Automatic Seizure Detection in Multichannel EEG using McCIT2FIS Approach

S. Dora*, B. Rangarajan†, K. Subramanian‡ and S. Suresh§
School of Computer Engineering
Nanyang Technological University
Singapore, 639798
Email: *shirin0021@e.ntu.edu.sg, †badrinarayanan@ntu.edu.sg, ‡kartick1@e.ntu.edu.sg, §ssundaram@ntu.edu.sg

***Abstract*—In this paper, an automatic seizure detection technique using multichannel EEG is proposed based on Meta-cognitive Complex-valued Interval Type-2 Fuzzy Inference System (McCIT2FIS). A wavelet chaos theory based feature extraction is employed to extract the features from EEG signal as it can handle the non stationarity in data and Sparse Multinomial Logistic Regression via Bayesian L1 Regularisation (SBMLR) based feature selection is employed to select the most discriminative features. McCIT2FIS is employed to classify the samples as either interictal or ictal EEG segment as it has been shown to be capable of handling noisy data by virtue of Interval Type-2 fuzzy sets, and is good at classification because of its ability to handle complex-valued data. Further, we have also shown that the feature selected using SBMLR can be successfully mapped back to the channels allowing us to identify the epileptogenic regions of the brain. The performance of the McCIT2FIS was also compared with the support vector machines and the results indicate that McCIT2FIS is better capable of detecting seizure based on EEG signals.**

***Index Terms*—Seizure detection, electroencephalogram, interval type-2, noisy data, complex-valued fuzzy system**

I. INTRODUCTION

Epilepsy is one of the most commonly occurring neurological disorder, affecting nearly 1% of the world's population. It is characterized by the occurrence of repeated episodes of abnormal electrical activity in the brain called seizures. An Electroencephalogram (EEG) signal is an important tool for diagnosis of epilepsy as it accurately represents the brain's electrical activity in time, frequency and space domain.

A seizure can be easily detected by visual inspection of an EEG signal but the unpredictable nature of seizures makes it difficult for an epileptic person to lead a normal life. Further, such a visual inspection often requires careful investigation of a large amount of EEG data by a trained neurologist. This can be a time consuming process and may lead to erroneous conclusions. Hence, there is a need for developing automated tools for detecting a seizure.

Gotman proposed the first seizure detection algorithm [1] in 1982 which tried to solve the seizure detection problem using mimetic techniques. It systematically searched for clinical signs of an epileptic seizure i.e. excessive synchronous electrical activity, in an EEG recording. A seizure was declared when the extent of synchronicity was more than a predefined threshold. Gotman's technique had no way to distinguish the rhythmic activity caused by a seizure from that caused by natural artifacts like sleep spindle, etc. To overcome this drawback, many approaches have been proposed which target epilepsy detection as a pattern recognition problem using features extracted by signal processing techniques.

The early methods utilized features extracted using Fourier transform [2] and autocorrelation function [3] for seizure detection. Both these techniques are unable to handle the non-stationarity associated with EEG signals. This led to the development of techniques which utilized features extracted using wavelet transform [4] as it offered better localization across time and frequency. In [5], a wavelet based feature extraction technique was proposed which also used nonlinear measures of chaos in the signal like Lyapunov exponent and correlation dimension of multiple EEG subbands. It used a combination of statistical and chaos theory features, extracted from multiple subbands of a single channel EEG signal to train a classifier to categorize an input signal as an ictal or interictal EEG segment. This technique requires a trained clinician to first identify the epileptogenic regions of interest so as to extract features from the corresponding channels before any classification can be performed.

Different machine learning classifiers have been used for seizure detection including support vector machines [6], [7], artificial neural networks [2], [8] and neuro fuzzy inference systems [9]. Neuro-fuzzy inference systems unite the plasticity of neural networks and interpretability of fuzzy systems into a single classifier and have been shown to be possess greater learning potential [10]. But, these systems employ type-1 fuzzy sets which cannot handle non stationary data efficiently. On the other hand, a type-2 fuzzy set [11], [12], a generalization over the type-1 fuzzy set, has a secondary membership function which allows it to handle greater level of non-stationarity in data.

In this paper we have used a Meta-Cognitive Complex-valued Interval Type-2 Fuzzy Inference System (McCIT2FIS) [13] to develop a patient specific seizure detector for multichannel EEG as it can effectively model non stationary data. McCIT2FIS employs a self regulated learning algorithm which allows it to learn each sample in an appropriate manner. Such an approach prevents over training, thereby allowing greater generalization. Further, it has been shown that a complex-valued neural networks have two orthogonal decision boundaries [14] which enables them to solve more complex

