Multivariate Discriminant Analysis of Multiparametric Brain MRI to Differentiate High Grade and Low Grade Gliomas - A Computer-Aided Diagnosis Development Study

Füsun Çıtak Er*, Zeynep Firat, Ilhami Kovanlikaya, Ugur Ture, Esin Ozturk-Isik *Member, IEEE*

Abstract— **The aim of this study is to investigate the predictive capacity of multiparametric magnetic resonance imaging (MRI) findings using multivariate discriminant analysis. Preoperative clinical findings and multiparametric MRI, including diffusion weighted MR imaging, diffusion tensor imaging, perfusion MR imaging and MR spectroscopic imaging, were used as predictors to distinguish high grade from low grade gliomas. Principal component analysis was performed prior to discriminant analysis for dimensional reduction. Linear and quadratic discriminant analysis were performed and compared based on sensitivity and specificity analysis. The sensitivities of linear and quadratic discriminant analysis were 76.5% and 83.5%, respectively. Their specificities were 68.5% and 46.5%, respectively. Quadratic discriminant analysis provided a better discrimination than linear discriminant analysis for this dataset. This study is a model for a computer aided diagnosis system for glioma grading.**

I. Introduction

THE most frequent primary tumor of central nervous system is glioma that originates from the glial cells, which are the support cells of neurons. Accurate tumor classification provides staging information for the modern clinical practice of neuro-oncology [1]. The World Health Organization (WHO) glioma classification system provides a malignancy scale ranging from one to four [2]. Grade I is the least advanced disease that is probably cured following surgical resection alone. Although grade II has low-level proliferative activity, it often shows local recurrence. The lesions with histological evidence of malignancy are graded as grade III, while cytologically malignant, mitotically active, necrosis-prone neoplasms typically associated with rapid pre- and postoperative disease evolution and a fatal outcome are graded as grade IV [2].

Developing computer aided diagnosis (CAD) systems is an interdisciplinary approach that combines data mining, machine learning and image processing techniques to assist radiologists for interpreting medical images, including computed tomography, X-ray, MRI and ultrasound.

A CAD system design is highly disease dependent that needs extensive work on both predictors of the disease and classification techniques. Previous studies have developed CAD systems for grading brain tumors using several predictors. A previous study reported 85% sensitivity in discrimination of low and high grade gliomas using multiparametric MRI (mpMRI) without diffusion weighted imaging [3]. Awasthi et al. made discriminant function analysis based on dynamic contrast-enhanced (DCE) perfusion derived indices and immunohistochemical markers to classify low and high grade gliomas, and reported an overall sensitivity of %92.1 [4]. Valéry et al. studied correlation between surface of tumor volume (STV) and clinical parameters [5]. Yei et al. studied staging of gliomas using data mining techniques [6].

In this study, a CAD system was designed based on preoperative mpMRI of brain to grade gliomas. The predictive capacity of indices of mpMRI including diffusion weighted MR imaging, diffusion tensor imaging, perfusion MR imaging and MR spectroscopic imaging sequences were assessed. Principal Component Analysis (PCA) method was used to uncorrelate data for dimensional reduction. Then, Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) statistical methods were used to model data to predict the final grade. Performances of the models were compared using sensitivity and specificity analysis.

II. Material and Methods

A. Subjects

Thirty-one newly-diagnosed brain tumor patients (2 GrI, 9 GrII, 7 GrIII and 13 GrIV; mean age=48.39±14.32) were scanned on a 3T MR scanner (Philips, Netherlands) with an eight-channel head coil. All patients underwent routine brain tumor protocol including; T2 weighted (T2w) turbo spin echo (TSE), T2w fluid attenuated inversion recovery (FLAIR), diffusion weighted MR imaging (b=1000), diffusion tensor imaging (DTI), pre- and post-contrast 3D T1w turbo field echo (TFE), perfusion MR imaging and MR spectroscopic imaging sequences. Fractional anisotropy (FA), cerebral blood volume (CBV), mean transit time

Manuscript received July 1, 2013. This work was supported in part by Philips Healthcare through our research agreement.

F. C. E. Author is with the Department of Genetics and Bioengineering, Yeditepe University, 34755 Istanbul, Turkey (phone: +(90)216-578-0000 (3175); fax: +(90)216-578-0299; e-mail: fusun.er@std.yeditepe.edu.tr).

Z. F. is with the Department of Radiology, Yeditepe University Hospital, 34752 İstanbul, Turkey (e-mail: zeyfirat@yahoo.com).

