MRI-SPECT Data Fusion for Temporal Lobe Epilepsy Surgery Candidate Selection

Mostafa Ghannad-Rezaie[a], Hamid Soltanian-Zadeh[a,b], Kourosh Jafari-Khouzani[b], Mohammad-Reza Siadat[b], Reza A. Zoroofi[a], Kost V. Elisevich[c]

[a]*Department of Electrical and Computer Engineering,*
University of Tehran, Tehran 14395 Iran
m.ghannad@ece.ut.ac.ir, zoroofi@ut.ac.ir

[b]*Radiology Image Analysis Lab., Henry Ford Health System, Detroit, MI 48202 USA*
{hamids, siadat,kjafari}@rad.hfh.edu

[c]*Department of Neurosurgery Henry Ford Health System, Detroit, MI 48202 USA*
nskoe@neuro.hfh.edu

ABSTRACT
This paper presents a data fusion algorithm in a decision-support system to identify potential candidates for surgery in temporal lobe epilepsy. To this end, multimodality images including magnetic resonance imaging (MRI) and single photon emission computed tomography (SPECT) are used to predict surgery outcome. Effective features such as hippocampus structure and texture are extracted and combined to make reliable decisions. The experimental results using a support vector machine classifier show that the proposed approach may reliably predict the surgery outcome.

KEY WORDS
Data fusion, temporal lobe epilepsy, decision support systems, and MRI.

1. Introduction

Epilepsy is recognized as an organic process of the brain. More formally, epilepsy is an occasional, excessive, and disorderly discharge of nerve tissue, seizure, which sometimes can be detected by electroencephalographic (EEG) recording. It is a complex disease caused by a variety of pathological processes that result in treatment selection difficulties. Pharmacotherapy or surgical treatments are the neurologist alternatives. Selection of the most appropriate treatment could change the patient's life.

Despite optimal pharmacotherapy about 20–30% of patients do not become seizure-free [1]. For some of these patients, surgery is a therapeutic option. Success of resective surgery has been estimated to increase from 43% to 85% during the period 1986–1999 [1, 2]. A recent prospective randomized controlled trial of surgery for temporal lobe epilepsy showed that 58% of patients randomized to surgery were seizure-free compared to 8% of medical group [2].

Surgery is considered a valuable option for medically intractable epilepsy even in the absence of a proven drug resistance; in addition, surgical outcome may be greatly influenced by the presence of selected prognostic indicators [3]. However, there are still uncertainties on who are the best surgical candidates, i.e., those who most likely will present good surgical outcome.

In a recent narrative literature review of temporal resections, good surgical outcome was associated with a number of factors (hippocampal sclerosis, anterior temporal localization of interictal epileptiform activity, etc.) [2]. However, the published results were frequently confusing and contradictory, thus preventing inferences for clinical practice. Methodological issues were indicated by the authors as the most likely explanation of the conflicting literature reports [3].

For this reason, a quantitative review of the available literature has been undertaken in [2] to assess the overall outcome of epilepsy surgery and to identify the factors better correlating to seizure outcome. The aim of the study was to perform a meta-analysis of the results of published observational studies and assess the prognostic significance of selected variables outlining the characteristics of the clinical condition, the correlations between the epileptogenic and functional lesion, and the type of surgical procedure.

In such a complex problem, computer aided systems may help neurologists to make a more reliable decision. Using a database of other epilepsy cases, a soft computing algorithm may locate similar cases and with regard to previous experiences, propose a conclusive and supported suggestion for neurologist for upcoming cases. The most frequent factors among patients with similar surgery results are more likely to have effect on the decision.

Several modalities are now available for detecting the structural and the functional abnormalities of a seizure focus. This article presents techniques that can be used to integrate the data derived from different imaging modalities. In particular, an approach is described for integrating MRI and SPECT data for determining epilepsy surgery candidates.

The rest of the paper is organized as follows. Section II and III are devoted to data fusion and feature selection strategies. In Section IV, HBIDS and the proposed training method are described. Simulation results and comparison of alternative methods are presented in Section V. Finally, the paper is concluded in Section VI.