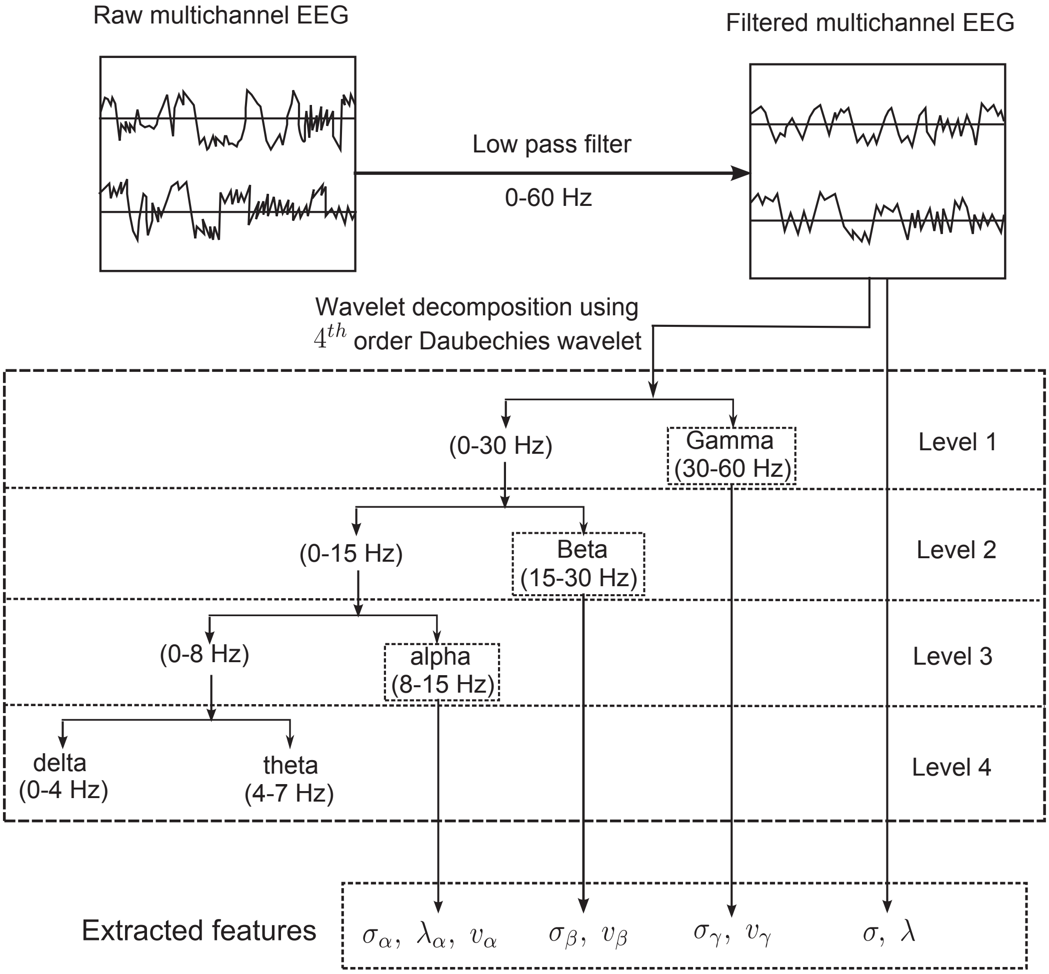

Fig. 1. Overview of the feature extraction process

real-valued problems inherently. It has also been shown that a complex-valued neural network has higher computational power than a real valued network [15]. The features for the multichannel EEG seizure detector were obtained by using the wavelet based feature extraction technique described in [5], individually for each channels. The ability to handle multichannel EEG alleviates the need for the intervention of a trained clinician. Further, a Sparse Multinomial Logistic Regression via Bayesian L1 Regularisation (SBMLR) [16] based feature selection is employed to select few features which better highlight the distinction between interictal and ictal EEG segments. We have shown that by mapping the selected features to the corresponding channels it is possible to identify the prospective epileptogenic regions of interest.

The performance of the proposed McCIT2FIS based seizure detector is evaluated on Upenn and Mayo seizure detection challenge data set [17], [18]. We have chosen the Patient 1 in our study to highlight the advantage of the classifier in seizure detection. The performance of the system is compared against standard support vector machine classifier [19] using the features extracted using wavelet chaos theory as well as a subset of extracted features chosen using SBMLR. The performance comparison with stratified 5-fold cross-validation indicates improved performance of McCIT2FIS using the selected features.

This paper is organized as follows. Section II describes the data set followed by an explanation of the feature extraction and feature selection technique. Section III describes the architecture and learning algorithm of McCIT2FIS classifier. Experiments and results are explained in the Section IV followed by a discussion in Section V. Finally, a conclusion of this paper is summarized in the Section VI.

II. Material and methods

A. EEG data set

The EEG data, used for evaluations, in this paper consists of intracrannial Electroencephalogram (EEG) data hosted by UPenn - Mayo clinic seizure detection challenge on Kaggle.com [17], [18]. It may be noted that all the simulations have been performed using the data for Patient 1 to understand the performance of McCIT2FIS on seizure detection. Since, this is just an initial study, we have only reported results for only Patient 1. It was recorded at a frequency of 500 Hz or 5000 Hz and the number of electrodes used for each patient vary from 16 to 72. Each sample consists of EEG recordings of one second duration and can be classified as ictal (collected during a seizure) or interictal (collected atmost an hour before or after a seizure) segment.

B. Wavelet Chaos Theory based Feature Extraction

Many feature extraction techniques have been proposed in literature which can be broadly classified under Fourier spectral analysis [2], autoregressive parameteric techniques [3] and discrete wavelet transform [4]. It has been shown that the features extracted using discrete wavelet transform provide a much better representation of an EEG signal due to their ability to handle non-stationarity in the data [4], [20], [21]. Further, it has been shown in literature that the nonlinear dynamics of the human brain are much less chaotic during the preictal phase [22], [23], [24]. Hence, we have used a combination of statistical and chaos theory features from a segment of EEG data for the purpose of seizure detection [5]. It may be noted that the feature extraction approach described in [5] was used for single channel EEG recordings whereas in this paper we have extended the proposed approach to multichannel EEG segments. This alleviates the need for the intervention of a trained clinician to identify the epileptogenic regions of interest for feature extraction.

