Classification of Pre-Clinical Seizure States using Scalp EEG Cross-Frequency Coupling Features

Daniel Jacobs, Trevor Hilton, Martin del Campo, Peter L Carlen, and Berj L Bardakjian, *Member, IEEE*

Abstract— Objective: **This work proposes a machine-learning based system for scalp EEG that flags an alarm in advance of a clinical seizure onset.** *Methods:* **EEG recordings from 12 patients with drug resistant epilepsy were marked by an expert neurologist for clinical seizure onset. Scalp EEG recordings consisted of 56 seizures and 9.67 hours of interictal periods. Data from 6 patients was reserved for testing, and the rest was split into training and testing sets. A global spatial average of a cross-frequency coupling (CFC) index, $\bar{I}_{cfc}$, was extracted in 2s windows, and used as the feature for machine learning. A multistage state classifier (MSC) based on random forest algorithms was trained and tested on this data. Training was done to classify 3 states: interictal baseline, and segments prior to and following EG onset. Classifier performance was assessed using receiver-operating characteristic (ROC) analysis.** *Results:* **The MSC produced an alarm 45±16s in advance of clinical seizure onset across seizures from the 12 patients. It performed with a sensitivity of 87.9%, a specificity of 82.4%, and an area-under-the-ROC (AUC) of 93.4%. On patients for which it received training, performance metrics increased. Performance metrics did not change when the MSC used reduced electrode ring configurations.** *Conclusion:* **Using the scalp $\bar{I}_{cfc}$, the MSC produced an alarm in advance of clinical seizure onset for all 12 patients. Patient-specific training improved the specificity of classification.** *Significance:* **The MSC is non-invasive, and demonstrates that CFC features may be suitable for use in a home-based seizure monitoring system.**

Index Terms—**Epilepsy, clinical seizure onset, early detection, cross-frequency coupling features, machine learning, EEG, random forest, multi-stage state classification.**

I. INTRODUCTION

EPILEPSY affects over 50 million individuals in the world [1]. Despite major advances in recent years, pharmacotherapy controls seizures in only 70% of the epileptic population [2]. The remaining 30% of patients have what is termed medically refractory, or drug-resistant epilepsy [3]. The primary treatment for these patients is resective surgery, but only a fraction (perhaps as low as 6%) of patients are eligible

[4]. Moreover, in patients that do undergo resective surgery, only half of them will achieve long-term seizure freedom [5]. As such, there is a significant number of people living with the risks of epilepsy, the most feared of which may be sudden unexpected death in epilepsy (SUDEP) [6]. There exists a need for seizure detection and monitoring devices to alert caregivers to the occurrence of seizures in patients whose disease is not controlled [7].

False positives in these systems are a considerable problem. Pioneering seizure detection work using electroencephalograms (EEG) by Gotman [8] demonstrated low specificity; only 22% of detections were seizures. This was later improved to a sensitivity of 76% with a false detection rate of 0.84 per hour [9]. Subsequent studies have incorporated machine learning techniques in order to handle larger feature sets, but in several cases they were unable to outperform older systems [10], [11], or the specificity of the system was not reported [12]. Since the late 1990's, state of the art systems focus on patient specificity: they are able to achieve a sensitivity of 100%, with false detection rates around 0.02 per hour [13]. Patient specific algorithms have achieved these kinds of results using various wavelet-based, non-linear, and spectral features in combination with support vector machines (SVM) [14], recurrent neural networks [15], and logic based algorithms [16]. However, it remains a challenge to detect seizures using features that can generalize across patient datasets, and still provide a low rate of false alarms.

In addition to generalizing across patients, a desirable function of detection systems is to detect seizures as early as possible [17]. Early detection approaches are often capable of detecting seizures within several seconds of onset; published examples of detection latencies include 10s [18], 9.3s [13], 8s [14], and 7s [16]. Others have detections prior to the seizure onset, for example 4-10s [19]. A recent canine study was able to forecast seizure onset with a mean Time-in-Warning of 0.1 for 1-hr preictal time [20].

A novel feature which could allow improvements for early detection of seizure onset is pathological cross-frequency coupling (CFC). Phase-amplitude CFC was shown to be elevated near seizure initiation in the iEEG recordings of children with refractory epilepsy [21]. CFC of delta with high frequency oscillations (HFO) during seizures in the iEEG of patients with extra-temporal lobe epilepsy could discriminate between seizure and non-seizure generating tissues [22]. Delta-HFO CFC was also shown to discriminate between epileptogenic and non-epileptogenic tissues in the brain recorded during non-seizure periods when patients were experiencing non-REM sleep [23]. Delta-HFO CFC measures have also been shown to

identify seizure states, as well as the transitions between them using unsupervised machine learning [24] and have been used as features to detect seizures in long term iEEG recordings [25]. The ability of CFC to discriminate epileptic activity both spatially and temporally in the iEEG suggests that it may also be a viable feature for detecting seizure state transitions leading up to seizure in the scalp EEG.

This paper presents a state-transition detection system for scalp EEG in the form of a multistage state classifier (MSC). The novel aspects of this work are 1) the proposed MSC structure, and 2) the use of CFC features for machine learning classification of seizure states. In Fig. 1, we show an example of the MSC's action. In the following sections we will describe the patient dataset, the machine learning-based pipeline, and the specific signal features that make this possible.

II. Materials and Methods

A. Patient Recordings of Seizure States

In this section, we first describe the patients selected for this study, and then explain how the data was organized for use in the machine learning pipeline.

1) Patient Overview: Patient recordings were collected from 12 subjects who underwent pre-surgical evaluation at the Toronto Western Hospital Epilepsy Monitoring Unit. The dataset is divided into two subsets: Group 1 (N=6, patients 1-6) with recordings split between both training and testing, and Group 2 (N=6, patients 7-12) reserved exclusively for testing purposes. Patients were selected on the availability of recorded interictal and clinical seizures separated by hours. Patients in Set 1 had the additional constraint of available simultaneous iEEG, scalp EEG, and video recordings. The simultaneous iEEG was used to define electrographic (EG) seizure onset only, for machine learning training purposes. Electrographic seizure onset was defined by an expert neurologist at the first signs of epileptiform activity within the recorded EEG alone. Clinical seizure onset was defined by the same observer using visual changes in patient behavior such as loss of motor control and convulsions. All available patients meeting these criteria were included. Informed consent was obtained from each patient and the ethics committees of the affiliated institutions approved this study. The dataset is summarized in Table I.

A total of 56 seizures and 9.67 hours of interictal recordings were collected. All available seizure and interictal data that was free from large-scale movement artifacts were included in this study. Seizure epochs consist of the complete seizures events with some baseline prior and subsequent. Sampling rates were between 500 to 1,024 Hz, depending on the patient, and all EEG recordings were FIR notch filtered at 60 Hz and harmonics up to the Nyquist frequency in each patient. All recordings used a global reference at the Fpz electrode location. All patients' scalp EEG used 25 channel recordings placed according to the 10-20 system, as well as channels recording left and right sphenoidal, supraorbital, electroocu-logram (EOG), and an electrocardiogram (ECG) for a total of 32 channels. The exception was patient 4, in which case the EEG used 19 channel recordings placed according to the 10-20 system as well as EOG and ECG recording channels.

In order to keep the channels involved in analysis consistent across patients, only the 19 EEG channels common across all patients were used. EOG and ECG recordings were not included in the analysis. The iEEG recordings were used only to identify electrographic seizure onsets.

In all seizures from the dataset, EG and clinical seizure onset were marked by an expert neurologist. Patient 4 exhibited clear EG onset, but since the seizures were nocturnal and non-convulsive, no clear clinical onset was apparent.

