

Non-linear Least Squares Features Transformation for Improving the Performance of Probabilistic Neural Networks in Classifying Human Brain Tumors on MRI

Pantelis Georgiadis¹, Dionisis Cavouras², Ioannis Kalatzis², Antonis Daskalakis¹, George Kagadis¹, Koralia Sifaki³, Menelaos Malamas³, George Nikiforidis¹, and Ekaterini Solomou⁴

¹ Medical Image Processing and Analysis Group (M.I.P.A.), Laboratory of Medical Physics, School of Medicine, University of Patras, Rio, GR-26500 Greece, pgeorgiadis@med.upatras.gr

² Medical Signal and Image Processing Lab, Department of Medical Instruments Technology, Technological Educational Institute of Athens, Ag. Spyridonos Street, Aigaleo, 122 10, Athens, Greece

³ 251 General Hellenic Airforce Hospital, MRI Unit, Katehaki, Athens, GR-11525, Greece

⁴ Department of Radiology, School of Medicine, University of Patras, Rio, GR-26503 Greece

Abstract. The aim of the present study was to design, implement, and evaluate a software system for discriminating between metastases, meningiomas, and gliomas on MRI. The proposed classifier is a modified probabilistic neural network (PNN), incorporating a second degree least squares features transformation (LSFT) into the PNN classifier. Thirty-six textural features were extracted from each one of 75 T1-weighted post-contrast MR images (24 metastases, 21 meningiomas, and 30 gliomas). Classification performance was evaluated employing the leave-one-out method and for all possible textural feature combinations. LSFT enhanced the performance of the PNN, achieving 93.33% in discriminating between the three major types of human brain tumors, against 89.33% scored by the PNN alone. Best feature combination for achieving highest discrimination power included the mean value and entropy, which reflect specific properties of texture, i.e. signal strength and inhomogeneity. LSFT improved PNN performance, increased class separability, and resulted in dimensionality reduction.

1 Introduction

One of the most promising techniques for generating useful information for brain tissue characterization is Magnetic Resonance Imaging (MRI). MRI, in comparison with other diagnostic imaging modalities, such as computerized tomography, provides superior contrast and resolution for different brain tissues [1]. Additionally, MR images encapsulate valuable information regarding numerous tissue parameters (proton density, spin-lattice (T1) and spin-spin (T2) relaxation times, flow velocity, and chemical shift), which lead to more accurate brain tissue characterization. These unique advantages have characterized MRI as the method of choice in brain tumor studies [2]. In order to extract the diagnostic information of different parameters reflected in MRI, image analysis techniques have been employed.

Several studies have proposed various approaches to characterize human brain tumors, using either MR textural [3-5], spectroscopic [6-10], or a combination of both textural and spectroscopic parameters [11-12]. These studies have formulated the discrimination between several tumor types as a two-class problem, employing either statistical analysis (linear discriminant analysis, analysis of variance, hierarchical ascending classification) or state of art classifiers, such as support vector machines.

The aim of the present study was to improve tumor classification accuracy, as compared to previous studies, in distinguishing between three major categories of human brain tumors: metastases, meningiomas, and gliomas. We considered a three-class problem formulation, we employed textural features, and we conditioned these features by means of least squares features transformation (LSFT) to boost the performance of the Probabilistic Neural Network (PNN) classifier.

2 Materials and Methods

2.1 Data Acquisition

A total number of 75 T1-weighted post-contrast (Gadolinium) MR images were obtained from the Hellenic Airforce Hospital with verified intracranial tumors, using a SIEMENS-*Sonata* 1.5 Tesla MR Unit. The image dataset comprised 24 metastases, 21 meningiomas, and 30 gliomas. From each case, only T1-weighted post-contrast (Gadolinium) images, with Spin Echo (SE) sequence, Echo Time (TE = 15ms), and Repetition Time (TR = 500ms), were used for further analysis. Transverse images were selected through the tumor's center by an expert radiologist (M.M.).

2.2 Feature Extraction

Utilizing these images, the radiologist specified Regions Of Interest (ROIs) that included tumor regions, using custom developed software (Fig. 1). From each ROI, a series of 36 textural features were generated; 4 features were extracted from the ROI's histogram, expressing information concerning the frequency of appearance of

Table 1. Textural features extracted

Methods	Features
Histogram (1st order statistics)	Mean Value, Standard Deviation, Skewness, Kurtosis
Mean and range of 0°, 45°, 90° and 135° co-occurrence matrices (2nd order statistics)	Angular Second Moment, Contrast, Correlation, Sum Of Squares, Inverse Difference Moment, Sum Average, Sum Variance, Sum Entropy, Entropy, Difference Variance, Difference Entropy
Mean and range of 0°, 45°, 90° and 135° run-length matrices (2nd order statistics)	Short Run Emphasis, Long Run Emphasis, Gray Level Non Uniformity, Run Length Non Uniformity, Run Percentage

each gray level; 22 features were extracted from the gray level co-occurrence matrices [13], reflecting the frequency of appearance of each gray level at adjacent pixels; 10 features were extracted from the run-length matrices [14], describing the frequency of appearance of a set of consecutive pixels having the same gray value (Table 1).