The Figure 1 shows an overview of the feature extraction process. The multichannel EEG segments are low pass filtered to 0 to 60 Hz so as to discard any noise. The filtered EEG segment is decomposed into five subbands, using the Level 4 Daubechies wavelet transform, which consist of frequencies in the range delta (0-4 Hz), theta (4-7 Hz), alpha (8-15 Hz), beta (15-30 Hz) and gamma (30-60 Hz). The filtered EEG along with the subbands are then used to obtain the statistical as well as nonlinear features for seizure detection. The statistical features consists of standard deviation of the filtered EEG singal (σ) and; the alpha (σ_α), beta (σ_β) and gamma (σ_γ) subbands. The nonlinear features extracted consist of the correlation dimension computed from the alpha (υ_α), beta (υ_α) and gamma (υ_γ) subbands; and the Lyapunov exponent computed from the filtered EEG (λ) and alpha(λ_α) subband.

C. Sparse Multinomial Logistic Regression via Bayesian l1 Regularisation (SBMLR) based Feature Selection

SBMLR [16] is a state-of-the-art feature selection method which realizes a sparse feature selection by using a Laplace prior. Here, the logistic regression is used to build a regularized regression model. SBMLR is shown to provide better classification accuracy with biomedical data [16]. Also, SBMLR is shown to improve the generalization performance [25]. In this study, we have employed SBMLR to select the most discriminative features of the multichannel EEG.

III. META-COGNITVE COMPLEX-VALUED INTERVAL TYPE-2 FUZZY INFERENCE SYSTEM

McCIT2FIS learns to classify a given EEG segments as ictal or interictal based on the features extracted using wavelet-chaos analysis of the EEG segments. A given EEG segment, $(\mathbf{x^t}, c^t)$, consists of an m-dimensional feature vector $\mathbf{x^t} = [x_1^t, \ldots, x_i^t, \ldots, x_m^t]$ with class label c^t where $c^t \in [1, 2]$. The coded class label for a given samples are given by $\mathbf{y^t} = [y_1^t, y_2^t]$ where y_g^t is defined as

$$y_g^t = \begin{cases} 1+i & g = c^t \\ -1-i & otherwise \end{cases} \quad g = 1,\ 2. \tag{1}$$

The aim of McCIT2FIS is to approximate the functional relationship that maps a given EEG sample to the corresponding coded class label. To approximate this mapping closely McCIT2FIS automatically determines required number of rules during the training phase. Without loss of generality we will assume that the network has grown to $(K-1)$ rules after presentation of $(t-1)$ samples. Next the architecture and the learning algorithm of McCIT2FIS will be described assuming $(\mathbf{x^t}, c^t)$ as the current training sample.

A. McCIT2FIS architecture

McCIT2FIS employs a six layered architecture utilizing the type-2 Takagi-Sugeno-Kang type fuzzy inference mechanism as shown in the Figure 2. Next, we will describe the functioning of each layer in detail.

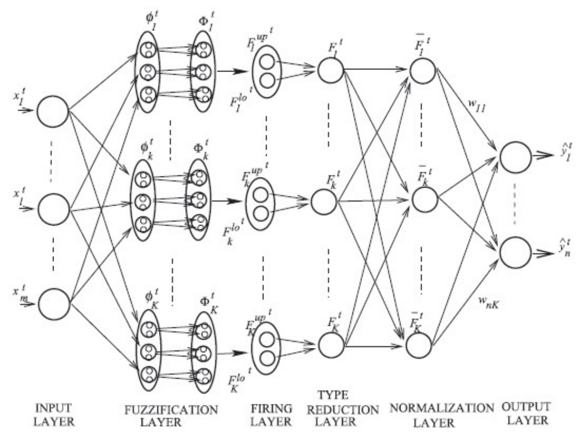

Fig. 2. Cognitive component of McCIT2FIS

Input layer consists of m nodes. This layer propagates the m dimension network input to the fuzzification layer. Hence, the output of l^{th} node in this layer is given as

$$u_l = x_l^t \tag{2}$$

Fuzzification layer consists of $2*K$ nodes which compute the interval type-2 membership of input recieved from input layer using a q-Gaussian membership function. The membership of the i^{th} input feature for k^{th} node is given as

$$\phi(u_l, {}^g\mu_{lk}, {}^g\sigma_k, q) = (1 + (1-q)d(u_l, {}^g\mu_{lk}, {}^g\sigma_k))^{\frac{1}{(1-q)}}, \quad g \in [1,\ 2]. \tag{3}$$

where,

$$d(u_l, \mu_{lk}, \sigma_k) = \frac{(u_l - \mu_{lk})^H (u_l - \mu_{lk})}{2\sigma_k^2}.$$

where ${}^1\mu_{lk}$ and ${}^2\mu_{lk}$ represent the limits of the complex valued centers with spreads ${}^1\sigma_k$ and ${}^2\sigma_k$. H is the Hermitian operator on a complex number. The output of k^{th} node in this layer can be represented as an interval $[[\left(\phi_{lk}^{lo}\right)^t, (\phi_{lk}^{up})^t]$ where $\left(\phi_{lk}^{lo}\right)^t$ and $(\phi_{lk}^{up})^t$ are respectively defined as

