MACHINE-LEARNING-BASED CLASSIFICATION OF GLIOBLASTOMA USING DYNAMIC SUSCEPTIBILITY ENHANCED MR IMAGE DERIVED DELTA-RADIOMIC FEATURES

A Thesis
Presented to
The Academic Faculty

By

Jiwoong Jason Jeong

In Partial Fulfillment
Of the Requirements for the degree
Master of Science in Medical Physics

Georgia Institute of Technology

May, 2018

MACHINE-LEARNING-BASED CLASSIFICATION OF GLIOBLASTOMA USING DYNAMIC SUSCEPTIBILITY ENHANCED MR IMAGE DERIVED DELTA-RADIOMIC FEATURES

Approved by:

Dr. C-K. ChrisWang, Co-Advisor
School of Mechanical Engineering
Georgia Institute of Technology

Dr. Xiaofeng Yang, Co-Advisor
Department of Radiation Oncology
Emory University School of Medicine
School of Mechanical Engineering
Georgia Institute of Technology

Dr. Eric Elder
Department of Radiation Oncology
Emory University School of Medicine
School of Mechanical Engineering
Georgia Institute of Technology

Dr. Tian Liu
Department of Radiation Oncology
Emory University School of Medicine

Date Approved: April 23, 2018

Acknowledgements

I would like to thank my co-advisor, Dr. Xiaofeng Yang, for giving me the opportunity to work on this project and taking his time mentoring me throughout this process as well as giving me the ability to work in the Emory environment. I would also like to thank my other co-advisor, Dr. C-K Chris Wang for generously guiding me and giving me advice in many aspects in school and research opportunities.

Additionally, I would like to thank my committee members, Dr. Eric Elder and Dr. Tian Liu, for taking time out of their busy schedule to be in this committee as well as mentoring me with my work at Emory.

I would also like to thank my mom and dad for their continual trust and support.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AUC	Area Under the Curve
CBF	Cerebral Blood Flow
CBV	Cerebral Blood Volume
CNS	Central Nervous System
CT	Computed Tomography
DICOM	Digital Imaging and Communications in Medicine
DSC	Dynamic Susceptibility Contrast Enhanced
DVH	Dose Volume Histogram
DWI	Diffusion Weighted Imaging
FLAIR	FLuid Attenuated Inversion Recovery
FOV	Field Of View
GBM	Glioblastoma
GE	Gradient Echo
GLCM	Gray Level Run Length
IBEX	Imaging BioMarker Explorer
IRB	Institutional Review Board
LASSO	Least Absolute Shrinkage and Selection Operator
MRI	Magentic Resonance Imaging
NIFTI	Neuroimaging Informatics Technology Initiative
NSCLC	Non-Small Cell Lung Cancer
NTCP	Normal Tissue Complication Probability
PET	Positron Emission Tomography
QUANTEC	Quantitative Analysis of Normal Tissue Effects in the Clinic
ROC	Receiver Operating Characteristic
ROI	Region Of Interest
SE	Spin Echo
SVM	Support Vector Machine
TE	Echo time
TR	Repetition Time

CHAPTER 1 – INTRODUCTION

1.1 A Brief Introduction to Glioblastoma Multiforme

Grade IV gliomas, more commonly known as glioblastoma (GBM) multiforme, are the most aggressive and lethal brain tumors. Compared to lower grade (grades I, II, and III) gliomas with an average relative 2-year survival rate of 80%, GBMs have significantly lower relative 2-year survival of 30% with a median overall survival rate of 12-15 months (1, 2). The incidence of malignant gliomas is about 17,000 per year or 5 in 100,000. Overall, about 65% are grade IV GBMs which result in a mortality rate of over 10,000 deaths a year (3).

These poor outcomes stem from the uncooperative, heterogeneous nature of GBMs where some GBMs develop and progress from lower-grade gliomas while other GBMs are primary tumors (4). Much like other solid tumors, GBM develops a very heterogeneous pattern of mutations (5). Only recently has there been studies showing the predictive and prognostic power of genetic characterization of GBM (6, 7) and expression data correlated with response to these treatments (5, 8, 9). Even so, the genetic and molecular heterogeneity of GBMs still poses a significant hurdle to developing novel therapies and clinical decision making (10).

In general, surgery is the first treatment option for both high and low grade GBMs but the grade and severity of the tumor plays a significant role whether or not and how they should be treated. It is difficult to recommend a treatment option that balances quality of life and overall survival when low grade GBMs are asymptomatic and are found incidentally while high grade GBMs' progression free overall survival depends significantly on the extent of the resection (11). Furthermore, radiation and chemo therapy are often prescribed after surgery whose recommendations are historically based on a few clinical parameters like age (11). While there

are ongoing efforts to better understand the underlying heterogeneous biology of GBM with advances in treatment methods, it still remains as one of the most perverse tumor types.

1.2 Objectives

In recent years, to circumvent the complexities of intra-tumoral and inter-patient tissue characterization of GBM through biopsies, that are traditionally difficult to get a detailed characterization of (12), various imaging modalities have been used to characterize the phenotypic expression of the genetic heterogeneities through the radiomic approach (13, 14). Correlations to clinical outcomes have been observed in the breast (15), brain (16), head and neck (17) by different modalities like X-ray computed tomography (CT), positron emission tomography (PET), and magnetic resonance imaging (MRI).

The purpose of this thesis work is to improve the tissue characterization of these highly heterogeneous GBMs using delta-radiomic features of dynamic susceptibility contrast enhanced (DSC) magnetic resonance (MR) images, which are commonly used to derive blood perfusion of the targeted tumor, with a random forest machine learning approach. This method will be able to extract delta-radiomic features from DSC MRIs, select and apply the most salient features, and train a random forest to classify high or low grade GBMs with cross validation.

CHAPTER 2 – BACKGROUND

2.1 Principles of Dynamic Susceptibility Contrast Enhanced (DSC) Perfusion MRI

2.1.1 MRI and Perfusion Imaging

Since its inception, MRI has been used to assess the various abnormalities in the human body. MRI is especially useful for those abnormalities in the brain and central nervous system (CNS) where most contrast agents can't penetrate due to the blood-brain barrier. These abnormalities include primary tumors, secondary or distant metastases, infections, vascular and degenerative diseases of the brain and CNS. This is because of the paramagnetic properties MRI contrast agents, like gadolinium, that allow the user to emphasize the changes in T1 and T2 relaxation rates in MRIs as well as the innocuous pharmokinetitc properties of contrast drugs (18). Many studies have shown that using high contrast media in perfusion MRI, with optimized protocols, can significantly improve the detection, characterization, and monitoring of abnormalities in the CNS (19-22).

MRI is an exceptional candidate to fully characterize these highly heterogeneous glioblastomas. With its distinct advantages and the advances in visualization, resolution, and MR protocols (sequences) interrogating the morphologic and functional characteristics of these heterogeneous tumors has become increasingly wide spread. Although MRI can provide very detailed structural images, it is difficult to fully characterize the phenotypic features of tumors without functional information provided by MR perfusion imaging. Perfusion is defined as the steady-state delivery of blood to a tissue i.e. capillary blood flow. Since as early as the 1980s, several combinations of contrast agents and MR protocols have been used to non-invasively

measure perfusion using MRI with a key focus in the brain (23). In addition, other physiologic parameters can be derived from perfusion MR e.g. blood volume, blood velocity, and blood oxygenation to enhance our understanding of tumor dynamics. In order to add functional, phenotypic information, our proposed classification method was tested on dynamic susceptibility contrast-enhanced perfusion MRI.

2.1.2 Dynamic Susceptibility Contrast Enhanced (DSC) or Bolus Tracking MRI

The theory of DSC MRI is to extrapolate and derive the dynamics of blood flow using MR volumes over time. This is achieved through tracer kinetic analysis via contrast agent concentrations' relationship to the given MR pulse sequence. Bolus tracking using DSC imaging is the tracking of the kinetics of the bolus, e.g. the paramagnetic gadolinium-based chelate (contrast agent), using either T2 or T2* weighted MR sequences (24). As described by Villringer et. al (23) in 1988, the first-pass extraction of contrast agent is zero when the blood-brain barrier is healthy and intact and the intra vascular compartmentalization of the contrast agent creates strong, microscopic magnetic susceptibility gradients. These gradients cause a change in the dephasing of spins as the spins diffuse more quickly among these gradients resulting in signal loss in the mentioned T2 or T2* weighted images. For traditional T2 weighted, or spin echo (SE), MR sequences, the signal loss is minimized due to the refocusing of the dephasing spins whereas T2* weighted, or gradient echo (GE), MR sequences will generally experience a greater signal loss due to the presence of the microscopic field perturbers in the blood vessels.

A detailed analysis of the changes in transverse relaxation rate as a function of vessel size through SE or GE in DSC MRI by Weisskoff, Boxerman, Fisel, and others (25-27) is shown in Figure 1. Their results show that GE sequences are sensitive at all blood vessel sizes and SE are

mainly sensitive to blood vessel sizes comparable to capillaries which means that SE sequences require about double the dose of contrast than GE but for the benefit of preferential sensitivity in detecting changes in small vessel density. Studies have correlated micro vascular cerebral blood volume (CBV) seen by SE is about 45% of the total as observed by PET and GE MRI sequences and found an approximate linear relationship between DSC tissue contrast agent concentration and the change in T2 relaxation rate: $\Delta R_2(t) \propto C_t(t)$ where $C_t(t)$ is the contrast concentration in the tissue at time t. (28, 29). Using this relationship, various assumptions, and equations, which are outside the scope of this paper, one can derive CBV measurements, cerebral blood flow (CBF), flow heterogeneity, and other metabolic signatures (24) to characterize the tumor.

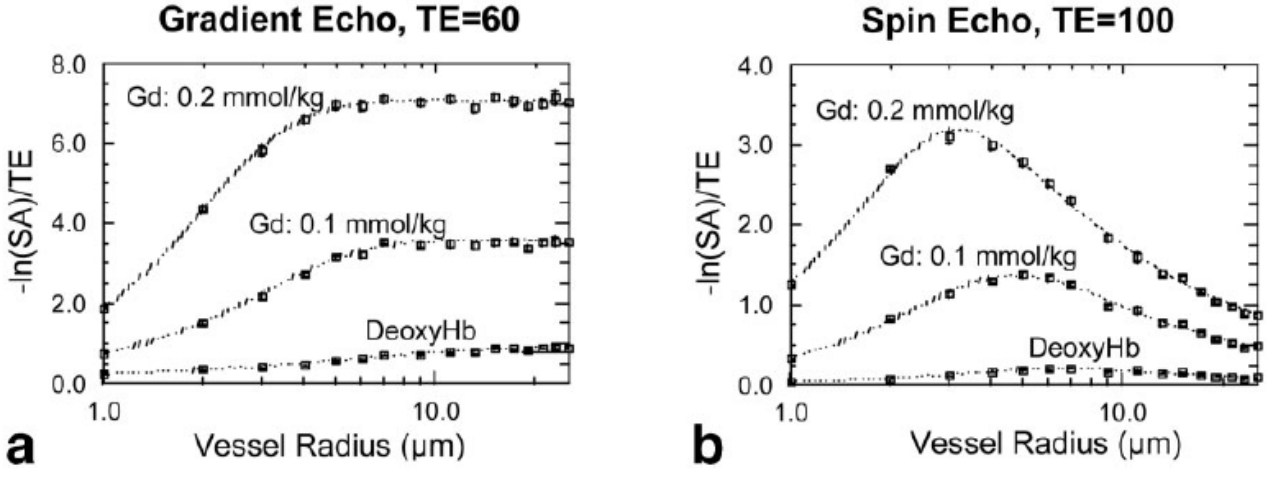

Figure 1. Transverse Relaxation Rates. The change in transverse relaxation rates for SE and GE sequences at different contrast concentration (24).

