable information when considered separately, hypotheses about the viability of tissue can be strengthened immensely if all images are considered concurrently. The radiologists' knowledge of how to combine the information contained in the images is summarized as a set of heuristic rules (Table 2).

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

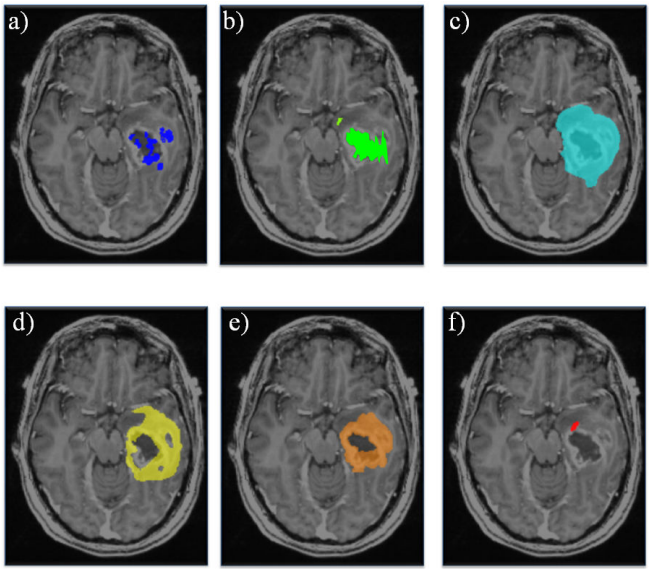

Figure 1. Radiologist generated Segmentations: a) SWI, b) Necrosis, c) FLAIR, d) Diffusion restriction, e) Enhancement, f) rCBV have been superimposed on T1-post contrast image.

B. Rules

In general, the GBM region of interest is composed of tissue in multiple states, including: necrotic tissue, actively growing tumor, surrounding edema, infiltrating tumor and normal, non-diseased brain [6]. Variances do occur, notably, necrotic tissue is occasionally absent or not visible, and although often actively growing tumor is shown as taking up contrast dye, at times significant tumor burden is identified in regions of tissue that do not take up Gadolinium contrast. While tissue in each of these states may be identifiable, the boundaries are not discrete and the classification not unambiguous. The infiltrative nature of gliomas instead results in a visible gradient boundary between regions. In order to represent this border uncertainty, we represent tumor location by labeling voxel areas with a Probability of Viable Tumor indication rather than an absolute classification.

Summarized in Table 2, the radiologist assigned probability of active tumor to each object map and combined overlapping object maps within a voxel. Normal brain denotes no incidence of tumor whereas highest represented the highest probability of tumor. Indeterminate signifies lack

Table 2. Assigning Probability of Viable Tumor

Probability of Viable Tumor	Criteria
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)

of knowledge about the presence of tumor in a particular voxel. Individually segmented MRI volumes were combined by the radiologist to produce labels suggesting a probability of tumor being found at each voxel location.

IV. Machine Learning

Machine learning techniques were developed to automatically reduce the high dimension source data and generate a tumor probability prediction. It is intended that the values produced by our machine learning algorithms should predict the radiologist assigned *Probability of Viable Tumor* in each subject.

Classifier models process voxel values from each of the eight MR data types, which were fused to form feature vectors. These models were trained and tested using combinations of truth labels, generated by combining manual segmentations based on the radiologist's rule set and estimate of the probability of active tumor, as summarized in Table 2.

A. *Approach*

Random forests (RF) [13] and linear classifiers were trained using the leave-one-out experimental paradigm. Given N labeled data sets, this approach uses N-1 data sets to train the classifier and predicts the labels of the N^{th} set. This process is repeated until the classifier has predicted all data sets.

The RF algorithm generates a predictive model by constructing an ensemble of decision trees (also known as classification or regression trees) [13] using the training data set. Each decision tree is constructed using a randomly drawn subset of the variables in the data and using a subset of examples from the training set. This creates a set of trees, each with a slightly different predictive model. For a given test input, random forests combines the votes of all trees to generate an overall prediction. The probability of correctness is then computed as the fraction of trees that voted yes. Our decision tree is constructed by partitioning the training data based on the values of the variable that best splits the data into homogenous subsets (called nodes) according to the value of the target variable [14, 15]. This process of partitioning is applied repeatedly using the rest of the variables until no further gain is attained. Each leaf node of the decision tree is associated with a constant prediction that is equal to the most frequently occurring value of the target variable.

