A Proposal to Automate Seizure Detection based on a Comparative Study of EEG Signal Analysis

Hrishikesh Telang
Student
Department of Computer
Engineering,
D. J. Sanghvi College of
Engineering,
Mumbai

Shreya More
Student
Department of Computer
Engineering,
D. J. Sanghvi College of
Engineering,
Mumbai

Yatri Modi
Student
Department of Computer
Engineering,
D. J. Sanghvi College of
Engineering,
Mumbai

Ruhina Karani
Assistant Professor
Department of Computer Engineering,
D. J. Sanghvi College of Engineering, Mumbai

ABSTRACT
Epilepsy is a chronic neurological disorder which is characterized by recurrent and sudden seizures. People with epilepsy suffer from multiple types of seizures and Electroencephalography is an important clinical tool for diagnosing, monitoring and managing neurological disorders related to epilepsy. EEG signals are most often used to diagnose epilepsy, as seizures cause anomalies in EEG readings. In today's world where adult life expectancy is rising and humans are living longer than ever before, the healthcare system generates vast amounts of data, including EEG signals. This paper examines the prospects and challenges faced in utilizing this data in order to optimize seizure detection in order to improve the patients' quality of life. This paper also explores how Machine Learning can be applied to extract features and analyze the EEG signals and propose methods to achieve high classification accuracy.

Keywords
EEG signal analysis, Epileptic Seizure Detection, Machine Learning, Feature Extraction, Wavelet Transform, Signal Preprocessing, Signal Classification, Bidirectional Neural Networks, Auto Regressive model, Approximate Entropy, Wavelet Packet Decomposition, Extreme Learning Machines.

1. INTRODUCTION
Epilepsy is a group of neurological diseases, the defining feature of the condition being unprovoked seizures. If left undetected, epileptic seizures can severely impact the lives of patients, causing fatal accidents and unexpected death. While the condition can improve with medication, for about 30% of patients, they are non-curative and patients can have seizures even after surgery[1].

Around 50 million people worldwide suffer from Epilepsy and the condition can affect a person at any age. While it is possible to visually inspect an EEG report and interpret it, a level of subjectivity is introduced to the process of seizure detection and it tends to be time-consuming and tedious as the EEG data lasts for several hours. Thus it would be preferable to automate the detection process and reduce the human effort required for it.

EEG signals are recorded by metal electrodes placed on the scalp of the patient. During seizures, the scalp EEG of patients with epilepsy is characterized by periodic EEG waveforms of high amplitude, reflecting abnormal discharge of a large group of neurons. Between seizures, epileptiform transient waveforms, which include spikes and sharp waves, are typically observed on the scalp EEG of such patients[10]. The patient's EEG signal contains enough information to be able to predict an epileptic seizure and it is possible to collect such features together and feed them to a classifier. However, several challenges can arise during the development of such a system. If we were to collect EEG data in real time, due to the sheer volume of the data being generated, storage and computation could be major problem. Secondly, not all seizures are epileptic and EEG signals are stationary signals. As a result, patterns that correspond to seizures are not uniform across patients. Thus, manually-extracted features might not scale well to new patterns of seizure activity, and supervised feature extraction may not be sufficient for learning algorithms.[2]

In this study we have examined recent developments in EEG signal analysis, seeking to provide an alternative technique that may prove to be more efficient than the individual methods, either in reducing the computational time, processing power or in improving the overall accuracy. It remains to be tested and developed as a synthesis of the proposed methodologies in this paper. However, we are confident that the results will not be futile.

The rest of the paper beyond this point is organized as follows:

Previous works in this field have been explored in the Literature Review section, in all the subsections of analysing EEG signals for epileptic seizure activity. The section, Proposed Methodology, that follows enumerates current research in each of the subdomains - Preprocessing, Feature extraction and Classification - and proposes to combine efficient methodologies from each to build a stronger and more reliable framework for automated seizure detection in EEG signals, an aspect further elaborated in the Conclusion section that follows. Potential for further research has been discussed at the end of the paper.