2. Visual Data Fusion for Decision Making

2.1. Temporal Lobe Epilepsy Surgical Outcome

The surgical outcome can be quite variable from case to case [4]-[5]. In most successful surgeries, the seizures completely disappear with non–disabling simple seizures during the first two years, and convulsions may appear only when medications are withdrawn. In some other cases, the primary seizure disappears but rarely some disabling seizures during the first two years may occur. Other patients may experience worthwhile seizure reduction and prolonged seizure-free intervals amounting to half of the follow-up period. In the worse case, there is no significant seizure reduction. Therefore, prediction of the success of surgery is quite important in deciding whether the surgery is the best treatment. The main contribution of this paper is to provide an image fusion approach to predict the usefulness of surgery. The first step towards this goal is to identify effective features.

Outcome after temporal lobectomy has been studied intensively in recent years. Reliable outcome data are the linch-pin on which surgical candidacy, patient counselling, and postoperative management are based, yet findings frequently conflict, making them difficult to interpret. This is not surprising, as research in this field is characterized by a variety of methodologic approaches. Accurate and effective interpretation of the body of knowledge regarding outcome after temporal lobectomy can be facilitated by an understanding of the approach to and practice of research in this area.

Findings for each risk factor or preoperative variable analyzed in each article can be aggregated into five major groups according to general subject. These comprised clinical factors, electroencephalography (EEG), preoperative magnetic resonance imaging (MRI), preoperative functional imaging such as SPECT, operative factors, and histopathology findings.

2.2. Data Fusion

In a decision support structure, clinical features are combined to reach the final conclusion. If the features have a common unit, called commensurate features in the data fusion literature, the combination can be accomplished by the traditional weight summation. For epileptic clinical feature combination, feature vectors have no common unit, thus in the literature, some feature vector concatenation algorithms have been proposed. To provide an explicit control over how much each vector contributes to the final decision, usually a weight vector is applied to the concatenated feature vector.

$$\overline{F} = \left[w_1 \overline{f_1} \vdots w_2 \overline{f_2} \vdots \cdots \vdots w_n \overline{f_n} \right] \quad (1)$$

The weight vector is selected subject to classification performance optimization [7]. In the current case, the weight vector is used to maximize between classes to within classes distance ratio:

$$\begin{cases} \overline{W} = \arg\max_{w}(Dis(\overline{\overline{F}}_1^{\overline{w}}, \overline{\overline{F}}_2^{\overline{w}})) \\ where: \overline{W} = [w_1, ..., w_n]; subject: \left\| \overline{W} \right\| = 1 \end{cases} \quad (2)$$

where F_i^w represents set of feature vectors in the *i*-th class weighted by W and *Dis* is between class to within class ratio function where the Euclidean distance is used as the feature space distance measure [7]. For example, the first class could be appropriate surgery candidates while the second one is non-surgical treatment preferred patients.

2.3. Hippocampus

In temporal lobe epilepsy, hippocampus is the key structure in MRI. An expert marked a region of interest (ROI) around hippocampus area using brain landmarks. ROI is a polygon with more than four vertices:

$$P^k : \{ p_i = (x_i, y_i) \mid i = 1..n, n > 4 \} \quad (3)$$

The mesial border of the ROI was chosen as a fixed landmark obtained by extending a straight line from the ventral most point of the circular sulcus to the upper portion of the parahippocampal gyrus. The lateral limit of the ROI extended to the centre of the temporal lobe white matter (Fig. 2-a).

