Soft Computing Approaches to Computer Aided Decision Making for Temporal Lobe Epilepsy

M. Ghannad Rezaie[a], H. Soltanian-Zadeh[a,b], M.-R. Siadat[b], K.V. Elisevich[c]

[a]*Department of Electrical and Computer Engineering, University of Tehran, Tehran 14395, Iran*

[b]*Radiology Image Analysis Lab., Henry Ford Health System, Detroit, MI 48202, USA*

[c]*Department of Neurosurgery, Henry Ford Health System, Detroit, MI 48202,USA*

Emails: m.ghannad@ece.ut.ac.ir, hamids@rad.hfh.edu, siadat@rad.hfh.edu, nskoe@neuro.hfh.edu

***Abstract* – This paper presents and compares soft computing approaches for prediction of surgery outcome in temporal lobe epilepsy. Because of a wide range of effective parameters in epilepsy and unclear exact contribution of each, determination of the best treatment is difficult. We have implemented and compared data fusion methods and decision support algorithms to overcome this difficulty. Our simulation studies and experimental results using HBIDS (Human Brain Image Database System) data show the power of LS-SVM (Least Squared Support Vector Machine) classifiers for this purpose.**

I. 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 epilepsy surgery has been estimated to increase from 43% to 85% during the period 1986–1999 [2]. Data from multiple sources suggest that 55–70% of patients undergoing temporal resection and 30–50% of patient undergoing extra-temporal resection become completely seizure-free [1]. A recent prospective randomized controlled trial of surgery for temporal lobe epilepsy showed that 58% of patients randomized to surgery was 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, absence of preoperative generalized seizures, and absence of seizures in the first post-operative week) [2]. However, the published results were frequently confusing and contradictory, thus preventing inferences for clinical practice. Methodological issues (e.g., sample size, selection criteria, and methods of analysis) 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 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. In this paper, a computerized treatment decision support algorithm for temporal lobe epilepsy is presented. Also, popular decision support algorithms such as support vector machines, Bayesian network, and nearest neighbour have been compared.

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

II. Data Fusion for Decision Making

A. *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 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 a data fusion approach to evaluate the usefulness of surgery by comparing patients with similar medical and clinical conditions. The first step towards this approach is to identify effective features.

B. *Features Selection*

Based on the review of many articles in [2] meeting all the eligibility criteria, febrile seizures, mesial temporal sclerosis, tumours, abnormal MRI, EEG/MRI concordance, and extensive surgical resection were the strongest prognostic indicators of seizure remission (positive predictors); by contrast, postoperative discharges and intracranial monitoring predicted an unfavourable prognosis (negative predictors).

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].

We examined pre-operative interictal and ictal electroencephalographic (EEG) findings, age of onset, gender, duration of epilepsy, risk factors, family history, physical findings, age at operation, side of operation, and pathology of resected tissue in order to determine if any of these factors were associated with outcome. All of the features are combined in a decision support system (see Fig. 1). In the following section, some details of this structure are illustrated.

C. *Feature Vector Construction and 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.

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

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:

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

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].

D. *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 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. In surgery decision-making problem, both rates must be considered; however, false alarm rate (plan surgery for a patient who does not need it) is more likely to be of concern. In this paper, performance of decision systems is evaluated using the sum of false alarm and false rejection rates.

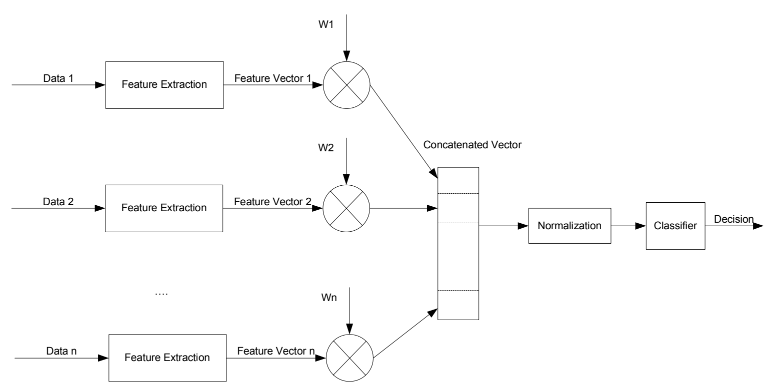

Fig. 1. Non-commensurate data fusion diagram.