I. K. is with the Department of Radiology, Weill Cornell Medical College, NY 10021 USA (e-mail: ilk2002@med.cornell.edu).

U. T. is with the Department of Neurosurgery, Yeditepe University Hospital, 34752 İstanbul, Turkey (e-mail: uture@yeditepe.edu.tr).

E. O. I. is with the Department of Biomedical Engineering, Yeditepe University, 34755 Istanbul, Turkey (e-mail: esin.ozturk@yeditepe.edu.tr).

978-1-4799-3163-7/13/$31.00 ©2013 IEEE

(MTT), cerebral blood flow (CBF), apparent diffusion coefficient (ADC), and peak heights and ratios of N-acetyl aspartate (NAA), creatine (Cr) and choline (Cho) were calculated. NAA over Cr (NAA/Cr) and Cho over Cr (Cho/Cr) ratios were also computed. For each patient, the voxel displaying highest choline level or CBV value was determined as the first region of interest (ROI) that had tumor. A second ROI that was 1 cm away from the tumor margin was defined as the border ROI. A third ROI that was 2 cm away from the tumor margin was defined as the surrounding tissue. Finally, a fourth ROI was defined at the symmetrical normal region. The volume of each ROI was 1-2 cm^3. Regions of interest were placed on all MR images by an experienced Radiologist manually while keeping the location and size of the ROI same.

All features were normalized with the corresponding normal region features. Glioma grades I and II were grouped as low grade, and grades III and IV were grouped as high grade. In total, twenty-two features, which were age of patient, FA, CBV, MTT, CBF, ADC, Cho/Cr and NAA/Cr of first, second and third ROIs normalized with fourth ROI's values were used as predictors, and their distributions in high and low grade gliomas are summarized in Table I.

TABLE I
PREDICTORS OF GRADE

Predictor	Low Grade	High Grade
Age	42.82±10.81	51.45±15.32
1st ROI - FA	0.32±0.20	0.36±0.22
2nd ROI - FA	0.76±0.30	0.58±0.19
3rd ROI - FA	0.80±0.39	0.78±0.51
1st ROI - CBV	0.94±0.88	2.46±2.08
2nd ROI - CBV	1.10±0.67	1.42±1.09
3rd ROI - CBV	1.08±0.74	1.45±2.02
1st ROI - CBF	1.05±0.86	2.58±2.34
2nd ROI - CBF	1.09±0.61	1.82±1.85
3rd ROI - CBF	1.09±0.67	1.87±2.16
1st ROI - MTT	0.85±0.58	0.93±0.35
2nd ROI - MTT	1.04±0.17	0.92±0.31
3rd ROI - MTT	0.97±0.36	0.89±0.34
1st ROI - ADC	2.01±0.62	1.44±0.62
2nd ROI - ADC	1.10±0.54	1.37±0.37
3rd ROI - ADC	1.13±0.27	1.36±0.47
1st ROI – Cho/Cr	2.08±0.43	2.90±2.55
2nd ROI – Cho/Cr	1.45±0.49	2.09±1.62
3rd ROI – Cho/Cr	1.17±0.30	1.59±1.19
1st ROI – NAA/Cr	0.55±0.49	0.50±0.72
2nd ROI – NAA/Cr	0.75±0.35	0.99±1.37
3rd ROI – NAA/Cr	0.91±0.37	1.14±1.54

1st ROI, 2nd ROI and 3rd ROI are regions of interests that belongs to tumor, border of tumor and surrounding tissue of tumor regions normalized with normal tissue values, respectively.

Low grade and high grade values are presented as mean±std.

B. Methods

1) Preprocessing of data: The data was multivariate and PCA is a maximum variance projection method, so the data was scaled using unit variance scaling method, and then the means were centered to zero [7].

2) Principal Component Analysis: Principal Component Analysis (PCA) produces a set of uncorrelated variables, called principal components, from correlated variables [8]. PCA is recommended to be performed prior to any kind of multivariate data analysis [8]. Locating possible outliers in the data, clustering data into subgroups or verifying results of clustering programs, providing variance-covariance matrix to a discriminant analysis, uncorrelating data, and reducing dimensionality are some functions of PCA [8].

PCA commonly uses Hotelling's T^2 method [9] to determine outliers, which is a statistical technique to calculate distance of data from the center of the cluster.

3) Linear Discriminant Analysis and Quadratic Discriminant Analysis: Discriminant analysis is a multivariate technique to predict the group of an observation based on a set of predictors [8]. Discriminant analysis is similar to regression analysis but differs in output. While regression analysis is able to predict the value of a dependent variable, discriminant analysis predicts class membership of an individual observation [8].