2. LITERATURE REVIEW

The process of EEG signal analysis involves three major stages - signal preprocessing, feature extraction and classification, prevailing findings on which have been described below.
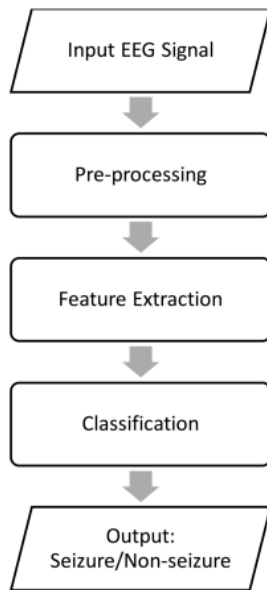

Fig 1: Steps to be followed for EEG signal classification as seizure/ non-seizure.

The initial step of multichannel signal classification, signal preprocessing, is the de-noising and segmentation of the raw EEG signal, followed by different methods of change-points detection. This stage is fundamental for further data processing both for one-dimensional and multidimensional signals. To ensure fast processing and reliable feature extraction, certain pre-processing steps and fast computational tools are necessary. Traditionally, Digital filters are used to remove noise from the observed signal and to reduce its undesirable frequency components[3].

The next stage in EEG signal analysis is Feature Extraction that classically involves linear analysis, nonlinear dynamic analysis, time domain analysis, frequency-domain analysis, time-frequency analysis, phase synchronization, etc. Further modifications to these have allowed fourier transform to evolve into Fast Fourier Transform (FFT), Short Time Fourier Transform (STFT) and Discrete Wavelet Transform (DWT). Spectral leakages, poor performance and limited time frequency resolution gave rise to the theory of Wavelet Transforms (WT). WT offer a method for multiscale representation and analysis of signals, providing relevant information in detecting epileptogenic spikes and non-spikes. They differ from the traditional Fourier techniques in localising information to time-frequency planes and are capable of trading one variant of signal for another, rendering them particularly useful in analysing non-stationary EEG signals. [4] A WT is defined as follows:

$$WT(\tau, a) = (1/\sqrt{a}) \int_{-\infty}^{\infty} x(t) \, h^t(\frac{t-a}{\tau}) \, dt$$

The final stage of EEG signal analysis involves classification, early literature on which cites usage of decision functions, trainable classification networks, distance functions, syntactic methods and their hybrids for signal classification. Features transformed into dimensionless standard scores between

classes within each individual's data, helped avoid the obscuring effect of latency due to differences between individuals' BEP signals.

Intracranial EEG signals have previously been classified as epileptic and nonepileptic primarily using supervised machine learning techniques like linear discriminant analysis (LDA) and support vector machines (SVM). A study found that SVM classifier gives higher classification accuracy than LDA classifier for both time- and time-frequency domain features [5]. This indicates that analysing the signals in the non-linear domain is more relevant than analysing in the linear domain.

Single channel EEG signals were analysed for seizure detection in [6] and [7]. An accuracy of over 90% was achieved in the former after experimenting with multiple classification primitives and SVM ensemble parameters. An ensemble of classifiers is formed when the individual decisions from each classifier in a collection of classifiers are combined in some way and used to classify test examples. Ensembles are known to improve performance and the best result was obtained by the assignment of segments to individual SVM classifiers and combining their results using the method of majority voting.

3. PROPOSED METHODOLOGY

3.1. Signal Preprocessing

EEG activity is quite small, measured in microvolts (mV). Due to such a small range, EEG signals are susceptible to internal and external disturbances. These include muscle noise, ocular artefacts, power line noise, baseline wander and motion artefacts. In order to eliminate such disturbances, a set of 0.5 Hz high-pass and 50 Hz notch Finite Impulse Response (FIR) filters to eliminate baseline drift and power line interference were used in [8]. In the study, they let each of the 14 Formula channels of EEG from their database pass through two separate 64-tap FIRs having 16-bit coefficients with quantization selected to maximize dynamic range. A window-based method was used to design both filters.

