Detection of Seizures in EEG Signal using Weighted Locally Linear Embedding and SVM Classifier

Yaozhang Pan, Shuzhi Sam Ge*
and Abdullah Al Mamun
Social Robotics Lab
Interactive Digital Media Institute
National University of Singapore
Singapore, 119077
E-mail: {yaozhang.pan,samge,eleaam}
@nus.edu.sg

Feng Ru Tang
Epilepsy Research Lab
National Neuroscience Institute
Singapore, 308433
email: Feng_Ru_Tang@nni.com.sg

***Abstract*—To diagnose the structural disorders of brain, electroencephalography (EEG) is routinely used for observing the epileptic seizures in neurology clinics, which is one of the major brain disorders till today. In this work, we present a new, EEG-based, brain-state identification method which could form the basis for detecting epileptic seizure. We aim to classify the EEG signals and diagnose the epileptic seizures directly by using weighted locally linear embedding (WLLE) and support vector machine (SVM). Firstly, we use WLLE to do feature extraction of the EEG signal to obtain more compact representations of the internal characteristic and structure in the original data, which captures the information necessary for further manipulations. Then, SVM classifier is used to identify the seizures onset state from normal state of the patients.**

***Index Terms*—seizures detection, weighted distance measurement, locally linear embedding, weighted locally linear embedding**

I. Introduction

Epileptic seizures are manifestations of epilepsy, a serious brain dynamical disorder and as one of the most widespread neurological disease suffered by approximately 1 % of the world population. This disorder is characterized by sudden recurrent and transient disturbances of mental function and movements of the body that result from abnormal synchronized firing of a large number of neurons. The large synchronized event is known as a seizure, paroxysmal discharge, or ictal event.

EEG is an important clinical tool for diagnosing, monitoring and managing neurological disorders related to epilepsy. The presence of epileptiform activity in the EEG confirms the diagnosis of epilepsy. During seizures, the scalp EEG of patients with epilepsy is characterized by high amplitude synchronized periodic EEG waveforms, reflecting abnormal discharge of a large group of neurons. Between seizures, epileptiform transient waveforms, which include spikes and sharp waves, are typically observed on the scalp EEG of such patients.

Initially, the epilepsy research community shared low confidence in the prediction of seizures cause that seizures were highly random phenomena without any prior indication of occurrence. Over time, many research groups have observed clinical symptoms and quantitative measures that foreshadow seizure onsets. Prospective studies on prediction of epileptic seizures have been published in peer-reviewed journals to date. These methods include time-domain analysis of EEG signal by statistical analysis and characteristics computation [1]–[3], frequency-domain analysis by decomposing the EEG signal into components of different frequencies [4], [5], non-linear dynamics and chaos theory [6], [7], Lyapunov exponent [8] and intelligent systems such as artificial neural network (ANN) and other artificial-intelligence structures [9].

In this work, we will take an approach called locally linear embedding (LLE) and Weighted Locally Linear Embedding (WLLE), which is developed to do the internal feature extraction. LLE is an unsupervised learning algorithm that computes low-dimensional, neighborhood-preserving embedding of high-dimensional inputs [10]. Unlike clustering methods for local dimensionality reduction, LLE maps its inputs into a single global coordinate system of lower dimensionality, and its optimizations do not involve local minima. By exploiting the local symmetries of linear reconstructions, LLE is able to learn the global structure of data set.

The weighted locally linear embedding (WLLE) is an extension algorithm we have developed especially for signals processing of deformed distribution. Deformation of a distribution may caused by attraction, repulsion, strengthening effect and weakening effect between data. In this case, standard Euclidean metric is not good enough for measuring the similarity of two data. As mentioned in [10], LLE cannot give good performance for the data set which is not sufficient and well-sampled. This may be mainly caused by the limitation of the distance measures, Euclidean metric, utilized by LLE. To combat this problem, we construct a new distance measure for similarity measurement of data in deformed distribution, which can improve the neighbor selection procedure by avoiding data overlapping and isolation. Based on the weighted distance measurement, we propose WLLE, an unsupervised learning algorithm for preserving neighborhood relationship, clustering input space, and extracting the features clusters simultaneously for input data. The extracted features will then serve as the possible inputs to the pattern recognition and classification algorithm.

