2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Detection of Tonic-Clonic Seizures using Wavelet Entropy of Scalp EEG

Joseph Mathew, *Member, IEEE,* Subha Ramakrishnan Manuskandan, N. Sivakumaran, and P.A. Karthick.

Abstract— **Epilepsy is the most common chronic neurologic disorder characterized by the recurrence of unprovoked seizures. These seizures are paroxysmal events that result from abnormal neuronal discharges and are categorized into various types based on the clinical manifestations and localization. Tonic-Clonic seizures (TCSZ) may lead to injuries, and constitute the major risk factor for sudden unexpected death in epilepsy (SUDEP), especially in unattended patients. Therapeutic decisions and clinical trials rely on Video EEG which is not practical outside of clinical setting. In this study, wavelet entropy of scalp EEG signals are utilized to discriminate the seizures with and without clinical manifestations. The scalp EEG records from the publically available Temple University Hospital (TUH) dataset are considered for this work. A seven-level, fourth order Daubechies (db4) wavelet is utilized for the decomposition of first four seconds of scalp EEG during seizures. The entropy is extracted from the resultant coefficients and are used to develop SVM based models. Most of the extracted features found to have significant differences ($p<0.05$). The results show that polynomial SVM model achieves an accuracy of 95.5%, positive predictive value (PPV) of 99.4%, negative predictive value (NPV) of 91.57% and F-Score of 95.9%. Therefore, the proposed approach could be a support in detecting life-threatening seizures.**

I. Introduction

Epilepsy is a chronic disease, defined by unusual enhanced synchrony of neurons called seizure [1]. The infrequent and uncertain occurrence of seizure activity causes disability and impairment to patients [2]. The patients experience physical and physiological complications during seizure and their conditions become worse when the seizure is uncontrolled. Tonic-Clonic seizure (TCSZ) is such an uncontrolled seizure for which there is a prevalence of sudden unexpected death [3]. These complications impedes the quality of life. Some of the difficulties in managing epilepsy can be improved by the ability to detect these clinical seizures. This rapid and accurate information can enhance the therapeutic and diagnostic applications [4].

In the literature, several methods have been proposed to automatically detect seizure episodes. Researchers have used nonlinear and non-stationary approaches that include time, frequency, and time-frequency domain approaches for the detection of epileptic discharges.[4] In addition, the machine learning models such as fuzzy logic, Artificial Neural Network (ANN) [5] and Adaptive Neuro-Fuzzy Inference System (ANFIS) [6] have employed to discriminate seizure and normal brain activities[7] [8]. Ictal and interictal states of epilepsy have been classified using ANN and Support Vector Machines (SVM) [9]. SVM algorithm have been utilized to develop models from the time and frequency domain features for the detection of electro clinical seizure such as generalized TCSZ [4]. Recently, wavelet transform based statistical features have been employed to differentiate the seizure onset and spread regions from the intracranial EEG signals [10].

In most of the researches, the detection models are generated using signal processing and machine learning approaches mainly to distinguish between the ictal and interictal and, normal and abnormal conditions of brain activities. There are only a few research works focused on the classification of seizure types [11]. In this work, an attempt has been made to differentiate the TCSZ and electrographic seizures from the scalp EEGs. A fourth order Daubechies (db4) wavelet transform is used to decompose the scalp EEG signals into seven levels. Then the entropy is extracted from the wavelet coefficients to analyze the complexity of these signals. A total of eight features are employed to build a Support Vector Machine (SVM) classification model for the detection of TCSZ.

II. Methods

The scalp EEG signals used in this study are taken from publicly available Temple University Hospital (TUH). Database. It is one of a large database that contains many seizure types. Among this, Focal Non-specified Seizure (FNSZ), Generalized Non-specified Seizure (GNSZ) and TCSZ were selected from consecutive 28 patients. TCSZ is electro clinical seizure that has severe physical and physiological complications during seizure. Whereas FNSZ and GNSZ are considered as electrographic seizures (EGSZ) that do not exhibit any clinical manifestations during the seizure [12].