While the above methods work on an EEG dataset, it would be much more efficient to try and remove as many disturbances as possible in real time, i.e., when the data is being collected. There can also be a significant number of external factors which add a lot of noise and interference to the recorded EEG signals including eyeball movement and repeated blinks of the eye. It would be much more efficient to remove some of these interferences in real time, saving a lot of effort in the feature extraction stage of detecting a seizure. In order to remove these patient generated interferences in real-time, [9] created a broad 5 step process. This includes:

1. Passing the raw EEG samples through the digital bandpass filter to remove slow baseline drift (noise).

2. Determine the blink threshold (Vt) for specific subject in brief training session.

3. Compare the absolute sample value with Vt.

4. If the value is exceeded then remove N samples from the vicinity of zero crossing (N/2 on either side of threshold crossing).

In order to overcome the disadvantages produced by noisy signals, one of the earliest studies on EEG signals suggests use of Wavelet Transforms in the pre-processing stage[10]. Based on the experimental results they obtained, a potential plan of action could include the use of Artificial Neural Networks in conjunction with wavelet transforms to detect

unwanted spikes in signals. The two ways to accomplish this are using a set of parameters mined from the signal, or using the raw signal itself. For the former, success would depend heavily on proper selection of parameters and it is difficult to predict in advance which parameters to choose. However using a sliding window of small size to give the raw EEG signal as input to the ANN was found to be a better approach in [10]. To increase the size of input data, we can increase the size of the sliding time window, however training the network on this data then becomes difficult and a computationally intensive process.

Consequently, using Wavelet Transforms for preprocessing is extremely efficient, as it allows for a larger input window without increasing the training time. The successful results obtained with WT preprocessing can be attributed to the variable frequency representation and data compression properties of WTs. Such a preprocessing can also be used in other signal processing applications of ANNs to reduce data input size and/or improve performance.

In the study, the classification accuracy of each processed wavelet was tested by first extracting a fixed amount of coefficients and giving these as input to the ANNs. Different 'scales' were used, where each scale had a specific number of coefficients (8 or 24) which were obtained as a result of applying wavelet transforms using Debauchie's wavelets. Classification accuracy was tested not only for individual scales, but for a combination of different scales and the relationship between the coefficients and scales was determined.[10] The most important factor in performance was found to be the proper selection of the scale or the combination of scales.

The methodology proposed in the paper combines the aforementioned concepts to create a pre-processing technique that uses Wavelet Transforms to reduce input size without decreasing the amount of information mined from the signal. The use of the WT should decrease the input size of the data without compromising on its performance.

- Based on the length of the input signal, determine the number of segments of the input signals

- Choose an appropriate sliding window width.

- Apply Wavelet Transform to the chosen input data (Daub 4 and Daub 20 were used in [10]).

- Feed these segments to the Feature Extraction stage through the sliding window.

Wavelet Transforms, however, are computationally expensive and tough to implement in hardware. A real time cascaded moving average filter, as proposed in [11], is a computationally simpler denoising technique for EEG signals. It removes baseline wander and other high frequency artifacts. A real time implementation of the filter on a FPGA shows that less than 6% of the total resources available are utilized in it. This indicates the simplicity of the design of the filter. The computational efficiency of the algorithm used has been proved by testing the Signal to Noise Ratio(SNR), Peak Signal to Noise Ratio (PSNR) and Normalized Mean Square Error (NMSE).

In conclusion, using Wavelet Transforms along with a sliding window to give input to the Artificial Neural Network is a suitable pre-processing technique, while reducing its complexity can be a topic for future research.

3.2. Feature Extraction

A study introducing the use of WT as preprocessor for neural network detection of EEG spikes [4] involved a 16 channel recording of 4 patients sampled at a rate of 100 Hz. EEG experts labelled the recordings as spikes and non-spikes. WT of selected EEG data segments of 5.12 seconds were analysed using Daubechies series type 20 mother wavelet, their coefficients representing 320 ms and 640 ms of the EEG data. According to the paper, due to a variable window size, WT exhibited excellent compression and detection capabilities despite the abstruse nature of artifacts like EMG. The result was obtained as an eight scale multi-resolution display, from which a total of 16 coefficients were taken from levels 2 and 3. It was further processed through a feedforward perceptron network with a backpropagation learning algorithm with a different number of hidden neurons. Amongst a total of 1189 training segments and 1489 testing segments, the model achieved an overall performance between 88.65% and 90.45%, wherein models containing 7 and 8 hidden neurons provided the best results, which was a very promising start.