SVM classifier is an unsupervised approach. An unsupervised approach allows for uniform treatment of seizure detection and prediction, and offers many advantages for implementation [2]. In this paper, we use a simple SVM classifier as standard classification to demonstrate the detection performance of our proposed algorithm.

The rest of the paper is organized as follows. The experiment method for recording the long term EEG signal from mice is described in Section II. Then the preprocessing of the raw EEG data, such as labeling and rearranging are presented in Section III. Detailed procedures of WLLE are demonstrated in Section IV. In Section V, simple support vector machine classifier is introduced to solve the linear separable problem. The results of simulation studies are presented in Section VI, which demonstrate great advantages of the feature extraction algorithm WLLE. At last, conclusion is given in Section VII.

II. Obtaining EEG Data Sets

Male Swiss mice weighing 25-30g were used for the study according to our established procedures [11], [12]. Mice were given a single subcutaneous injection of methyl-scopolamine nitrate (1mg/kg) 30 min before the injection of either saline in the control or pilocarpine in the experimental group. In the latter group, the mice received a single i.p. injection of 300mg/kg pilocarpine and experienced acute status epilepticus (SE). All experiments were approved by the Tan Tock Seng Hospital, National Neuroscience Institute, Institutional Animal Care and Use Committee. In the handling and care of all animals, the guidelines for animal research of NIH were strictly followed. Efforts were made throughout the study to minimize animal suffering and to use the minimum number of animals.

For EEG and video monitoring, 14 experimental mice at 2 months after inducing pilocarpine and 2 age-matched control mice were used. EEG data were recorded continuously (24 hrs/day) using TSE EEG Telemetry System (TSE Technical and Scientific Equipment GmbH, Germany) according to the protocols used in our previous study [11]. In brief, two leads of the transmitter were connected to two tiny screws which were fixed on the scull 2.3mm posterior the bregma and 2mm lateral to the midline, and then consolidated with cyanoacrylate and dental acrylic cement (Boswarth Company, USA) under deep anesthesia with hydrochloride (40mg/kg). The transmitter has a built in 1000-times amplification capacity with input voltage range of $3\mu V$ to $0.7mV$, and its frequency ranges from 0.4 to 60Hz. The signal from transmitter was transmitted wirelessly to the receiver which was then digitized (sampling frequency 50Hz) and stored in a personal computer. EEG and digital video acquisition monitoring (Chateau digital surveillance network system, Chateau Technical Corp, Singapore) were done continuously 24 hrs/day, and all data were stored in computer for further processing and analysis.

III. Data Preprocessing

All recorded EEG data were first assessed manually which contains 2 steps: the observation of the raw EEG signal; and the careful watching of the subjects appearance on simultaneous recorded video. For the video assessment, we watched the entire recorded video and marked the EEG recording time where we founded any obvious behavior changes which can be ranked as stages 4 and 5 according to revised criteria of kindling model [13]: rearing (stage 4) and falling or lose of posture (stage 5). Through the manual observation, we confirmed about the obvious seizures onset in the EEG signal and the raw EEG data were labeled according to the time and duration of those epilepsy occurrences.

After assessing and labeling the raw EEG data manually, we preprocess the raw data for further operations, such as feature extraction and classification. To evaluate most features, it is important to maintain stationarity of the data segment. Statistical tests reveal quasi-stationarity of the EEG signal anywhere from 1s to several minutes [14]. Since seizures spread so quickly, a displacement as small as possible that does not provide too much variability is desired. Such the raw data are divided into consecutive 1 s epochs. The epoch-divided sections of each channel of the EEG signal $x(n)$, $n = 1, 2, \cdots, (N-1)O + D$ were arranged as the columns (pattern vectors) of an $D \times N$ matrix X, where D is the dimension (number of samples) of each pattern vector, N is the number of patterns, and O is the delay between patterns.