FNSZ is specifically localized in the hemispheric or focal part of the brain and appeared in broad range of seizure etiologies. The primary indicator of seizure event is morphology of EEG signals and the seizure will show a spike and slow wave or poly spike and slow wave complex [13]. GNSZ follow the same morphology, evolution, and frequency descriptors of FNSZ, but they cover a larger area of the skull.

Joseph Mathew, N. Sivakumaran and P.A Karthick are with the Department of Instrumentation and Control Engineering, National Institute of Technology, Tiruchirappalli, Tamil Nadu, India.
Subha Ramakrishnan Manuskandan is with Karuvee Innovations Private Limited.

Corresponding author email is jchackompally@gmail.com

TCSZ is a type of electro clinical seizure, characterized by muscle tension and stiffening followed by violent convulsions and jerking [13]. Thus, this type of seizure can't be identified without clinical manifestations.

The monopolar channels were optimized and converted to bipolar using TCP montage [14]. Only the seizures that were recorded at the sampling rate of 400 Hz were considered in this work. There were 64 seizures fulfilled this criteria.

A. Wavelet Transform (WT)

Wavelet transform is employed to utilize the non-stationary property of scalp EEG signals. In wavelet transform (WT), the selected EEG signal is considered as the root node of the tree and the signal is decomposed by wavelet function into number of subspaces [15]. The continuous wavelet transform of a signal $x(t)$ is computed by multiplying the integration of the signal with the scaled and shifted version of mother wavelet ψ and is expressed as

$$CWT(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\psi\left(\frac{t-b}{a}\right) dt \qquad (1)$$

where a, and b, are scaling and shifting parameters respectively.

One of the most reliable and digital implementation of wavelet transform is discrete wavelet transform (DWT). This method is based on a filter bank decomposition. It uses set of high-pass and low-pass filters that compute wavelet detail coefficients (*D1-D7*) and approximation coefficient (A1-A7) respectively at decomposition level *j*. The DWT that uses power of two is very convenient and efficient for discrete signals such as EEG [14].

B. Feature Extraction

In the feature construction, DWT coefficients generated at each level are utilized to construct the features. The decomposition levels are fixed to obtain different sub-band frequencies such as 100-200 Hz (D_1), 50-100 Hz (D_2), 25-50 Hz (D_3), 12.5-25 Hz (D_4), 6.25-12.5 Hz (D_5), 3.125-6.25 Hz (D_6), 1.56-3.125Hz (D_7), and 0-1.56Hz (A_7). Entropy measures the degree of randomness of signals. In order to analysis characteristic nature of the EEG signals, non-linear feature (entropy) is derived from the wavelet coefficients and employed for the classification [16]. The expression for the entropy is

$$Entropy = \sum_{i} P(Di)\ log_2\ P(Di) \qquad (2)$$

Where *P (Di)* is the probability of the magnitude of wavelet coefficients.

C. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a supervised classification method and is considered for high dimensional binary classification problems [17]. This approach determines best decision boundary by estimating margin between the two classes. In this study, a linear, polynomial and radial basis function (RBF) kernel are used to construct the decision boundary. The kernel functions are defined in the Table I [18]. The order of the polynomial kernel function for kernel is set as three.

The sensitivity, specificity, positive prediction value (*PPV*), negative prediction value (*NPV*), accuracy and F-score are used to evaluate performance of the model.

$$\text{Sensitivity(SN)} = TP/(TP + FN) \qquad (3)$$

$$\text{Specificity(SP)} = \text{TN}/(\text{TN} + \text{FP}) \qquad (4)$$

$$\text{Precision (PPV)} = \text{TP}/(\text{TP} + \text{FP}) \qquad (5)$$

$$NPV = TN/(TN + FN) \qquad (6)$$

$$\text{Accuracy(AC)} = \frac{\text{TP} + \text{TN}}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \qquad (7)$$

$$F - Score(FS) = \frac{TP}{TP + 1/2(FP + FN)} \qquad (8)$$

where, *TP*=True Positive, *TN*= True Negative, *FP*=False Positive, and *FN*= False Negative.