Discriminant analysis provides a function called discriminant function to classify data. The dataset is projected into the discriminant function by minimizing the intra-class distance while maximizing the inter-class distance.

Linear discriminant analysis and quadratic discriminant analysis are two types of discriminant analysis that differs in the resultant discriminant function that is linear or quadratic.

Figure 1 demonstrates an example of linear discrimination of a two-dimensional data that belongs to two classes marked with squares and circles. m_{1b} and m_{1a} are the means of the first class marked with squares before and after the projection, while m_{2b} and m_{2a} are that of the other class [10].

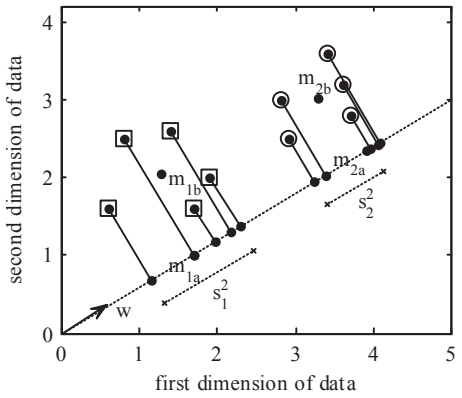

Fig. 1. Two-dimensional data that belongs to one of the two classes marked with squares and circles. m_{1b} and m_{2b} represent the mean values of classes 1 and 2 before projection, while m1a and m_{2a} are the means values of projected data onto a projection vector w. The scatter ranges of projections are marked with s_1^2 and s_2^2.

All the discriminant analysis software was developed in MATLAB version R2009b (The MathWorks Inc., Natick, MA) using Statistics Toolbox functions. Ten-fold cross-validation technique was used to separate training and validation datasets [11]. One fold was used as the validation dataset to test the system, while the other nine folds were used as the training dataset. This procedure was repeated for each of the ten folds. The average of these ten test results was determined as the validation result of ten-fold cross validation. The random nature of the k-fold cross validation technique and our small data size resulted in a large variance amongst the results. To overcome this problem, ten-fold cross validation process was iterated ten times, and was averaged for the final result.

Training dataset was dimensionally reduced using PCA. New set of dimensionally reduced data was used to train LDA and QDA classification methods. Then, LDA and QDA classifications were tested by the validation dataset. The projection functions of LDA and QDA were calculated. Sensitivity and specificity of the LDA and QDA were compared for performance analysis of these classification methods.

III. Results

A. Principal Component Analysis

Table II shows the eigenvalues calculated by PCA analysis. Even with the first and second principal components, the system reached a 93.30% cumulative sum. In this study, five principal components were selected to reach 98%. Remaining predictors were ignored.

TABLE II
RESULTS OF PCA ANALYSIS

Predictor	Eigenvalue	Cumulative Sum (%)
1st Principal	206.4626	0.8816
2nd Principal	12.0528	0.9330
3rd Principal	6.4204	0.9605
4th Principal	3.7241	0.9764
5th Principal	2.7255	0.9880
6th Principal	0.7603	0.9912
7th Principal	0.4393	0.9931
8th Principal	0.3619	0.9947
9th Principal	0.2507	0.9957
10th Principal	0.2131	0.9966
11th Principal	0.1999	0.9975
12th Principal	0.1476	0.9981
13th Principal	0.1325	0.9987
14th Principal	0.1074	0.9991
15th Principal	0.0515	0.9994
16th Principal	0.0390	0.9995
17th Principal	0.0350	0.9997
18th Principal	0.0266	0.9998
19th Principal	0.0226	0.9999
20th Principal	0.0161	1.0000
21th Principal	0.0048	1.0000
22th Principal	0.0039	1.0000

1st ROI, 2nd ROI and 3rd ROI are regions of interests that belongs to tumor, border of tumor and surrounding tissue of tumor regions normalized with normal tissue values, respectively.

Low grade and high grade values are presented as mean±std.

Figure 2 shows the second principal component against the first principal component of each data point, in which high and low grade classes are marked with circles and squares. Although 10th data point was marked as an outlier by Hotelling's T^2 technique, it was near the center of the data, and it was not eliminated due to the small size of this dataset.