$$X = \begin{bmatrix} x(1) & x(O+1) & \cdot & \cdot & x[(N-1)\cdot O+1] \\ x(2) & x(O+2) & \cdot & \cdot & x[(N-1)\cdot O+2] \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x(D) & x(O+D) & \cdot & \cdot & x[(N-1)\cdot O+D] \end{bmatrix} \quad (1)$$

Motivated by analysis on stationarity and redundancy in [14], we divide the EEG signal into 1 s epoches, say, $D = 50$ since the sampling rate is 50 Hz. We choose $O = N/2$ (overlap of half of the samples between consecutive pattern vectors). Such short and overlapped windows ensure that all transient events will be completely represented and dominant in at least one of the patterns, yet they are long enough to characterize the main rhythms of the on-going EEG signal.

IV. Feature Extraction of EEG Data by Weighted Locally Linear Embedding

A number of promising feature extraction methods, each with different theoretical bases, have demonstrated usefulness in seizure prediction. Iasemidis and Sackellares applied nonlinear dynamical techniques, principal Lyapunov exponent (PLE), for prediction seizures beginning [15]. Some considered exhaustive search and genetic approach for feature selection [16], however, the optimal results can only obtained by testing 850 features using genetic algorithm and 4300 features using the exhaustive search, the computational complexity is really a problem. In this paper, we proposed an unsupervised learning algorithm, WLLE, for internal feature extraction of the EEG signal. The WLLE algorithm is motivated by LLE algorithm.

Based on the weighted distance measurement, we modify the LLE algorithm to form the new algorithm, WLLE, to improve the dimension reduction and internal feature extraction performance especially for the deformed distributed data. By taking into account the distribution information surrounding each prototype to optimize the distance measure, we can improve the neighbor finding procedure of LLE algorithm and avoid the redundancy, overlapping and isolation due to improper neighbor selection. Better neighbors selection will make the dimension reduced representations more accurate to represent the internal feature, characteristic, and structure of the high dimension data.

A. Weighted Locally Linear Embedding

The assumption behind manifold learning algorithms is that the data points $X = \{x_1, x_2, \ldots, x_N\}$ in the high-dimensional space R^D are generated by varying a few parameters over a smooth low-dimensional manifold.

For data points $X = \{x_1, x_2, \ldots, x_N\}$ in the high dimensional space R^D, the goal of dimension reduction for feature extraction is to calculate a low dimensional embedding of the data into R^d where $d << D$.

We attempt to express data point number x_i as a linear combination of its k nearest neighbors $x_j,\ j = 1, 2, \ldots, k$.

$$\hat{x}_i = \sum_{j \in \Omega_i} w_{ij} x_j, \tag{2}$$

where Ω_i is the neighborhood of sample x_i. In the original LLE algorithm standard Euclidean metric based KNN is used to select the nearest neighbors. However, in this work, we utilize the weighted distance measures in order to improve feature extraction performance.

B. Neighborhood Selection Based on Weighted Distance Measurement

In this paper, we propose a novel dimension reduction algorithm, weighted locally linear embedding (WLLE) which use weighted distance measurement to improve the dimension reduction and internal feature extraction performance especially for the deformed distributed data.

1) Deformed Distribution: In the data manipulation like nearest neighbor searching, each datum can be regarded as the center of a probability distribution and the similarity of its neighbors to the datum can be measured by Euclidean distance with the assumption that samples are well-distributed. However, because of the attraction, repulsion, strengthening effect and weakening effect between data, the standard normal distributions will be greatly deformed. Obviously, neglecting such a deformation and still using the standard Euclidean distance to measure the similarity will lead to performance decline. As mentioned in [17], the data set should be sufficient and well-sampled, otherwise the performance of LLE algorithm will not be good enough. For example, as illustrated in Fig. 1, the samples are not well-distributed, data density changes sharply within a small area, the query point is marked by a cross, and its neighbors marked by circles. We use ϵ-neighborhoods algorithm to finding nearest neighbors of the query point from its neighbors. For this deformed distribution data set, ϵ-neighborhoods method based on standard Euclidean distance measurement selects neighbors from a single direction, and these neighbors are closely gathered. Obviously, if we use these chosen neighbors to reconstruct the query point, the information captured in this direction will have serous redundancy; at the same time, no information from other directions are reserved for query point reconstruction. These chosen neighbors cannot represent and reconstruct the query point well, most internal features and intrinsic structure will be lost after dimension reduction by LLE.