2) Creation of Training and Testing Datasets: Recordings for patient Group 1 were divided between training and testing sets as described in Table I. Half of recorded seizures from each patient, and a total of 2.75 hours of interictal recordings, were set aside for training. This training set was later randomly subdivided into validation subsets during a 5-fold cross-validation process. The remaining seizures and 2.41 hours of interictal recordings were set aside for testing. All recordings from patient Group 2 were set aside for testing. In total, 33 seizures and 6.92 hours of interictal recordings were for used testing.

All interictal test segments were extracted in roughly 10-minute epochs. For seizure test epochs, the largest of the per-patient average EG-to-clinical onset duration was 66s. Therefore, seizure epochs contain the entire seizure with an additional 66s of recording leading up to the clinical seizure onset. For both training and testing sets, the long epochs were segmented into 2s non-overlapping windows for feature extraction.

B. Global Cross-Frequency Coupling Feature

The feature vector extracted from the 2s scalp EEG windows is a global spatial average (across EEG channels) of a CFC measure. Feature extraction was performed in MATLAB (The MathWorks, Natick, MA, U.S.A.) following the process described below.

1) Continuous Wavelet Transforms: Time-frequency spectrograms were obtained by applying a continuous wavelet transform (CWT), specifically a complex Morlet wavelet, on the EEG time-series recordings $x(t)$,

$$W(s, \tau) = \int_t x(t)\phi_{s,t}^*(t)dt \tag{1}$$

where,

$$\phi_{s,t}^*(t) = \frac{1}{\sqrt{s}}\phi_o\left(\frac{t-\tau}{s}\right) \tag{2}$$

is the Morlet basis function with $*$ denoting the complex conjugate. The Morlet is defined as,

$$\phi_o(t) = \frac{1}{\sqrt{2\pi}}e^{i\omega_c t - \frac{t^2}{2}} \tag{3}$$

The scales were transformed to frequencies f from the angular frequency ω_c using the relation $\omega_c = 2\pi f s = 5.1$ rad s^{-1}. The CWT produces a complex valued coefficient matrix,

$$W(f, t) = w(f, t) + j\widetilde{w}(f, t) \tag{4}$$

from which the amplitude and phase of the signal is obtained over time, per the sampling rate, and frequency ranges per the
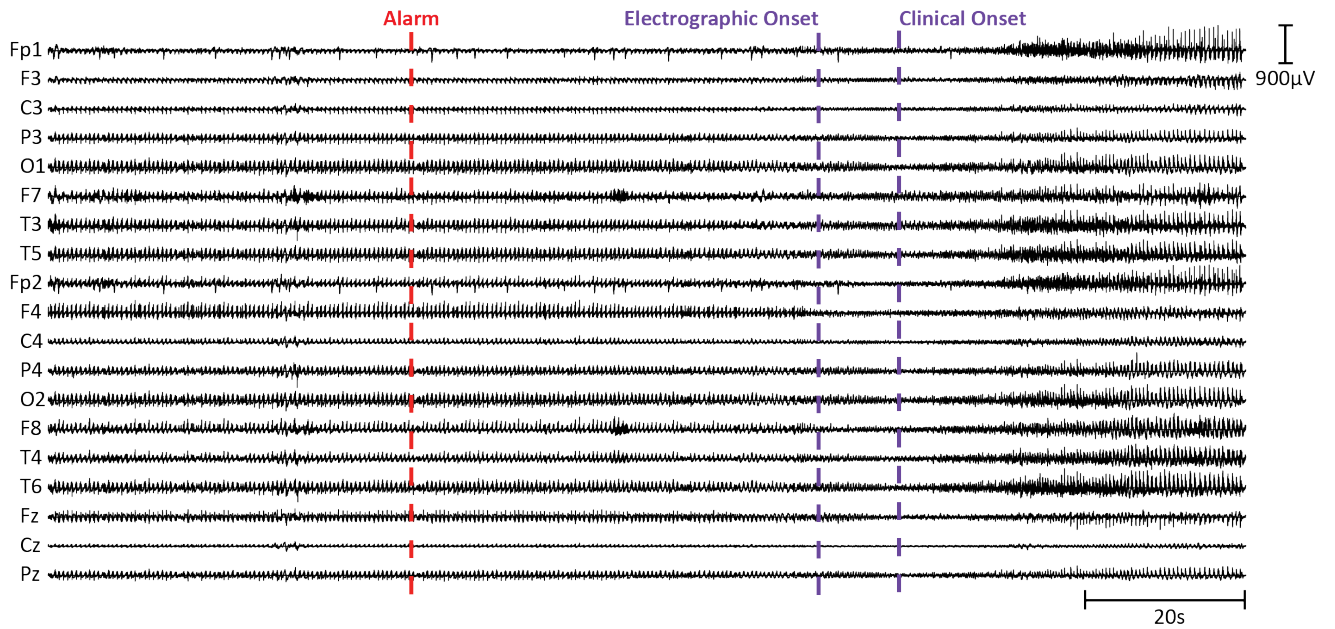
Fig. 1. The alarm (vertical red line) marked by automated multistage state classification (MSC) of the EEG recording of patient 1, seizure 2. The alarm is triggered solely from scalp EEG features. The alarm output by the MSC precedes electrographic and clinical seizure onset times (vertical purple lines) by 51s and 61s respectively. The onset timings were determined by an expert neurologist. Recordings were made at 500Hz, referenced to the Fpz electrode, and notch filtered at 60Hz and harmonics up to the Nyquist frequency.

desired bands. For this work, a Morlet CWT with bandwidth of 6Hz was used to obtain spectrograms over 2 frequency ranges for each 2s epoch: a low frequency range denoted $f_L \in (1, 1.1, 1.2, ..., 10)Hz$ and a high frequency range denoted $f_H \in (20, 21, 22, ..., 150)Hz$. The low frequency range was selected to cover the delta and theta ranges as in [24], while the high frequency range was cut off at 150 Hz to encompass the maximal extent of the 3 dB-below-maximum of the average CFC values, whose computation is discussed below. CWTs were computed with a buffer of 2.87 seconds on either side of the epoch to mitigate the end effects introduced by the CWT filtering operation at the low frequency scales. A CFC measure can now be computed from the complex wavelet coefficients of f_ϕ and f_A for each 2 second epoch.

2) Global Index of Cross-Frequency Coupling: The CFC measure selected is Tort's modulation index [28], from here on referred to as the index of cross-frequency coupling (I_{cfc}). As described by Tort, I_{cfc} assesses coupling between an amplitude envelope time series, here $A(t, f_H)$, and an instantaneous phase time series, here $\phi(t, f_L)$. These are derived from the respective complex wavelet coefficients over each 2s window t,

$$A(t, f_H) = |w(t, f_H) + j\widetilde{w}(t, f_H)| \quad (5)$$

$$\phi(t, f_L) = \arctan \frac{\widetilde{w}(t, f_L)}{w(t, f_L)} \quad (6)$$

The $\phi(t, f_L)$ is binned into $N = 18$ bins of 20° each, and the mean of $A(t, f_H)$ is computed in each bin, and normalized by the sum over all bins according to:

$$p_j(t, f_H, f_L) = \frac{\langle A(t, f_H) \rangle_j}{\sum_{k=1}^{N} \langle A(t, f_H) \rangle_k} \quad (7)$$