To obtain the anatomical features, this polygon is used to generate the nearest ellipse covering hippocampus in each slice on each side (Fig. 2-b). Each ellipse is characterized by three parameters (a, b and θ):

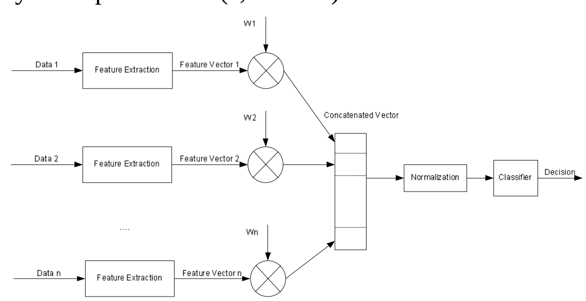


Fig. 1. Non-commensurate data fusion diagram.

$$E_{a,b,\theta} : (\frac{\cos^2\theta}{a^2} + \frac{\sin^2\theta}{b^2})\tilde{x}^2 + (\frac{\sin^2\theta}{a^2} + \frac{\cos^2\theta}{b^2})\tilde{y}^2 + 2\sin\theta\cos\theta(\frac{1}{a^2} - \frac{1}{b^2})\tilde{x}\tilde{y} = 1, \quad (4)$$

$$where \begin{cases} \tilde{x} = x - x_0 \\ \tilde{y} = y - y_0 \end{cases}$$

and x_0 and y_0 are the mean of the polygon points (Figure 3). A principle component analysis could obtain these parameters easily.

These ellipses represent 3D model of hippocampus. Lower number of parameters in this model eases shape encoding, feature extraction and information processing in difference patients. The parameters are normalized to the total volume of the brain.

3. Feature Selection and Classification

3.1. Features Selection

Based on a review of many articles, abnormal hippocampus anatomy in MRI, tissue properties, functional properties and difference between two epileptic sides are the main visual factors in surgery decision.
Actually, firm conclusions cannot be drawn for the extent of resection, EEG/MRI concordance, and post-operative discharges for the heterogeneity of study results [6].
In a previous work [11], we examined non-visual information fusion, where personal information, patient history, and diagnosis information were the main features. This paper is devoted to the imaging features.
Using the elliptic 3D-model, most important hippocampus anatomical features are calculated. These features include hippocampus total normalized surface area, average of absolute difference between *a* parameters and also *b* parameters in different slices and average of *a/b* ratio. In [12], a wavelet based hippocampus texture representation is described. According to the results, energy levels of D_{20} wavelet bases with two levels of decomposition are the most discriminative features (four features for each slice). In addition, entropy features described in [12] are used, but in the second level of importance. For functional features, averages of SPECT intensity and variance vector are considered. In this method, the mean intensities of the ictal and interictal SPECT images are normalized to a standard value. Finally, the energy features of the difference between the two hippocampi images are used to quantify the difference between the normal and abnormal sides.

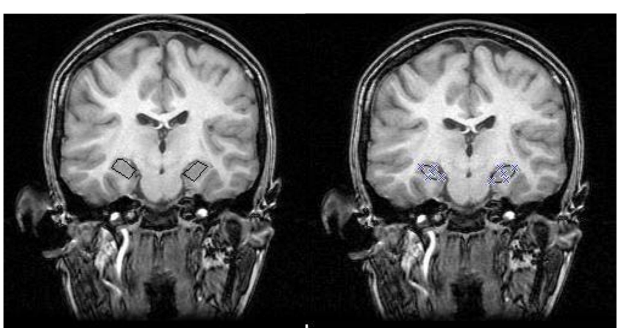


Fig. 2. A. Hippocampus polygon ROI is marked by an expert in the MRI-T1 coronal image. B. Nearest ellipse to ROI is obtained.

All of the features are concatenated using II.B process into a feature vector with 19 elements.

3.2. Feature Reduction

Obviously, the above described feature vector is too long. To find most important components, some features are combined using linear regression to reduce the feature vector dimension. Orthogonal least squared process is used to combine features [7] and extract the best five combined features.

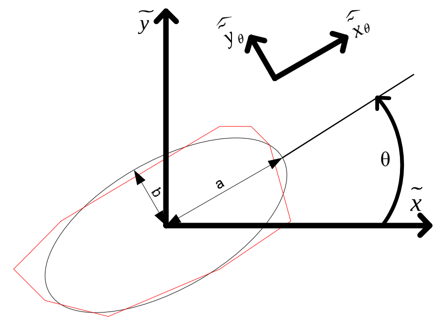


Fig. 3. Proposed algorithm finds the nearest ellipse to a polygon using principle algorithm anlysis.