To solve this problem, we introduce the weighted distance measurement motivated by [18]. The main idea of the weighted distance measurement is giving a different but appropriate distance scale to each prototype to make the distance measure more reasonable for representing the global distribution of the data set. Fig. 1 shows the advantages of this scaled adaptive distance measurement. The previous redundancy and deficiency problem can be solved.

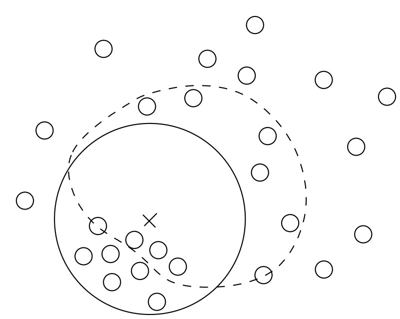

Fig. 1. Select nearest neighbors using ϵ-neighborhoods algorithm by Euclidean distance (solid line) and weighted distance (dash line).

To find the suitable adaptive weight scaling, we first construct a simple yet effective transformation to simulate the possible deformation of data distribution. We would demonstrate the definition of deformed distribution.

Definition 1: (Deformed Distribution) Consider a d-dimensional random vector $Y = (Y_1, Y_2, \ldots, Y_d)^T$ that takes a standard D dimensional normal distribution $N(0, I)$, that is, it has a probability density function

$$f(y) = \frac{1}{(2\pi)^{d/2} e^{1/2 y^T y}} \tag{3}$$

Let a random vector X be defined by the transformation

$$X = (a + b\frac{Y^T \tau}{\|Y\|})Y \tag{4}$$

where Y denotes the original well-distributed data set, $a > b \geq 0$ are the parameters reflect overall scale and orientation of distribution, τ is a normalized vector denoting the deformation orientation, $\|Y\| = \sqrt{Y^T Y}$, and X represent the deformed distribution with parameters a and b in the direction τ, denoted as $X = D_d(a, b, \tau)$ [18].

2) Weighted Distance Measurement: As mentioned above, a deformed distribution can be viewed as an eccentric distribution that biases towards a specific direction. It can be obtained from a standard normal distribution by the transformation $X = Y(a + b\cos\theta)$. In this case, the Euclidean distance is not suitable to describe the similarity directly, since the data are not well-distributed. Instead, we firstly restore the deformation by an inverse transformation $Y = X/(a + b\cos\theta)$, and then measure the distance. We call this weighted distance measurement. This weighted distance redresses the deformation and should be more suitable to describe the similarity for data set that are not well-distributed.

Definition 2: (Weighted Distance) Assume that $x_0 \in R^d$ is the center of a deformed distribution $D_d(a, b, \tau)$. The weighted distance from a point $x \in R^d$ to x_0 is defined to be

$$\text{Dist}(x_0, x) = \frac{\|x - x_0\|}{a + b\frac{(x-x_0)^T\tau}{\|x-x_0\|}} \tag{5}$$

or

$$\text{Dist}(x_0, x) = \|x - x_0\|/(a + b\cos\theta) \tag{6}$$

where θ is the angle between vectors $x - x_0$ and τ, and $\frac{1}{a+b\cos\theta}$ is the weight of the distance from x to x_0 [18].

One disadvantage of the weighted distance measurement is just a weighted distance, but not a metric, since $\text{Dist}(x_0, x)$ may not equal to $\text{Dist}(x, x_0)$ under the definition of weighted distance.

3) Parameter Estimation for Weighted Distance: To facilitate parameter estimation for weighted distance, we first present some properties.

Theorem 1: If random vector $X = D_d(a, b, \tau)$, then $E(X) = c_1 b\tau$ and $E(\|X\|) = c_2 a$, where c_1 and c_2 are constants.

$$c_1 = 2^{1/2}\frac{\Gamma((d+1)/2)}{\Gamma(d/2)d} \tag{7}$$

$$c_2 = 2^{1/2}\frac{\Gamma((d+1)/2)}{\Gamma(d/2)} \tag{8}$$

where Γ is the Gamma function $\Gamma(k) = \int_0^\infty t^{k-1}e^{-t}dt$, $(k > 0)$ [18].