Now, p_j represents a discrete probability density value, where j indicates the phase bin number which is associated with f_L. An entropy measure,

$$H(t, f_H, f_L) = -\sum_{j=1}^{N} p_j(t, f_H, f_L) \log(p_j(t, f_H, f_L)) \quad (8)$$

can be determined and normalized to obtain the I_{cfc},

$$I_{cfc}(t, f_H, f_L) = \frac{H_{max} - H(t, f_H, f_L)}{H_{max}} \quad (9)$$

where H_{max} is the maximum possible entropy value, which for a uniform distribution has a value $H_{max} = \log(N)$. The global I_{cfc}, denoted $\bar{I}_{cfc}$, gives a global measure of the CFC characteristics observed over multiple EEG recording channel locations. For a multivariate time series $x_m(t)$ having $m = 1, 2, ..., M$ channels:

$$\bar{I}_{cfc}(f_A, f_\phi) = \langle I_{cfc}(f_A, f_\phi) \rangle_M \quad (10)$$

where $\langle ... \rangle_M$ denotes the spatial average over M channels. $\bar{I}_{cfc}$ computed over the frequency ranges f_H and f_L results in a 131x91 comodulogram, where each $\{f_H, f_L\}$ pair corresponds to a feature, of which there are $P = 11,921$ in total. The comodulogram can be denoted as a feature vector x_p, where $\{x_p \in [0, 1], p = 1, 2, ..., P\}$.

An example of the variation in the $\bar{I}_{cfc}$ comodulogram leading up to and during Seizure 1 of Patient 1 is shown in Fig. 2.

3) $\bar{I}_{cfc}$ Feature Vector Thresholding: A final step in the feature extraction, taken to reduce computational burden during training and testing of the MSC, is to apply a binary threshold to the $\bar{I}_{cfc}$. An example of this process is shown in Fig. 3. The original $\bar{I}_{cfc}$ comodulogram is shown in Fig.

TABLE I
PATIENT DATASET

Patient No.	Sampling Rate (Hz)	Training		Testing	
		No. of Seizures	Interictal Data (hrs.)	No. of Seizures	Interictal Data (hrs.)
1	500	1	0.50	1	0.46
2	500	10	0.39	10	0.17
3	500	2	0.36	2	0.49
4	1000	7	0.50	6	0.50
5	1000	2	0.50	1	0.33
6	1000	1	0.50	1	0.46
7	500			3	1.47
8	500			4	1.00
9	1024			1	0.23
10	512			2	0.87
11	512			1	0.27
12	512			1	0.66
Totals		**23**	**2.75**	**33**	**6.92**

* Recordings for Patients 7 - 12 were used for testing only in the proposed machine-learning system.

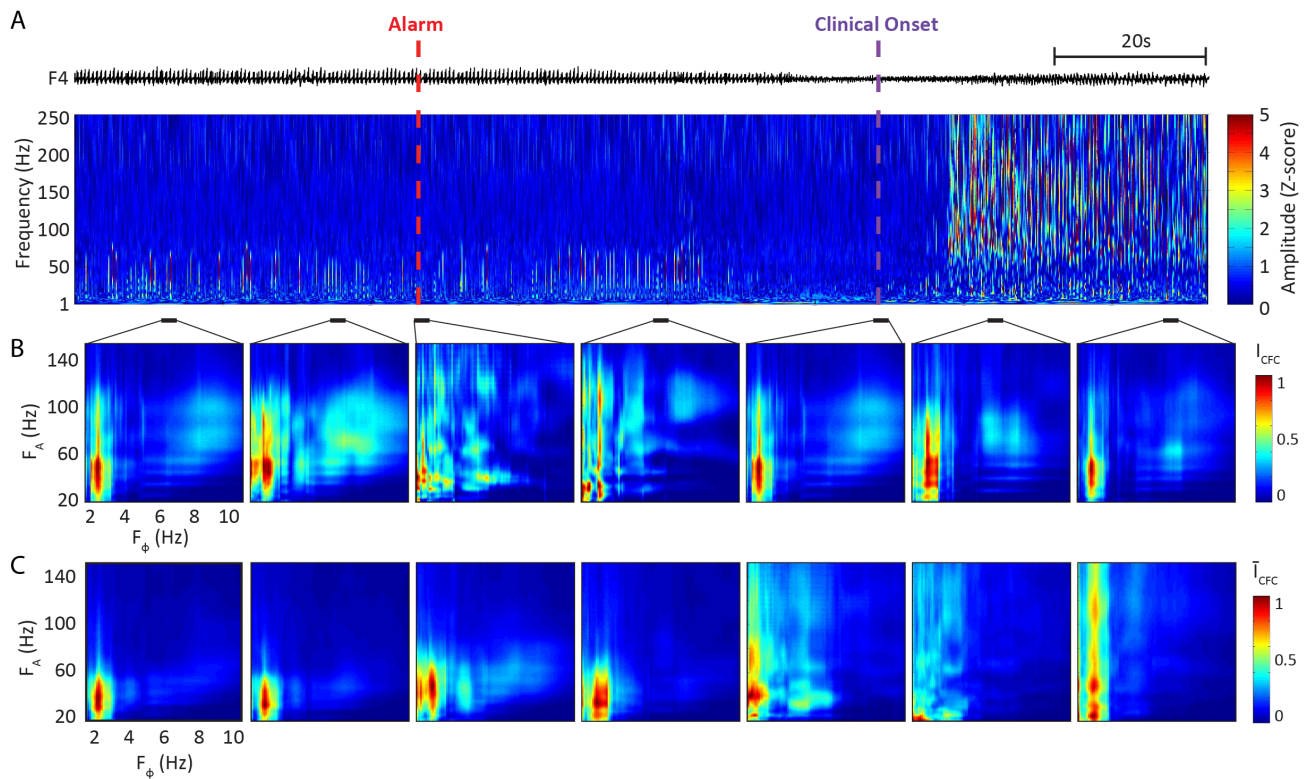No.: Number; sec.: seconds; hrs.: hours.

Fig. 2. Variation of global cross-frequency coupling in scalp EEG over time. (A) Example EEG channel F4 from patient 1, seizure 2, and the corresponding continuous wavelet transform. The wavelet is z-score normalized over time within each frequency band. (B) The cross-frequency coupling index comodulogram at example 2s windows for channel F4. (C) The global average of cross-frequency coupling index across all channels. This feature is thresholded and used for machine learning classification.

3 (A). As shown in Fig. 3 (B), the distribution of values in the $\bar{I}_{cfc}$ is often skewed. The elements of x_p can be sorted in ascending order, and a threshold th_{cfc} can be selected at the nth quantile of their distribution. All elements above threshold are set to 1, all elements below are set to 0. This thresholding procedure was carried out over two regions, as shown in Fig. 3 (C), in order to preserve some of the high frequency (60Hz+) content which would otherwise be eliminated. The exact value of th_{cfc} is determined for each classifier within the MSC during training via 5-fold cross-validation (CV), an example
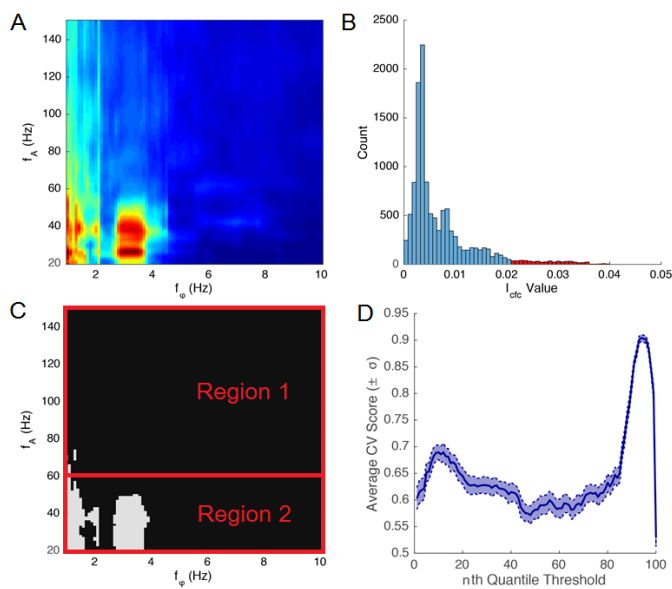
Fig. 3. Feature vector for state classification. (A) I_{cfc} is calculated for all selected channels, and the cross-channel spatial average produced, referred to as $\bar{I}_{cfc}$. An example of $\bar{I}_{cfc}$ is shown for a 2s window from patient 1. (B) The distribution of values contained in the $\bar{I}_{cfc}$ are displayed in a histogram. A threshold th_{cfc} is applied; red bars indicate values that are above th_{cfc} in the distribution. (C) $\bar{I}_{cfc}$ after thresholding at the th_{cfc} quantile, performed in two separate regions in order to preserve high-frequency CFC information. This matrix of binary values is reshaped into the feature vector for machine learning. (D) The th_{cfc} value is selected through 5-fold cross-validation (CV) of the classifier during training.