2.2 Mechanics of Machine Learning in Radiation Oncology

In radiation oncology, the idea of modeling and predicting radiation therapy response has been observed as early as 1944 in the isoeffect curves of Strandquist to the contemporary Quantitative Analysis of Normal Tissue Effects in the Clinic (QUANTEC) toxicity tables (30). Recently, interest in medical modeling for prognostic and therapeutic predictions has increased proportionally due to the increased availability of electronic patient data, images, and reports. The sheer volume of data being collected now has encouraged the use of data mining and machine learning methods to uncover previously hidden patterns and make accurate predictions. Although machine learning is actively used in various applications, such as speech recognition, email filtering (spam or junk email), and targeted advertising (commonly used by Google and Facebook), clinical adaptation has been slow due to the high barrier of understanding the complex models and inherent concerns and considerations about the machine learning method.

The first consideration that needs to be addressed is the type of input data or predictive features. There are many types of features available such as patient age and total dose delivered to the tissue for treatment and are divided mainly into two types: dosimetric and non-dosimetric predictors. Non-dosimetric predictors are the clinical parameters such as age, sex, histopathology, and genomic indicators. Dosimetric predictors are the variable data that can directly relate to the amount or delivery of radiation such as dose volume histograms (DVH) and threshold volumes or doses for specific tissue or area observed in QUANTEC tables. Conventionally, dosimetric parameters have been used to predict normal tissue complication probability (NTCP), like radiation pneumonitis, and tumor control probability (31-33). All of these features are sources for potentially useful information and QUANTEC has accordingly

suggested that to provide a comprehensive model, both dosimetric and non-dosimetric predictors should be used (34).

The second consideration is the selection of features. Although more and more information is available, this does not mean that all of the information collected is useful towards an application. With the large amount of data available, it becomes more and more difficult to select the most salient and informative features. This can be done through some statistical feature reduction method, such as the estimation of feature redundancy (35), an iterative wrapping approach (36), or a feature regularization method such as the least absolute shrinkage and selection operator (LASSO) (37). Feature reduction methods allow the researcher to select and focus on the most important features as well as reduce the computational load of the machine learning process.

The last consideration is the generalizability of the predictions. One of the most common difficulties to a predictive model is overfitting, i.e. the model becomes too specific to the training or input data and can not be generalized to the population in focus. Overfitting can be reduced by increasing the sample size. However, it is difficult to get a sample that is both large and specific enough for the general patient population logistically and statistically. Solving the overfitting bias is especially difficult problem for the validation of the latest techniques, modalities, treatments, and trials that have very small patient populations. A solution to the problem of overfitting, in the case of machine learning is cross-validation. Cross-validation is a method of measuring or seeing how generalizable the sample data is to general data. This is achieved by separating the data available into training and testing sets where the machine learning model is trained with the training, “internal”, data and then validated with the testing, “external”, data. This can be achieved in a few different ways but the two most common cross-validation methods are the k-fold cross validation and the hold-out methods.

Taking these considerations in mind, machine learning methods such as logistic regression, support vector machine (SVM), artificial neural networks, and random forests can be applied. In general, logistic regression is a method that maps a linear combination of predictors to a probability on a sigmoidal logistic function and is used mostly for clinical questions with simple binary answers with few independent predictors (38). SVM are used for binary classification as well but for more complex, non-linearly separable, data (39). SVM transforms and manipulates the data in such a way that so that the data can be separated linearly, a simple case would be to square the distance between data points. Artificial Neural Networks are more complex and weights connections and relationships between features in multiple layers to perform predictions much like the human brain (40). Lastly, the random forest is a set, or ensemble of decision trees where each tree is built to split observations with similar response variables as seen in Figure 2 (41). All of these machine learning methods, and more, are being used in some capacity in radiation oncology for tumor localization, image processing, normal tissue toxicity, and survival analysis.

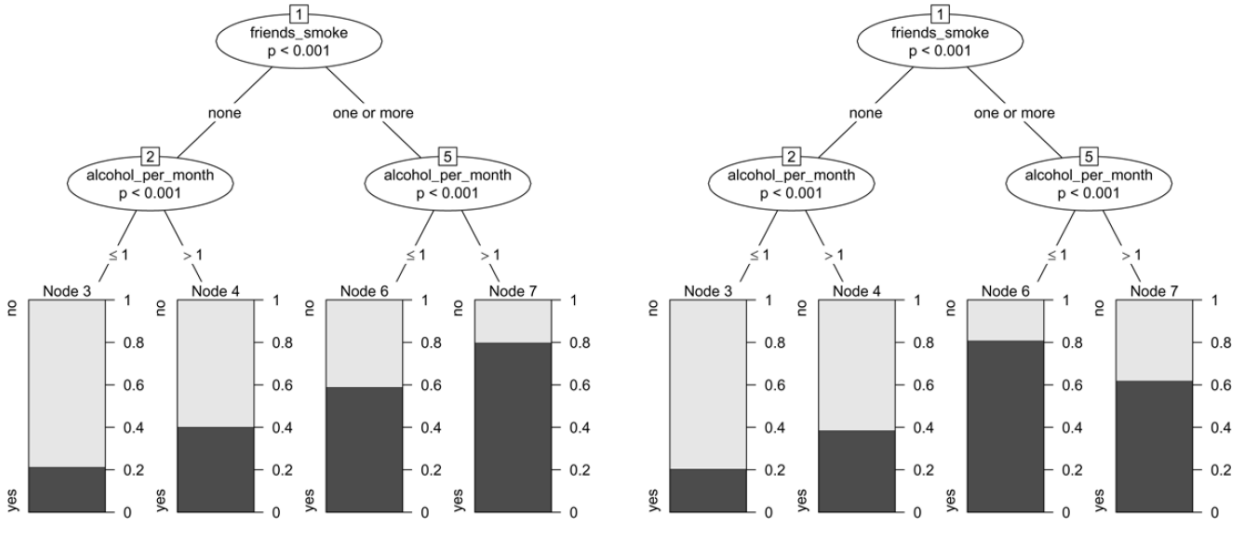

Figure 2. Random Forest Splitting. Two simple random forest decision trees where there isn't an interaction between smoking and alcohol on the left and there is on the right (41).

2.3 Introduction to Radiomics

Radiomics is a nascent field in quantitative imaging that uses advanced algorithms and considerable computing power to describe tumor phenotypes, monitor treatment response, and assess normal tissue toxicity quantifiably. Remarkable interest has been drawn to the field due to its noninvasive nature and potential for accurately diagnosing and predicting patient prognosis. It involves the high-throughput extraction of advanced quantitative features invisible to the naked eye to objectively and quantitatively describe tumor phenotypes in medical imaging. These radiomics features are 'mined' from medical images using advanced algorithms that examine multiple dimensions of the image such as physical (shape and size), textural (the spatial arrangement of voxels), histogram-based, and filtered-based features.

While current treatment decisions are based on a wide variety of traditional diagnostic tests, in recent years, radiomics has gained ground as a method in which one can predict and associate clinical outcomes of patients with cancers like glioblastoma multiforme (GBM) (16), breast cancer (15), renal cell cancer (42), and head and neck cancer (17). Any medical image can be used to derive radiological features and associate them with clinical value like disease diagnosis, progression, and overall survival. Simply put, radiomics is an advanced computational identification, diagnosis, and/or prediction of patient response through medical image features. As such, it has three key components that make up the workflow of radiomics: 1) image, 2) analysis, and 3) validation.

The first component of the workflow is the image. This involves the acquisition of the CT, MR, or PET image and segmentation of the region of interest (ROI). Then radiomic features can be extracted from the segmented ROI through various algorithms that interpret the image data. These feature extraction methods can be grouped as volume based features (like size,

shape, and sphericity), statistics based features (from the histogram-based first order to matrix-based higher order texture analysis, as observed in studies performed on MRI (43) and ultrasound (44)), model based features, and fractal dimensions. Volume based features have been correlated with patient outcome (45) but in recent years, texture analyses has been rising in popularity as they have been shown to predict chemotherapy response in non-small cell lung cancer (NSCLC) (46), differentiate prostate cancer by Gleason score (47), and show prognostic power in GBM (48). These features and combinations of them can be applied to even more detailed clinical applications such as assessing tumor heterogeneity in glioblastomas (13) and tumor phenotype information (14, 49). The last part of the workflow is the modeling and validation of extracted features. These features have to be able to predict patient outcomes like NTCP and tumor control probability (31-33) consistently for it to have clinical value. This requires the use of some statistical or machine learning method to test and validate the predictive power of these features.

CHAPTER 3 – DESIGN OF OUR CLASSIFICATION METHOD

3.1 Patient Data

For our classification method, we obtained our patient data from the Department of Radiology and Imaging Sciences and Winship Cancer Institute at Emory University. The use of the data was approved by the Emory institutional review board (IRB) with written informed consent obtained from the study subjects. Clinical MRI data from twenty-five patients who had biopsy or tumor resection after MRI exams were selected for retrospective analysis. Histo-pathological analysis of these patients revealed that there were 13 high grade and 12 low grade GBMs.

3.2 MRI protocols

All patients underwent MRI exams on a 3T MRI scanner (Magnetom TrioTim; Siemens, Germany) with a routine brain tumor protocol which included the following MR sequences: T1-weighted axial, sagittal, coronal, T2 weighted (axial) fluid attenuated inversion recovery (FLAIR), diffusion weighted imaging (DWI) and DSC MRI. Axial T2 FLAIR images, which were used for segmentation, were obtained using the following parameters: repetition time (TR)/echo time (TE) = 6000/93 ms, flip angle = 130°, inversion time = 2030 ms, slice thickness = 3 mm, field of view (FOV) = 128 × 128 mm2, and 25 axial slices (thickness = 1.875 mm).

DSC MRI data were acquired using single-shot echo planar sequence with repetition time (TR) = 45 ms, echo time (TE) = 2000 ms, measurement = 50-70, FOV: 22cm, matrix, 128 × 128. A bolus of 0.05–0.15mmol/kg gadolinium (Gd) contrast agent was injected at a rate of 3 mL/s at

20 s after starting the dynamic data acquisition. Total scanning time for DSC MRI scan was around 2 min.