For an arbitrary sample $x_i \in D$, we assume that it represents a deformed distribution and is the origin of this deformed distribution. Then, we use its k-nearest neighbors $X_i = \{x_{i1}, x_{i2}, \ldots, x_{ik}\}$ to estimate the parameters of the deformed distribution, that is, a_i, b_i and τ_i.

First, we calculate the difference between a sample and all its k-nearest neighbors, $V_i = \{v_{i1}, v_{i2}, \ldots, v_{ik}\}$, where $v_{ij} = x_{ij} - x_i$, $j = 1, 2, \ldots, k$. Then, we use $\hat{G}_i$ and $\hat{L}_i$, which are the center of mass and the averaged vector length of V_i:

$$\hat{G}_i = \sum_{j=1}^{k} v_{ij}/k, \qquad \hat{L}_i = \sum_{j=1}^{k} \|v_{ij}\|/k \tag{9}$$

to estimate $E(X)$ and $E(\|X\|)$, respectively. According to Theorem 1, we obtain an estimation to a_i, b_i and τ_i:

$$\hat{a}_i = \frac{\hat{L}_i}{c_2}, \qquad \hat{b}_i = \frac{\|\hat{G}_i\|}{c_1}, \qquad \hat{\tau}_i = \frac{\hat{G}_i}{\|\hat{G}_i\|} \tag{10}$$

C. Optimization of Reconstruction Weights

After properly select the neighbors, the next step of WLLE is to obtain the optimal reconstruction weights for each neighbor point. The optimal weight matrix w_{ij} for data reconstruction can be obtained by minimizing the approximation error cost function

$$\epsilon(W_i) = \sum_i d_W\left(x_i, \sum_{j\in\Omega_i} w_{ij}x_j\right)^2, \tag{11}$$

subject to the constraints

$$j \notin \Omega_i \Rightarrow w_{ij} = 0 \tag{12}$$

$$\sum_{j\in\Omega_i} w_{ij} = 1, \tag{13}$$

where $w_i = [w_{i1}, \ldots, w_{ik}]$ is the weights connecting sample x_i to its neighbors. The function $d_W(\cdot,\cdot)$ is an appropriate distance measure. The first constraint says that only data points in the neighborhood of data point x_i should be used in the reconstruction of $\hat{x}_i$, while the second constraint imposes invariance to translation.

D. Mapping to Low-dimensional Embedding

The final step of WLLE is to compute a low dimensional embedding of the high dimensional inputs x_i based on the reconstruction weights w_{ij}. The high dimensional data are mapped into the low dimensional space R^d by requiring reconstruction to work as well as possible. This leads to another minimization problem [19], [20]. The low dimensional outputs y_i, $i = 1, 2, \ldots, N$ are found by minimizing the cost function,

$$\Phi(Y) = \sum_i d_W\left(y_i, \sum_{j\in\Omega_i} w_{ij}y_j\right)^2, \tag{14}$$

where $Y = [y_1, \ldots, y_N]$ consists of the data points embedded into the low dimensional space. This minimization problem is not well-posed without further constraints. Zero mean and unity covariance is used in the LLE algorithm to make the problem well-posed. In other words Y should obey the constraints

$$\sum_{i=1}^{N} y_i = \mathbf{0} \tag{15}$$

$$\frac{1}{N}YY^T = \mathbf{I}, \tag{16}$$

where $\mathbf{0}$ is a vector with all elements being zero, and $\mathbf{I}$ is an identical matrix. The first constraint is to assure that coordinates y_i can be translated by a constant displacement without affecting the cost, while the second constraint imposes unit covariance of the embedding vectors.

V. Support Vector Machine Classifier

SVM are widely used for learning classifiers and regression models. Its theoretical support is from statistical learning theory. The SVM empirically works very well, at least for some classes of problems, provide a good generalization performance on pattern classification problems despite the fact that it does not incorporate problem-domain knowledge.