One study [12] ranked individual features based on their importance in seizure prediction, and proposed that as a first step before customising an EEG based seizure warning system. In the study, ten features were extracted from each level of wavelet decomposition. Using these features, the data was classified by a multilayer perceptron (MLP), wherein the feature set was distributed into a certain number of classes and probabilities of in-class scatter, between-class scatter and separability matrices were calculated for each of them. Based on the effect of every feature vector on separability, the three most effective features were found to correspond to the lower frequency components - number of zero crossings, standard deviation and Shannon's entropy. Comparing features based on their effect on separability of classes, summed over all decomposition levels, the number of extremes, number of zero crossings and energy of the signal were found to be the most efficient, with an accuracy of 94%, contrasting the lower bound range at 89%.

More recently, as a specialised technique for WT, a multivariate empirical wavelet transform was proposed in [13] that builds signal adaptive wavelet based filters. The subband signals, called modes, have a tightly packed frequency support centered around a specific frequency. A boundary search method is applied to the mean spectrum magnitude to generate an adaptive wavelet filter bank. Hilbert transform is applied to extract instantaneous amplitudes and frequencies from filtered subband signals. A ten-fold cross validation method is used to evaluate the performance of their feature extraction method, calculated based on three parameters - sensitivity, specificity and accuracy. Having trained and tested their model on the CHB-MIT scalp EEG database's 23 patients excluding the twelfth patient, it achieved a maximum average sensitivity of 97.91% and a maximum average specificity of 99.57% using five selected EEG channels. Having tested the EEG records on six types of classifiers, the study found Random Forest classifier to be superior than the other five classifiers.

A study [3] proposed two combination strategies based on three common techniques for feature extraction - Auto Regressive (AR) model, Approximate Entropy (ApEn) and Wavelet Packet Decomposition (WPD). It was suggested that AR be combined with either ApEn or WPD for a highly efficient mechanism for feature extraction. In an AR model, observations from previous timesteps are used as inputs to a

regression equation to predict the result at the following timestep, generally expressed as,

$$X_t = \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t$$

Where X_t is the time series to be predicted, X_{t-i} are the past observations, φ_i are their respective coefficients or parameters of the model and ε_t is the white noise for an AR of order p, expressed as AR(p). The accuracy of estimates improve with an increase in the order of an AR, which in the paper has been varied from 1 to 12 with differing results.

Approximate Entropy (ApEn) is a statistic quantifying the amount of regularity and complexity of fluctuations in a time series data. For an EEG signal of length N, in a real m-dimensional space R^m, the time series sequence forms another sequence of (N-m+1) vectors of length m. So from a time series data *u(1), u(2), u(3), ..., u(N),* representing N equally spaced raw data values, a time sequence *x(1), x(2), ..., x(N-m+1)* is generated, defined by: *x(i) = [u(i), u(i+1), ... , u(i+m-1)]*. For each *i, 1≤i≤(N-m+1)* the sequence *x(1), x(2), ..., x(N-m+1)* is used to construct $C_i^m = (number\ of\ x(j)\ such\ that\ d[x(i), x(j)] \leq r)/(N - m + 1)$

Where $d[x, x^*] = max_a|u(a) - u^*(a)|$ and represents the distance between vectors x(i) and x(j) given by the maximum distance between their scalar components denoted by *u(a)*. Another variable $\phi^m(r)$ is defined as follows

$$\varphi^m(r) = (N - m + 1)^{-1} \sum_{i=1}^{N-m+1} \log (C_i^m(r))$$

Approximate Entropy is then defined in terms of $\phi^m(r)$ and $\phi^{m-1}(r)$ as
$ApEn = \phi^m(r) - \phi^{m-1}(r)$
ApEn quantifies the likelihood of not repeatedly encountering similar patterns of observation in a time series data, resulting in a small ApEn value for a time series containing many repetitive patterns and a larger ApEn value for a less predictable one. Besides being less computationally demanding, Approximate Entropy is also less affected by noise.