3.3 Software

In our method, several software were used to convert, extract, and analyze the patient images. The two software we used the most was Matlab and IBEX (Imaging Biomarker Explorer). Matlab (matrix laboratory) is a proprietary programming language developed by MathWorks and is used worldwide for many applications from plotting data, manipulating matrices, creating graphical user interfaces, and more. IBEX is an open infrastructure software developed at the University of Texas MD Anderson Cancer Center to quickly support the extraction of features in the radiomic workflow. Other software we used were ImageJ, to convert image file formats and image registration, and DicomBrowser, for DICOM (Digital Imaging and Communications in Medicine) metadata, or tag, information processing.

3.4 Image Preparation

Before extracting any features, we outlined a general plan of how we should proceed. Our plan was to use the image with the most contrast between the tumor and the normal tissue and the highest resolution, T2 FLAIR, to contour the tumor. Then we would use these “true” contours to extract radiomic features from the DSC perfusion images. However, many of the T2 FLAIR and DSC perfusion images we obtained from radiology were unfortunately not registered to each other. In order to use contours derived from T2 FLAIR images, we registered all DSC

perfusion volumes to the T2 FLAIR images using ImageJ and the Align3 TP plugin. A sample registration is shown in Figure 3.

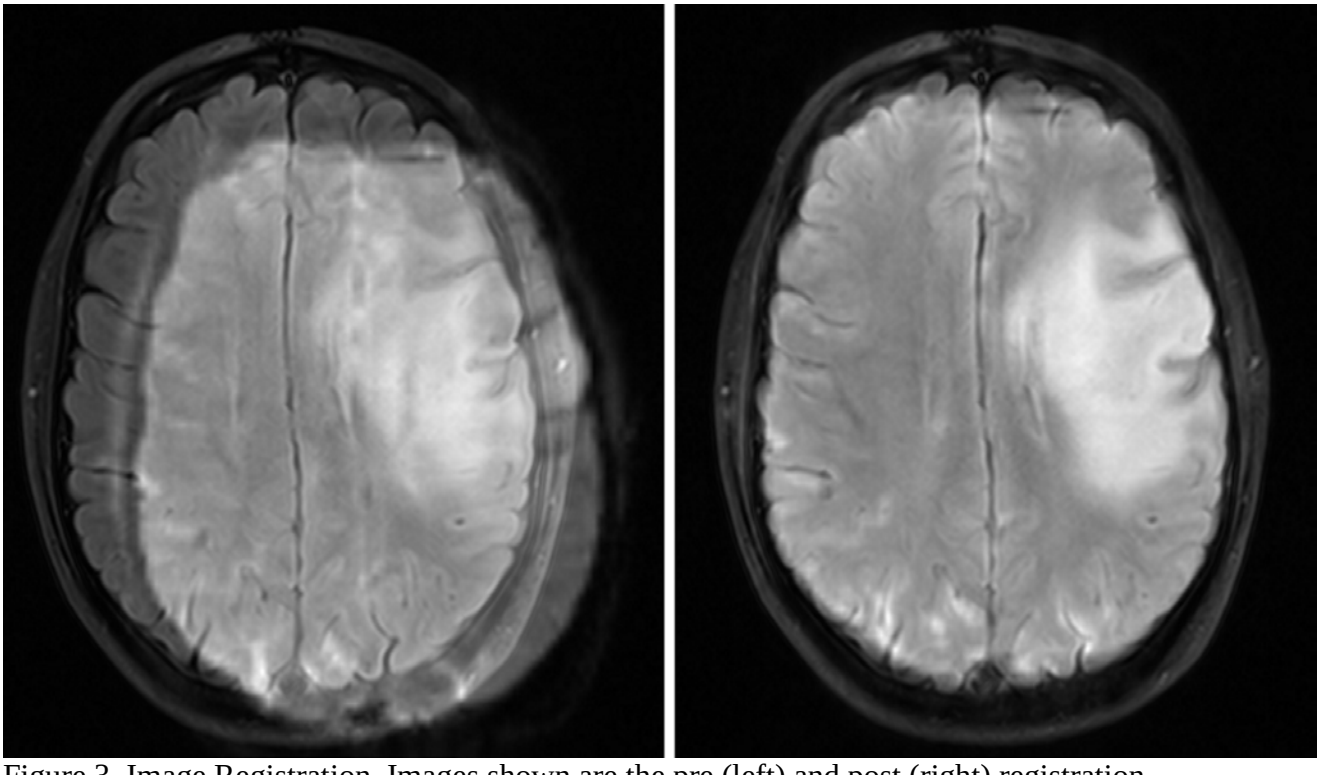

Figure 3. Image Registration. Images shown are the pre (left) and post (right) registration images.

Additionally, the initial batch of images provided to us by radiology was in NIFTI (Neuroimaging Informatics Technology Initiative) format, a format initiated by the National Institute of Mental Health and the National Institute of Neurological Disorders and Stroke. This format is used often in fMRI to make the dissemination of neurological images, informatics tools, and research easier than the previously used Analyze 7.5 format. However, our feature extraction tool, IBEX, was designed to be a more general and centralized radiomic software that uses the two common medical image formats, DICOM and Pinnacle. In order to load these DSC

perfusion images into IBEX, we used ImageJ and the TudorDICOM tools plugin to convert the registered nifti images to individual DICOM image slices. Unfortunately, the conversion process automatically deletes DICOM tags, especially those required for IBEX import like acquisition number, slice number, slice positioning, and etc. This required the use of DicomBrowser and an in-house Matlab code to put back correctly the required DICOM tags to the newly registered DSC images.

3.5 Image Contour

Once all images were registered and able to be loaded into IBEX, we contoured the tumor in it using T2 FLAIR images. For each contour, the window levels were set to have the best visual contrast between the normal brain tissue and tumor and were contoured in IBEX slice by slice. In order to derive delta-radiomic features, we exported the contours into Matlab and flipped the contours across the midline of the brain to obtain normal tissue contours. Once flipped, these contours were translated to best match the anatomical features of the opposing hemisphere's tumor and trimmed down if any part of the contour was outside of the brain, crossed the midline, or dug into the tumor. After both contours were drawn and shaped to satisfaction, it was applied to all 50-70 DSC perfusion volumes for feature extraction. An example of the tumor contour, normal tissue contour, and contour trimming is seen in Figures 4 and 5.

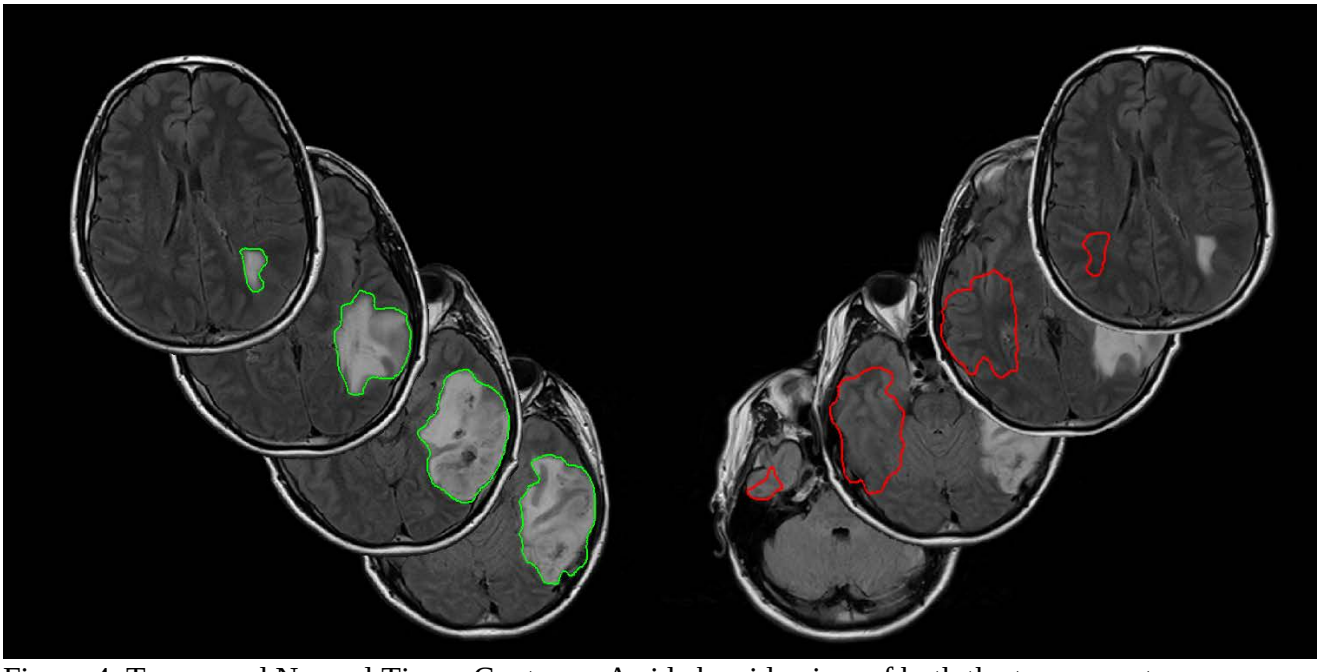

Figure 4. Tumor and Normal Tissue Contours. A side by side view of both the tumor contour (left) and the transformed normal tissue contour (right).

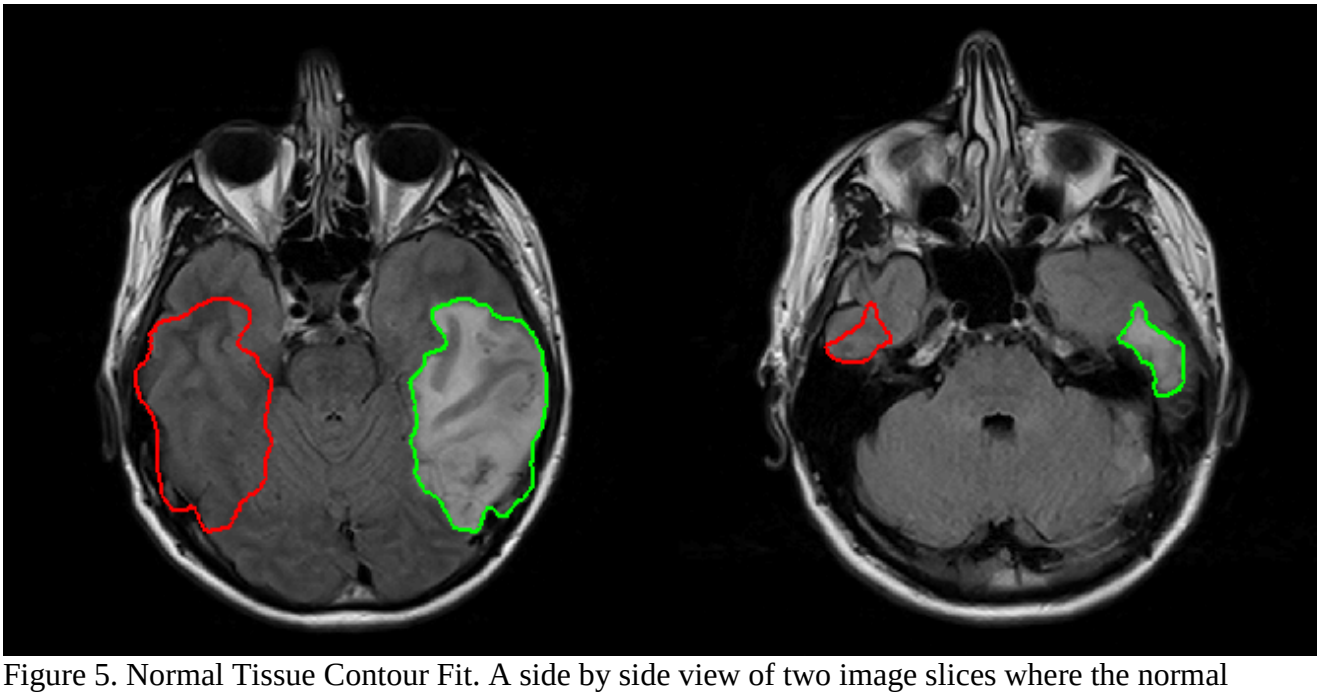

Figure 5. Normal Tissue Contour Fit. A side by side view of two image slices where the normal tissue contour (red) was trimmed to fit the patient anatomy resulting in a smaller contour than the tumor (green).