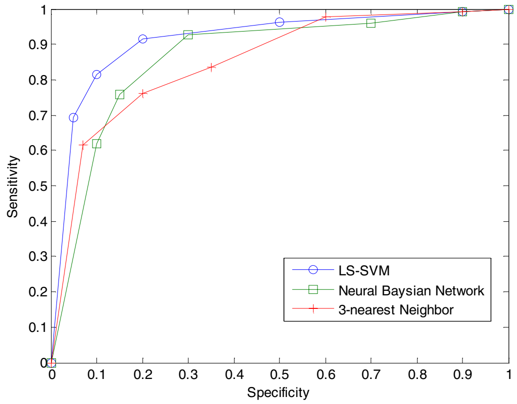

Fig. 2. Receiver-operating characteristic (ROC) curves plotting sensitivity versus specificity for different classifiers. The area under ROC curve for LS-SVM, Bayesian network and 3-nearest network are 0.913, 0.876 and 0.855, respectively.

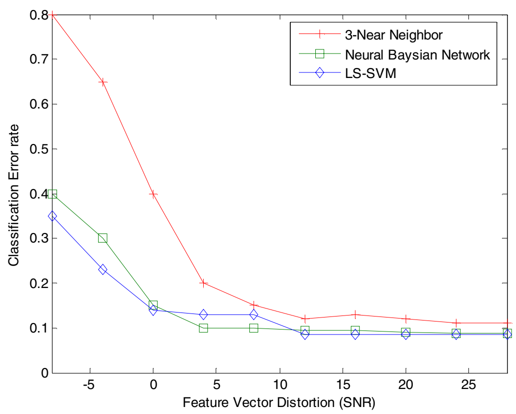

Fig. 3. Feature vector disturbance effect on classification error when the sum of false alarm and false rejection rates are minimal.

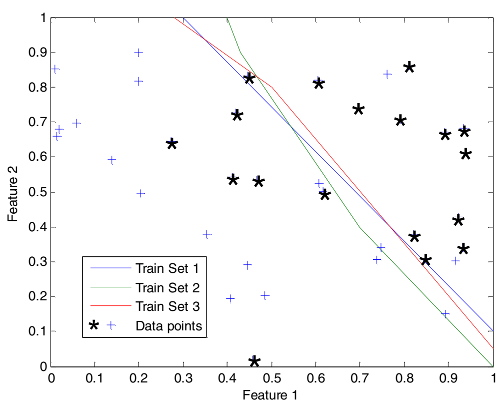

Fig. 4. SVM classifier in feature #1-feature #2 normalized space.

III. Simulation Setup and Database Description

A. *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 the first phase of the EEG study, the non-visual feature extractor is an expert or specialist. The experts do this routinely in the clinic based on well-defined standards. For un-structured text information, the wrapper is the expert or trained nurse. The structured data such as patient's personal information do not need to be analyzed by the wrapper, so they are directly stored in the database.

For each patient, the database contains the personal information, such as sex and age, diagnostic information, such as seizure type and epilepsy location, EEG result, suggested treatment and surgery outcome as well as visual information [10].

In this research, only the non-visual information of each patient has been used. Each patient's data is represented by a concatenated feature vector with seven elements and a value that represents the surgery outcome (for the patients with only pharmacotherapy, surgery is assumed to be unnecessary). The features include age, sex, weight, genetic background, abnormality of EEG, seizure type diagnosis, epilepsy location, tumour diagnosis, previous treatment history, and surgery decision. Some of the factors such as age and sex have been combined to present a single feature (weighted fusion). Also a ranking table has been used to encode medical and drug history information [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.

B. *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. The train and test vectors are normalized classification.

IV. Simulation Results and Discussions

In this research, three conventional classifiers used in decision support systems have been compared (k-nearest neighbor [7], Bayesian network [11], and support vector machine (SVM)). Primary simulations of the k-nearest neighbor algorithm for k = [1…5] show that the best results are obtained using k=3. For the SVM, RBF implementation of least square (LS-SVM) has been applied [12]. Training process is terminated in SVM and Bayesian network based on the testing curve to prevent over-training.