$$(\phi_{ik}^{up})^t = \begin{cases} \phi(u_l, {}^1\mu_{lk}, {}^1\sigma_k, q) & if\ ||u_l|| < ||^1\mu_{lk}|| \\ \phi(u_l, {}^2\mu_{lk}, {}^2\sigma_k, q) & if\ ||u_l|| > ||^2\mu_{lk}|| \\ 1 & otherwise \end{cases} \tag{4}$$

$$\left(\phi_{lk}^{lo}\right)^t = \begin{cases} \phi(u_l, {}^2\mu_{lk}, {}^2\sigma_k, q) & \text{if } ||u_l|| \le \frac{||^1\mu_{lk}|| + ||^2\mu_{lk}||}{2} \\ \phi(u_l, {}^1\mu_{lk}, {}^1\sigma_k, q) & otherwise \end{cases} \tag{5}$$

where $||.||$ is the L^2-norm.

Firing Layer also consists of $2 * K$ nodes and it converts the output of fuzzification layer to a interval type-1 fuzzy set. The output of this layer if given as

$$\Phi_k^t = [\left(F_k^{lo}\right)^t, (F_k^{up})^t]$$
$$\text{where, } \left(F_k^{lo}\right)^t = \prod_{l=1}^{m} \left(\phi_{lk}^{lo}\right)^t$$
$$\text{and } \left(F_k^{up}\right)^t = \prod_{l=1}^{m} \left(\phi_{lk}^{up}\right)^t \tag{6}$$

Type reduction layer consists of K nodes and converts the output of firing layer to a simple membership value and its output is given as

$$F_k^t = \frac{\left(F_k^{lo}\right)^t + (F_k^{up})^t}{2} \tag{7}$$

Normalization layer consists of K nodes that normalized the membership value obtained from type reduction layer and its output is given as

$$\bar{F}_k^t = \frac{F_k^t}{\sum_{p=1}^K F_p^t} \tag{8}$$

Output layer returns the predicted output according to a weighted average of the normalized membership values as

$$\hat{y}_g^t = \sum_{k=1}^{K} w_{gk} \bar{F}_k^t, \tag{9}$$

where w_{gk} is the complex valued output weight connecting the k^{th} rule node to the g^{th} output node. The predicted class label ($\hat{c}^t$) is determined as

$$\hat{c}^t \;=\; arg \max_{g=1,2} \mathbb{R}(\hat{y}_g^t) \tag{10}$$

The operator $\mathbb{R}(\hat{y}_g^t)$ returns the real part of the complex valued $\hat{y}_g^t$. Next, we will briefly describe the projection based learning algorithm used for McCIT2FIS.

B. Projection based Learning algorithm

The Projection Based Learning (PBL) algorithm is employed to estimate the complex valued weights in the output layer such that the sum of squared error is minimized. The PBL learning rule estimates the weights by simultaneously minimizing the hinge loss error for all the samples. Therefore the optimal weights are given by

$$W^* = arg \min_{\mathbf{W}} \frac{1}{2} \sum_{t=1}^{N} \sum_{g=1}^{2} \left(e_g^t\right)^2 \tag{11}$$

where N is the total number of training samples. The above equation can be solved using the Wirtinger Calculus and the solution is given by

$$\sum_{k=1}^{K} \sum_{t=1}^{N} \bar{F}_k^t \bar{F}_l^t w_{gk}^H = \sum_{t=1}^{N} \bar{F}_l^t \left(y_h^t\right)^H \tag{12}$$

This can be written in matrix form as

$$AW^H = B \tag{13}$$

and the optimal weights are given by

$$W = (A^{-1}B)^H \tag{14}$$

Next, the self-regulatory learning algorithm of McCIT2FIS will be described briefly.

C. Self-regulatory learning algorithm of McCIT2FIS

The learning algorithm of McCIT2FIS self-regulates its learning process by learning each training sample according to the novel knowledge present in the sample with respect to the network. It estimates the absolute maximum hinge-loss error (E_M^t) and the knowledge potential (ψ) [13] for each sample to choose a suitable strategy for learning a given sample. These measures are given as,

$$e_g^t = \begin{cases} y_g^t - \hat{y}_g^t & \mathbb{R}\left(y_g^t\right) \times \mathbb{R}\left(\hat{y}_g^t\right) < 1 \\ 0 & otherwise \end{cases} \quad g \;=\; 1,\ 2. \tag{15}$$

and

$$E_M^t = \arg\max |\mathbf{e}^t|. \tag{16}$$

Knowledge pontential (ψ) is given as,

$$\psi^t = \frac{\sum_{k=1}^K F_k}{K}. \tag{17}$$

Based on these measures, the learning algorithm can choose to delete a sample, reserve a sample or learn a sample.