This RF based analysis is complemented by assessing the baseline predictive ability of a linear classifier on our data. The linear model simplifies feature space partitioning by assuming that voxels belonging to different tissue types, each represented by its feature vector, can be categorized using a D-dimensional hyper-plane. Where D is equal to eight in this work and matches the length of the feature vector. This simplification provides some resistance to over-fitting the training data, but often neglects to capture complex intricacies present in real data.

To construct tumor probability estimates 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 diseased.

B. *Results*

We compare the predictive power of our Random Forests method to that of a linear classifier, trained with identical data. While both methods highlight areas including active tumor, the RF classifier generated a multi-parametric probability map that more accurately predicted radiologist generated segmentations and tumor extent than did the linear classifier. Figure 2 highlights results attained using regression analysis, which compares radiologist generated ground truth (a), RF (b), and the linear model (c).

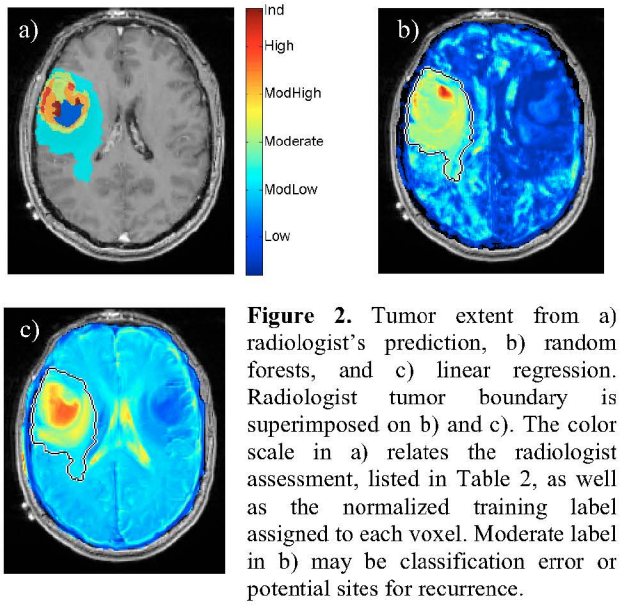

Figure 2. Tumor extent from a) radiologist's prediction, b) random forests, and c) linear regression. Radiologist tumor boundary is superimposed on b) and c). The color scale in a) relates the radiologist assessment, listed in Table 2, as well as the normalized training label assigned to each voxel. Moderate label in b) may be classification error or potential sites for recurrence.

Receiver Operating Characteristic (ROC) analysis [16] was used to compare the predictions of the random forests classifier and a linear regression-based classifier relative to the manual segmentation standard. All voxels in each multi-parametric image volume were analyzed. Figure 3 illustrates the average ROC curves for all subjects and demonstrates that the RF classifier is a significantly better predictor of radiologist generated segmentations and tumor extent than the linear classifier.

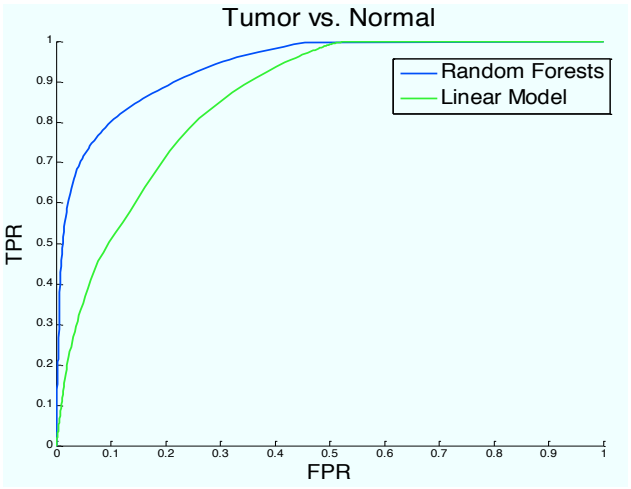

Figure 3. ROC curves for random forests (blue) and linear model (green) predictions of tumor extent.

Table 3 summarizes the results of ROC analysis for individual subjects. Differences in AUC between Random Forests and Linear classifiers were determined to be normally distributed with a Shapiro Wilk W test (P = 0.53). The mean difference was tested against a hypothesized difference of zero with the paired t test and was significantly different (P = 0.0002). Random Forests results in an AUC that is 16% larger than the AUC for the linear classifier.