3.6 Feature Extraction

Due to the nascent nature of radiomic analysis of DSC perfusion images, we elected to extract as many features as possible in IBEX. This included using the feature categories provided in IBEX such as, Shape, Intensity Histogram, Gray Level Cooccurence Matrix, and etc. Some feature categories were excluded due to the inability to extract some of the features with some DSC perfusion images on select patients. This was to ensure that the same features were extracted for all patient DSC volumes, both tumor and normal tissue, to apply our classification method to. As a result, nine feature categories and their corresponding features, as listed in Table 1, were selected and extracted for a total of 1689 features.

Table 1. List of Feature Categories and Names

Feature Category	Feature Names		
Shape	Compactness1	SurfaceArea	MeanBreadth
	Compactness2	SurfaceAreaDensity	Orientation
	Max3DDiameter	Mass	Roundness
	SphericalDisproportion	Convex	NumberOfObjects
	Sphericity	ConvexHullVolume	NumberOfVoxel
	Volume	ConvexHullVolume3D	VoxelSize
IntensityDirect	Energy	InterQuartileRange	MeanAbsoluteDeviation
	RootMeanSquare	GlobalEntropy	MedianAbsoluteDeviation
	Variance	GlobalUniformity	LocalEntropy/Range/StdMax
	Kurtosis	GlobalMax	LocalEntropy/Range/StdMin
	Skewness	GlobalMin	LocalEntropy/Range/StdMean
	Range	GlobalMean	LocalEntropy/Range/StdMedian
	Percentile	GlobalMedian	LocalEntropy/Range/StdStd
	Quantile	GlobalStd	
IntensityHistogram	Kurtosis	InterQuartileRange	Dissimilarity
	Skewness	AutoCorrelation	Entropy
	Range	ClusterProminence	Homogeneity2
	Percentile	ClusterShade	InformationMeasureCorr1
	PercentileArea	CluseterTendency	
	Quantile	DifferenceEntropy	
GradientOrientHistogram	InterQuartileRange	MedianAbsoluteDeviation	Quantile
	Kurtosis	Percentile	Range
	MeanAbsoluteDeviation	PercentileArea	Skewness
GrayLevelCooccurenceMatrix25* GrayLevelCooccurenceMatrix3**	InformationMeasureCorr2	SumAverage	Correlation
	InverseDiffMomentNorm	SumEntropy	Energy
	InverseDiffNorm	SumVariance	Homogeneity
	InverseVariance	Variance	
	MaxProbability	Contrast	
NeighborIntensityDifference25* NeighborIntensityDifference3**	Busyness	Complexity	TextureStrength
	Coarseness	Contrast	
GrayLevelRunLengthMatrix25*	GrayLevelNonuniformity	LongRunLowGrayLevelEmpha	ShortRunEmphasis
	HighGrayLevelRunEmpha	LowGrayLevelRunEmpha	ShortRunHighGrayLevelEmpha
	LongRunEmphasis	RunLengthNonuniformity	ShortRunLowGrayLevelEmpha
	LongRunHighGrayLevelEmpha	RunPercentage	
*25 = Feature calculated from all 2D slices **3 = Feature calculated from the 3D image matrix*			

Once all features were extracted, various feature transformation methods were applied to calculate delta-radiomic features and normalize the data. To extract delta-radiomic features, three methods were used: subtraction (absolute difference) to the baseline, feature change over time, absolute difference and normalization (relative difference) to normal tissue features. As delta-radiomics is a very new concept, we decided to use these subtraction and normalization derived

delta-radiomic features to observe which features optimized our classification method. Additional normalization methods were applied again, to: (a) observe if normalization of features improved our classification method and (b) reduce inter-patient feature variances. A list of the final feature transformation methods used and corresponding descriptions of what they are and how they were derived are shown in Table 2.

Table 2. List of Final Feature Transformation Methods

Method Name:	Method Description:
Method 1	Raw tumor features
Method 2	Delta radiomic features via subtraction of baseline features
Method 3	Delta radiomic features via subtracting normal tissue from tumor features
Method 4	Delta radiomic features via dividing tumor features by normal tissue features
Method 5	Method 3 normalized by subtracting the baseline
Method 6	Method 3 normalized by dividing with the baseline
Method 7	Method 4 normalized by dividing with the baseline

3.7 Feature Selection

A feature selection method was applied once all features were extracted. Feature selection is a commonly applied step among radiomic studies to reduce computation burden and extract the most salient and informative features. The first part of our feature selection was to extract features that showed a significant difference between high and low grade gliomas. We used a combination of LASSO (least absolute shrinkage and selection operator) and two sample T-test to achieve this. LASSO is a commonly used statistical regression analysis that is used to increase the prediction accuracy of statistical models and selects unique and significant

predictors, in our case features. A two sample T-test is a statistical test used to determine if two population means are equal.

In our method, we used the Two Sample T-test (2Samp) to select features that are significantly different at the 0.05 significance level using HG patients and LG patients as the two samples and then used the first pass of LASSO to additionally select and regularize features. After both feature selections are finished, only those features that are selected to be significant by both tests, L2S, were chosen to be used in our method. Table 3 shows the results of our initial predictions using just the three significant-feature selection methods. We can see that although there isn't a clear and significant difference between the three feature selection methods, in four out of the seven methods, the combination of LASSO and Two Sample T-test had the highest average accuracies before removing any redundant features or optimizations to the random forest algorithm.

Table 3. Table of Initial Prediction Accuracies per Method per Feature Selection Method

Method:	Feature selection:	HG Accuracy:	LG Accuracy:	AVG Accuracy:
Method 1	LASSO	0.588180485	0.463721318	0.525950901
	2Samp	0.611237262	0.442014356	**0.526625809**
	L2S	0.64024795	0.411843434	0.526045692
Method 2	LASSO	0.64016413	0.395136249	0.51765019
	2Samp	0.634646299	0.405232623	0.519939461
	L2S	0.645606326	0.422515749	**0.534061037**
Method 3	LASSO	0.730536244	0.527881171	0.629208707
	2Samp	0.709297821	0.520521244	0.614909532
	L2S	0.74628706	0.518226683	**0.632256872**
Method 4	LASSO	0.814984581	0.479001233	0.646992907
	2Samp	0.802242374	0.53103346	**0.666637917**
	L2S	0.804618342	0.474291773	0.639455057
Method 5	LASSO	0.863513936	0.571756815	**0.717635376**
	2Samp	0.847720376	0.513321342	0.680520859
	L2S	0.853592728	0.57425941	0.713926069
Method 6	LASSO	0.871931855	0.506695656	0.689313756
	2Samp	0.875245396	0.488748332	0.681996864
	L2S	0.879281474	0.519995879	**0.699638677**
Method 7	LASSO	0.875763259	0.497065025	0.686414142
	2Samp	0.878690205	0.490281614	0.68448591
	L2S	0.886931796	0.520348719	**0.703640257**

*L2S = LASSO + 2Sample T-Test

Once the most significant features have been selected, redundant features were removed. This was achieved using a correlation-matrix-based selection with variable correlation limits. The correlation-matrix-based selection removes redundant features using a correlation matrix, where any feature correlation matrix-column's column-wise average absolute correlations above a set limit is removed. During this elimination process, any columns with non-finite numbers e.g. infinite and NaN values, were removed automatically. This process removes the significant but redundant features leaving us with non-redundant and significant features. A visual representation of removing redundant features at various correlation limits are shown in Figure 6.

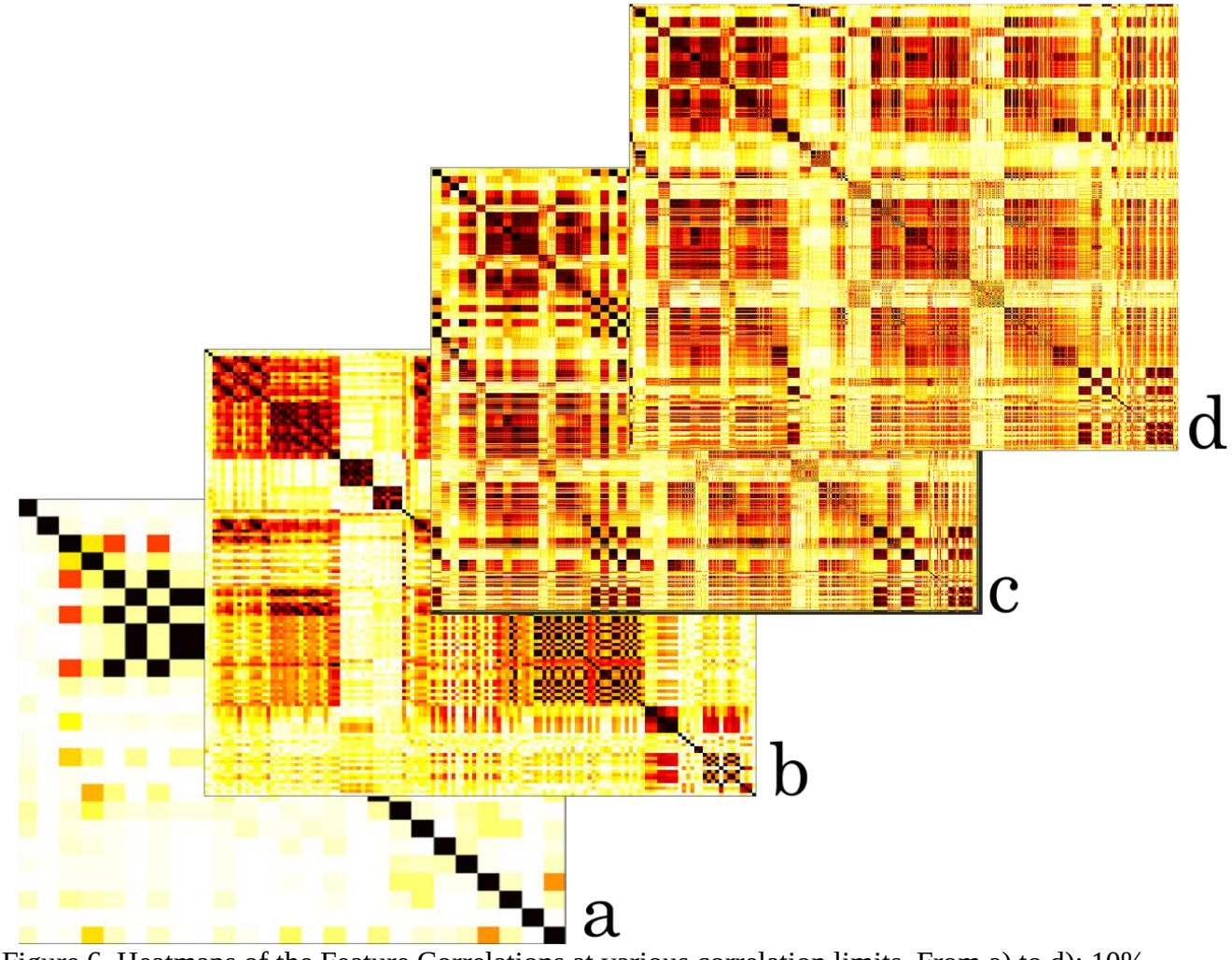

Figure 6. Heatmaps of the Feature Correlations at various correlation limits. From a) to d): 10%, 30%, 50%, and 100% correlation limits.