TABLE I. KERNEL FUNCTIONS

Linear SVM	$k(x,y) = x'y$,
Polynomial SVM	$k(x,y) = (1 + x'y)^d$
RBF SVM	$k(x,y) = \exp(-\|x - y\|^2)$
where x and y represent data and d is the order of polynomial function	

III. RESULTS AND DISCUSSIONS

The total number of onset channels considered in this study is 559. Among this, 83 channels are associated with TCSZ and the remaining channels are of electrographic seizures namely FNSZ and GNSZ. The representatives of sample signal from each classes are shown in Fig 1. The distributions of wavelet entropy extracted from seven decomposition levels are presented in Fig 2. The median of EGSZ are found to be lower for all decomposition levels except D_7 and A_7 (D_1 (TCSZ-19.89, EGSZ- 19.25), D_2 (TCSZ-20.22, EGSZ-19.11), D_3 (TCSZ-20.42, EGSZ-18.57), D_4 (TCSZ-20, EGSZ-19.15), D_5 (TCSZ-20.44, EGSZ-19.22), D_6 (TCSZ- 20.42, EGSZ-19.06), D_7 (TCSZ 19.23, EGSZ- 20.45), A_7 (TCSZ-19.07, EGSZ- 20.42)). The high value of entropy indicates that the signals are more random with higher complexity. Therefore, these results suggest the EEG signals associated with the TCSZ are less random, more regular with lower complexity than electrographic seizures.

The Wilcoxon test is used to test the statistical relationship of extracted features and shows that all features expect D_2 are significantly different ($p < 0.05$). From the Fig 2, it is clearly understood that most of the features are found overlapping and a linear separation is not possible with any of the single feature. Therefore, combination of all features were considered for the classification task. The features from the majority class (EGSZ) and minority class are labeled as 1 and -1 respectively. In order to address the issues of data imbalancing, a widely adopted and perhaps the most straightforward method such as resampling was applied [19]. This involves in removing samples from the majority class (under-sampling) and/or adding more examples from the minority class (over-sampling). In this study, we adopted underdamping method that randomly selects 120 samples from the majority class in each classification task. This was repeated for 100 times and mean value of model performances are presented in the Table II. 10-fold cross-validation was

also applied to validate the test samples. In this cross-validation, the samples are randomly partitioned into 10 equal set of subsamples. Of the 10 subsample set, single subsample set is retained as validation data as test model and remaining 9 sub sample sets are used as training. We repeat this procedure 10 times reserving a different tenth for testing.

Nonlinear classifiers such as linear SVM, polynomial SVM and RBF SVM have been employed to distinguish these seizure types. 10-fold cross-validation method was selected for the evaluation of test samples. It is found that linear SVM yielded poor performances, with very low specificity. The accuracy of linear SVM model is found to be 52.6%. It appears that this model performs similar to the random predictor and is not capable of detecting TCSZ.