3.3. Classification

Medical classification accuracy studies often yield continuous data based on predictive models for treatment outcomes. The sensitivity and specificity of a diagnostic test depends on more than just the "quality" of the test--they also depend on the definition of what constitutes an abnormal test. A popular method for evaluating the performance of a diagnostic test is the receiver operating characteristic (ROC) curve analysis [8]. ROC is a plot of the true positive rate against the false positive rate for the different possible cut-points of the classifier. Each point of the ROC curve is obtained by finding the true positive rate when the decision threshold is selected based on a specific false alarm rate.
The area under the ROC curve represents accuracy of a classifier. In medical problems, false alarm rate as well as false rejection rate should be lower that pre-specified limits. The trade off between false alarm rate and false rejection rate is problem specific.

4. Dataset Preparation

4.1. HBIDS

Human brain image database system (HBIDS) is under development for epilepsy patients at Henry Ford Health System, Detroit, MI [9, 10]. The proposed HBIDS will examine surgical candidacy among temporal lobe epilepsy patients based on their brain images and other data modalities. Moreover, it can discover relatively weak correlations between symptoms, medical history, treatment planning, outcome of the epilepsy surgery, and brain images. The HBIDS data include modalities such as MRI and SPECT along with patient's personal and medical information and EEG study [10]. The data has been de-identified according to HIPPA regulations [9].

For each patient, the database contains the MRI T1, T2, FLAIR, co-registered SPECT images, and 3D hippocampus model as well as non-visual information [10].
In this research, only the visual information has been used. Each patient's data is represented by a feature vector with five elements and a value that represents the surgery outcome (for the patients with only pharmacotherapy, surgery is assumed to be unnecessary). The features include hippocampus anatomical, texture and functional features. [6]. In some cases, patients' information is not complete. In the training phase, missing data are filled by the average of the other patients in the same class. In the testing phase, they are filled by the average of the entire available data. Thirty-five patients with temporal lobe epilepsy who have undergone temporal lobectomies at Henry Ford Health System are selected for the study. The initial pre-surgical evaluation of the epileptic patients includes history and neurological examination.

4.2. Training Method

For most efficient use of the data, training and test sets are not separated. In each training epoch, 4/5 of the patients are randomly selected to train the classifier. The rest of the patients (1/5) are used to test. The final classifier is the average of many training processes. This training strategy provides maximum database usage efficiency at the cost of higher computational complexity. In this experiment, more than 50 train-test sets are used. The training process terminates when the classifier's mean squared error of the test-set increases in the two last epochs.

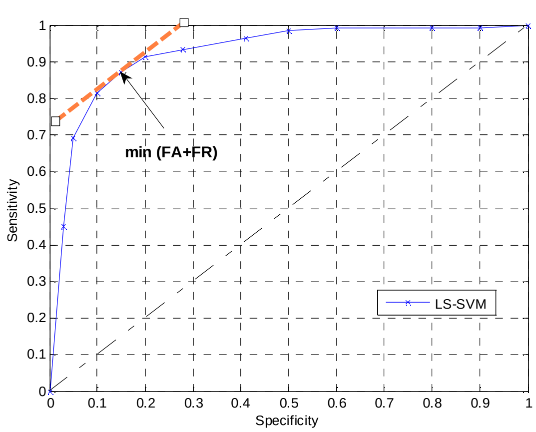


Fig. 4. Receiver-operating characteristic (ROC) curve plotting sensitivity versus specificity for different classifiers. The area under ROC curve for LS-SVM is 0.921.