of which is shown in Fig. 3 (D).

C. Machine Learning Classification Approach

The general machine learning approach to supervised detection is described in Fig 4 (A), where three main stages of the project are outlined. Firstly, the EEG data is separated into training and testing sets, as discussed earlier. Secondly, features from the training set of EEG are extracted (in this case, the $\bar{I}_{cfc}$) and used to train and cross-validate the MSC. Finally, the trained MSC is assessed on the set of data reserved for testing. The overall structure of the MSC, and the training and testing approaches taken, are outlined in this section. The MSC was implemented in Python using Scikit-learn machine learning libraries. [29]

1) State Labels for Training: The state labels used for training purposes are described in Fig 4 (B). Feature vectors extracted from interictal baseline windows are labeled I_I. Feature vectors extracted from windows within the 10s preceding EG are labeled S_1. Feature vectors extracted from windows within the 10s following EG are labeled S_2.

The rationale for these states are as follows. Seizure onset is most clearly defined in the EEG by the EG. While the EG may not be clearly visible in the scalp EEG, it is well defined by the simultaneously recorded iEEG, and used clinically to help delineate the epileptogenic zone (EZ). The clinical rationale for selecting EEG segments pre and post EG is twofold: firstly, by identifying a state before EG onset and distinguishing it from baseline, it might be used for early detection, minimizing alarm latency. Secondly, by combining information about two

state transitions, the rate of false positive alarms may be reduced. The rationale for the 10s window length is from previous work [30] [24] from our group, where 10s was found to contain sufficient information to localize the EZ.

2) MSC Structure and Training: The structure of the MSC is described in Fig. 4 (C). The MSC consists of a three random forest (RF) classifiers, with basic logical decision thresholds governing internal state transitions. An RF classifier is an ensemble learning method that averages the output of many tree-structured classifiers produced from the same training data set. The result is a low-noise, low-bias classifier [31]. The general RF algorithm is described in [32].

The MSC operates as follows. It starts in the internal baseline state, I_I. The $\bar{I}_{cfc}$ feature vector is passed through the first RF classifier, called $I_I S_1$, which makes a decision if the vector belongs to the I_I or S_1 state. If the classifier output is below a set decision threshold th_{roc}, the MSC remains in the I_I internal state, waits for new input, and repeats. If the classifier output exceeds th_{roc}, the MSC moves to the internal S_1 state. A feature vector from the next 2s window is then tested by the second RF classifier $I_I S_2$, which tests for I_I vs S_2. If this classification falls below th_{roc}, the system falls back to $I_I S_1$. If it passes, the third RF classifier $S_1 S_2$ tests for S_2 vs S_1, on the basis that S_2 should only occur following S_1, and S_2 must be distinct from S_1. This multistage process was intended to increase the specificity of the system, and to incorporate a-priori information about the temporal arrangement of S_1 and S_2 states. Only if the $S_1 S_2$ classification exceeds th_{roc} is an alarm activated. The system is subsequently reset to the original state. The value of th_{roc} is determined at the testing stage via receiver-operating characteristic (ROC) analysis, and the same th_{roc} is applied to each of the RF classifiers.

During training of the MSC, 2 parameters were optimized. Multi-iteration 5-fold ROC-based cross validation over the training set was performed for this purpose. Firstly, the optimal number of estimators in the ensemble for each RF was found to be 200 for $I_I S_1$ and $I_I S_2$, and 150 for $S_1 S_2$. Secondly, the optimal value for th_{cfc} was found to be the 94th, 95th and 53rd quantiles for $I_I S_1$, $I_I S_2$, and $S_1 S_2$ respectively. (This optimization process was shown for $I_I S_2$ in Fig 3 (D).)

3) MSC Testing Performance Assessment: MSC testing was conducted using the interictal epochs (each roughly 10mins in duration) and seizure epochs (each consisting of the seizure plus 66s leading up to clinical onset). Each non-overlapping 2s window within these epochs was tested. For the interictal epochs, an alarm is a False Positive (FP), and the lack alarm is a True Negative (TN). For the seizure blocks, an alarm is a True Positive (TP), and the lack of alarm is a False Negative (FN).