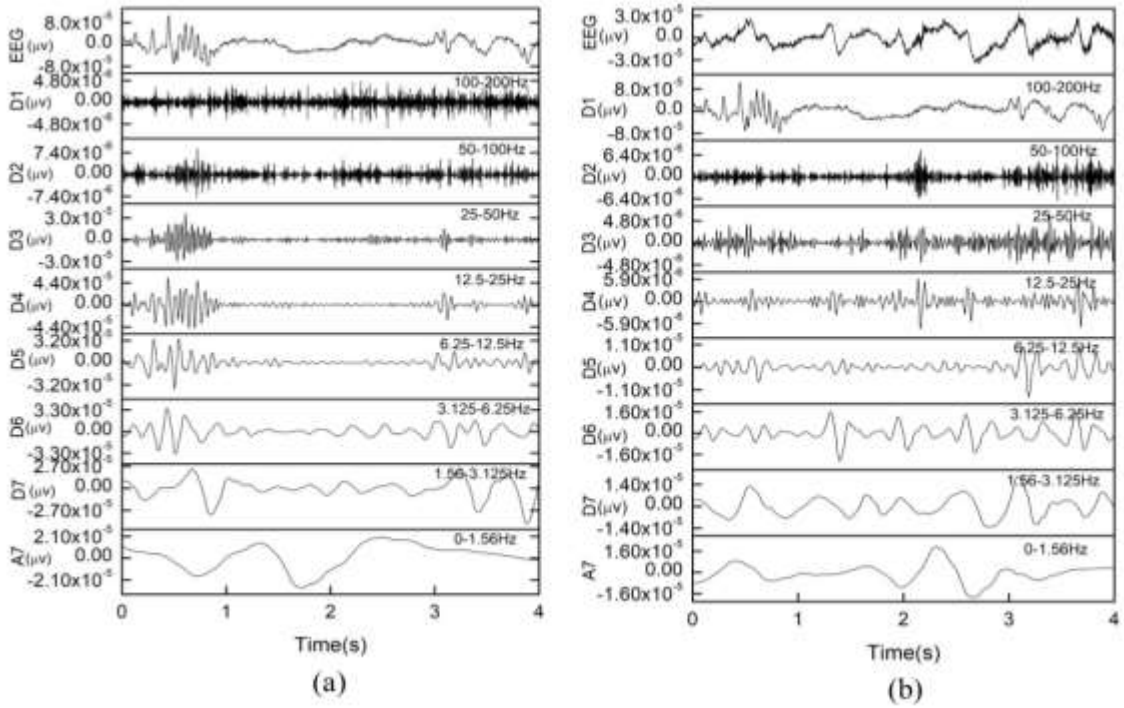

Figure 1 : Decomposition levels of representative signals (a) TCSZ, (b) EGSZ

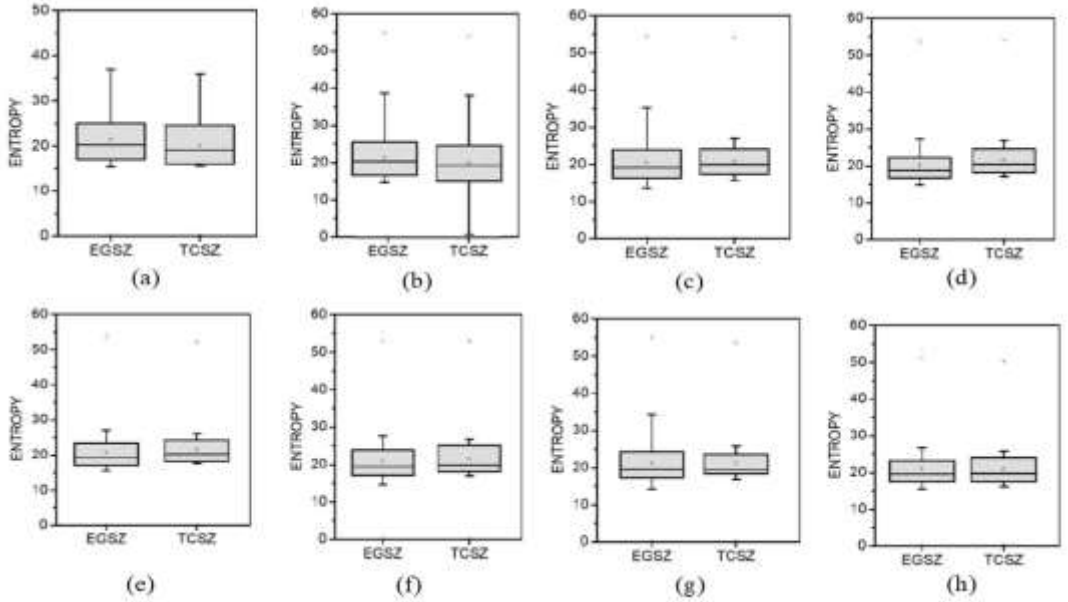

Figure 2: Distribution of wavelet energy in different sub-bands. a) 0-1.56Hz, b) 1.56-3.125Hz, c) 3.125-6.25Hz, d) 6.25-12.5Hz, e)12.5-25Hz, f) 25-50Hz, g) 50-100Hz, and h)100-200Hz.

The performance of polynomial SVM and RBF SVM models are found to be more reliable in detecting the seizure classes. The RBF SVM model yielded an accuracy of 94.4% and F score of 95.1%.The specificity and precision are also found to be 99.5% and 99.6% respectively. The polynomial SVM correctly identified 112 EGSZ channels (total- 120) and 81 TCSZ channels (total- 83). This was repeated for a 100 iterations to obtain average performance. Thus, Polynomial model yielded accuracy of 95.5%, sensitivity of 92.7%, specificity of 99.3% and F scores of 95.9%. It implies that these nonlinear models are capable of separating the two classes at higher feature space.

TABLE II. PERFROMANCE METRICS

Classifier	AC (%)	SN (%)	SP (%)	PPV (%)	NPV (%)	FS (%)
RBF SVM	94.4	91.0	99.5	99.6	87.9	95.1
Polynomial SVM	95.5	92.7	99.3	99.4	91.6	95.9
Linear SVM	52.6	55.3	42.6	73.1	18.1	62.6

In this work, wavelet decomposition is utilized to analyze the nonstationary variations of the scalp EEGs during clinical and non-clinical seizures. These results suggest that the wavelet entropy-based Polynomial SVM detects the TCSZ channels with a maximum accuracy of 95.5%. However, the model is developed using 10-fold cross-validation. In order to enhance the reliability of our proposed approach, the machine learning model based on leave-one-subject-out cross-validation will be developed in future.

IV. CONCLUSION

