Classification of Phosphorus Magnetic Resonance Spectroscopic Imaging of Brain Tumors Using Support Vector Machine and Logistic Regression at 3T

Fusun Citak Er*, *IEEE Member*, Gokce Hale Hatay*, Emre Okeer, Muhammed Yildirim, Bahattin Hakyemez, and Esin Ozturk-Isik, *IEEE Member*

***Abstract*— This study aims classification of phosphorus magnetic resonance spectroscopic imaging (^{31}P-MRSI) data of human brain tumors using machine-learning algorithms. The metabolite peak intensities and ratios were estimated for brain tumor and healthy ^{31}P MR spectra acquired at 3T. The spectra were classified based on metabolite characteristics using logistic regression and support vector machine. This study showed that machine learning could be successfully applied for classification of ^{31}P-MR spectra of brain tumors. Future studies will measure the performance of classification algorithms for ^{31}P-MRSI of brain tumors in a larger patient cohort.**

I. Introduction

Phosphorus magnetic resonance spectroscopic imaging (^{31}P-MRSI) is an MR imaging modality that detects signals coming from phosphorus containing metabolites and it can provide information regarding the energy metabolism, oxygen state and pH within a given region of interest. Phosphocreatine (PCr), phosphocholine (PC), phosphoethanolamine (PE), glycerophosphocholine (GPC), glycerophosphoethanolamine (GPE), inorganic phosphate (Pi), and three different peaks for the ATP molecule, γ-ATP, α-ATP, and β-ATP peaks, are the major metabolites observed with ^{31}P-MRS. Phosphocreatine (PCr) is a marker of phosphorilative energy metabolism and is considered as the reference peak in ^{31}P spectra. PC and PE are the phosphomonoesters (PME) that are produced from choline and ethanolamine phosphorylation during membrane synthesis. GPC and GPE are the phosphodiesters (PDE) that are formed during membrane degradation. β-ATP peak is used to estimate the ATP level, because this peak does not experience contamination from AMP and ADP. In addition to the metabolic information, ^{31}P MRS can also be used to estimate the intracellular pH level using the frequency difference between the inorganic phosphate and PCr peaks [1].

Several studies have reported ^{31}P-MRSI peak differences between brain tumors and healthy tissue. A reduction of PCr/β-ATP ratio [2] and an increase in Pi/PCr ratio [1] has been reported in ischemic tissue. Additionally, a 20-70 percent decrease in metabolite intensities, and a significant increase in pH levels have been reported in brain tumors imaged at 1.5T [1]. Maintz et al. similarly reported an alkaline environment (pH=7.16) and a decrease in PCr and PDE peaks, and in PDE/α-ATP ratio in meningiomas [2]. In patients diagnosed with a grade II or grade III gliomas, a slight alkalization (pH = 7.09) and more than a twofold reduction in the PDE/α-ATP ratio was observed [2]. Moreover, a recent study reported an increased ratio of phosphocholine to glycerophosphocholine (PC/GPC) with increasing glioma grade [3]. Additionally, an increase in PC/GPC ratio has been reported to indicate tumor progression after antiangiogenic therapy and has been associated with shorter overall survival [4]. An elevation in the ratio of PE/GPE in normal appearing tissue was also considered as indicative of tumor infiltration [4].

There have been several studies aiming to classify brain tumors based on MR spectroscopic imaging using machine-learning algorithms. Support vector machine, logistic regression, and linear discriminant analysis algorithms have been successfully implemented for classifying proton high-resolution magic angle spinning spectroscopy (^{1}H HRMAS) of recurrent low-grade gliomas in terms of their malignancy transformation [5]. Logistic regression analysis was performed to identify the relationship between choline over N-acetyl aspartate (Cho/NAA), histological parameters, and the degree of tumor infiltration, and it was shown that high-grade and low-grade gliomas exhibited different spectroscopic patterns [6]. A sensitivity of 93.8% and a specificity of 85.7% were reported for discriminating progressive tumors from non-progressive tumors based on ln(Cho/Cr) and ln(Cho/NAA) using logistic regression [7]. Logistic regression analysis was also employed to differentiate tissue characteristics (spectroscopically normal, pure tumor, mixed tumor, and radiation necrosis) based on ^{1}H-MRSI parameters [8]. Support vector machines were used to discriminate gliomas and meningiomas from healthy

* These authors contributed equally to this work, and they are co-first authors.