Finally, an additional positive-feature selection was performed after initial performance evaluations. These positive-features were defined as those features that were added with each increasing correlation limit that positively affected the outcome of overall prediction accuracy. This positive-feature selection was applied with the expectation: 1) that it would improve prediction accuracies by removing features that reduced the accuracy of our method; 2) reduce the computational load by significantly reducing the number of features used to train and test the model; 3) extract and observe radiomic and delta-radiomic signatures, combinations of features, and feature groups that are the most informative in classifying GBMs.

3.8 Random Forest and Performance Evaluation

After feature selection, a random forest machine learning method was used to classify patients into high or low grade gliomas. Specifically, we chose to use the Matlab function, fitensemble(), with a semi-optimized template tree. At each correlation limit, the corresponding features and labels (of high or low grade patients) for each method are supplied to the random forest to train and test it, resulting in various performance metrics such as prediction accuracy (predicted label), prediction confidence (score), and receiver operating characteristic (ROC) curves.

To evaluate the performance of our classification method to the larger population and reduce the overfitting bias, we chose to use a leave-one-out cross validation on the random forest. This involved our method to be trained and tested over multiple iterations, each with a new and distinct test patient. After the leave-one-out cross validation testing, we evaluated the performance of our machine-learning based classification method by calculating the accuracy of our predictions, average confidence of the prediction, and calculating the area under the curve (AUC) of the receiver operating characteristic (ROC) curve while stratifying these values by true high or low grade patients.

Additionally, improve our classification method further, we used a best-majority voting method, using the top four performing methods for each patient to classify the patient as HG or LG, and calculated these performance parameters again to reduce any strong biases from poor performers (50).

3.9 Flowchart of our proposed method

The following figure is the flowchart summary of our proposed classification method:

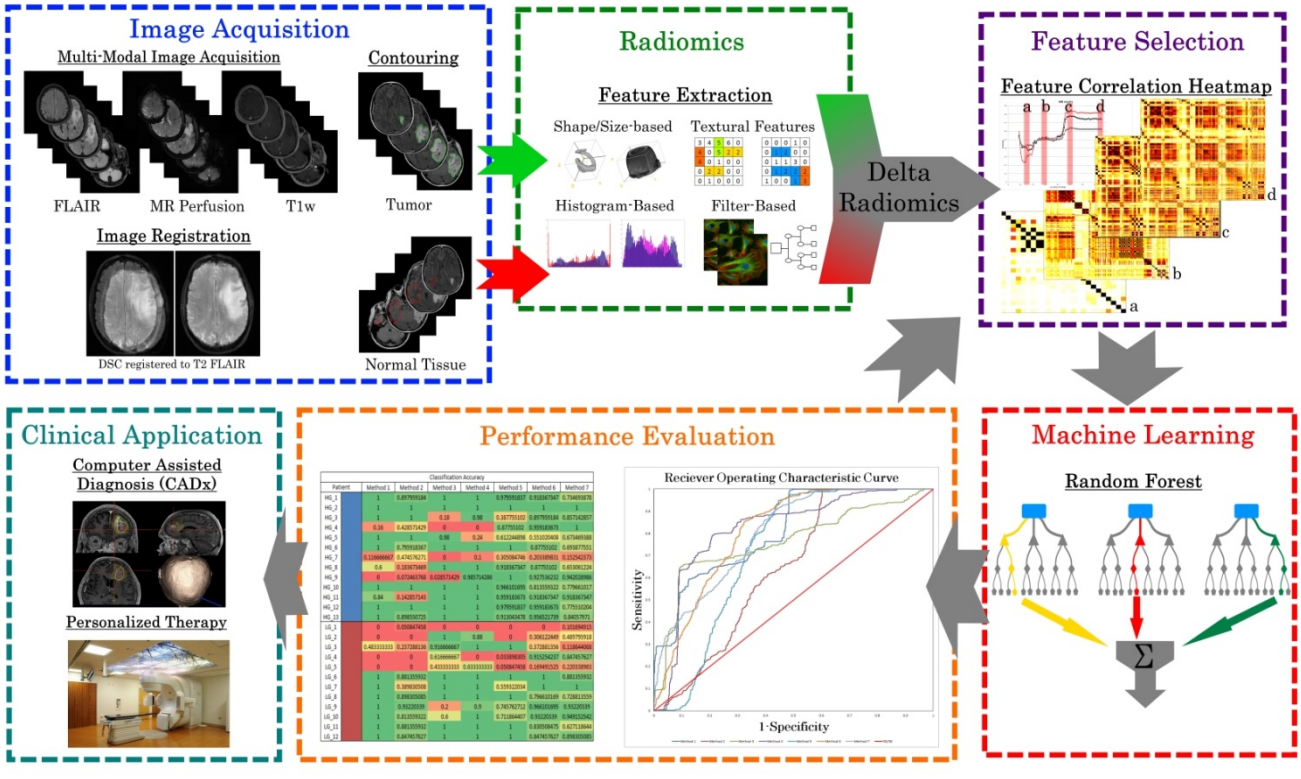

Figure 7. Flowchart of our proposed classification method

In summary, our method involves first acquiring the images and then registering the DSC images to the T2 FLAIR image. Next, contours of the tumor and normal tissue are derived from T2 FLAIR and are used to extract radiomic features. Afterwards, various feature transformation methods are derived such as raw tumor features and absolute difference and relative difference delta-radiomic features. Once all feature transformation methods are derived, feature selection is applied to each method to extract the non-redundant and informative features. These features are

partitioned and used to train and test a random forest with a leave-one-out cross validation. Finally, performance metrics are derived to: (a) evaluate the classification method and (b) to extract positive features that will be applied again to the random forest to evaluate for any improvements. The ultimate goal of our proposed classification method is to apply it in the clinical setting to accurately classify glioblastoma grades and assist the physician in their diagnosis.

CHAPTER 4 – RESULTS

Table 4. Table of Best Results for each Feature Transformation Method

L1O Best							
	Method 1 (71)	Method 2 (24)	Method 3 (522)	Method 4 (87)	Method 5 (679)	Method 6 (159)	Method 7 (526)
Average Accuracy	0.758133333	0.784231122	0.74407619	0.81	0.821176164	0.775928494	0.74123211
HG Accuracy	0.774615385	0.815669434	0.697582418	0.82974359	0.949458725	0.860501607	0.784385247
LG Accuracy	0.740277778	0.750172951	0.794444444	0.788611111	0.68220339	0.684307621	0.694482878
AUC	0.821896449	0.313580922	0.737472724	0.834576473	0.79671371	0.675886834	0.765721286
Confidence	0.658903571	0.66867749	0.621714965	0.681012237	0.658371101	0.653995065	0.593938476

Table 5. Table of Results for each Method using Positive Features

L1O Positive-Features only							
	Method 1 (277)	Method 2 (331)	Method 3 (326)	Method 4 (96)	Method 5 (384)	Method 6 (275)	Method 7 (256)
Average Accuracy	0.688	0.59307095	0.718209524	0.788761905	0.720008422	0.759872869	0.712632608
HG Accuracy	0.747435897	0.68419772	0.706813187	0.792747253	0.838347356	0.835399292	0.770841118
LG Accuracy	0.623611111	0.494350282	0.730555556	0.784444444	0.59180791	0.678052577	0.649573389
AUC	0.810037691	0.64107123	0.762604642	0.822122595	0.725481289	0.811005414	0.802247043
Confidence	0.679771511	0.576209987	0.631756438	0.676562977	0.602979133	0.624922407	0.597808952

Finally, with the use of our best-majority voting method, the best resulting accuracies were 0.9016±0.1911, 0.9498±0.0913, and 0.8494±0.2548 for all, HG, and LG patients respectively with an AUC of 0.9380.

4.1 Result Description

Table 4 and 5 shows the results of our proposed classification method. Both tables show the average classification accuracy, HG accuracy, LG accuracy, the area under the curve (AUC) of the receiver operating characteristic (ROC) curve and the confidence of the predictions. Next to each method, the number of features used for classification for each method is listed.

CHAPTER 5 – DISCUSSION

5.1 Comparison of Results

5.1.1 Results across correlation limits

In comparing the various methods of feature manipulation and correlation limits, we can observe a few trends appearing. For example, the first two feature transformation methods, 1 and 2, show that peak performances occur at lower correlation limits 28.75% and 13.75% respectively while other features show peak performances at higher correlation limits as shown in Figure 8. Additionally, we can see that the delta-radiomic features derived from the difference between tumor and normal tissue, normalized or not, performed better than those that were not.