The overall performance of the MSC is determined primarily through receiver operator characteristic (ROC) curves. These curves permit assessment of the system's optimal performance as determined by sensitivity, specificity, and accuracy. Sensitivity, i.e. the proportion of seizure events detected, is defined as:
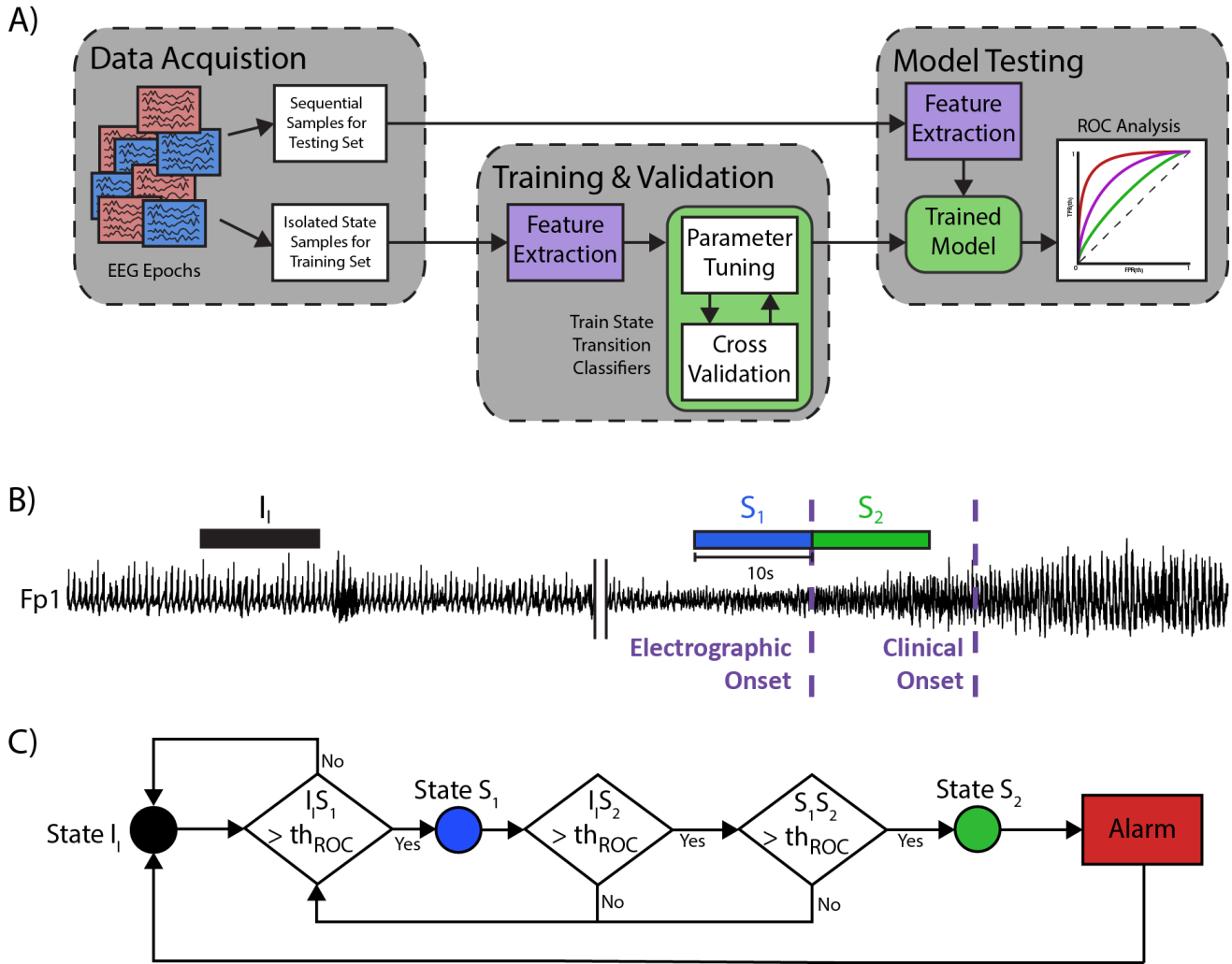
Fig. 4. (A) Strategy of the MSC project. Firstly, the dataset is divided into training (Group 1) and testing (Group 1 and Group 2) samples. Secondly, the MSC is optimized using features from the training set. Thirdly, the model is assessed on the testing set. (B) Training of the MSC involves labeling of samples. S_1 and S_2 states are defined as the 10s immediately prior to and following the neurologist-identified electrographic onset. The interictal state I_I is viewed as baseline. (C) The proposed MSC state logic. The MSC first loops for I_I to S_1 transitions. If the S_1 state is achieved, the system tests the subsequent 2s window for both I_I to S_2 and S_1 to S_2 transitions. This multistage process is designed to reduce false positives and include a-priori information about the temporal arrangement of S_1 and S_2 states. Each decision represents the output of a RF classifier (diamond boxes). An alarm is raised only when transitions occur sequentially from the I_I to S_1 to S_2 states. After an alarm, the system reverts to I_I. The optimal value for decision threshold th_{roc} is selected by receiver-operating characteristic (ROC) analysis.

$$Sensitivity = \frac{\sum TP}{\sum TP + \sum FN}, \qquad (11)$$

1-specificity, i.e. the proportion of interictal segments that are falsely classified as seizures, is defined as:

$$1 - Specificity = \frac{\sum FP}{\sum FP + \sum TN}. \qquad (12)$$

Accuracy, i.e. the proportion of correct classifications, is defined as:

$$Accuracy = \frac{\sum TP + \sum TN}{\sum TP + \sum FP + \sum FN + \sum TN}. \qquad (13)$$

In order to produce the ROC curve, th_{roc} was applied across the MSC from [0, 1] in 0.05 increments. For each threshold level, a ($Sensitivity$, $Specificity$) pair is computed. In

addition, three metrics for assessing model performance can be extracted from the ROC. Firstly, the area-under-the-curve (AUC). Secondly, a specificity factor ξ_l which represents the best $Specificity$ for 100% $Sensitivity$. Thirdly, we propose a specificity factor ξ_e which represents how quickly the model reaches optimal $Sensitivity$ as a function of $Specificity$, defined as follows:

$$Sensitivity = 1 - exp(-\frac{1 - Specificity}{1 - \xi_e}). \qquad (14)$$

The exact value of ξ_e can be found by curve fitting the exponential expression to the points of the ROC curve, and the goodness-of-fit is determined by the R^2 value.

Due to the number of seizures (positive events) in this study, the ROCs demonstrate discrete upward jumps. To report the performance of the MSC with greater confidence, optimal $Sensitivity$, $Specificity$, and $Accuracy$ values are reported

from the ξ_e curve-fit. In all cases in this study, these values are more conservative representations of MSC performance than their counterparts directly from the original ROC.

III. RESULTS

A. Global CFC is a seizure state marker

An example of the variation in $\bar{I}_{cfc}$ between training states is shown in Fig. 5. CFC between delta (2-4Hz) and low gamma (20-50Hz) rhythms is strongest in the seizure states. This feature forms the underlying basis for distinguishing interictal I_I vs S_1, S_2 states. The feature is much clearer in the iEEG due to the reduced noise in the recording, but is still visible in the scalp due to the averaging properties of the $\bar{I}_{cfc}$ computation. As seen in the iEEG, the pathological CFC extends beyond the low gamma into the HFO (>80Hz) range, but the HFO contribution cannot be detected at the scalp.

Differences between the S_1 and S_2 seizure states can be seen in two ways. Firstly, S_1 demonstrates a stronger CFC than S_2 in the delta-gamma range. Secondly, S_1 demonstrates a theta (4-8Hz) rhythm modulating the low gamma which is not seen in S_2. These features appeared consistently across the patient dataset, enabling the MSC to perform reasonably well even in the 6 patients for which it received no training.

B. Classifier Performance using All Channels

Performance of the MSC was assessed across various sub-groups of the testing set: Group 1 (for which the MSC received training), Group 2 (for which the MSC received no training), and All patients. The ROC curve for each group is shown in Fig 6. The MSC performed well for all groups, but best for Group 1, as indicated by the AUC = 0.99, 0.92, and 0.93 for Group 1, 2, and All respectively. A similar trend is shown in the ξ_e curve fit, for which ξ_e = 0.99, 0.90, 0.92 for Group 1, 2, and All respectively. Group 1 also achieved 100% sensitivity with the highest specificity, as shown by the ξ_l = 0.96, 0.83, 0.86 for Group 1, 2, and All respectively. Patient specific information in Group 1 seems to increase the specificity of the system, although performance is reasonable in Group 2 for which no training occurred.

For Group 1, the MSC achieved sensitivity of 97.5%, specificity of 95.0%, and an overall accuracy of 95.0%. For Group 2, sensitivity of 85.2%, a specificity of 79.9%, and an overall accuracy of 79.9%. For the All patient set, the MSC achieved a sensitivity of 87.9%, a specificity of 82.4%, and an overall accuracy of 82.4%. This comparison is summarized in Table II, alongside the ξ_e and ξ_l from the ROC curves for each set. The seizure alarm time does not appear to differ significantly between Group 1 vs Group 2 vs All.

The alarm times produced by the MSC prior to clinical seizure onset are shown in Table III for All patients. These alarm times are reported for the th_{roc} that corresponds to ξ_l=0.86, where 100% sensitivity is achieved for greatest possible specificity across the All patients. For comparison, the alarm times prior to electrographic seizure onset are also shown for Group 1 patients, since EG onset (used only during the training process) was not a requirement for Group 2. Across all tested clinical seizures (N_s=33), alarm was sounded 45±16s (mean ± standard deviation) prior to clinical onset.