- **Sample deletion strategy:** If a sample is correctly classified and the E_M^t is less than the sample delete threshold then the given sample is not learnt as the network already contains the information present in the training sample.
- **Sample learning strategy:** Based on the information present in the training sample, the learning can choose to add a new rule to the network or update the existing

rules in the network. The learning algorithm chooses to add a new rule if the network prediction is incorrect or the error is greater than self-adaptive rule add threshold, and the knowledge potential is less than the novelty threshold, i.e., if

$$((c^t \neq \hat{c}^t \text{ OR } E_M^t > E_a) \text{ AND } \psi^t < E_n). \quad (18)$$

Each time a new rule is added to the network, the self-adaptive add threshold is updated as

$$E_a = \gamma E_a + (1-\gamma)E_M^t \quad (19)$$

where γ is the slope of adaptation. The centers for the new added rule are calculated as

$$^1\mu_K = 0.9\mathbf{u}^t \quad (20)$$

$$^2\mu_K = 1.1\mathbf{u}^t. \quad (21)$$

The width of the K^{th} rule is determined in accordance distance between the center of new rule from the nearest interclass (nrI) and intraclass (nrS) rules which are defined as

$$^g nrS = arg \min_{l=c} ||\mathbf{u}^t - \ ^g\mu^l||; \ g \in [1,2] \quad (22)$$

$$^g nrI = arg \min_{l \neq c} ||\mathbf{u}^t - \ ^g\mu^l||; \ g \in [1,2]. \quad (23)$$

The Euclidean distance between the newly added rule and nrS and nrI is given as

$$^g d_S = ||\mathbf{u}^t - \ ^g\mu_{nrS}|| \quad (24)$$

$$^g d_I = ||\mathbf{u}^t - \ ^g\mu_{nrI}|| \quad g \in [1,2] \quad (25)$$

When the new rule is closer to nrS, the width of the new rule is given as

$$^g\sigma_K = \kappa \times^g d_S \quad g \in [1,2] \quad (26)$$

where κ is chosen in the range $[0.5, 0.9]$. Similarly, if the new rule is closer to nrI, than the width of the new rule is given as

$$^g\sigma_K = \eta \times^g d_I \quad g \in [1,2] \quad (27)$$

where η is chosen in the range $[0.1, 0.5]$. It may be noted that the rule addition also requires an updation of the A and B matrix as

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{(K-1)\times(K-1)} + \left(\bar{F}^t\right)^T \bar{F}^t & \sum_{t=1}^{t} \bar{F}_K^t \bar{F}_p^t \\ \sum_{t=1}^{t} \bar{F}_K^t \bar{F}_p^t & a_{K\times K} \end{bmatrix} \quad (28)$$

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_{(K-1)\times 2} + \left(\bar{F}^t\right)^T \left(\mathbf{y}^t\right)^T \\ \sum_{t=1}^{N} \bar{F}_K^t \left(y_g^t\right)^H \end{bmatrix} \quad (29)$$

Based on re-estimated $\mathbf{A}$ and $\mathbf{B}$ matrices, the weight $\mathbf{W}$ is recalculated based on eq.(14).

The network parameters are updated when the sample is correctly classified but the error is greater than the self-adaptive parameter update threshold, i.e.,

$$c^t = \hat{c}^t \text{ AND } E_M^t > E_l. \quad (30)$$

Further, it may be noted that each time a parameter is updated the self-adaptive parameter update threshold is updated as

$$E_l = \gamma E_l + (1-\gamma)E_M^t \quad (31)$$

where γ is the slope of update. When a sample is used to update the network parameters the $\mathbf{A}$ and $\mathbf{B}$ matrix are updated as

$$\mathbf{A} = \mathbf{A} + \left(\bar{\mathbf{F}}^t\right)^T \left(\bar{\mathbf{F}}^t\right) \quad (32)$$

$$\mathbf{B} = \mathbf{B} + \left(\bar{\mathbf{F}}^t\right)^T \left(\mathbf{y}^t\right)^H. \quad (33)$$

and the updated output weights are calculated as

$$\mathbf{w}_{(K-1)} = \mathbf{w}_{(K-1)} + \mathbf{A}^{-1} \left(\bar{\mathbf{F}}^t\right)^T \left(\mathbf{e}^t\right)^H \quad (34)$$

- **Sample reserve strategy:** In case the criterion for both the above strategies is not satisfied then the sample is reserved for learning at a later time.

Refer [13] for a detailed explanation of McCIT2FIS.

IV. Experiments and Results

In this section, we evaluate McCIT2FIS for the problem of automatic seizure detection. Figure 3 shows the schematic diagram representing the seizure detection problem using multichannel EEG as described in this paper. Here, the features were extracted from the multichannel EEG using the wavelet chaos theory based feature extraction method. The data in this paper was obtained for 1 epileptic patient (Patient 1) from the Upenn and Mayo seizure detection challenge [26], [27] and the results of 5-fold cross-validation were reported. Note that the results of 5-fold cross validation were reported as the actual class labels for the test samples have not been released in the competetion. Further, we have also applied a feature selection techniques called SBMLR to select the most discriminative features for classification. The performance of McCIT2FIS was also compared with the performance of SVM for similar evaluations.

A. Seizure detection using all the features

After performing feature extraction using wavelet chaos theory for Patient 1 as described in the Section II-B, we have obtained 174 samples with 612 features. In this section, we analyze the performance of McCIT2FIS in comparison with SVM using all the 612 features as shown in the Table I. Here, the 'η_o' denotes the overall accuracy of the classifier and is defined as the ratio of correctly classified samples to total number of samples. It can be inferred from the table that, McCIT2FIS performs at par with the standard SVM classifier. Although McCIT2FIS provides results with accuracy close to that of a well-known classifier, the approach is difficult to implement in hardware due to the large number of features. In this study, we have 612 features corresponding to different channels. In case of an epileptic patient, its nor necessary that all the channels show synchronous activity. Hence, there exists a need to build an automated channel selection mechanism. We have addressed this problem by providing a more efficient method of choosing the channels in the following section.