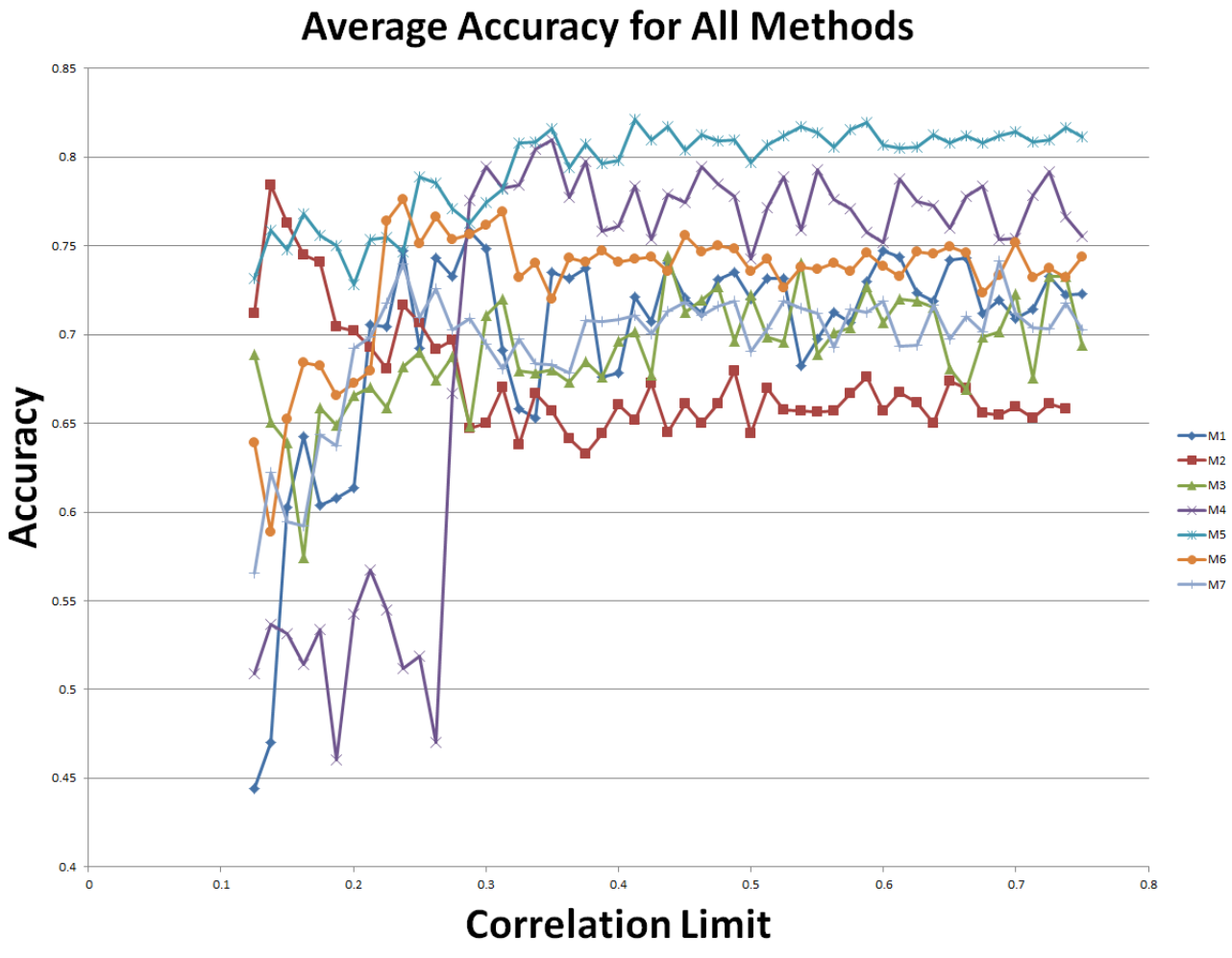

Figure 8. Performance Graph. Graphs of feature transformation method's accuracy across correlation limits

Looking at the peak accuracies for all feature transformation methods, we derived an accuracy heatmap by stratifying the accuracies by patients as shown in Figure 9. One patient, LG patient 1 (LG_1), was consistently performing terribly for all methods at an average accuracy across all methods of 0.068. This leads to two likely conclusions: that either this specific patient was not classified correctly or that our classification method suffers with this specific type of LG patient. While our method takes the glioma classifications provided to us as the ground truth, some cases are more difficult to classify than others without extensive genetic testing as GBMs are very heterogeneous. Genetic characterization of GBMs has largely been categorized into four

subtypes: classic, mesenchymal, neural, and proneural (8) which may make both the histo-pathologic and radiomic classification difficult. Additionally, when inquired about the correct classification of a few poor performing patients earlier in the study, we received a response saying that two patients were misclassified and was actually high grade patients. This justifies and strengthens the credibility of our classification method suggesting that it can highlight suspicious and, in some cases, misclassified patients.

Classification Accuracy							
Patient	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6	Method 7
HG_1	1	0.653061224	1	1	1	0.979591837	0.693877551
HG_2	1	0.775510204	1	1	1	1	1
HG_3	1	0.285714286	0.04	1	1	0.836734694	0.816326531
HG_4	1	0.87755102	0	0	0.918367347	1	0.979591837
HG_5	0.68	0.857142857	1	0.62	0.714285714	0.714285714	0.653061224
HG_6	1	1	1	1	1	0.979591837	0.959183673
HG_7	0.45	1	0	0.166666667	0.86440678	0.457627119	0.050847458
HG_8	0	0.591836735	1	1	0.93877551	0.693877551	0.93877551
HG_9	0	0.985507246	0.028571429	1	0.971014493	0.811594203	0.536231884
HG_10	1	0.728813559	1	1	1	0.966101695	0.898305085
HG_11	1	0.87755102	1	1	0.979591837	0.897959184	0.959183673
HG_12	0.94	1	1	1	1	0.979591837	0.755102041
HG_13	1	0.971014493	1	1	0.956521739	0.869565217	0.956521739
LG_1	0	0.406779661	0	0	0.016949153	0	0.050847458
LG_2	1	0.612244898	1	0.98	0	0.448979592	0.469387755
LG_3	1	1	0.883333333	0.866666667	1	0.559322034	0.305084746
LG_4	0.95	0.169491525	0.416666667	0.266666667	0.508474576	0.898305085	0.796610169
LG_5	0.616666667	0.745762712	0.666666667	0.583333333	0.152542373	0.101694915	0.288135593
LG_6	1	1	1	1	1	0.983050847	0.966101695
LG_7	1	0.728813559	1	1	1	1	0.983050847
LG_8	1	0.949152542	1	1	1	0.813559322	0.847457627
LG_9	0.15	0.542372881	0.6	0.766666667	0.677966102	0.93220339	0.949152542
LG_10	0.166666667	0.847457627	0.966666667	1	0.830508475	0.847457627	0.915254237
LG_11	1	1	1	1	1	0.779661017	0.86440678
LG_12	1	1	1	1	1	0.847457627	0.898305085

Figure 9. Best Accuracy Heatmap. An accuracy heatmap for each method's best performance

Another observation is that normalized and non-normalized delta-radiomic features' HG and LG accuracies are different. For LG patients, the difference between accuracies of delta-radiomic features (method 3 and 4), versus normalized-delta-radiomic features (method 5, 6, and 7), were significantly different 0.7915±0.0029 and 0.6870±0.0054 respectively. While it wasn't as significant, HG patients' accuracies between the two groups were different as well with 0.7637±0.0661 and 0.8648±0.0675 for delta-radiomic and normalized-delta-radiomic features respectively. These results suggest that normalizing delta-radiomic features increase the accuracy of HG patient classifications at the cost of significantly reducing the accuracy of LG patients'.

Looking at Figure 10 and at the AUC values in our results, we can see that the performance of the normalized raw features was very poor at an AUC of 31.36%, worse than pure random, 50% with more false positives than true positives. However, other methods, including raw features, had good performances with an average AUC of 0.7720±0.0540 meaning that the diagnostic ability of our binary classification method is quite good.

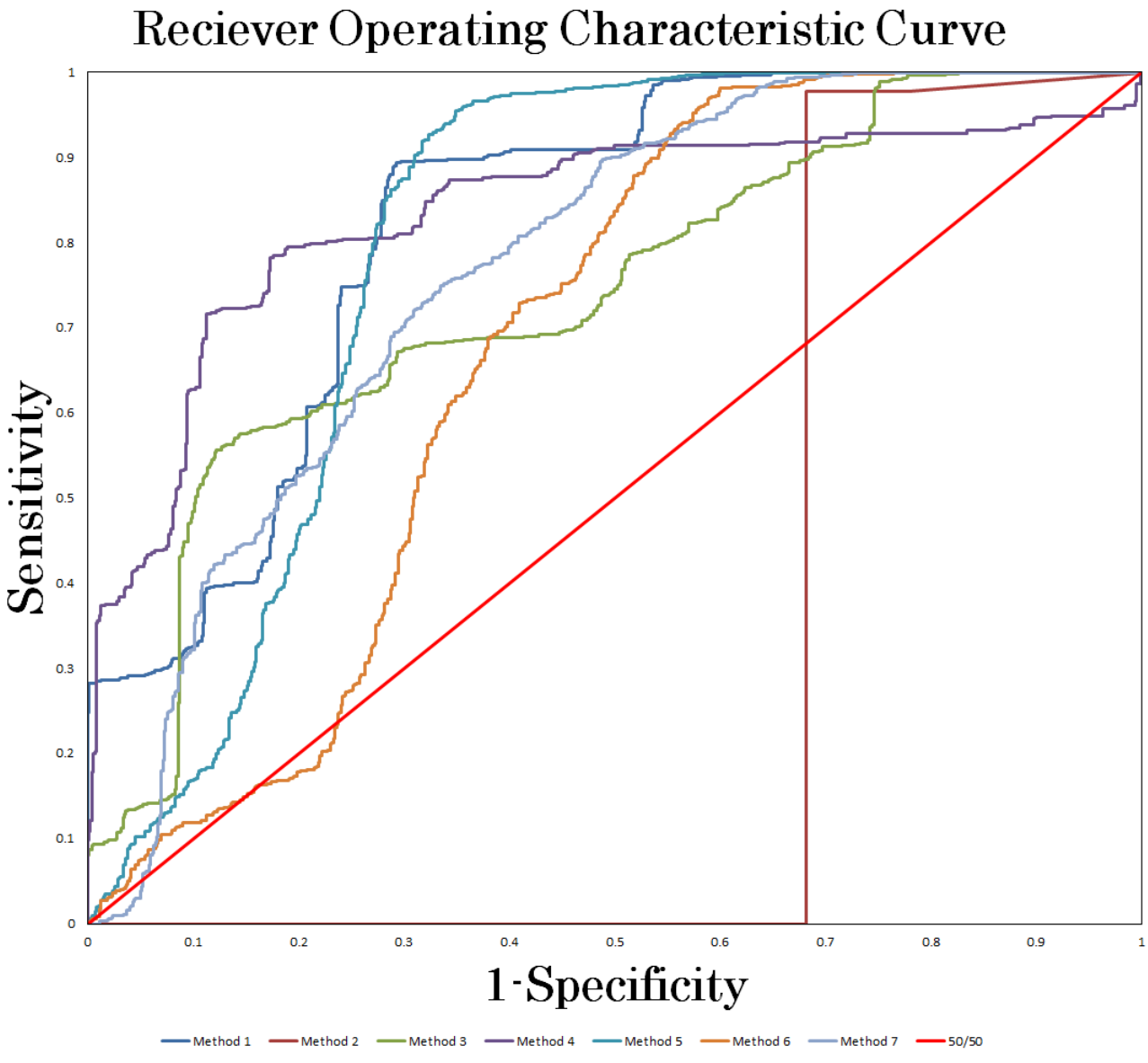

Figure 10. Best ROC Curves. Reciever operating characteristic curves for each method's best performing correlation

5.1.2 Results of using Positive Features

Figure 11 and 12 shows the heatmap and ROC curves respectively for each method using the positive features. As seen on the results, tables 4 and 5, there isn't a significant difference between the peak performing correlations for each feature transformation method and positive

features with a change in average accuracy, HG accuracy, LG accuracy, and AUC of -0.0649±0.0591, -0.0480±0.0485, -0.0832±0.0802, and 0.0612±0.1234 respectively. However, looking at the ROC curves, we can observe that using positive features significantly increased the ROC of Method 2 from 0.3136 to 0.6411. An explanation of the reduction of average accuracy in the positive features is the way positive features were calculated. Positive features were calculated to be any added features that increased the overall accuracy of the classification over a wide correlation limit range. However, the peak performances of many methods were less than the full range of correlation limits tested, resulting in redundant or slightly confounding features being added into the positive feature set as seen in table 4 and 5, reducing the performance slightly.