The objective of SVM is to find an optimal hyperplane that correctly classifies points as much as possible and separates the points of two classes as far as possible. Given the training sample $\{x_i, d_i\}_{i=1}^N$, where $x_i \in R^m$ is the i^{th} sample and $d_i \in \{-1, 1\}$ is the corresponding desired output. Using a discriminant function

$$g(x) = w^T x + b \tag{17}$$

and choose a weight vector $w \in R^m$ and bias $b \in R$ such that

$$\begin{cases} g(x_i) = w^T x_i + b \geq 0, for\ d_i = +1 \\ g(x_i) = w^T x_i + b < 0, for\ d_i = -1 \end{cases} \tag{18}$$

The constrained optimization problem can be defined as follows,

Definition 3: Given the training sample $x_i, d_{i}{}_{i=1}^N$, find the optimum values of the weight vector w, and bias b such that they satisfy the constrains

$$d_i g(x_i) = d_i(w^T x_i + b) \geq 1, for\ I = 1, 2, \ldots, N \tag{19}$$

and the weight vector w minimizes the cost function

$$\Phi(w) = \frac{1}{2} w^T w \tag{20}$$

Using the method of Lagrange multipliers, the optimization problem can be solved and the solution for w can be obtained by

$$w = \sum_{i=1}^{N} \alpha_i d_i x_i \tag{21}$$

And the bias term can be solved to be $b = d_i - w^T x_i$.

VI. Simulation Study

In this section we present the simulation study for demonstrating the performance of seizure detection approach we proposed. Details on the effects of different feature extraction methods and classification methods are given below.

Long term EEG signal (24hrs), recorded from both normal mouse and pilocarpine treated mouse who are freely moving were used to test our seizure detection approach. The sampling rate is 50Hz, and it is large quantity of data in a whole day's record. We choose a couple of epoches contains both normal EEG and epileptic EEG, and label them by experts based on the observation of EEG signal and the video records.

In this simulation, we select two 600 sec (30000 samples) epoches as the train data to verify the effectiveness of our detection algorithm, one is in normal state and another one is in seizure. We also choose one epoch of 600 sec (30000 samples) in seizure onset as test data for validation.

For all the EEG signals, we first rearrange signals to 1 s windows with 0.5 s overlap. Each data set contains 30000 samples are rearranged into 1199 pattern vectors with dimension of 50. Then we extract features from the rearranged pattern vectors by LLE and WLLE, respectively. Here we use the LLE algorithm to do a comparison to show the advantages of WLLE. The parameters chosen in this simulation for LLE and WLLE are: number of neighbors $K = 4$; embedding dimensionality $d = 2$. Through the dimension reduction algorithm LLE and WLLE, we map the high dimensional pattern vectors (50-dimensional) into low dimensional space (2-dimensional), while keep the intrinsic structure and internal features of the data set.

The 2D embedding of the original 50D pattern vectors through WLLE is shown in Fig. 2 and the 2D embedding obtain through LLE is shown in Fig. 3. From these two figures, we can see the differences of the performances of the two algorithms. The 2D embedding by WLLE shows obvious clusters for all the three data sets, but those by LLE are overlapped together, which means difficulties for classify. The reason why WLLE can improve the clustering properties of its embedding lies in that by using weighted distance measure, we actually transferred the original data set from a space to another space, which may make the original linear inseparable problem become linear separable. This is also the reason why the embedding of WLLE seems a strange mapping different from traditional LLE.

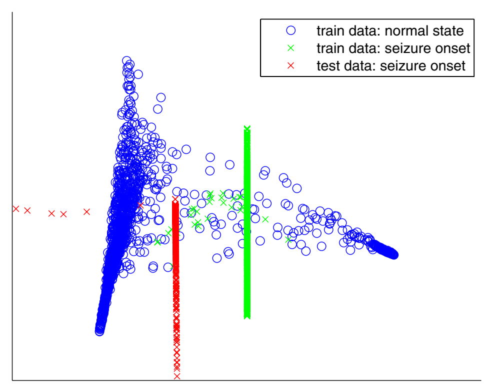

Fig. 2. Feature Extraction of EEG data by WLLE

The 2D embedding shown in Fig. 2 have a clear cluster structure thus becomes a linear separable problem can be easily classified by SVM. Fig. 4 shows the classification result and the hyperplane for classify. The result of this experiment, such as accuracy rate (AR), false positives (FP) and true negatives (TN) are shown in Table I.