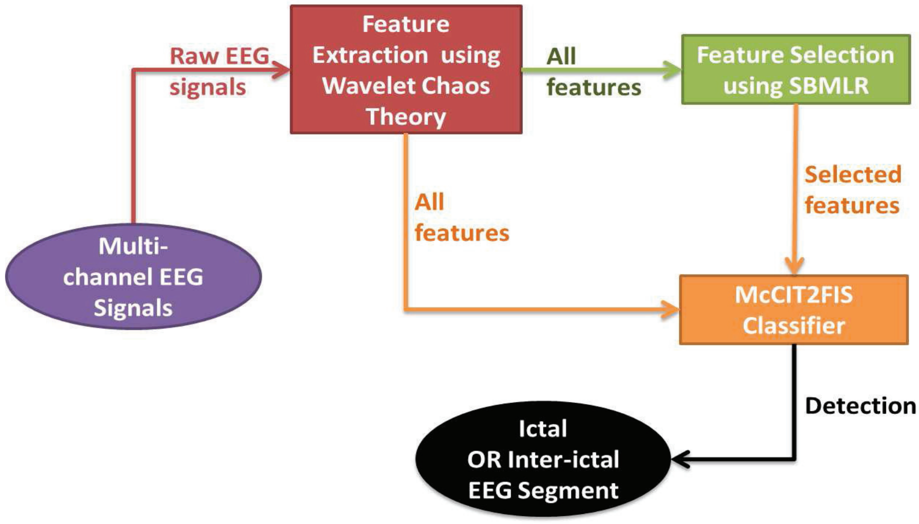

Fig. 3. Schematic diagram of automatic seizure detection using multichannel EEG

TABLE I
PERFORMANCE OF SVM AND MCCIT2FIS USING ALL THE 613 FEATURES

SVM(613 features)			
Fold#	Training η_o	Testing η_o	Support vectors
1	100	97.14	25
2	100	100	23
3	100	88.57	18
4	100	88.57	52
5	100	97.06	26
McCIT2FIS(613 features)			
Fold#	Training η_o	Testing η_o	Rules
1	99	97	26
2	99	100	21
3	99	83	18
4	99	91	22
5	99	94	22

TABLE II
PERFORMANCE OF SVM AND MCCIT2FIS USING ONLY 8 FEATURES

SVM(8 features)			
Fold#	Training η_o	Testing η_o	Support vectors
1	97.84	97.14	56
2	97.12	100	56
3	100	100	5
4	97.12	100	33
5	97.86	97.06	21
McCIT2FIS(8 features)			
Fold#	Training η_o	Testing η_o	Rules
1	96	100	7
2	96	100	20
3	96	97	13
4	99	100	22
5	96	100	19

B. Towards efficient channel selection in multichannel EEG using SBMLR based feature selection

Sparse multinomial logistic regression via bayesian l1 regularisation (SBMLR) [16] is a well-known feature selection technique as described in the Section II-C. We have employed SBMLR to Patient 1 data set and obtained eight most discriminative features. The Table II shows the performance of McCIT2FIS and SVM using the eight features. It can be noted that, McCIT2FIS clearly outperforms SVM providing better accuracy with less number of rules except for the third fold, where SVM performs better when compared to McCIT2FIS. The usage of self-regulated learning strategies in McCIT2FIS is evident from the optimal training and better testing accuracies achieved using fewer rules. As described in this study, we have obtained eight different features using SBMLR based feature selection technique. These features correspond to eight different channels in the EEG segment which show considerably different electrical activity across ictal and interictal EEGs. The Figure 4 shows the plot of ictal and interictal EEG segments of one second duration. The axis label for the y-axis shows the channel for which the EEG was plotted. It can be observed from the figure that the channels 'LFG26' and 'LFG18' show high frequency synchronous electrical activity (hallmark of an epileptic seizure). Further the channels 'LFG42', 'LFG16', 'LFG20' and 'LFG58' are synchronous with low frequency activity and gradually approaching synchrony with the channels 'LFG26' and 'LFG18'. This highlights the efficacy of the SBMLR technique in choosing the channels with synchronous activity for identifying the occurrence of a seizure.

V. DISCUSSION

This paper introduces an automatic seizure detection in multichannel EEG using McCIT2FIS approach. In a epileptic patient, a seizure is marked by excessive synchronous electrical activity across multiple channels. The Figure 5 shows one second of electrical activity across two channels for an epileptic patient during a seizure. It can be observed from the figure that there is considerable synchronous electrical activity