Classification Accuracy								
Patient		Method 1	Method 2	Method 3	Method 4	Method 5	Method 6	Method 7
HG_1		1	0.897959184	1	1	0.979591837	0.918367347	0.734693878
HG_2		1	1	1	1	1	1	1
HG_3		1	1	0.18	0.98	0.387755102	0.897959184	0.857142857
HG_4		0.16	0.428571429	0	0	0.87755102	0.959183673	1
HG_5		1	1	0.98	0.24	0.612244898	0.551020408	0.673469388
HG_6		1	0.795918367	1	1	1	0.87755102	0.693877551
HG_7		0.116666667	0.474576271	0	0.1	0.305084746	0.203389831	0.152542373
HG_8		0.6	0.183673469	1	1	0.918367347	0.87755102	0.653061224
HG_9		0	0.072463768	0.028571429	0.985714286	1	0.927536232	0.942028986
HG_10		1	1	1	1	0.966101695	0.813559322	0.779661017
HG_11		0.84	0.142857143	1	1	0.959183673	0.918367347	0.918367347
HG_12		1	1	1	1	0.979591837	0.959183673	0.775510204
HG_13		1	0.898550725	1	1	0.913043478	0.956521739	0.84057971
LG_1		0	0.050847458	0	0	0	0	0.101694915
LG_2		0	0	1	0.88	0	0.306122449	0.489795918
LG_3		0.483333333	0.237288136	0.916666667	1	1	0.372881356	0.118644068
LG_4		0	0	0.616666667	0	0.033898305	0.915254237	0.847457627
LG_5		0	0	0.433333333	0.633333333	0.050847458	0.169491525	0.220338983
LG_6		1	0.881355932	1	1	1	1	0.881355932
LG_7		1	0.389830508	1	1	0.559322034	1	1
LG_8		1	0.898305085	1	1	1	0.796610169	0.728813559
LG_9		1	0.93220339	0.2	0.9	0.745762712	0.966101695	0.93220339
LG_10		1	0.813559322	0.6	1	0.711864407	0.93220339	0.949152542
LG_11		1	0.881355932	1	1	1	0.830508475	0.627118644
LG_12		1	0.847457627	1	1	1	0.847457627	0.898305085

Figure 11. Positive Feature Accuracy Heatmap. A heatmap for each method using its positive features

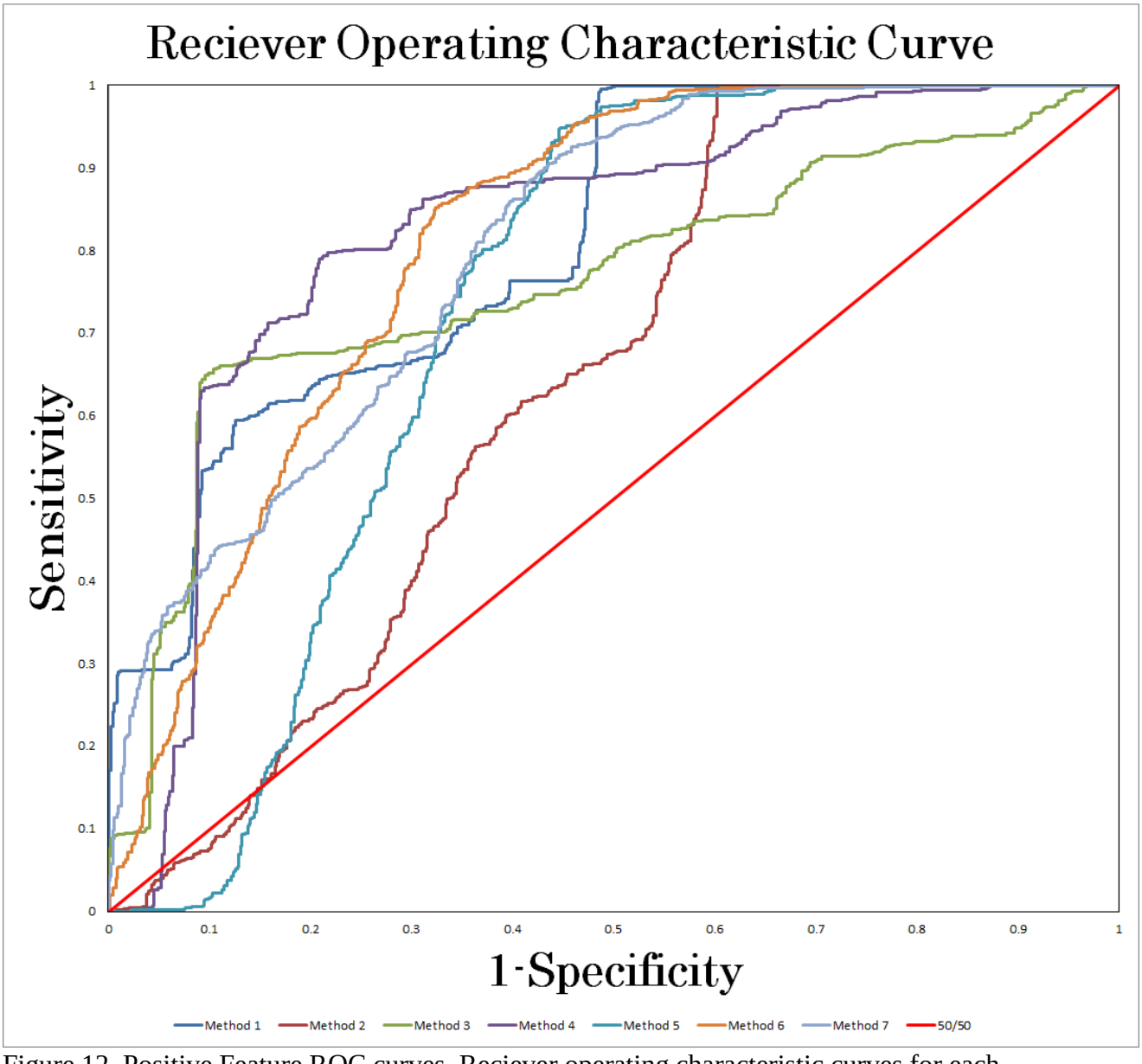

Figure 12. Positive Feature ROC curves. Reciever operating characteristic curves for each method using its positive features

5.2 Comparison of Feature Transformation Methods

Out of the seven feature transformation methods, the method with the best overall accuracy was method 5 (82.12%), a normalized delta-radiomic feature method. However,

considering the discrepancies between HG and LG accuracies, we consider method 4 (81%), a relative delta-radiomic feature method, to be better overall for its relatively high accuracy in both HG and LG patient classification. However, using the best-majority voting method, our proposed method can take advantage of the differences in the seven methods as discussed before to achieve better performances across the board.

When comparing methods and the number of features used for each method, we observed that normalizing features to itself, especially by division (method 2, 4, and 6), significantly reduced the number of features used for our classification method than their non-normalized counterparts. This is observed in method 1 and 2 (71 vs 24 features), method 3 and 4 (522 vs 87), and method 5 and 6 (679 vs 159). For each pair of methods compared, overall performance parameters did not change significantly but reduced the number of features as well as computational load.

5.3 Best feature groups for classification

A list of the most commonly used feature groups in all methods are listed in Table 6. From table 6, we can see that of the 69 features used in at least 4 feature transformation methods, 56 are Gray-Level-Cooccurance-Matrix (GLCM), textural, features (30 derived from 2D slices and 26 from a 3D image matrix). Overwhelmingly, over 80% of the strong and common features used in our methods are textural features as described by Haralick, Shanmugam (51) in 1973 and later confirmed for the use in radiomic applications by Ravanelli, Farina (52) in 2010 for non-small cell lung cancer (NSCLC) and Aerts, Velazquez (14) in 2014 for a more general application. These commonly used features may be the radiomic signature, or finger print, for

GBM classification in DSC and could allow future studies to test with and against these features to develop a robust classification or predictive model.

Table 6. List of Commonly used Features and Feature Groups

Methods Used in:	Feature Groups:	Feature:
6 Methods	GLCM (2D)	InverseVariance (2)
	GLCM (3D)	InverseVariance (1)
5 Methods	GLCM (2D)	InformationMeasureCorr2 (2) InverseVariance (6) MaxProbability (2)
	GLCM (3D)	Correlation (1) InverseVariance (3)
	Intensity Direct	60Percentile (1)
	Intensity Histogram	5PercentileArea (1)
4 Methods	GLCM (2D)	Contrast (2) Correlation (7) InformationMeasureCorr1 (2) InformationMeasureCorr2 (1) InverseDiffMomentNorm (1) InverseVariance (1) MaxProbability (4)
	GLCM (3D)	ClusterProminence (2) Contrast (2) Correlation (6) Homogeneity2 (2) InverseVariance (4) MaxProbability (4) SumEntropy (1)
	GLRLM (2D)	HighGrayLevelRunEmpha (1)
	Intensity Direct	GlobalMax (1) LocalStdMedian (1) 5Percentile (1) Range (1)
	Intensity Histogram	5Percentile (1) 60Percentile (1) Range (1)
	Neighbor Intensity Difference (2D)	TextureStrength (1)
	Neighbor Intensity Difference (3D)	Contrast (1)
*GLCM = GrayLevelCooccurenceMatrix **GLRLM = GrayLevelRunLengthMatrix		

CHAPTER 6 – CONCLUSION

In conclusion, our proposed method performed well, with an average of 90% accuracy in classifying both high and low grade GBMs based on the DSC MRI data. This study shows that delta-radiomic features of DSC MRI, specifically GLCM and other textural features, are highly correlated with GBM disease grades that may further elucidate the underlying tumor biology and response to therapy. It suggests that radiomic analysis and machine learning perform well together and can be generalized to a broader set of diseases and imaging modalities. Our method derived several delta-radiomic features using the difference between the normal and disease tissue as well as the change in features from the baseline. Our method will be tested further as more patients are introduced and clinical outcome information becomes available. With future advances in imaging, machine learning, and radiomic analysis, a more robust, accurate and validated classification or predictive model will be widely available to assist future oncologists deliver better and more personalized treatments.