A detection of epileptic seizures provides an alternative viable option to improve the quality of life in patients with drug resistant epilepsy. In this study, wavelet entropy features are proposed to differentiate the TCSZ and EGSZ from the scalp EEG signals. The signals are decomposed to seven level wavelet decomposition and entropy is extracted from the resultant coefficients. These features are used to design SVM based detection models. The performance of the proposed model is analyzed using different kernel tricks. The results show that the entropy is lower for the signals during TCSZ. It is found that most of the features exhibits significant difference between these two seizure types. The developed polynomial SVM model achieved a maximum accuracy of 95.5%, sensitivity of 92.7%, NPV of 91.6% and F-score of 95.9%, whereas RBF SVM yielded an optimum specificity of 99.5%, and precision of 99.6% in detecting these complex clinical manifestations. Therefore, the proposed approach could be used to detect the life threatening seizures in and outside of clinical settings.

ACKNOWLEDGMENT

We thank Temple University Hospital (TUH) for providing TUH seizure dataset.

REFERENCES

[1] R. S. Fisher,W. Emde Boas , W. Blume, C. Elger, P. Genton, P. Lee, and J. Engel, "Epileptic Seizures and Epilepsy : Definitions Proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE)," *Epilepsia*, vol. 46(4), pp. 470-472, 2005.

[2] I. Kiral-Kornek, S. Roy, E. Nurse, B. Mashford, P. Karoly, T. Carroll, D. Payne, S. Saha, S. Baldassano, T. O'Brien, and D. Grayden, "Epileptic seizure prediction using big data and deep learning: toward a mobile system," *EBioMedicine*. vol 27, pp.103-111, 2018.

[3] L. Hussain, S. Saeed, I.A. Awan, and A. Idris, "Multiscaled complexity analysis of EEG epileptic seizure using entropy-based techniques," *Archives of Neuroscience*, vol. 5(1), 2018.