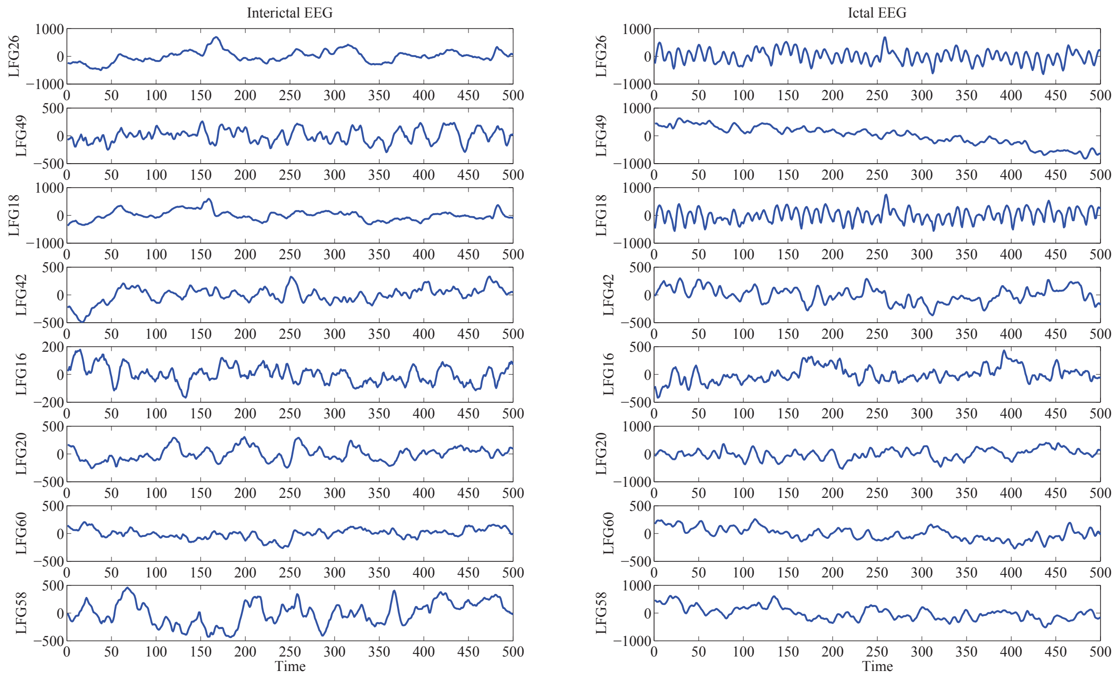

Fig. 4. Inter-ictal and Ictal EEG for the eight selected channels

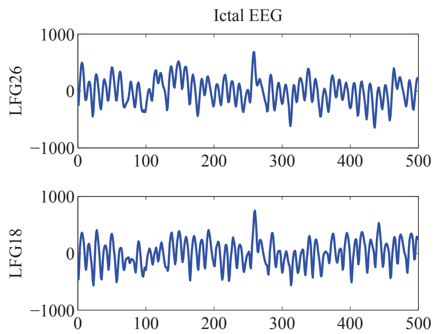

Fig. 5. EEG recording of two channels with synchronous activityduring a seizure

across the channels corresponding to the selected features during a seizure. This activity is marked by sudden changes the postsynaptic potential called spikes. Hence, there exists a need to identify these spikes in EEG for an accurate detection of seizures. It has been shown that spiking neural networks are inherently capable of handling data that is in the form of spikes. Similar to neural fuzzy approaches, several evolving approaches have been proposed by the author in the framework of spiking neural networks [28], [29]. In our future work, we intend to combine the interpretability of fuzzy systems and the ability of spiking neural networks to inherently handle the data that is in the form of spikes for better identification of seizures.

VI. Conclusion

An automatic seizure detection technique in multichannel EEG has been proposed in this paper. A combination of statistical and chaos theory based nonlinear EEG features were extracted for classification. SBMLR is applied on the extracted features to choose a few discriminative features. The McCIT2FIS classifier was used to classify the samples as either interictal or ictal EEG segment and the results were further compared with SVM. Initially, both the classifiers were evaluated using all the features. Further, the classifiers were compared with reduced number of features. The results clearly highlight that McCIT2FIS outperforms SVM when reduced number of features were chosen.

Acknowledgement

The authors wish to extend their thanks to the Ministry of Education (MoE), Singapore, for providing financial support through tier I (No. M4011269) funding to conduct this study.

References

[1] J. Gotman, "Automatic recognition of epileptic seizures in the EEG," *Electroencephalography and clinical Neurophysiology*, vol. 54, no. 5, pp. 530–540, 1982.

[2] V. Srinivasan, C. Eswaran, and N. Sriraam, "Artificial neural network based epileptic detection using time-domain and frequency-domain features," *Journal of Medical Systems*, vol. 29, no. 6, pp. 647–660, 2005.

[3] A. Subasi, E. Erçelebi, A. Alkan, and E. Koklukaya, "Comparison of subspace-based methods with AR parametric methods in epileptic seizure detection," *Computers in Biology and Medicine*, vol. 36, no. 2, pp. 195–208, 2006.

[4] H. Adeli, Z. Zhou, and N. Dadmehr, "Analysis of eeg records in an epileptic patient using wavelet transform," *Journal of neuroscience methods*, vol. 123, no. 1, pp. 69–87, 2003.

[5] H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr, "A wavelet-chaos methodology for analysis of EEGs and EEG subbands to detect seizure and epilepsy," *Biomedical Engineering, IEEE Transactions on*, vol. 54, no. 2, pp. 205–211, 2007.

[6] C. H. Seng, R. Demirli, L. Khuon, and D. Bolger, "Seizure detection in EEG signals using support vector machines," in *Bioengineering Conference (NEBEC), 2012 38th Annual Northeast*, 2012, pp. 231–232.

[7] A. Temko, E. Thomas, W. Marnane, G. Lightbody, and G. Boylan, "EEG-based neonatal seizure detection with support vector machines," *Clinical Neurophysiology*, vol. 122, no. 3, pp. 464–473, 2011.