Wavelet Packet Decomposition (WPD) helps analyse non-stationary EEG signals by passing discrete time signals through more filters than the Discrete Wavelet Transform (DWT). For n levels of decomposition, the WPD produces 2^n sets of nodes in contrast to (3n+1) sets of nodes for a DWT. A WPD recursively decomposes the signal into an approximation coefficient and a detail coefficient, both of which it further decomposes to construct a complete wavelet packet tree.

The paper [14] combines the three feature extraction methods described above and proposes two combinational methods to extract features from an EEG signal. The first one combines AR and ApEn in which they calculate 6 groups of AR coefficients for 6 channels in the same segment, generating (6*p) features from an AR of order p. ApEn is then used to calculate entropies of each channel and extract 6 features which are then combined with the AR features to generate a (6*p + 6) length feature vector to be fed to the classifier. The second method combines AR and WPD by first decomposing segments of EEG signals into subbands using WPD, from which 2^n sets of wavelet coefficients are calculated. For each set of coefficients its p order AR is calculated for 6 channels in the same segment, resulting in a $(6 * p * 2^n)$ dimensional feature vector. On testing both the feature extraction techniques on a Support Vector Machine, it

was found that on varying p from 1 to 5 both the methods are similar in classification accuracy. However, AR+WPD has a superior performance, with a peak at 100% in one of the subjects, when p ranges from 6 to 12, although AR+ApEn performs better than AR on average.

Considering the variety of effective approaches discussed above, we propose a hybrid technique to use an AR model along with WPD due to its superior performance at higher values of p, although it may be computationally expensive. Generating accurate results is of utmost importance in a medical diagnosis and it mustn't be compromised at any cost. Further, the WPD can be augmented by using empirical transforms as suggested in [13]. Combining the AR+WPD model with the feature ranking mechanism proposed in [12] would lead to a more relevant set of features for seizure prediction. Opting for fewer features would generate more precise measures of seizure prediction, helping enhance overall performance of the proposed model.

3.3. Classification
We propose two methods for classification. The first one is based on recent research conducted on Extreme Learning Machines while the other one uses a bidirectional RNN. ELM are Single Hidden layer Feedforward Neural Networks (SLFNs). The hidden neurons serve as a bridge between the input and the output neurons for a direct transfer of the information. The steps are as follows:[15] Randomly assign the bias b and input weight w.
1. Calculate the output matrix at the hidden layer
   H=g(w.x+b)
2. Calculate the output weight
   B=H'T

Where H' is the Moore-Penrose generalized matrix of H.[16]

Feedforward neural networks have comparatively slower learning speeds and this can prove to be a serious bottleneck in their applications. Extreme Learning Machine(ELM) is a recently proposed methodology for classification. It randomly chooses the input weights and analytically determines the output weights for Single Hidden Layer Feedforward Neural Networks(SLFN). It was found to have better generalization and testing rate than conventional SVMs. It was also able to train SLFNs faster than classical learning algorithms due to a high learning rate [16]. ELMs also increase the speed of learning and were shown to achieve a classification accuracy of 94.85%.
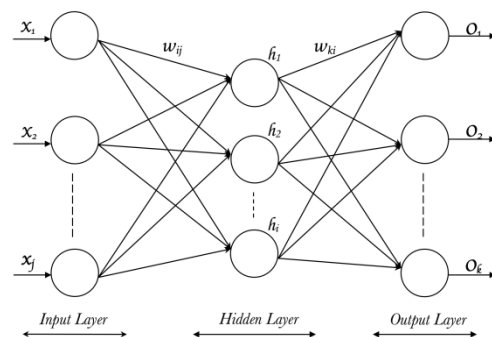

Fig 2: Extreme Learning Machine

EEG signals of epilepsy patients contain spikes, whose discriminatory features can be detected using convolutional neural networks. Rectifiers (ReLU) have been used over sigmoid functions as activation functions for various reasons. It reduces sparsity in hidden units, and most importantly,

don't face the gradient vanishing problem as faced by sigmoid and tanh function. [17] Activation functions also face disadvantages like dead gradient (large gradients deactivating neurons). This problem is more pronounced in activation functions like the sigmoid function and ReLU. In [18] they use leaky ReLU, where leaky denotes the additional slope in the function. The study used a neural network with multiple convolutional layers with different filter sizes. The first 3 layers were convolutional layers having different filter sizes, while the fifth layer is a binary logistic layer which makes the final classification decisions. The fourth layer, the fully connected layer, accepts inputs from the 3 previous layers which are merged. The activation function is applied after every layer. To evaluate performance, the receiver operating characteristic (ROC) curve was calculated and the area under the ROC curve (AUC) was calculated as benchmark criterion. It managed to outperform all other classifiers.