This research was supported by TUBITAK Career Development Grant 112E036 and Marie Curie International Reintegration Grant (IRG) FP7-PEOPLE-RG-2009 256528.

F. Citak Er is with the Department of Genetics and Bioengineering, Yeditepe University, Istanbul, 34755 TURKEY. (e-mail: fusun.er@gmail.com).

G.H. Hatay is with the Department of Electrical and Electronics Engineering, Yeditepe University, Istanbul, 34755 TURKEY. (e-mail: gokcehale.hatay@std.yeditepe.edu.tr).

E. Okeer and B. Hakyemez are with the Department of Radiology, Uludag University, Bursa, 16120 TURKEY. (e-mail: emreokeer@gmail.com, bhakyemez@uludag.edu.tr).

M. Yildirim is with the MR Development, Advanced Diagnostic Imaging, Philips Healthcare, Best, and Biomedical NMR, Eindhoven University of Technology, Eindhoven, THE NETHERLANDS. (e-mail: muhammed.yildirim@philips.com).

E. Ozturk-Isik is with the Biomedical Engineering Department, Yeditepe University, Istanbul, 34755 TURKEY. (corresponding author, phone: (+90)216-578-0000; fax: (+90)216-578-0400; e-mail: esin.ozturk@yeditepe.edu.tr).

tissue based on ^{1}H-MRSI metabolic markers [9]. Devos et al. compared the performances of linear discriminant analysis, and least squares support vector machines (LS-SVM) with a linear kernel and with a radial basis function kernel based on MRSI data to differentiate low-grade versus high-grade tumors, low-grade versus high-grade gliomas and gliomas versus meningiomas with area under the curve (AUC) results of higher than 0.91, 0.94, and 0.99, respectively [10].

In this study, we aim to classify brain tumors based on phosphorus MR spectroscopic imaging acquired at a high field strength of 3T using logistic regression, and support vector machine with a linear and non-linear kernel, which has not been previously reported.

II. Materials and Methods

A. Subjects

Three healthy volunteers (average age = 39.7±12.9 years) and 10 brain tumor patients (2 NHL, 2 GBM, 1 oligodendroglioma grade 1, 4 grade 2, 1 metastasis, 1 anaplastic oligodendroglioma, average age=48.9±11.8 years), who provided informed consent in accordance with the Ethics Review Board regulations of our institute, were scanned on a 3T clinical MR scanner (Philips Medical Systems, Best, Netherlands), equipped with a dual channel ^{31}P/^{1}H quadrature head coil. A T1 weighted fast field echo (FFE) or a T2 weighted spin echo (SE) MR images were acquired and used as the anatomical reference for subsequent spectroscopic data acquisition. Two-dimensional ^{31}P-MRSI datasets were acquired with a pulse and acquire (PA) sequence (TR=4.8s, NSA=4, 3000Hz, 1024 points, FOV=160x160mm, 20x20x40mm voxel size, scan time=16 min). There was a timing delay of 1.4072 ms between the RF pulse and data acquisition, which resulted in a first-order phase error. MR spectroscopic data were processed using AMARES within jMRUI software [11]. Two consecutive 10Hz Lorentzian filters were used to apodize the signal, and first order phase error was corrected. jMRUI was used to overlay the anatomical MR images with the spectra, and 42 healthy and 44 tumor voxels were determined from all datasets. After pre-processing in jMRUI, spectra were read in MATLAB (The Matworks Inc., Natick, MA) with in-house MR spectra reader scripts. The frequency ranges of all the ^{31}P metabolites were determined, and the peak heights and ratios of these metabolites were calculated for each voxel. The metabolite intensities were normalized with the PCr peak height of the same voxel. Logistic regression, and support vector machine with a linear and non-linear kernel were used to classify tumor voxels based on their phosphorus MR spectroscopic imaging characteristics.

B. Logistic Regression

Logistic regression fits a probabilistic model or a hypothesis function h, to a data set, S, defined as,

$$S = \{(x_i, y_i)\}_{i=1,n}, x_i \in R^f, \; y_i \in \{0, +1\}, and \quad (1)$$

$$h_\theta(x) = g(\theta^T x + \theta_0), \quad (2)$$

where, f is the number of features, n is the number of samples, x are feature vectors of size f, and y are the outputs of feature vectors x [12]. The hypothesis function, h, calculates the probability of y being a +1 for a given input vector x ($0 \leq h_\theta(x) \leq 1$).