C. Classification using Reduced Channel Sets

A desirable characteristic for the MSC is to function on a reduced channel set for purposes of mobility and ease of use. Accordingly, the performance of the MSC was explored as a function of scalp EEG channel coverage that might occur in a headband-type device. The set of 19 channels was reduced to two new circumferential ring subsets of 10 and 8 channels respectively. The CZ electrode was not included. These ring configurations are shown in Fig 7, along with the ROC performance for each configuration.

The AUC values for the three configurations differ by only 0.06%. The ξ_e specificity factor varies only by 0.012. Detection latencies (reported for the th_{roc} corresponding to ξ_l for each channel subset) are shown in Table IV, along with sensitivity, specificity and accuracy metrics. Sensitivity differs by 1.9%, specificity by 1.5%, and accuracy by 1.5%. The pre-clinical seizure alarm times do not appear to be significantly different. Overall, performance does not appear to be influenced by the choice of electrode subset.

IV. DISCUSSION

A. Existing EEG-based State Detection

Computerized monitoring of EEG states in a clinical setting was achieved with reasonable success as early as the early 1990's [34]. These systems suffered from low specificity and sensitivity. In order to combat these issues, systems began to employ patient-specific tuning [13]. In order to further address false positive rates, some systems are invasive [35]. Such systems are generally viewed as obtrusive, which may limit their adoption [17]. The MSC which we propose does not have these restrictions: it functioned well for Group 2 patients that the system never trained for, and uses a convenient set of scalp channels. Patient-specific information, however, is very helpful to decrease the false positive rates, as seen by the performance for Group 1 patients in Fig. 6.

Table V contains a summary of recent seizure detection methods. Contrasting these studies with the current work, a number of points come to light. Firstly, the number of patients (N_p = 12) and seizures (N_s = 56) used in this study is rather low, but still within the range of previously published work. The relative dominance of the number of seizures from Patients 2 and 4, which contribute large portions of both training and testing data, is balanced by the patient variability (N_p = 12). Secondly, the sensitivity achieved by the MSC (87.9%) is comparable to other studies, and the specificity (82.4%) is on the low range. However, both sensitivity and specificity are comparable with the inclusion of patient-specific information (97.5%, 95% respectively). Thirdly, the detection ahead of clinical seizure (45 ± 16s) permits a window for intervention which other detection methods may not offer.

B. $\bar{I}_{cfc}$ in Scalp EEG

The proposed MSC has achieved a high sensitivity and specificity by making use of the $\bar{I}_{cfc}$ feature. CFC can be involved in normal brain processes such as memory encoding [36] or auditory processing [37]. However, it can also
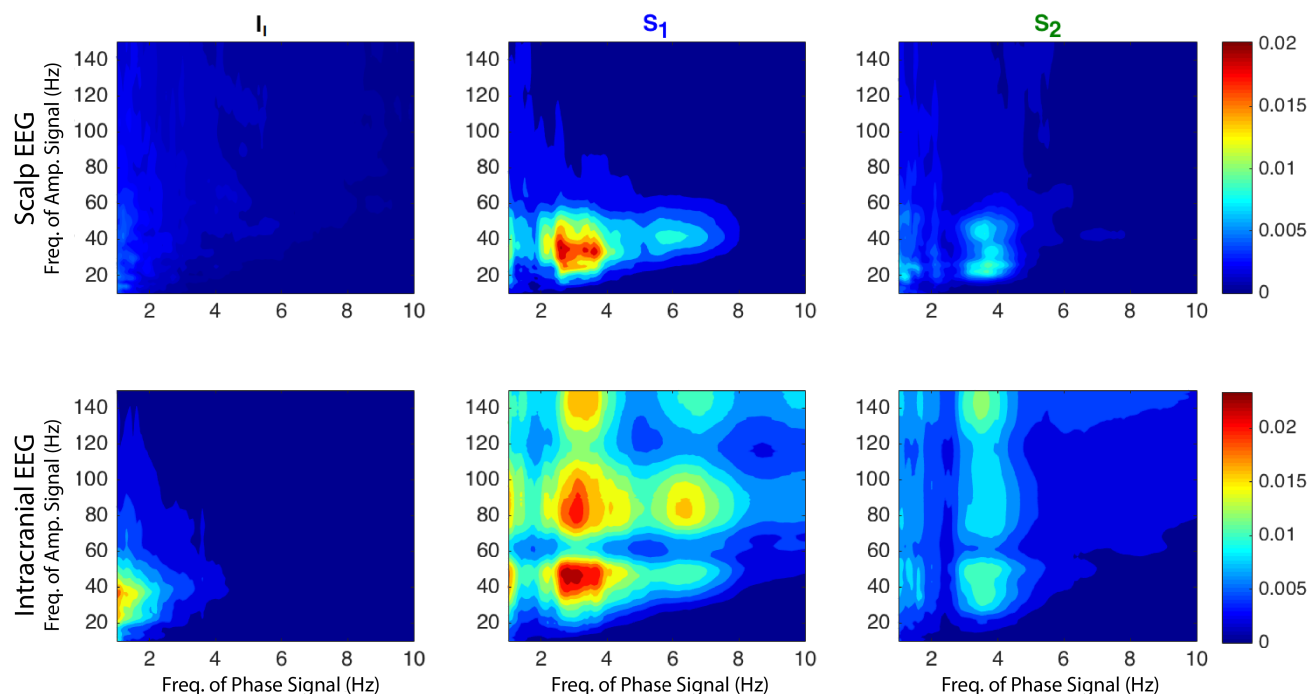
Fig. 5. $\bar{I}_{cfc}$ is shown for three ten second time segments corresponding to the three epileptic states of interest and in the two domains of interest. From left to right: interictal (I_I), pre-EG seizure onset (S_1), and post-EG seizure onset (S_2). The top and bottom panels show $\bar{I}_{cfc}$ characteristics in the scalp EEG and iEEG domains, respectively. In the EEG, $\bar{I}_{cfc}$ was the color bars on the right for each domain. Each comodulogram depicts CFC characteristics using the same frequency ranges for amplitude modulated (vertical axis) and phase modulating (horizontal axis) rhythms.

TABLE II
MSC PERFORMANCE METRICS ACROSS TEST SETS

Patient Group	Pre-Clinical Seizure Alarm Time (s)*	AUC (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	ξ_e	ξ_l
Group 1	48 ± 16	98.8	95.0	97.5	95.0	0.99	0.96
Group 2	47 ± 15	91.6	79.9	85.2	79.9	0.90	0.83
All	45 ± 16	93.4	82.4	87.9	82.4	0.92	0.86

* mean ± standard deviation

be involved in pathology: in epilepsy, CFC has been tied to seizure states and epileptogenic tissue [21] - [25]. The global averaging across channel locations of the scalp CFC map, and the subsequent thresholding by th_{cfc}, appear to be an appropriate selection process for identifying seizure state transitions in scalp EEG. By training the MSC on the S_1 and S_2 states, as seen in Fig 4 (B), CFC that is specific to seizure states appears to be highlighted.

It is important to note that the proposed MSC functions on the scalp EEG with no preprocessing beyond the notching out of electrical line noise and harmonics. This might lead to concerns about spurious I_{cfc}, which can result from spikes and fast activities in pre-ictal and ictal periods [26]. The spatial averaging inherent to the computation of the global CFC index (Equation 10) as a biomarker may reduce the influence of spurious CFC from individual channels. The statistical relevance of I_{cfc} features to the EZ was assessed by two previous studies from our group: firstly, in surrogate analysis on iEEG [22], and secondly, by noise simulation in scalp EEG [30]. In the current study, while the relationship between

fast activities and the measured $\bar{I}_{cfc}$ is not directly assessed, the performance metrics for the MSC on Group 1 patients especially would indicate that the machine-learning system is able to differentiate between spurious and physiologically-relevant CFC.