Recurrent Neural Networks (RNNs) can be used to automate the classification process. RNNs containing feedback connections along with feed forward connections allow the neural network to store past knowledge while processing data. [19] proposes a bidirectional RNN (BRNN) which will also factor in future context in addition to past knowledge and improve on the RNN architecture.

The proposed BRNN maintains two recurrent layers, one in the backward direction and one in the forward direction. The outputs of the two layers are then combined in a final output layer. The common output layer combines outputs from each directional RNN to produce the final output for the Formula element in the data sequence. This bidirectional setup ensures constant amount of contextual influence in processing elements of the input data sequence.
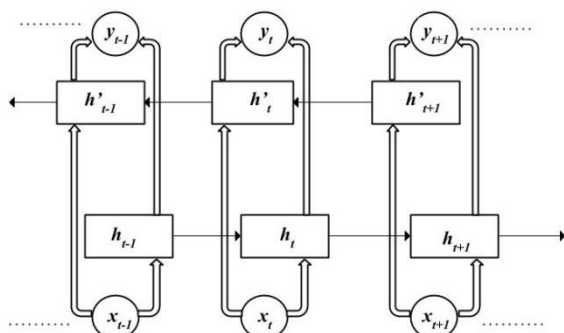

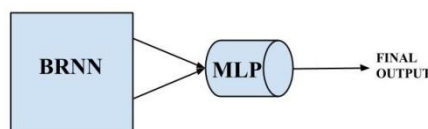**Fig 3: Bidirectional Recurrent Neural Network**

Fig 4: Output of BRNN is fed to the MLP

Its architecture is built upon that of a cellular Artificial Neural Network and preserves most of its properties. The cells are able to communicate with their neighbours and each input element of the input pattern is processed simultaneously. The outputs of these cells is combined and fed to a multilayered perceptron (MLP) to obtain the final output. They use a Jacobian free Unscented Kalman Filter (UKF) for training, as standard backpropagation would complicate the training process due to the complex architecture. The model successfully detected seizures with 100% accuracy and a

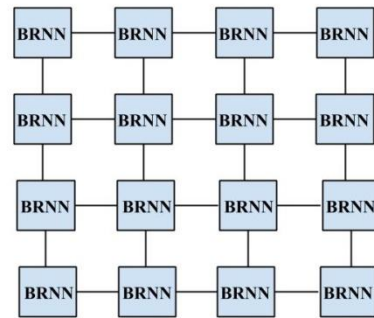mean 7 second delay, significantly improving upon other findings.

Fig 5: Cellular Neural Network where each element is a Bidirectional Recurrent Neural Network (BRNN)

4. CONCLUSION

This paper proposes a methodology for epileptic seizure detection through EEG signal analysis. The preliminary stage involves reducing the input size of the signal dataset by using Wavelet Transforms along with a sliding window, without compromising on the quality of data extracted from it in the preprocessing stage. In the next stage the use of Auto Regressive model in conjunction with Approximation Entropy has been proposed for feature extraction. The features are to be ranked and the ones most suited to detecting seizures in EEG signals are to be selected, leading to a reduced overhead in terms of generating feature vectors, which would then be more specialised to the task of seizure detection. The final stage of the proposed methodology advocates for use of BRNNs and ELMs along with the Leaky ReLU activation function for providing high classification accuracy for the input. The paper has been written with the hope that this methodology helps improve seizure detection and improves the quality of life of epilepsy patients.