The sigmoid function was used as the transfer function, g, denoted as,

$$g(\theta^T x + \theta_0) = \frac{1}{1+e^{-(\theta^T x + \theta_0)}}. \quad (3)$$

The possible values for y were set as 0 for healthy and +1for tumor voxels. A cut-off value, b, should be determined to use the hypothesis function as a two-valued discrete function for class membership prediction purposes. Then, the hypothesis function, h, can be re-written as,

$$h(x) = \begin{cases} 0, & g(\theta^T x + \theta_0) < b \\ +1, & g(\theta^T x + \theta_0) \geq b \end{cases} \quad (4)$$

In this study, Statistics Toolbox 7.1 (R2009a) of MATLAB (The Mathworks Inc., Natick, MA) was used to calculate the optimal coefficients vector θ. The optimal cutoff point, b, was found using receiver operating characteristic (ROC) curve in MATLAB.

C. Support Vector Machine

Support vector machine (SVM) is a machine learning method, which aims to find a hyperplane, h, that would be used for predicting the class membership of a new dataset [12]. Finding the optimal hyperplane is an optimization problem, such that for each member of training data, $\{(x_i, y_i)\}_{i=1,n}$, the margin between the groups given as,

$$margin = \frac{2}{||w||}, \quad (5)$$

is maximized subject to,

$$y_i * (w.x_i - b) \geq 1, \quad (6)$$

where, w is the normal vector of the optimal separating hyperplane, x are feature vectors of samples, b is a scalar representing the distance of the normal vector to the origin, and y denotes the class of the feature vector, x. In this study, the possible classes were labeled as -1 for healthy and +1 for tumor voxels.

If the training data is not linearly separable, a non-linear kernel can be used to project the data to a high-dimensional space [13]. In this study, support vector machine classification was used with a linear or a polynomial kernel in MATLAB. Three-order polynomial kernel was defined as,

$$K(x, y) = (x^T.y + c)^d, \quad (7)$$

where, d was the degree of the polynomial function and c was a scalar. The optimum value for the box constraint c for the soft margin was calculated by employing a fit function that minimizes misclassification rate using four-fold cross validation.

D. Support Vector Machine-Based Recursive Feature Elimination

Recursive feature elimination (RFE) enhances the performance of the classifiers when the size of the features is high. RFE is an iterative method, where the feature having the lowest score based on a classifier model is eliminated with each iteration, and the iterations are stopped when more

efficient features are left. In this study, support vector machine with linear kernel was used to calculate the scores for the features. Lagrange multipliers, α, of the support vectors were used to calculate the scores of a feature j, as,

$$r_j = \left(\sum_{i=1}^{k} \alpha_i y_i x_{i,j}\right)^2, \quad (8)$$

where, k is the number of support vectors, and $x_{i,j}$ is the value of feature, j, for sample, i. The elimination procedure was repeated five times [14].

E. Performance Comparison of Classifiers

Support vector machine with a linear kernel (SVM_LN), support vector machine with a polynomial kernel (SVM_PL), and logistic regression (LR) were used to distinguish normal and tumor voxels based on ^{31}P-MRSI data. Four-fold cross validation was used to measure their relative performances. Cross validation gives different results at each trial due to random partitioning of data. To reduce variability, forty rounds of cross-validation were performed using different partitions, and the validation results were averaged. The ninety percent confidence intervals of accuracy were also computed.

III. Results and Discussion

Figure 1 shows an example ^{31}P-MRSI data acquired from a thirty-seven years old female patient diagnosed with a grade II oligodendroglioma. Tumor was hyperintense in the T2 weighted spin echo image. There were lower PCr, and higher PME and PDE peaks within the tumor region.

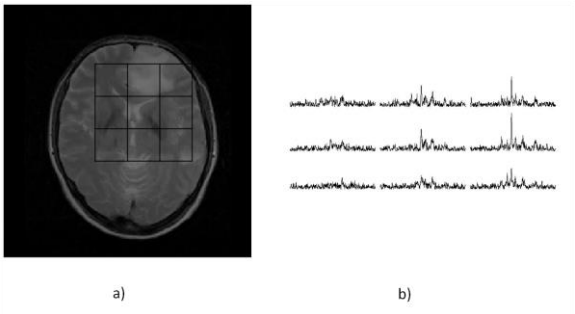

Figure 1. An example ^{31}P-MRSI data acquired from a thirty-seven years old female patient diagnosed with a grade II oligodendroglioma. a) T2 weighted spin echo (SE) MR image, and b) ^{31}P-MRSI data within the tumor region.