[4] P.A. Karthick, H. Tanaka, H.M. Khoo, and J. Gotman, "Prediction of secondary generalization from a focal onset seizure in intracerebral EEG," *Clin Neurophysiol*, vol. 129(5), pp. 1030-1040, 2018.

[5] D. Nabil, R. Benali, and F.B. Reguig, "Epileptic seizure recognition using EEG wavelet decomposition based on nonlinear and statistical features with support vector machine classification," *Biomed Tech (Berl)*, vol. 65(2), pp.133-148, 2020.

[6] A. Subasi, "Application of adaptive neuro-fuzzy inference system for epileptic seizure detection using wavelet feature extraction," *Comput. Biol. Med*., vol. 37(2), pp. 227-244, 2007.

[7] L. M. Patnaik, and O.K. Manyam, "Epileptic EEG detection using neural networks and post-classification," *Comput Meth Prog Bio.*, vol. 91(2), pp. 100-109, 2008.

[8] Y. Kumar, M.L. Dewal, and R.S. Anand, "Epileptic seizures detection in EEG using DWT-based ApEn and artificial neural network". *Signal, Image Video P.*, vol. 8(7), pp.1323-1334, 2014.

[9] A. Hamad, E.H. Houssein, A.E. Hassanien, and A.A. Fahmy, "Feature extraction of epilepsy EEG using discrete wavelet transform." *12th international computer engineering conference (ICENCO)*,2016, pp. 190-195.

[10] P. A. Karthick, H. Tanaka, H.M. Khoo, and J. Gotman, "Could we have missed out the seizure onset: A study based on intracranial EEG". *Clin Neurophysiol*. vol. 131(1), pp.114-26. 2020.

[11] U. Asif, S. Roy, J. Tang, and S. Harrer, "SeizureNet: a deep convolutional neural network for accurate seizure type classification and seizure detection". *arXiv preprint arXiv:1903.03232.* 2019.

[12] V. Shah, E. Von Weltin, S. Lopez, J.R. McHugh, L. Veloso, M. Golmohammadi, I. Obeid, and J.V. Picone, "The Temple University Hospital Seizure Detection Corpus," *Front.Neuroinfor.*, vol. 12, p.83. 2018.

[13] D. Ochal, S. Rahman, S. Ferrell, T. Elseify, I. Obeid, and J. Picone, "The Temple University Hospital EEG Corpus : Annotation Guidelines," *Institute for Signal and Information Processing Report,* vol. 1(1). 2020.

[14] V. Shah, M. Golmohammadi, S. Ziyabari, E. Von Weltin, I. Obeid, and J Picone, "Optimizing channel selection for seizure detection," *IEEE Signal Processing in Medicine and Biology Symposium (SPMB),* 2017, pp. 1-5.

[15] H.U. Amin, A.S. Malik, R.F. Ahmad, N. Badruddin, N. Kamel, M. Hussain, W.T. Chooi, "Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques," *Australas Phys **Eng** Sci Med,* vol. 38(1), pp.139-149, 2015.

[16] P.A. Karthick, K.R. Wan, R. Yuvaraj, A.A. See, N.K. King, and J. Dauwels, "Detection of subthalamic nucleus using time-frequency features of microelectrode recordings and random forest classifier". *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),* 2019, pp. 4164-4167.

[17] V. N. Vapnik, "An overview of statistical learning theory," *IEEE Trans Neural Netw*, vol. 10, pp. 988- 99, 1999.

[18] N. Christianini, and J. C. Shawe-Taylor. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge, UK: Cambridge University Press, 2000.

[19] J. Mathew, C.K. Pang, M. Luo, and W.H. Leong, W.H., "Classification of imbalanced data by oversampling in kernel space of support vector machines," *IEEE Trans Neural Netw Learn Syst*., vol. 29(9), pp.4065-4076, 2017.