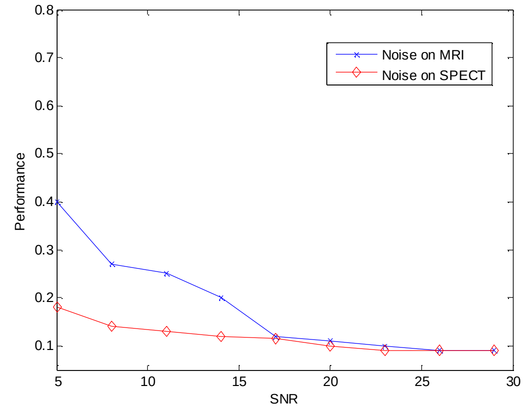
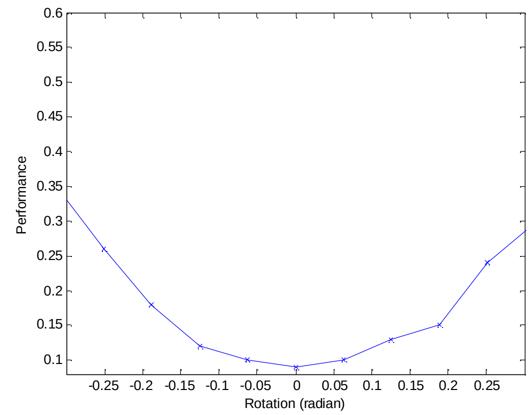
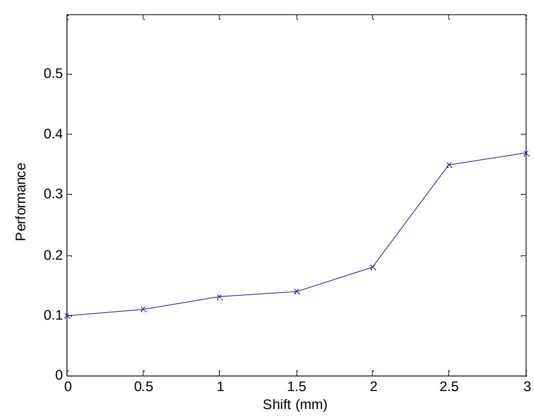


Fig. 5. Feature vector disturbance effect on classification error when the sum of false alarm and false rejection rates are minimal.



(A)



(B)

Fig. 6. Registration accuracy effect on performance; A. Rotation effect and B. shift effect.

5. Experimental Results and Discussions

Averages of six slices (range = 5 to 7) were used to calculate left and right temporal lobe WM FLAIR relaxation times (WM-FLAIR) and left and right hippocampal FLAIR relaxation times (Hippo-FLAIR). The study is conducted on 16 patients from the HBDIS datasets.

For each slice, FLAIR values were measured by placing its largest possible ROI within the anatomical boundaries of the hippocampus and temporal lobe WM on coronal slices.

In this research, a support vector machine (SVM) classifier is used. Training process is terminated in SVM network based on the testing curve to prevent over-training. To verify the classifier's accuracy, ROC curve are generated and shown in Fig. 4. In each case, four points of the ROC curves are calculated. The area under ROC curve for LS-SVM is 0.921. The lowest sum of false alarm and false rejection rates is around 10% classification error (Fig. 4).

Due to the usual artifacts in medical features, medical decision support algorithms should have good disturbance robustness. To this end, using the ROC curve, the decision threshold corresponding to the minimum summation of the false alarm and false rejection rates is chosen (Fig. 4). A Gaussian white noise is added to MRI and SPECT images. The error curves are obtained by averaging the results of 10 experiments (Fig. 5). Note that the white noise with higher that -15dB power may significantly decrease classifier performance. This noise affects texture and anatomical features. SPECT image noise did not destroy performance because the functional features are more robust against white noise and also there are less important in the final decision. Based on the above simulation results, SVM seems to be a good alternative for epilepsy prediction problem in high MRI SNR condition.