[8] A. Petrosian, D. Prokhorov, R. Homan, R. Dasheiff, and D. Wunsch, "Recurrent neural network based prediction of epileptic seizures in intra- and extracranial eeg," *Neurocomputing*, vol. 30, no. 1, pp. 201–218, 2000.

[9] A. Subasi, "Application of adaptive neuro-fuzzy inference system for epileptic seizure detection using wavelet feature extraction," *Computers in Biology and Medicine*, vol. 37, no. 2, pp. 227–244, 2007.

[10] K. Subramanian, S. Suresh, and N. Sundararajan, "A metacognitive neuro-fuzzy inference system (McFIS) for sequential classification problems," *Fuzzy Systems, IEEE Transactions on*, vol. 21, no. 6, pp. 1080–1095, 2013.

[11] A. Das, K. Subramanian, and S. Suresh, "A computationally fast interval type-2 neuro-fuzzy inference system and its meta-cognitive projection based learning algorithm," in *Neural Networks (IJCNN), 2014 International Joint Conference on*, 2014, pp. 1510–1516.

[12] K. Subramanian, A. K. Das, S. Sundaram, and S. Ramasamy, "A meta-cognitive interval type-2 fuzzy inference system and its projection based learning algorithm," *Evolving Systems*, vol. 5, no. 4, pp. 219–230, 2014.

[13] K. Subramanian, R. Savitha, and S. Suresh, "A Metacognitive Complex-Valued Interval Type-2 Fuzzy Inference System," *Neural Networks and Learning Systems, IEEE Transactions on*, vol. 25, no. 9, pp. 1659–1672, 2014.

[14] T. Nitta, "Orthogonality of decision boundaries in complex-valued neural networks," *Neural Computation*, vol. 16, no. 1, pp. 73–97, 2004.

[15] ——, "The computational power of complex-valued neuron," in *Artificial Neural Networks and Neural Information Processing-ICANN/ICONIP 2003*, 2003, pp. 993–1000.

[16] G. C. Cawley, N. L. C. Talbot, and M. Girolami, "Sparse multinomial logistic regression via bayesian l1 regularisation," in *NIPS*, 2006, pp. 209–216.

[17] J. B. Wagenaar, B. H. Brinkmann, Z. Ives, G. A. Worrell, and B. Litt, "A multimodal platform for cloud-based collaborative research," in *Neural Engineering (NER), 2013 6th International IEEE/EMBS Conference on*, 2013, pp. 1386–1389.

[18] M. Stead, M. Bower, B. H. Brinkmann, K. Lee, W. R. Marsh, F. B. Meyer, B. Litt, J. V. Gompel, and G. A. Worrell, "Microseizures and the spatiotemporal scales of human partial epilepsy," *Brain*, vol. 133, no. 9, pp. 2789–2797, 2010.

[19] C. Cortes and V. Vapnik, "Support vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[20] Y. U. Khan and J. Gotman, "Wavelet based automatic seizure detection in intracerebral electroencephalogram," *Clinical Neurophysiology*, vol. 114, no. 5, pp. 898–908, 2003.

[21] N. Hazarika, J. Z. Chen, A. C. Tsoi, and A. Sergejew, "Classification of EEG signals using the wavelet transform," in *Digital Signal Processing Proceedings, 1997. DSP 97., 1997 13th International Conference on*, vol. 1, 1997, pp. 89–92.

[22] L. D. Iasemidis and J. C. Sackellares, "The evolution with time of the spatial distribution of the largest Lyapunov exponent on the human epileptic cortex," *Measuring chaos in the human brain*, pp. 49–82, 1991.

[23] L. D. Iasemidis, D. S. Shiau, J. C. Sackellares, P. M. Pardalos, and A. Prasad, "Dynamical resetting of the human brain at epileptic seizures: application of nonlinear dynamics and global optimization techniques," *Biomedical Engineering, IEEE Transactions on*, vol. 51, no. 3, pp. 493–506, 2004.

[24] E. Bullmore, M. Brammer, G. Alarcon, and C. Binnie, "A new technique for fractal analysis applied to human, intracerebrally recorded, ictal electroencephalographic signals," *Neuroscience letters*, vol. 146, no. 2, pp. 227–230, 1992.

[25] K. Watanabe and T. Kurita, "Locality preserving multi-nominal logistic regression," in *International Conference on Pattern Recognition*, 2008, pp. 1–4.

[26] J. Wagenaar, B. Brinkmann, Z. Ives, G. Worrell, and B. Litt, "A multimodal platform for cloud-based collaborative research," in *Proceedings of IEEE Neural Engineering*, 2013, pp. 1386–1389.

[27] M. Stead, M. Bower, B. H. Brinkmann, K. Lee, W. R. Marsh, F. B. Meyer, B. Litt, J. V. Gompel, and G. A. Worrell, "Microseizures and the spatiotemporal scales of human partial epilepsy," *Brain*, p. awq190, 2010.

[28] S. Dora, S. Sundaram, and N. Sundararajan, "A Sequential Learning Algorithm for a Minimal Spiking Neural Network (MSNN) Classifier," in *Proceedings of the International Joint Conference on Neural Networks*, 2014, pp. 2415–2421.

[29] S. Dora, S. Ramasamy, and S. Sundaram, "A basis coupled evolving spiking neural network with afferent input neurons," in *Proceedings of the International joint conference on neural networks*, 2013, pp. 1–8.