All features were normalized to zero mean and unit standard deviation [15], according to relation (1)

$$x_i' = \frac{x_i - m}{std} , \tag{1}$$

where x_i and x_i' are the i -th feature values before and after the normalization respectively, and m and std are the mean value and standard deviation of feature x over all patterns (from all three tumor categories) respectively. Extracted feature vectors were fed into the classifier.

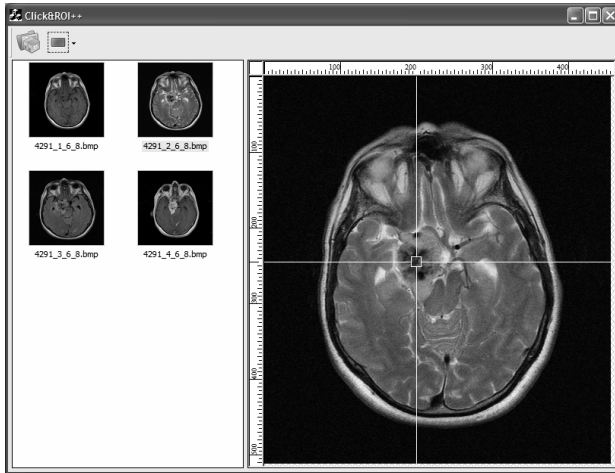


Fig. 1. Custom developed software used to extract ROIs and textural features from the MR images

2.3 Least Squares Feature Transformation – Probabilistic Neural Network Classifier (LSFT-PNN)

Due to the small size of our dataset the probabilistic neural network (PNN) classifier was chosen. The PNN is a non-parametric feed-forward neural network classifier, with low computational demands and fast in training. The discriminant function of PNN [16] is given by:

$$g_i(\mathbf{x}) = \frac{1}{(2\pi)^{d/2} \sigma^d} \frac{1}{N_i} \sum_{j=1}^{N_i} \exp \left[-\frac{(\mathbf{x} - \mathbf{x}_{ij})^T (\mathbf{x} - \mathbf{x}_{ij})}{2\sigma^2} \right], \tag{2}$$

where \mathbf{x} is the pattern to be classified, \mathbf{x}_{ij} are the training patterns, σ is the spread of the Gaussian activation function, N_i is the number of training patterns in class i , and d is the dimensionality of pattern vectors.

Training patterns \mathbf{x}_{ij} , prior to entering the PNN classifier, were transformed by means of a non-linear least squares feature transformation (LSFT) technique, to render classes more separable by clustering the patterns of each class around arbitrary pre-selected points. The proposed LSFT method is an extension of the linear least squares mapping technique, introduced by [17]. Initially, pattern vectors were extended with second degree elements. Accordingly, if $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_d]$ is a pattern vector, where d is the input space dimensionality, then vector \mathbf{x} was extended with the second degree elements x_i^2 and $x_i x_j$, where $i, j = 1, 2, \dots, d$ and $i \neq j$, resulting in the following extended pattern vector:

$$\hat{\mathbf{X}} = [x_1 \ x_2 \ \dots \ x_d \ x_1^2 \ x_2^2 \ \dots \ x_d^2 \ x_1 x_2 \ x_1 x_3 \ \dots \ x_1 x_d \ x_2 x_3 \ x_2 x_4 \ \dots \ x_2 x_d \ \dots \ x_{d-1} x_d] \quad (3)$$

The dimensionality of the extended pattern vector ($\hat{\mathbf{X}}$) is equal to [15]:

$$\hat{d} = \frac{(d + 2)!}{d!2!} - 1. \quad (4)$$

For the formulation of the LSFT 3-class problem, let space \mathbf{S} , with dimensionality equal to the number of classes ($K=3$), and let $\mathbf{P}_i = [p_{i1} \ p_{i2} \ p_{i3}]$, $i=1,2,3$, be arbitrary defined points in space \mathbf{S} , corresponding to each class i . A transformation \mathbf{T} is sought such that the total mean square error between the transformed extended vectors ($\mathbf{T}\hat{\mathbf{X}}_{ij}$) and \mathbf{P}_i is minimized as follows:

$$\nabla_{\mathbf{T}} \left[\sum_{i=1}^K \left(\frac{1}{N_i} \sum_{j=1}^{N_i} (\mathbf{T}\hat{\mathbf{x}}_{ij} - \mathbf{P}_i)' (\mathbf{T}\hat{\mathbf{x}}_{ij} - \mathbf{P}_i) \right) \right] = 0, \quad (5)$$

where K is the number of classes, N_i is the number of patterns of class i , and $\hat{\mathbf{x}}_{ij}$ are the 2nd degree extended training patterns of class i .