To compare the classifiers' accuracy, ROC curves for the three classifiers are generated and shown in Fig. 2. In each case, four points of the ROC curve are calculated. The area under ROC curve for LS-SVM, Bayesian network and 3-nearest network are 0.913, 0.876 and 0.855, respectively. The results show high sensitivity of 3-nearest neighbour. In contrast with SVM, the nearest neighbour algorithm depends on a few feature vectors, which are in the vicinity of the target point. Thus, it has high sensitivity, especially in a problem with a small number of samples.

Due to the usual artifacts in medical features, medical decision support algorithm should have good disturbance robustness. Although some features such as sex and age are naturally artifact-free, the performance of the algorithms can be presented in terms of the signal to artifact or signal to noise ratio (see Fig. 3). 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. The same training set is used for all classifiers. The error curves are the average of the results from 10 training processes. In high error conditions, SVM has the best performance. Also in some middle conditions, Bayesian network has shown a better performance, but only in a small interval.

Finally, SVM generalization capability is significant. Fig. 4 compares SVM classifier surface projection onto feature1-feature2 space for three different training sets. Closeness of these curves supports the SVM generalization. Based on the above simulation results, SVM seems to be the best alternative for epilepsy prediction problem among the methods studied.

In the future work, more complex features, such as MRI and SPECT images, will be used to reach more accurate decisions. Also, the authors plan to predict other important indices such as surgery risk and the best surgery time by utilizing neuro-fuzzy classifiers and expert systems. To this end, more advanced classifier evolutions may be required [11].

V. Conclusion

In this paper, we have proposed a data fusion algorithm to predict temporal lobe epilepsy surgery outcome. This method concatenates clinical date believed to have strong contributions to the surgery outcome. The simulation of the weighted vector concatenating data fusion and LS-SVM classifier on HBIDS has shown that the algorithm predicts whether the surgery is the best solution for a patient in more than 90% of the cases. Also, the ROC analysis and feature vector distortion studies have shown that the SVM method is more reliable that the other classifiers compared for this particular problem.

Acknowledgment

The authors appreciate kind cooperations of Dr. Brien Smith, Dr. Farshad Fotouhi, and Mr. Kourosh Jafari-khouzani in HBIDS development, data collection, and data interpretation.

References

[1] P. Kwan, M.J. Brodie, "Early Identification of Refractory Epilepsy," *Neurology J. in England*, vol. 342, issue 5, pp. 314–319, 2000.

[2] C. Tonini, E. Beghi, A.T. Berg, et al., "Predictors of Epilepsy Surgery Outcome: a Meta-analysis," *Elsevier Trans. on Epilepsy research*, vol. 62, pp. 75-87, 2004.

[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," *J. on Seizure*, vol. 9, pp. 407-411, 2000.

[4] C. Kilpatrick, T. O'Brien, Z. Matkovic, "Preoperative Evaluation for Temporal Lobe Surgery," *J. of Clinical Neuroscience*, vol. 10, issue 5, pp. 535-539, 2003.

[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," *J. of Epilepsy*, vol. 8, Issue 2, pp. 139-145, 1995.

[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 Trans. on Epilepsy research*, vol. 55, pp. 1-8, 2003.

[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," *J. 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 conf.*, pp. 487-498, 2003.

[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, pp. 98-102, 2004.

[11] J. Huang, j. Lu, and C.X. Ling, "Comparing naive Bayes, decision trees, and SVM with AUC and accuracy," *IEEE International Conf. on Data Mining*, pp. 553-556, 2003.

[12] C. Lu, T. Van Gestel, J.A.K. Suykens, S. Van Huffel, I. Vergote, and D. Timmerman, "Preoperative Prediction of Malignancy of Ovarian Tumors using Least Squares Support Vector Machines," *Artificial Intelligence in Medicin*e, vol. 28, issue 3, pp. 281-306, 2003.