Table 3. ROC Analysis (AUC)

Subject	Random Forests	Linear Classifier
W001	0.914	0.699
W010	0.925	0.759
W015	0.870	0.790
W019	0.952	0.810
W025	0.941	0.841
W029	0.921	0.721
W034	0.924	0.732
Overall	0.921	0.765

V. Conclusion

The infiltrative nature of gliomas makes assessment of tumor burden a challenge, and multi-parametric imaging markers may offer a method to improve measures of tumor invasion and, ultimately, extent of resection.

By enhancing our multi-parametric approach with quantitative methods for image co-registration, image processing and analysis, and subsequent image segmentation, we laid the foundation for a Machine Learning methodology to facilitate the integration of multi-parametric imaging sequences in consideration of tumors such as malignant gliomas. In this methodology we eliminate manual segmentation and generate a probability map that incorporates contrast enhancement with additional MRI markers to produce a composite image that predicts the probability of viable tumor and tissue type.

It is our expectation that these methods provide a foundation for subsequent studies that facilitate the integration of multi-modality imaging into the clinical management of primary brain tumors and ultimately improve patient's clinical outcomes as a result.

References

[1] M. J. McGirt, K. L. Chaichana, M. Gathinji *et al.*, "Independent association of extent of resection with survival in patients with malignant brain astrocytoma," *Journal of Neurosurgery,* vol. 110, no. 1, pp. 156-162, 2009.

[2] M. D. Prados, and V. Levin, "Biology and treatment of malignant glioma," *Semin Oncol,* vol. 27, no. 3 Suppl 6, pp. 1-10, Jun, 2000.

[3] M. A. Hammoud, R. Sawaya, W. Shi *et al.*, "Prognostic significance of preoperative MRI scans in glioblastoma multiforme," *Journal of Neuro-Oncology,* vol. 27, no. 1, pp. 65-73, 1996.

[4] E. Zacharaki, R. Verma, S. Chawla *et al.*, "Towards predicting neoplastic recurrence with multi-parametric MR."

[5] K. M. McMillan, B. P. Rogers, C. G. Koay *et al.*, "An objective method for combining multi-parametric MRI datasets to characterize malignant tumors," *Medical physics,* vol. 34, pp. 1053, 2007.

[6] J. J. Corso, E. Sharon, S. Dube *et al.*, "Efficient multilevel brain tumor segmentation with integrated bayesian model classification," *Medical Imaging, IEEE Transactions on,* vol. 27, no. 5, pp. 629-640, 2008.

[7] X. Hu, K. K. Wong, G. S. Young *et al.*, "Support vector machine multiparametric MRI identification of pseudoprogression from tumor recurrence in patients with resected glioblastoma," *Journal of Magnetic Resonance Imaging,* vol. 33, no. 2, pp. 296-305, 2011.

[8] A. R. Padhani, and K. A. Miles, "Multiparametric Imaging of Tumor Response to Therapy," *Radiology,* vol. 256, no. 2, pp. 348-364, August 1, 2010, 2010.

[9] ACRIN. "ACRIN 6686 (RTOG 0825) Advanced MRI Imaging Manual" January 18, 2013, 2013; http://www.acrin.org/Portals/0/Protocols/6686/6686%20MRmanual%2010012009.pdf.

[10] P. J. Basser, "Inferring microstructural features and the physiological state of tissues from diffusion-weighted images," *NMR in Biomedicine,* vol. 8, no. 7, pp. 333-344, 1995.

[11] E. C. Wong, R. B. Buxton, and L. R. Frank, "Quantitative imaging of perfusion using a single subtraction (QUIPSS and QUIPSS II)," *Magnetic Resonance in Medicine,* vol. 39, no. 5, pp. 702-708, 2005.

[12] S. Metz, C. Ganter, S. Lorenzen *et al.*, "Phenotyping of tumor biology in patients by multimodality multiparametric imaging: relationship of microcirculation, alphavbeta3 expression, and glucose metabolism," *J Nucl Med,* vol. 51, no. 11, pp. 1691-8, Nov, 2010.

[13] L. Breiman, "Random Forests," *Machine Learning,* vol. 45, no. 1, pp. 5-32, 2001.

[14] J. R. Quinlan, "Induction of decision trees," *Machine learning,* vol. 1, no. 1, pp. 81-106, 1986.

[15] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning, Data Mining, Inference and Prediction*, New York: Springer, 2001.

[16] J. A. Hanley, and B. J. McNeil, "A method of comparing the areas under receiver operating characteristic curves derived from the same cases," *Radiology,* vol. 148, no. 3, pp. 839-843, September 1, 1983, 1983.