Assuming equal *a-priori* probabilities for each class i , relation (5) results to:

$$\mathbf{T} = \left[\sum_{i=1}^K \left(\frac{1}{N_i} \sum_{j=1}^{N_i} \mathbf{P}_i \hat{\mathbf{x}}_{ij}' \right) \right] \left[\sum_{i=1}^K \left(\frac{1}{N_i} \sum_{j=1}^{N_i} \hat{\mathbf{x}}_{ij} \hat{\mathbf{x}}_{ij}' \right) \right]^{-1}. \quad (6)$$

Transformation matrix \mathbf{T} is a $K \times \hat{d}$ matrix, so the decision space dimensionality is equal to the number of brain tumor classes.

Following the LSFT procedure, patterns $\hat{\mathbf{x}}_{ij}$ were fed into the PNN classifier, resulting in the final discriminant function of the LSFT-PNN classifier:

$$g_i(\mathbf{x}) = \frac{1}{(2\pi)^{\hat{d}/2} \sigma^{\hat{d}}} \frac{1}{N_i} \sum_{j=1}^{N_i} \exp \left[- \frac{(\mathbf{T}\hat{\mathbf{x}} - \mathbf{T}\hat{\mathbf{x}}_{ij})' (\mathbf{T}\hat{\mathbf{x}} - \mathbf{T}\hat{\mathbf{x}}_{ij})}{2\sigma^2} \right]. \quad (7)$$

2.4 Feature Selection and System Evaluation

Best features combination was determined employing the robust but time consuming exhaustive search method [15], which involves designing the classifier by means of every possible feature-combination, each time evaluating the classifier's performance and finally selecting that feature combination that demonstrated the highest classification accuracy with the smallest number of features.

Classifier performance was evaluated employing the leave-one-out (LOO) method [15]. Accordingly, the LSFT-PNN classifier was designed by all but one pattern-vector, which was considered as unknown and it was classified. The process was repeated, each time leaving-out a different pattern-vector, until all pattern-vectors were thus classified to one of three classes. In this way, the classifier was evaluated by pattern-vectors not involved in its design. This type of classifier training required several hours of processing time.

All algorithms were implemented using custom software developed using Microsoft's Visual C++ 6. The feature generation and classification tasks were performed on a typical desktop computer (Pentium 4 @ 3GHz with 1GB of RAM) running Windows 2000 Professional operating system.

3 Results and Discussion

The LSFT-PNN classification scheme was optimized with respect to parameter settings and available feature data. The spread of Gaussian function in equation (7) was experimentally set equal to $\sigma=0.3$. Best feature vector, used for the optimal design of the LSFT-PNN classifier, comprised the mean value, the mean entropy, and the mean and range values of the difference entropy (see Fig. 2). These features reflect the signal strength and degree of inhomogeneity of the ROIs' gray-tones. The overall classification accuracy for directly discriminating the three types of human brain tumors was 93.33% (Table 2). Figure 2 shows the scatter diagram of the transformed optimum LSFT features combination, with black-filled labels indicating misclassified patterns.

Table 2. LSFT-PNN truth table for three-class problem

	Metastases	Meningiomas	Gliomas	Accuracy (%)
Metastases	21	1	2	87.50
Meningiomas	1	20	0	95.24
Gliomas	1	0	29	96.67
				93.33

The performance of the LSFT-PNN algorithm was tested against that of the PNN classifier ($\sigma=0.3$), used in the LSFT-PNN technique, as well as with the Quadratic Bayesian (QB) [15] and the kNN ($k=3$) [15] classifiers. The classifiers were trained in a similar manner to the LSFT-PNN classifier, employing in its design the LOO procedure and the exhaustive search method. The best feature combination of the