C. Early Seizure State Detection

The proposed MSC generated an alarm 43 ± 17s in advance of the clinical seizure onset. This early detection capability is quite remarkable given that it does not appear to change with patient-specific information, which does help to reduce false positive rates. Many published seizure prediction systems lean towards patient-specific tuning of the feature set in order to achieve success. However, in this case, the $\bar{I}_{cfc}$ appears to capture something common across the entire patient dataset. We propose that shifts in the $\bar{I}_{cfc}$ mapping of whole-scalp EEG are being detected before the visual changes in EEG that occur at electrographic onset. The global aspect of the $\bar{I}_{cfc}$ captures a trend across the entire scalp which appears to
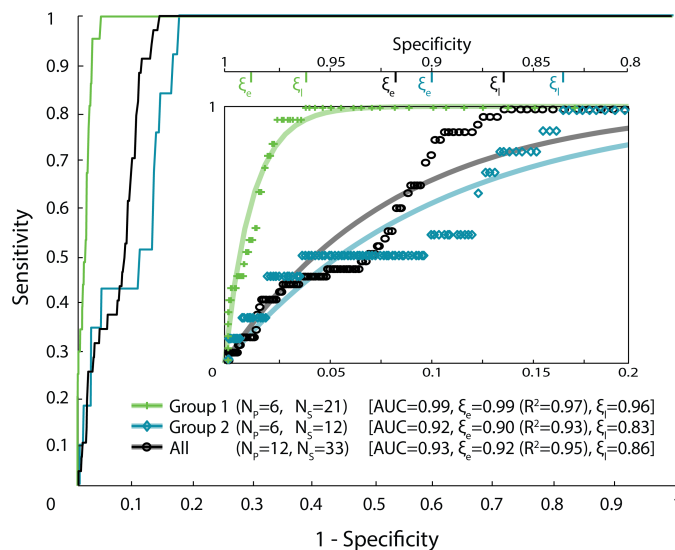
Fig. 6. MSC performance across testing subsets. The green curve represents performance on Group 1, for which the system received training. The turquoise curve represents performance on Group 2, whose seizures were not used for training. The black curve represents performance on all testing data. *Inset:* Each ROC represented by (sensitivity, specificity) points instead of a continuous curve. Faded lines represent ξ_e curve fits for each ROC. ξ_e and ξ_l values are indicated on the specificity axis above the inset. N_p = number of patients, N_s = number of seizures
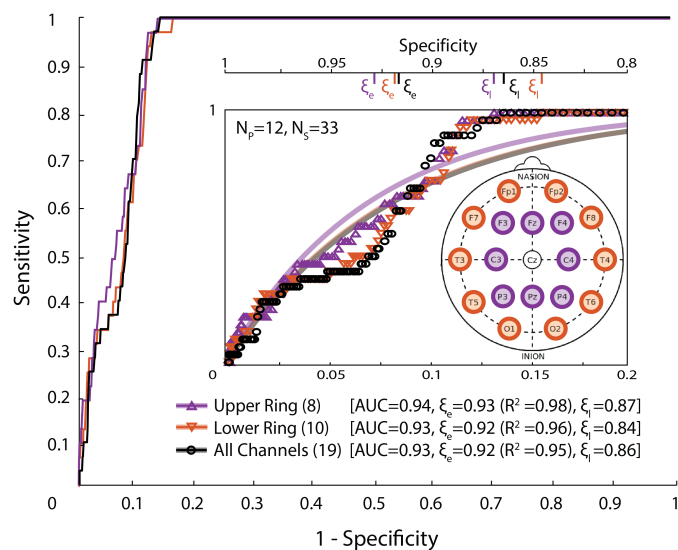

Fig. 7. MSC performance using information from different channel sets. The purple ROC represents performance using an upper ring of 8 channels (Fz, F3/4, C3/4, P3/4, Pz). The orange, a lower ring of 10 channels (Fp1/2, F7/8, T3/4, T5/6, O1/2). The black, All channels. *Inset:* Each ROC represented by (sensitivity, specificity) points instead of a continuous curve. Faded lines represent ξ_e curve fits for each ROC; Lower and All overlap. ξ_e and ξ_l values are indicated on the specificity axis above the inset. Overall, performance does not appear to change significantly across channel subsets. N_p = number of patients, N_s = number of seizures

to demonstrate comparable alarm times vs using 19 channels. This suggests that, with further testing involving long-term patient recordings, this MSC strategy could be valid for a home-based wearable EEG device.

V. CONCLUSION

It is possible to detect a pre-clinical seizure state transition using CFC features from the scalp EEG. A global index of cross-frequency coupling, $\bar{I}_{cfc}$, was used as the feature for a multistage state classifier (MSC) based on random forest algorithms. The system was able to provide alarms 45±16s ahead of seizure onset with a sensitivity of 87.9%, and a specificity of 82.4%. On patients for which it received training, sensitivity and specificity increased to 97.5% and 95.0% respectively. Even though the MSC was trained on a standard 10-20 system set of electrodes, the performance was not affected significantly when tested on reduced subsets of electrodes. This gives rise to the possibility of implementation in a headband-type wearable device for home-monitoring applications that would enhance the quality of life for epileptic patients.

ACKNOWLEDGMENT

REFERENCES

[1] A. K. Ngugi et al, "Estimation of the burden of active and life-time epilepsy: a meta-analytic approach," *Epilepsia*, vol. 51, no. 5, 2010.

TABLE III
SEIZURE ALARM INTERVALS PER PATIENT

Patient	No. of seizures	Mean Alarm to Clin. Onset (s) *	Mean Alarm to EG Onset (s) * ‡
1	1	61	47
2	10	41	36
3	2	53	35
4	6	non-convulsive†	60
5	1	30	20
6	1	55	35
7	3	43	
8	4	35	
9	1	58	
10	2	59	
11	1	57	
12	1	58	
Overall:	**33**	**45 ± 16 (33)**	**43 ± 17 (21)**

* population mean ± standard deviation (No. of seizures)

‡ Only scalp EEG used for patients 7-12.

† Patient 4's seizures were nocturnal and non-convulsive, so only EG onset was defined. A 66-second segment leading up to EG onset was used for testing.

No. = number; Clin. = clinical; EG = electrographic

contain useful information regarding state transitions involved with an impending seizure event.

For the pre-clinical seizure alarm times of the reduced channel sets, as summarized in Table IV, it is encouraging that using only 8 channels of scalp EEG, the MSC is able

TABLE IV
MSC PERFORMANCE METRICS FOR REDUCED CHANNEL SETS

Channel Set	Pre-Clinical Seizure Alarm Time (s)*	AUC (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	ξ_e	ξ_l
Upper Ring (8)	51 ± 11	93.9	83.9	89.8	83.9	0.93	0.87
Lower Ring (10)	54 ± 7	93.3	82.9	87.8	82.9	0.92	0.84
All Channels (19)	45 ± 16	93.4	82.4	87.9	82.4	0.92	0.86

* mean $\pm$ standard deviation

TABLE V
COMPARISON OF SCALP EEG-BASED SEIZURE DETECTION RESULTS

Reference	No. of patients	No. of seizures	Patient Specific	Sensitivity (%)	Specificity	Detection Delay
Wang et al. (2016) [39]	10	44	Yes	91.44	99.34%	n/a
Zabihi et al. (2016) [40]	24	161	Yes	88.27	93.21%	n/a
Sierra-Marcos et al. (2015) [41]	98	268	No	76.1	0.24/hr	n/a
Kelly et al. (2010) [42]	47	287	No	79.5	2/24hr	n/a
Zandi et al. (2010) [16]	14	63	Yes	90.5	0.51/hr	7s
Kuhlmann et al. (2009) [43]	21	88	Yes	80	0.60/hr	16.9s
Saab and Gotman (2005) [18]	44	195	No	76	0.34/h	9.8s

[2] J. A. French, "Refractory epilepsy: clinical overview," *Epilepsia*, vol. 48, no. s1, 2007.