Table I shows the scores of each feature calculated at five iterations of RFE. The feature having the least score is shown in bold, which is then eliminated at that given iteration. At the end of five iterations, seven features, which were GPC, PE, gATP, GPC+GPE, PC+PE, PC/GPC, and PE/GPE, were determined as the most effective features for classification.

The distribution of ^{31}P-MRSI peak intensities in normal versus tumor voxels is shown in Figure 2, where the circles are the mean, the boundaries of the boxes are the 25th and 75th percentiles, and the dots are the outliers of the data. Tumor voxels displayed a higher GPC, PE, γ-ATP, GPC+GPE, PC+PE, and PE/GPE than normal voxels.

Table II summarizes the average performance measurements of forty four-fold cross validations for each of the three classification models. Support vector machine with linear kernel, support vector machine with polynomial kernel, and logistic regression were able to distinguish normal voxels from tumor voxels based on their phosphorus MR spectroscopic peak parameters, with 80.74%, 77.69%, and 90.51% sensitivities, respectively.

TABLE I. Scores and Eliminated Features of SVM-RFE

	Iterations of SVM-RFE				
	1st Iteration	***2nd Iteration***	***3rd Iteration***	***4th Iteration***	***5th Iteration***
GPC	83.48	94.71	94.76	94.75	96.81
GPE	55.01	64.55	**64.49**	-	-
Pi	**35.15**	-	-	-	-
PC	37.23	**45.05**	-	-	-
PE	71.14	82.82	82.95	83.06	85.45
gATP	74.39	85.00	85.10	85.17	85.61
aATP	68.53	80.47	80.42	80.51	**81.25**
bATP	64.87	75.14	74.96	**75.10**	-
GPC+GPE	274.03	315.65	315.60	316.20	322.08
PC+PE	211.28	250.05	249.30	249.52	255.60
PC/GPC	209.35	254.43	253.96	253.96	258.45
PE/GPE	600.18	686.24	690.04	689.56	705.55

Eliminated features are shown in bold, which are indicated with a (-) in the subsequent iterations.

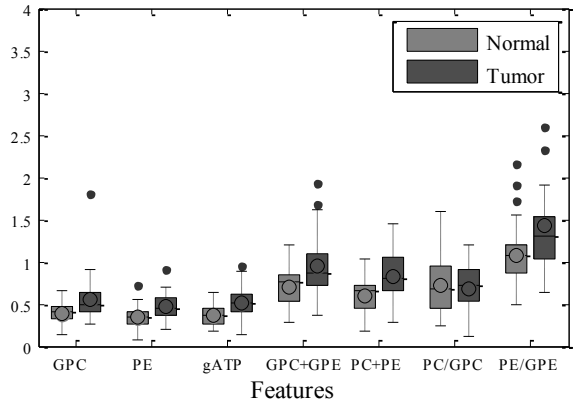

Figure 2. The distribution of the ^{31}P-MRSI data in normal versus tumor voxels.

TABLE II. Performance Results For three classifier Models

	Performance Evaluations (%)			
	Accuracy	***Sensitivity***	***Specificity***	***CI****
SVM_LN	74.80	80.74	68.37	[73.78 - 75.81]
SVM_PL	73.84	77.69	70.20	[72.99 - 74.68]
LR	81.71	90.51	72.56	[80.98 - 82.43]

* 90% confidence intervals for accuracy results.

Once the classifier accuracies were determined, the final classifier models were built using both the training and test datasets [5]. The ROC curve for the final model of logistic regression is shown in Figure 3. The optimum threshold value, b, was found as 0.538 at the corner of the ROC curve resulting in a false positive rate of 0.114 and a true positive rate of 0.786. The coefficient vector, θ, of the probability function, g, for the logistic regression model based on the optimal threshold value, b, was calculated as,

$$\theta = [1.66; 0.58; 15.19; 9.64; -7.77; 1.31; 7.19], \quad (9)$$

with a scalar $\theta_0 = -19.16$.