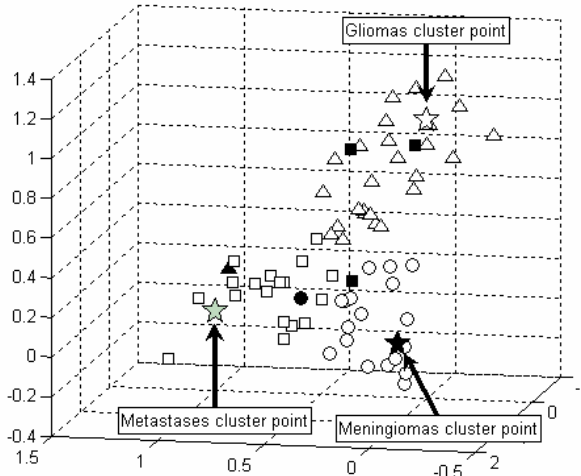


Fig. 2. Scatter diagram in the decision space of the optimum LSFT feature combination (metastasis (□), meningiomas (○) and gliomas (Δ) classes; misclassified patterns of LSFT-PNN classification: ■, ● and ▲ respectively)

Table 3. PNN truth table for 3-class problem

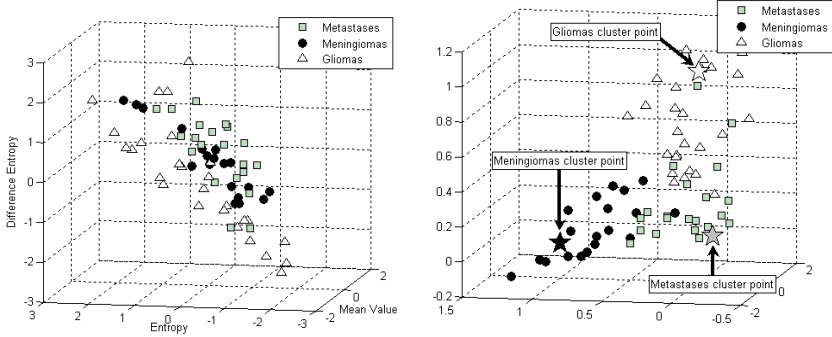
	Metastases	Meningiomas	Gliomas	Accuracy (%)
Metastases	20	2	2	83.33
Meningiomas	2	19	0	90.48
Gliomas	2	0	28	93.33
				89.33

PNN was found to be the mean and range values of the correlation, the sum average, and the difference entropy. This features express measures of the gray-tone linear dependencies (correlation) [13], of the disperse of intensity values (sum average), and give an indication of the degree of inhomogeneity in the gray-tones of the ROIs (entropy). The performance of the PNN classification scheme is summarized in (Table 3). The overall accuracies of the QB and the kNN classifiers were 83.58% and 86.56% respectively. Best feature vector for QB was mean value, entropy, contrast, and sum of average and for kNN was angular second moment, sum of average, different entropy and short run emphasis.

The effectiveness of the LSFT transformation on the PNN’s classification accuracy was also tested by formulating three, 2-class discrimination tests, as shown in Table 4. As it can be observed, the LSFT transformation improved the performance of the PNN in all three problems. Judging from the results achieved in all tests conducted (Tables 2-4), it can be claimed that the LSFT-PNN outperforms the PNN. This may be attributed to the increased class separability that the LSFT procedure provides

Table 4. LSFT-PNN versus PNN overall accuracies for 2-class problems

	Metastases vs Meningiomas (%)	Gliomas vs Meningiomas (%)	Gliomas vs Metastasis (%)
LSFT-PNN	97.50	100.00	93.75
PNN	90.00	97.83	91.67

**Fig. 3.** Pattern vectors scatter diagram of best 3-feature combination, a/ before LSFT application and b/ after LSFT

(Fig. 3). Another advantage of the LSFT is the dimensionality reduction to the number of classes, which leads to more robust classification, independently of the number of features in the pre-transformed data.

The computational requirements of the LSFT-PNN classifier are comparable to those of the PNN, as the time required to perform the LSFT is gained in the classification step, due to the reduced feature dimensionality.

Only few studies were found in literature that employed pattern recognition techniques and texture analysis to classify human brain tumors. In a recent study [12], an SVM-based classification system achieved 95% overall accuracy in discriminating between gliomas and meningiomas. These results are comparable with the findings of the present study regarding the two-class problem formulation of these two types of brain tumors (Table 4). However, in another study [5], employing hierarchical ascending classification with correspondence factorial analysis, discrimination accuracy between different tumor types ranged between 49% (tumors Vs oedemas) and 63% (benign Vs malignant tumors), employing textural features related to gray-level homogeneity (gray-level distribution) and the existence of structures within the lesion (run length distribution). In an older study [3], discriminant analysis and the k-nearest neighbor classifier were used to differentiate between human brain tumor and oedematous tissue, achieving maximum overall accuracy equal to 95% and using textural features associated with dependencies between neighboring gray levels (correlation and run length mean gray level). Similar features were also found to give optimum discrimination accuracies in the present work.