[3] P. Kwan and M. J. Brodie, "Definition of refractory epilepsy: defining the indefinable?," *The Lancet.Neurology*, vol. 9, Jan 2010.

[4] D. J. Englot et al, "Epilepsy surgery trends in the united states, 1990-2008," *Neurology*, vol. 78, Apr 17 2012.

[5] J. F. Tellez-Zenteno et al, "Long-term seizure outcomes following epilepsy surgery: a systematic review and meta-analysis," *Brain : a journal of neurology*, vol. 128, May 2005.

[6] R. Surges et al, "Sudden unexpected death in epilepsy: risk factors and potential pathomechanisms," *Nature Reviews Neurology*, vol. 5, no. 9, 2009.

[7] A. V. de Vel et al, "Non-eeg seizure-detection systems and potential sudep prevention: state of the art," *Seizure*, vol. 22, no. 5, 2013.

[8] J. Gotman, "Automatic recognition of epileptic seizures in the eeg," *Electroencephalography and clinical neurophysiology*, vol. 54, no. 5, 1982.

[9] J. Gotman, "Automatic seizure detection: improvements and evaluation," *Electroencephalography and clinical neurophysiology*, vol. 76, no. 4, 1990.

[10] W. Webber et al, "An approach to seizure detection using an artificial neural network (ann)," *Electroencephalography and clinical neurophysiology*, vol. 98, no. 4, 1996.

[11] N. Pradhan et al, "Detection of seizure activity in eeg by an artificial neural network: A preliminary study," *Computers and Biomedical Research*, vol. 29, no. 4, 1996.

[12] W. Weng and K. Khorasani, "An adaptive structure neural networks with application to eeg automatic seizure detection," *Neural Networks*, vol. 9, no. 7, 1996.

[13] H. Qu and J. Gotman, "A patient-specific algorithm for the detection of seizure onset in long-term eeg monitoring: possible use as a alarm device," *IEEE transactions on biomedical engineering*, vol. 44, no. 2, 1997.

[14] A. Shoeb et al, "Patient-specific seizure onset detection," *Epilepsy & Behavior*, vol. 5, no. 4, 2004.

[15] G. R. Minasyan et al, "Patient-specific early seizure detection from scalp electroencephalogram," *Journal of clinical neurophysiology*, vol. 27, Jun 2010.

[16] A. S. Zandi et al, "Automated real-time epileptic seizure detection in scalp eeg recordings using an algorithm based on wavelet packet transform," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 7, 2010.

[17] D. F. T. Quiroga et al, "Patient and caregiver view on seizure detection devices: a survey study," *Seizure*, vol. 41, pp. 179–181, 2016.

[18] M. Saab and J. Gotman, "A system to detect the onset of epileptic seizures in scalp eeg," *Clinical Neurophysiology*, vol. 116, no. 2, 2005.

[19] R. Meier et al, "Detecting epileptic seizures in long-term human eeg: a new approach to automatic online and real-time detection and classification of polymorphic seizure patterns," *Journal of clinical neurophysiology*, vol. 25, Jun 2008.

[20] E. B. Assi et al, "Functional-genetic Scheme for Seizure Forecasting in Canine Epilepsy ," *IEEE Transactions on Biomedical Engineering*, 2017.

[21] G. M. Ibrahim et al, "Dynamic modulation of epileptic high frequency oscillations by the phase of slower cortical rhythms," *Experimental neurology*, vol. 251, 2014.

[22] M. Guirgis et al, "Defining regions of interest using cross-frequency coupling in extratemporal lobe epilepsy patients," *Journal of neural engineering*, vol. 12, no. 2, 2015.

[23] N. Ellenrieder at al, "Interaction with slow waves during sleep improves discrimination of physiologic and pathologic high-frequency oscillations (80–500 hz)," *Epilepsia*, 2016.

[24] M. Guirgis et al, "Modulated high frequency oscillations can identify regions of interest in human ieeg using hidden markov models," in *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, IEEE, 2015.

[25] K. Edakawa et al, "Detection of epileptic seizures using phase-amplitude coupling in intracranial electroencephalography," *Scientific reports*, vol. 6, May 5 2016.

[26] E. M. Gerber et al, "Non-Sinusoidal Activity Can Produce Cross-Frequency Coupling in Cortical Signals in the Absence of Functional Interaction between Neural Sources", *PLoS One*, vol. 11, no. 12, Dec 2016.

[27] J. Aru et al, "Untangling cross-frequency coupling in neuroscience," *Curr. Op. in Neurobiology*, vol. 31, pp. 51-61, 2015.

[28] A. B. Tort et al, "Measuring phase-amplitude coupling between neuronal oscillations of different frequencies," *Journal of neurophysiology*, vol. 104, Aug 2010.

[29] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, 2011.

[30] L. Chunsheng et al, "Epileptogenic source imaging using cross frequency coupled signals from scalp eeg," *IEEE Transactions on Biomedical Engineering*, vol. 63, 2016.

[31] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, 2001.

[32] J. Friedman et al, *The elements of statistical learning*, vol. 1. Springer series in statistics Springer, Berlin, 2001.

[33] J. Theiler et al, "Testing for nonlinearity in time series: the method of surrogate data," *Physica D: Nonlinear Phenomena*, vol. 58, no. 1-4, 1992.

[34] B. Rosenblatt, J. Gotman, "Computerized EEG Monitoring," *Seminars in Pediatric Neurology*, vol. 6, June 1999.

[35] K. Gadhoumi et al, "Seizure prediction for therapeutic devices: A review," *Journal of Neuroscience Methods*, vol. 260, 2016.

[36] J. Lisman, "Relating Hippocampal Circuitry to Function: Recall of Memory Sequences by reciporcal Dentate-CA3 Interactions," *Neuron*, vol. 22, February 1999.

[37] R. T. Canolty et al, "High Gamma Power Is Phase-Locked to Theta Oscillations in Human Neocortex," *Science*, vol. 313, Sept 2006.

[38] H. T. Shiao et al, "SVM-Based System for Prediction of Epileptic Seizures from iEEG Signal," *IEEE Trans Biomed Eng*, vol. 64, May 2017.

[39] G. Wang et al, "Epileptic seizure detection based on partial directed coherence analysis," *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 3, 2016.

[40] M. Zabihi et al, "Analysis of high-Dimensional phase space via poincare section for patient-Specific seizure detection,"*IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 3, 2016.

[41] A. Sierra-Marcos et al, "Seizure detection with automated eeg analysis: a validation study focusing on periodic patterns,"*Clinical Neurophysiology*, vol. 126, no. 3, 2015.

[42] K. Kelly et al, "Assessment of a scalp eeg-based automated seizure detection system,"*Clinical Neurophysiology,*, vol. 121, no. 11, 2010.

[43] L. Kuhlmann et al, "Seizure detection using seizure probability estimation: Comparison of features used to detect seizures,"*Annals of Biomedical Engineering,*, vol. 37, no. 10, 2009.