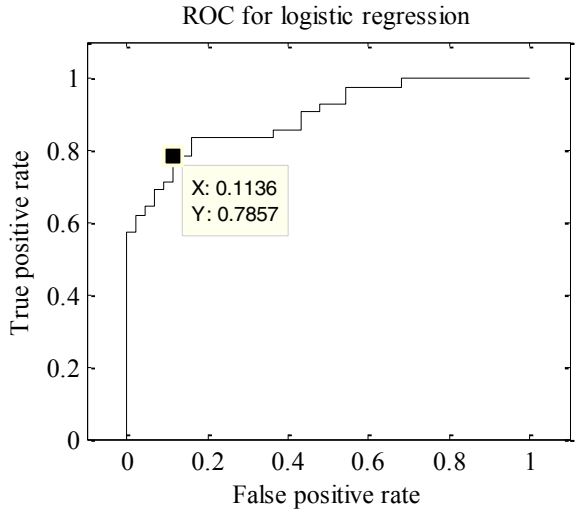

Figure 3. ROC curve for logistic regression method to classify tumor voxels based on ^{31}P-MRSI features using the whole dataset for training and then for testing. The optimal cutoff for the prediction threshold, b, was found at point [0.1136, 0.7857].

IV. Conclusion

The results of this study showed that support vector machine and logistic regression can be employed for defining classification models to discriminate brain tumor from normal tissue based on ^{31}P-MRSI data at 3T. The three classification methods were comparable in performance. Logistic regression resulted in a higher sensitivity, specificity and accuracy than both SVM methods. The main limitation of this study was small subject population resulting in rather low accuracy. Future studies will explore the applicability of machine learning algorithms for phosphorus MR spectroscopic data classification in a larger dataset. Additional studies will explore differentiating glioma subgroups using machine-learning algorithms based on ^{31}P-MRSI data.

Acknowledgment

This study was supported by TUBİTAK Career Development Grant 112E036 and EU Marie Curie IRG grant 256528.

References

1. Hubesch, B., et al., *P-31 MR spectroscopy of normal human brain and brain tumors.* Radiology, 1990. **174**(2): p. 401-9.
2. Maintz, D., et al., *Phosphorus-31 MR spectroscopy of normal adult human brain and brain tumours.* NMR Biomed, 2002. **15**(1): p. 18-27.
3. Elkhaled, A., et al., *Characterization of metabolites in infiltrating gliomas using ex vivo H high-resolution magic angle spinning spectroscopy.* NMR Biomed, 2014.
4. Hattingen, E., et al., *Phospholipid metabolites in recurrent glioblastoma: in vivo markers detect different tumor phenotypes before and under antiangiogenic therapy.* PLoS One, 2013. **8**(3): p. e56439.
5. Constantin, A., et al., *Identifying malignant transformations in recurrent low grade gliomas using high resolution magic angle spinning spectroscopy.* Artif Intell Med, 2012. **55**(1): p. 61-70.
6. Guo, J., et al., *The relationship between Cho/NAA and glioma metabolism: implementation for margin delineation of cerebral gliomas.* Acta Neurochir (Wien), 2012. **154**(8): p. 1361-70.
7. Lichy, M.P., et al., *[Application of (1)H MR spectroscopic imaging in radiation oncology: choline as a marker for determining the relative probability of tumor progression after radiation of glial brain tumors].* Rofo, 2006. **178**(6): p. 627-33.
8. Rock, J.P., et al., *Correlations between magnetic resonance spectroscopy and image-guided histopathology, with special attention to radiation necrosis.* Neurosurgery, 2002. **51**(4): p. 912-9.
9. Kounelakis, M.G., et al., *Revealing the metabolic profile of brain tumors for diagnosis purposes.* Conf Proc IEEE Eng Med Biol Soc, 2009. **2009**: p. 35-8.
10. Devos, A., et al., *The use of multivariate MR imaging intensities versus metabolic data from MR spectroscopic imaging for brain tumour classification.* J Magn Reson, 2005. **173**(2): p. 218-28.
11. Vanhamme, L., et al., *Time-domain quantification of series of biomedical magnetic resonance spectroscopy signals.* J Magn Reson, 1999. **140**(1): p. 120-30.
12. Alpaydın, E., *Introduction to Machine Learning*, 2004, Cambridge MA 02142, USA: The MIT Press. 409. p. 216-220.
13. Abe, S., *Support Vector Machines for Pattern Classification*, 2005, USA: Springer. p. 26.
14. Guyon, I., et al., *Gene selection for cancer classification using support vector machines.* Machine Learning, 2002. **46**(1-3): p. 389-422.