TABLE I
Classification Results

Data Set	AR	FP	TN
Train Data	90.66%	8.92%	0.42%
Test Data	76.4%	0.00%	23.6%

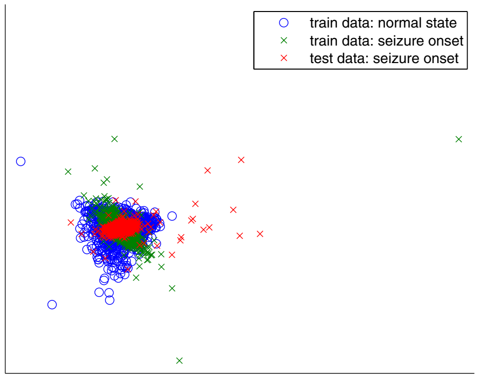

Fig. 3. Feature Extraction of EEG data by LLE

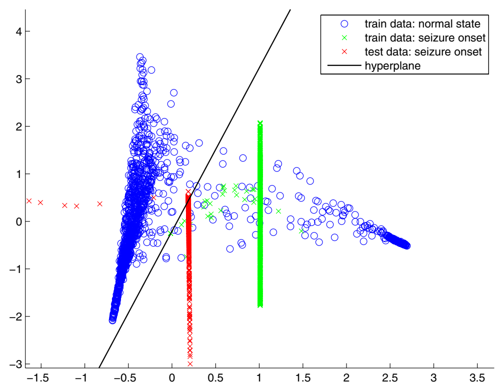

Fig. 4. Classification Result by SVM

VII. Conclusion

In this paper, we presented a new, EEG-based, brain-state identification method for detecting epileptic seizures. To classify the EEG signals and diagnose the epileptic seizures directly by using Weighted Locally Linear Embedding (WLLE) and support vector machine, we first used WLLE as feature extraction algorithm for the EEG signal to obtain more compact and low-dimensional representation of the internal characteristic and structure in the original data. The low-dimensional representatives which are linear separable and easy to classified, are then classified by SVM classifier. As such, seizures onset states are identified from normal states of the epilepsy patients. Real EEG signals recorded from mice were used in simulation study to show the advantages of WLLE feature extraction, and the classification accuracy rate of SVM was between 72.79% and 90.66% in simulation studies. This method holds promising to be translated to human experiments.

Acknowledgments

This study was partially supported by Singhealth Foundation Funds (NO: SHF/FG217P/2006) and National University of Singapore, URC Funding, Mind Robotics (NO: R263000490112).

References

[1] B. Litt, R. Esteller, and J. Echauz, "Epileptic seizures may begin hours in advance of clinical onset: a report of five patients," *Neuron*, vol. 30, pp. 51–64, 2001.

[2] A. B. Gardner, A. M. Krieger, G. Vachtsevanos, and B. Litt, "One-class novelty detection for seizure analysis from intracranial EEG," *Journal of Machine Learning Research*, vol. 7, pp. 1025–1044, 2006.

[3] Y. Pan, S. S. Ge, F. R. Tang, and A. A. Mamun, "Detection of epileptic spike-wave discharges using SVM," in *Proceedings of the 2007 IEEE International Conference on Control Applications*, Suntec City, Singapore, 1-3 October 2007, pp. 467–472.

[4] Y. Salant, I. Gath, and O. Henriksen, "Prediction of epileptic seizures from two-channel EEG," *Medical and biolodical Engineering and Computing*, vol. 36, pp. 549–556, 1998.

[5] P. J. Franaszczuk, G. K. Bergey, P. J. Durka, and H. M. Eisenberg, "Time-frequency analysis using the matching pursuit algorithm applied to seizures originating from the mesial temporal lobe," *Electroencephalography and Clinical Neurophysiology*, vol. 106, pp. 513–521, 1998.