4 Conclusion

The aim of this study was to design, implement, and evaluate a software pattern recognition system in discriminating between three categories of human brain tumors using the LSFT-PNN classification algorithm. Least squares features transformation was proved to improve the performance of the PNN classifier, mainly due to the increased class separability and to dimensionality reduction that the LSFT procedure provides.

Acknowledgments. Funding by the University of Patras Research Committee under the basic research program K. Karatheodoris, Project title “Computer Assisted Diagnosis of Brain Tumors based on Statistical Methods and Pattern Recognition Techniques” is gratefully acknowledged.

References

1. Shen, S., Sandham, W., Granat, M., Sterr, A.: MRI fuzzy segmentation of brain tissue using neighborhood attraction with neural-network optimization. *IEEE Trans Inf Technol Biomed.* 9, 459–467 (2005)
2. Soltanian-Zadeh, H., Peck, D.J., Windham, J.P., Mikkelsen, T.: Brain tumor segmentation and characterization by pattern analysis of multispectral NMR images. *NMR Biomed.* 11, 201–208 (1998)
3. Lerski, R.A., Straughan, K., Schad, L.R., Boyce, D., Bluml, S., Zuna, I.: MR image texture analysis—an approach to tissue characterization. *Magn Reson Imaging* 11, 873–887 (1993)
4. Schad, L.R., Bluml, S., Zuna, I.: MR tissue characterization of intracranial tumors by means of texture analysis. *Magn Reson Imaging* 11, 889–896 (1993)
5. Herlidou-Meme, S., Constans, J.M., Carsin, B., Olivie, D., Eliat, P.A., Nadal-Desbarats, L., Gondry, C., Le Rumeur, E., Idy-Peretti, I., de Certaines, J.D.: MRI texture analysis on texture test objects, normal brain and intracranial tumors. *Magn Reson Imaging* 21, 989–993 (2003)
6. Cho, Y.-D., Choi, G.-H., Lee, S.-P., Kim, J.-K.: ^1H -MRS metabolic patterns for distinguishing between meningiomas and other brain tumors. *Magnetic Resonance Imaging* 21, 663–672 (2003)
7. Tate, A.R., Majos, C., Moreno, A., Howe, F.A., Griffiths, J.R., Arus, C.: Automated classification of short echo time in in vivo ^1H brain tumor spectra: A multicenter study. *Magn Reson Med.* 49, 29–36 (2003)
8. Devos, A., Lukas, L., Suykens, J.A., Vanhamme, L., Tate, A.R., Howe, F.A., Majos, C., Moreno-Torres, A., van der Graaf, M., Arus, C., Van Huffel, S.: Classification of brain tumours using short echo time ^1H MR spectra. *J Magn Reson.* 170, 164–175 (2004)
9. Lukas, L., Devos, A., Suykens, J.A., Vanhamme, L., Howe, F.A., Majos, C., Moreno-Torres, A., Van der Graaf, M., Tate, A.R., Arus, C., Van Huffel, S.: Brain tumor classification based on long echo proton MRS signals. *Artif Intell Med.* 31, 73–89 (2004)
10. Menze, B.H., Lichy, M.P., Bachert, P., Kelm, B.M., Schlemmer, H.P., Hamprecht, F.A.: Optimal classification of long echo time in vivo magnetic resonance spectra in the detection of recurrent brain tumors. *NMR Biomed.* 19, 599–609 (2006)

11. Simonetti, A.W., Melssen, W.J.: Combination of feature-reduced MR spectroscopic and MR imaging data for improved brain tumor classification. *NMR Biomed.* 18, 34–43 (2005)
12. Devos, A., Simonetti, A.W., van der Graaf, M., Lukas, L., Suykens, J.A., Vanhamme, L., Buydens, L.M., Heerschap, A., Van Huffel, S.: The use of multivariate MR imaging intensities versus metabolic data from MR spectroscopic imaging for brain tumour classification. *J Magn Reson.* 173, 218–228 (2005)
13. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. *IEEE Trans Syst Man Cybern. SMC* 3, 610–621 (1973)
14. Galloway, M.M.: Texture analysis using grey level run lengths. *Comp. Graph. and. Image Proc.* 4, 172–179 (1975)
15. Theodoridis, S., Koutroumbas, K.: *Pattern recognition.* Academic Press, New York (1999)
16. Specht, D.F.: Probabilistic neural networks. *Neural Networks* 3, 109–118 (1990)
17. Ahmed, N., Rao, R.: *Orthogonal transforms for digital signal processing.* Springer, Heidelberg (1975)