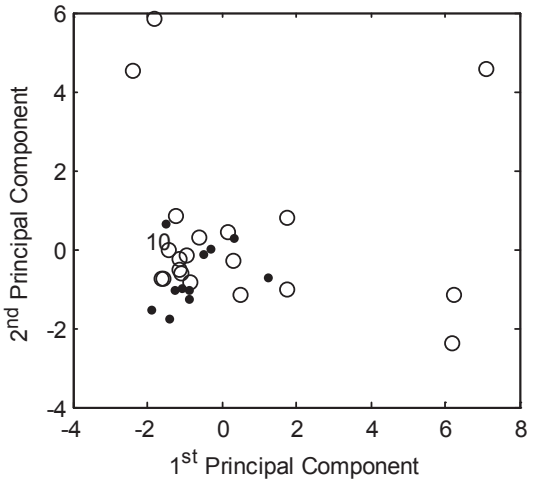

Fig. 2. The second principal component is plotted against the first principal component. The high grade and low grade classes are marked with circles and squares, respectively. 10th observation that lies under twice the standard deviation of the mean based on Hotelling's T^2 test was marked.

Table III shows the contribution of predictors on both the first and second principal components, divided into four quadrants. The first group lists the predictors that have positive coefficients for both principal components, which are NAA/Cr of all three ROIs, CBV of the tumor and border ROIs, MTT of the tumor and surrounding ROIs, ADC of the surrounding ROI, CBF of the tumor ROI, and age of the patient.

Cho/Cr of all three ROIs and MTT and ADC of the border ROI had negative coefficients for the first principal component, while they had positive coefficients for the second principal component.

FA of the surrounding ROI and ADC of the tumor ROI had negative coefficients for the first and second principal components. The last group lists the predictors that have positive coefficients for the first principal component and negative coefficients for the second principal component, which are FA of the tumor and border ROI, CBF of the border and surrounding ROIs, and CBV of the surrounding ROI. It was observed that CBF, CBV, and NAA/Cr from all three ROIs had positive coefficients for the first principal component, while NAA/Cr, Cho/Cr, and MTT had positive coefficients for the second principal component.

The largest coefficient in the first principal component belonged to FA of the tumor ROI. The largest coefficient in the second principal component belonged to MTT of the border ROI.

TABLE III
COEFFICIENTS OF PREDICTORS IN PRINCIPAL COMPONENTS

Predictor	1st Principal	2nd Principal
1st ROI - NAA/Cr	0.3806	0.3468
2nd ROI - CBV	0.3170	0.0947
3rd ROI - MTT	0.2730	0.3062
2nd ROI - NAA/Cr	0.2056	0.2088
3rd ROI - NAA/Cr	0.1989	0.1877
3rd ROI - ADC	0.1921	0.4933
1st ROI - CBV	0.1687	0.0599
1st ROI - MTT	0.1626	0.3772
1st ROI - CBF	0.1471	0.0149
Age	0.0092	0.0058
2nd ROI - MTT	-0.3617	1.1297
1st ROI - Cho/Cr	-0.0686	0.1874
2nd ROI - ADC	-0.0743	0.4234
2nd ROI - Cho/Cr	-0.1113	0.3057
3rd ROI - Cho/Cr	-0.1714	0.4106
3rd ROI - FA	-0.0825	-0.2996
1st ROI - ADC	-0.3295	-0.1264
1st ROI - FA	0.9893	-0.8001
2nd ROI - FA	0.2458	-0.8605
2nd ROI - CBF	0.2188	-0.0720
3rd ROI - CBF	0.2105	-0.0009
3rd ROI - CBV	0.1515	-0.0085

The coefficients of predictors over 1st and 2nd principal components are listed above.

Table IV lists the mean value of the first five principal components after dimension reduction using the PCA method.

TABLE IV
PRINCIPAL COMPONENTS OF THE DATASET AFTER PCA (MEAN±STD)

Principal Component	Low Grade	High Grade
1st Principal	-0.75± 0.91	0.41 ± 2.87
2ndPrincipal	-0.67 ± 0.78	0.37 ± 2.14
3rd Principal	-0.84 ± 0.93	0.46 ± 1.81
4th Principal	0.27 ± 1.12	-0.15 ± 1.65
5th Principal	0.34 ±1.36	-0.19 ± 1.20

Dataset contained twenty high grade data and eleven low grade data.

B. Linear Discriminant Analysis

An example of the function of projection vectors found by LDA analysis is given in Eq. 1,

$$f_{LDA}(a,b,c,d,e) = -0.6212 - 0.5310a - 0.26b - 1.1554c + 0.2014d + 0.2488e, \quad (1)$$

where, a is the 1st, b is the 2nd, c is the 3rd, d is the 4th, and e is the 5th principal components.