[6] M. W. Slutzky and D. J. Mogul, "Manipulating epileptiform bursting in the rat hippocampus using chaos control and adaptive techniques," *IEEE Tansactions on Biomedical engineering*, vol. 50, pp. 559–570, 2003.

[7] Y. Pan, S. S. Ge, and A. A. Mamun, "Nonlinear control of synaptic plasticity model for constraining bursting activity in epileptic seizures," in *Proceedings of the 2007 American Control Conference*, New York City, USA, July 11-13 2007, pp. 2012–2017.

[8] L. D. Iasemidis, J. C. Sackellares, H. P. Zaveri, and W. J. Williams, "Phase space topography and the Lyapunov exponent of electrocorticograms in partial seizures," *Brain topography*, vol. 2, pp. 187–201, 1990.

[9] A. B. Geva and D. H. Kerem, "Forecasting generalized epileptic seizures from the eeg signal by wavelet analysis and dynamic unsupervised fuzzy clustering," *IEEE Transaction of Biomed Engineering*, vol. 45, pp. 1205–1216, 1998.

[10] L. K. Saul and S. T. Roweis, "Think globally, fit locally: Unsupervised learning of low dimensional manifolds," *Journal of Machine Learning Research*, vol. 4, pp. 119–155, 2003.

[11] F. R. Tang, S. C. Chia, P. M. Chen, H. Gao, W. L. Lee, T. S. Yeo, J. M. Burgunder, A. Probst, M. K. Sim, and E. A. Ling, "Metabotropic glutamate receptor 2/3 in the hippocampus of patients with mesial temporal lobe epilepsy, of rats and mice after pilocarpine-induced status epilepticus," *Epilepsy Res.*, vol. 59, pp. 167–180, 2004.

[12] F. R. Tang, S. C. Chia, F. L. Jiang, D. L. Ma, P. M. Chen, and Y. C. Tang, "Calcium binding protein containing neurons in sclerotic mouse hippocampus with special reference to their afferents from the medial septum and the entorhinal cortex," *Neuroscience*, vol. 140, pp. 1467–1479, 2005.

[13] R. J. Racine, "Modification of seizure activity by electric stimulation, ii. motor seizure," *Electroencephalogr. Clin. Neurophysiol*, vol. 32, p. 281C294, 1972.

[14] S. M. Zoldi, A. Krystal, and H. S. Greenside, "Stationarity and redundancy of multichannel eeg data recorded during generalized tonic-clonic seizures," 1999. [Online]. Available: http://www.citebase.org/abstract?id=oai:arXiv.org:chao-dyn/9905004

[15] L. Iasemidis, H. Zaveri, J. Sackellares, W. Williams, and T. Hood, "Nonlinear dynamics of ECoG data in temporal lobe epilepsy," *Electroencephalography and Clinical Neurophysiology*, vol. 5, p. 339, 1988.

[16] M. D'Alessandro, G. Vachtsevanos, A. Hinson, R. Esteller, J. Echauz, and B. Litt, "A genetic approach to selection the optimal feature for epileptic seizure prediction," in *Proceedings of 23rd Annual EMBS International Conference*, Istanbul, Turkey, 25-28 October 2001.

[17] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, pp. 2323–2326, 2000.

[18] C. Y. Zhou and Y. Q. Chen, "Improving nearest neighbor classification with cam weighted distance," *Pattern Recognition*, vol. 39, pp. 1–11, 2006.

[19] S. S. Ge, F. Guan, A. P. Loh, and C. H. Fua, "Feature representation based on intrinsic structure discovery in high dimensional space," in *Proceedings of the 2006 IEEE International Conference on Robotics and Automation*, Orlando, Florida, USA, May 2006, pp. 3399–3404.

[20] F. GUAN, L. Y. LI, S. S. GE, and A. P. LOH, "Robust huamn detection and identification by using stereo and thermal images in human robot interaction," *International Journal of Information Acquisition*, vol. 4, no. 2, pp. 161–183, 2006.