REFERENCES

1. Surawicz TS, McCarthy BJ, Kupelian V, Jukich PJ, Bruner JM, Davis FG. Descriptive epidemiology of primary brain and CNS tumors: results from the Central Brain Tumor Registry of the United States, 1990-1994. Neuro Oncol. 1999;1(1):14-25.
2. Shapiro WR, Green SB, Burger PC, Mahaley MS, Jr., Selker RG, VanGilder JC, et al. Randomized trial of three chemotherapy regimens and two radiotherapy regimens and two radiotherapy regimens in postoperative treatment of malignant glioma. Brain Tumor Cooperative Group Trial 8001. J Neurosurg. 1989;71(1):1-9.
3. Narang S, Lehrer M, Yang D, Lee J, Rao A. Radiomics in glioblastoma: current status, challenges and potential opportunities. Translational Cancer Research. 2016;5(4):383-97.
4. Weber RG, Sabel M, Reifenberger J, Sommer C, Oberstrass J, Reifenberger G, et al. Characterization of genomic alterations associated with glioma progression by comparative genomic hybridization. Oncogene. 1996;13(5):983-94.
5. Brennan CW, Verhaak RG, McKenna A, Campos B, Noushmehr H, Salama SR, et al. The somatic genomic landscape of glioblastoma. Cell. 2013;155(2):462-77.
6. Nobusawa S, Watanabe T, Kleihues P, Ohgaki H. IDH1 mutations as molecular signature and predictive factor of secondary glioblastomas. Clin Cancer Res. 2009;15(19):6002-7.
7. Hartmann C, Hentschel B, Wick W, Capper D, Felsberg J, Simon M, et al. Patients with IDH1 wild type anaplastic astrocytomas exhibit worse prognosis than IDH1-mutated glioblastomas, and IDH1 mutation status accounts for the unfavorable prognostic effect of higher age: implications for classification of gliomas. Acta Neuropathol. 2010;120(6):707-18.
8. Verhaak RG, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell. 2010;17(1):98-110.
9. Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science. 2014;344(6190):1396-401.
10. Parker NR, Khong P, Parkinson JF, Howell VM, Wheeler HR. Molecular heterogeneity in glioblastoma: potential clinical implications. Front Oncol. 2015;5:55.
11. Oberheim Bush NA, Chang S. Treatment Strategies for Low-Grade Glioma in Adults. J Oncol Pract. 2016;12(12):1235-41.
12. Sottoriva A, Spiteri I, Piccirillo SG, Touloumis A, Collins VP, Marioni JC, et al. Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. Proc Natl Acad Sci U S A. 2013;110(10):4009-14.
13. Moton S, Elbanan M, Zinn PO, Colen RR. Imaging Genomics of Glioblastoma: Biology, Biomarkers, and Breakthroughs. Top Magn Reson Imaging. 2015;24(3):155-63.
14. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun. 2014;5:4006.
15. Yamamoto S, Maki DD, Korn RL, Kuo MD. Radiogenomic analysis of breast cancer using MRI: a preliminary study to define the landscape. AJR Am J Roentgenol. 2012;199(3):654-63.
16. Lee J, Narang S, Martinez JJ, Rao G, Rao A. Associating spatial diversity features of radiologically defined tumor habitats with epidermal growth factor receptor driver status and 12-month survival in glioblastoma: methods and preliminary investigation. J Med Imaging (Bellingham). 2015;2(4):041006.

17. Hatt M, Majdoub M, Vallieres M, Tixier F, Le Rest CC, Groheux D, et al. 18F-FDG PET uptake characterization through texture analysis: investigating the complementary nature of heterogeneity and functional tumor volume in a multi-cancer site patient cohort. J Nucl Med. 2015;56(1):38-44.
18. Caravan P, Ellison JJ, McMurry TJ, Lauffer RB. Gadolinium(III) Chelates as MRI Contrast Agents: Structure, Dynamics, and Applications. Chem Rev. 1999;99(9):2293-352.
19. Essig M, Lodemann KP, Le-Huu M, Bruning R, Kirchin M, Reith W. Intraindividual comparison of gadobenate dimeglumine and gadobutrol for cerebral magnetic resonance perfusion imaging at 1.5 T. Invest Radiol. 2006;41(3):256-63.
20. Essig M, Anzalone N, Combs SE, Dorfler A, Lee SK, Picozzi P, et al. MR imaging of neoplastic central nervous system lesions: review and recommendations for current practice. AJNR Am J Neuroradiol. 2012;33(5):803-17.
21. Lacerda S, Law M. Magnetic resonance perfusion and permeability imaging in brain tumors. Neuroimaging Clin N Am. 2009;19(4):527-57.
22. Shiroishi MS, Habibi M, Rajderkar D, Yurko C, Go JL, Lerner A, et al. Perfusion and permeability MR imaging of gliomas. Technol Cancer Res Treat. 2011;10(1):59-71.
23. Villringer A, Rosen BR, Belliveau JW, Ackerman JL, Lauffer RB, Buxton RB, et al. Dynamic imaging with lanthanide chelates in normal brain: contrast due to magnetic susceptibility effects. Magn Reson Med. 1988;6(2):164-74.
24. Ostergaard L. Principles of cerebral perfusion imaging by bolus tracking. J Magn Reson Imaging. 2005;22(6):710-7.
25. Weisskoff RM, Zuo CS, Boxerman JL, Rosen BR. Microscopic susceptibility variation and transverse relaxation: theory and experiment. Magn Reson Med. 1994;31(6):601-10.
26. Boxerman JL, Hamberg LM, Rosen BR, Weisskoff RM. MR contrast due to intravascular magnetic susceptibility perturbations. Magn Reson Med. 1995;34(4):555-66.
27. Fisel CR, Ackerman JL, Buxton RB, Garrido L, Belliveau JW, Rosen BR, et al. MR contrast due to microscopically heterogeneous magnetic susceptibility: numerical simulations and applications to cerebral physiology. Magn Reson Med. 1991;17(2):336-47.
28. Ostergaard L, Smith DF, Vestergaard-Poulsen P, Hansen SB, Gee AD, Gjedde A, et al. Absolute cerebral blood flow and blood volume measured by magnetic resonance imaging bolus tracking: comparison with positron emission tomography values. J Cereb Blood Flow Metab. 1998;18(4):425-32.
29. Simonsen CZ, Ostergaard L, Vestergaard-Poulsen P, Rohl L, Bjornerud A, Gyldensted C. CBF and CBV measurements by USPIO bolus tracking: reproducibility and comparison with Gd-based values. J Magn Reson Imaging. 1999;9(2):342-7.
30. Kang J, Schwartz R, Flickinger J, Beriwal S. Machine Learning Approaches for Predicting Radiation Therapy Outcomes: A Clinician's Perspective. Int J Radiat Oncol Biol Phys. 2015;93(5):1127-35.
31. Burman C, Kutcher GJ, Emami B, Goitein M. Fitting of normal tissue tolerance data to an analytic function. Int J Radiat Oncol Biol Phys. 1991;21(1):123-35.
32. Seppenwoolde Y, Lebesque JV, de Jaeger K, Belderbos JS, Boersma LJ, Schilstra C, et al. Comparing different NTCP models that predict the incidence of radiation pneumonitis. Normal tissue complication probability. Int J Radiat Oncol Biol Phys. 2003;55(3):724-35.
33. Willner J, Baier K, Caragiani E, Tschammler A, Flentje M. Dose, volume, and tumor control prediction in primary radiotherapy of non-small-cell lung cancer. Int J Radiat Oncol Biol Phys. 2002;52(2):382-9.

34. Marks LB, Yorke ED, Jackson A, Ten Haken RK, Constine LS, Eisbruch A, et al. Use of normal tissue complication probability models in the clinic. Int J Radiat Oncol Biol Phys. 2010;76(3 Suppl):S10-9.
35. Shannon CE. The mathematical theory of communication. 1963. MD Comput. 1997;14(4):306-17.
36. Kohavi R, John GH. Wrappers for feature subset selection. Artif Intell. 1997;97(1-2):273-324.
37. Tibshirani R. Regression shrinkage and selection via the Lasso. J Roy Stat Soc B Met. 1996;58(1):267-88.
38. Janssens GK, van Dam PA, van Goethem ML, Buytaert P. A logit model to evaluate the performance of diagnosis of solid palpable breast tumours. Eur J Obstet Gynecol Reprod Biol. 1990;34(1-2):157-65.
39. Noble WS. What is a support vector machine? Nat Biotechnol. 2006;24(12):1565-7.
40. Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain. Psychol Rev. 1958;65(6):386-408.
41. Strobl C, Malley J, Tutz G. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. Psychol Methods. 2009;14(4):323-48.
42. Shinagare AB, Vikram R, Jaffe C, Akin O, Kirby J, Huang E, et al. Radiogenomics of clear cell renal cell carcinoma: preliminary findings of The Cancer Genome Atlas-Renal Cell Carcinoma (TCGA-RCC) Imaging Research Group. Abdom Imaging. 2015;40(6):1684-92.
43. Bernasconi A, Antel SB, Collins DL, Bernasconi N, Olivier A, Dubeau F, et al. Texture analysis and morphological processing of magnetic resonance imaging assist detection of focal cortical dysplasia in extra-temporal partial epilepsy. Ann Neurol. 2001;49(6):770-5.
44. Chen DR, Chang RF, Kuo WJ, Chen MC, Huang YL. Diagnosis of breast tumors with sonographic texture analysis using wavelet transform and neural networks. Ultrasound Med Biol. 2002;28(10):1301-10.
45. Cannon GB, Pomerantz R. Cell-mediated immune responses--prognostic indicators of survival from breast cancer. Int J Cancer. 1989;44(6):995-9.
46. Cook GJ, Yip C, Siddique M, Goh V, Chicklore S, Roy A, et al. Are pretreatment 18F-FDG PET tumor textural features in non-small cell lung cancer associated with response and survival after chemoradiotherapy? J Nucl Med. 2013;54(1):19-26.
47. Wibmer A, Hricak H, Gondo T, Matsumoto K, Veeraraghavan H, Fehr D, et al. Haralick texture analysis of prostate MRI: utility for differentiating non-cancerous prostate from prostate cancer and differentiating prostate cancers with different Gleason scores. Eur Radiol. 2015;25(10):2840-50.
48. Cui Y, Tha KK, Terasaka S, Yamaguchi S, Wang J, Kudo K, et al. Prognostic Imaging Biomarkers in Glioblastoma: Development and Independent Validation on the Basis of Multiregion and Quantitative Analysis of MR Images. Radiology. 2016;278(2):546-53.
49. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer. 2012;48(4):441-6.
50. Dey A, Shaikh SH, Saeed K, Chaki N, editors. Modified Majority Voting Algorithm towards Creating Reference Image for Binarization2014; Cham: Springer International Publishing.
51. Haralick RM, Shanmugam K, Dinstein I. Textural Features for Image Classification. IEEE Transactions on Systems, Man, and Cybernetics. 1973;SMC-3(6):610-21.

52. Ravanelli M, Farina D, Morassi M, Roca E, Cavalleri G, Tassi G, et al. Texture analysis of advanced non-small cell lung cancer (NSCLC) on contrast-enhanced computed tomography: prediction of the response to the first-line chemotherapy. Eur Radiol. 2013;23(12):3450-5.