LDA results of the ten iterations are summarized in Table V. The average sensitivity and specificity of LDA were 76.5% and 65.5%, respectively.

TABLE V
TEST RESULTS OF LDA

	Sensitivity	Specificity
1st iteration	80%	65%
2nd iteration	65%	65%
3rd iteration	70%	70%
4th iteration	80%	70%
5th iteration	80%	75%
6th iteration	70%	70%
7th iteration	85%	70%
8th iteration	85%	65%
9th iteration	70%	65%
10th iteration	80%	70%
Average	76.5%	68.5%

Ten-fold cross-validation was repeated ten times and their averaged sensitivity and specificity values are shown.

C. Quadratic Discriminant Analysis

An example of the projection functions of QDA is given by Eq.5,

$$f_{QDA}(a,b,c,d,e) = 1.4894 - 0.8182a - 1.0546b - 0.9473c + 0.2503d + 0.3522e - 0.4789a^2 - 0.6379b^2 - 0.2361c^2 - 0.172d^2 + 0.0947e^2, \quad (5)$$

where, a is the 1st, b is the 2nd, c is the 3rd, d is the 4th, and e is the 5th principal components.

The projection functions of f_{QDA} classified this small validation data set slightly better than f_{LDA} for all iterations, as shown in Table VI. The average sensitivity of QDA was 83.5%.

TABLE VI
TEST RESULTS OF QDA

	Sensitivity	Specificity
1st iteration	85%	40%
2nd iteration	85%	50%
3rd iteration	80%	50%
4th iteration	85%	50%
5th iteration	80%	40%
6th iteration	80%	45%
7th iteration	90%	45%
8th iteration	85%	45%
9th iteration	80%	50%
10th iteration	85%	50%
Average	83.5%	46.5%

Ten-fold cross-validation was repeated ten times and their averaged sensitivity and specificity values are shown.

Additionally, the analysis was repeated without preprocessing the data, and dimension reduction with PCA. It was observed that the sensitivity and specificity of LDA were reduced to 39.5% and 57.5%, respectively. Similarly, sensitivity and specificity of QDA dropped to 64% and 44%, respectively. So, the classifiers performed better after PCA.

CONCLUSIONS

A computer aided diagnosis system for predicting the WHO glioma grade based on mpMRI was evaluated in this study. Unit variance scaling method and mean centering methods were used to normalize the data. PCA was used to eliminate redundant predictors. Parametric statistical methods of LDA and QDA were applied to predict glioma grade in a validation dataset. QDA performed better for this limited sample size. Future studies will look into the performance of different groups of features for the WHO glioma grade prediction in a larger patient cohort, and Support Vector Machine (SVM) algorithm will also be tested for glioma classification.

REFERENCES

1. Louis, D.N., E.C. Holland, and J.G. Cairncross, *Glioma classification: a molecular reappraisal.* Am J Pathol, 2001. **159**(3): p. 779-86.
2. Louis, D.N., et al., The 2007 WHO classification of tumours of the central nervous system. Acta Neuropathol, 2007. **114**(2): p. 97-109.
3. Zacharaki, E.I., et al., Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. Magn Reson Med, 2009. **62**(6): p. 1609-18.
4. Awasthi, R., et al., Discriminant analysis to classify glioma grading using dynamic contrast-enhanced MRI and immunohistochemical markers. Neuroradiology, 2012. **54**(3): p. 205-13.
5. Valery, C.A., et al., Extent of tumor-brain interface: a new tool to predict evolution of malignant gliomas. J Neurosurg, 2001. **94**(3): p. 433-6.
6. Yei, C., J. Yang, and D. Geng, *[Data mining in diagnostic knowledge acquisition from patients with brain glioma]*. Sheng Wu Yi Xue Gong Cheng Xue Za Zhi, 2002. **19**(3): p. 426-30.
7. Jackson, D. and R.D. Riley, A refined method for multivariate meta-analysis and meta-regression. Stat Med, 2013.
8. E.Johnson, D., Applied Multivariate Methods for Data Analysis. 1998: Duxbury Press.
9. Hotelling, H., *The Generalization of Student's Ratio.* Ann. Math. Statist., 1931. **2**(3): p. 360-378.
10. Alpaydın, E., *Introduction to Machine Learning*, in *MIT Press*. 2004, The MIT Press: Cambridge, Massachusetts London, England. p. 128-130.
11. Stone, M., *Cross-validatory Choice and Assessment of Statistical Predictions.* Journal of the Royal Statistical Society, 1974. **36**(2): p. 111-147.