Finally, we evaluate the performance of the algorithm in the presence of mis-registration. Fig. 6 shows the decision system performance when the images are shifted or rotated against detected hippocampus. The algorithm remains stable under 0.05 rad rotation and 0.5 mm shift. However, the results degrade with larger rotations or shifts. Hence, although the proposed algorithm is not very sensitive to the registration accuracy, accurate registration is crucial for obtaining the best performance from the system.

6. Conclusion

In this paper, we have proposed a data fusion algorithm to predict temporal lobe epilepsy surgery outcome. This method concatenates MRI and SPECT believed to have strong contributions to the surgery outcome. To extract anatomical properties of hippocampus, a 3D elliptic model is proposed. The results show that the algorithm predicts whether the surgery is the best solution for a patient in more than 90% of the cases. In addition, the ROC analysis and feature vector distortion studies have shown that the SVM method is robust for this particular application.

7. Acknowledgment

The authors appreciate kind cooperation of Dr. Brien Smith and Dr. Farshad Fotouhi in HBIDS development, data collection, and data interpretation.

References:

[1] P. Kwan, M.J. Brodie, Early Identification of Refractory Epilepsy, *Neurology Journal in England*, *342*(5), 2000, 314–319.

[2] C. Tonini, E. Beghi, A.T. Berg, et al., Predictors of Epilepsy Surgery Outcome: a Meta-analysis, *Elsevier Transaction on Epilepsy Research*, *62*(1), 2004, 75-87.

[3] M.D. Holmes, D. E. Born, R. L. Kutsy, A.J. Wilensky , G.A. Ojemann, and L.M. Ojemann, Outcome after Surgery in Patients with Refractory Temporal Lobe Epilepsy and Normal MRI, *Journal on Seizure*, *9*(1), 2000, 407-411.

[4] C. Kilpatrick, T. O'Brien, Z. Matkovic, Preoperative Evaluation for Temporal Lobe Surgery, *Journal of Clinical Neuroscience*, *10*(5), 2003, 535-539.

[5] B.P. Hermann, B. Connell, W.B. Barr, and A.R. Wyler, The Utility of the Warrington Recognition Memory Test for Temporal Lobe Epilepsy: Pre- and Postoperative Results, *Journal of Epilepsy*, *8*(2), 1995, 139-145.

[6] J. Janszky, R. Schulz, and A. Ebner, Clinical Features and Surgical Outcome of Medial Temporal Lobe Epilepsy with a History of Complex Febrile Convulsions, *Elsevier Transaction on Epilepsy research*, *55*(1) 2003, 1-8.

[7] R.O. Duda, P.E. Hart, and D.G. Stork, Pattern Classification. (John *Wily and Sons Inc.*, USA, 2001).

[8] K.H. Zou, F.S. Resnic, I. Talos, et al., A Global Goodness-of-fit Test for Receiver Operating Characteristic Curve Analysis via the Bootstrap Method, *Journal of Biomedical Informatic*s, available online on www.sciencedirect.com, 2005.

[9] M. Siadat, H. Soltanian-Zadeh, F. Fotouhi, and K. Elisevich, Multi-modality Medical Image Database for Temporal Lobe Epilepsy, *SPIE Medical Imaging Conference*, 2003, 487-498.

[10] M. Siadat, H. Soltanian-Zadeh, F. Fotouhi, K. Elisevich, Multimedia Human Brain Database System for Temporal Lobe Epilepsy with Content-Based Management, Navigation, and Retrieval, *BISTIC-BICON-NIH Biomedical Informatics for Clinical Decision Suppor*t Conference, 2004, 98-102.

[12] M. Ghannad-Rezaie, H. Soltanian-Zadeh, M R. Siadat, K.V. Elisevich, Soft Computing Approaches to Computer Aided Decision Making for Temporal Lobe Epilepsy, *To Be Presented in IEEE International Conference on Fuzzy Systems (NAFIPS-2005)*, , 2005.

[13] K. Jafari-Khouzani, M. R. Siadat, H. Soltanian-Zadeh, K. Elisevich, Texture Analysis of Hippocampus for Epilepsy, *Proceeding of SPIE*, 2003, 279-288.