4.1. Potential for further research

Machine Learning and Signal Processing play a huge role in EEG signal analysis, but there is still scope for improvement. Extreme Learning Machines[7] and Support Vector Machine ensembles[6], which were used with single channel EEG signals can be applied to multi-channel EEG signals as a part of further research. Whether or not using the leaky ReLU activation function along with the BRNN can lead to further improvement in training stage is another area for further research. Reducing the computational cost of performing Wavelet Transforms is another challenge, which if resolved, could lead to further performance improvement.

5. REFERENCES

[1] Mohammad-Parsa Hosseini, Mohammad R. Nazem-Zadeh, Dario Pompili, Hamid Soltanian-Zadeh, "Statistical validation of automatic methods for hippocampus segmentation in MR images of epileptic patients", *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014.

[2] Mohammad-Parsa Hosseini, Hamid Soltanian-Zadeh, Kost Elisevich, Dario Pompili, "Cloud-based deep learning of big EEG data for epileptic seizure prediction", *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2016.

[3] Ales Prochazka, Jaromir Kukal, Oldrich Vysata, "Wavelet transform use for feature extraction and EEG

signal segments classification", *3rd International Symposium on Communications, Control and Signal Processing, 2008. ISCCSP* 2008.

[4] Tulga Kalayci, Ozcan Ozdamar, Nurgiin Erdol, "The use of wavelet transform as a preprocessor for the neural network detection of EEG spikes", *Southeastcon '94. Creative Technology Transfer - A Global Affair., Proceedings of the 1994 IEEE.*

[5] Md. Khayrul Bashar, Faruque Reza, Zamzuri Idris, Hiroaki Yoshida, " Epileptic seizure classification from intracranial EEG signals: A comparative study EEG - based seizure classification", *IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES),* 2016.

[6] A. Maragakis and A. Rosengarten, "Seizure Detection with Single-Channel On-Scalp EEG using SVM-Ensembles," International Mathematics Research Notices, Vol. 2015, Article ID rnn000, 8 pages. doi:10.1093/imrn/rnn000, 2015.

[7] Sabrina Ammar, Mohamed Senouci,"Seizure detection with single-channel EEG using Extreme Learning Machine", *17th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)*, 2016.

[8] Wijesinghe, L.P, Wickramasuriya, D.S, Pasqual, A.A, "A generalized preprocessing and feature extraction platform for scalp EEG signals on FPGA", *IEEE Conference on Biomedical Engineering and Sciences (IECBES)*, 2014.

[9] Sanjay Gupta, Harvinder Singh, "Preprocessing EEG signals for direct human-system interface", *IEEE International Joint Symposia on Intelligence and Systems,* 1996.

[10] Tulga Kalayci, Ozcan Ozdamar, "Wavelet preprocessing for automated neural network detection of EEG spikes", IEEE Engineering in Medicine and Biology Magazine (Volume: 14, Issue: 2, Mar/Apr 1995).

[11] R R Sreekrishna, Saif Nalband, A. Amalin Prince, "Real time cascaded moving average filter for detrending of electroencephalogram signals", *International Conference on Communication and Signal Processing (ICCSP)*, 2016.

[12] Ataee, P., Avanaki, A.N., Shariatpanahi, H.F., Khoee, S.M, " Ranking features of wavelet-decomposed EEG based on significance in epileptic seizure prediction", *14th European Signal Processing Conference, IEEE*, 2006.

[13] Abhijit Bhattacharyya, Ram Bilas Pachori, "A Multivariate Approach for Patient-Specific EEG Seizure Detection Using Empirical Wavelet Transform", IEEE Transactions on Biomedical Engineering (Volume: 64, Issue: 9, Sept. 2017).

[14] Zhang, Y., Liu, B., Ji, X. et al., "Classification of EEG signals based on Autoregressive model and Wavelet Packet Decomposition", Neural Processing Letters,April 2017, Volume 45, Issue 2, pp 365–378.

[15] A. Garces, E. Correa, E. Laciar, D P. Hector, E V. Maximo, "An Automatic Sleep-Stage Classifier Using Electroencephalographic Signals", *International journal of Medical Science and technology*, pp. 13-21, 2008.

[16] Guang-Bin Huang, Qin-Yu Zhu, and Chee-Kheong Siew, "Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks", *Proceedings. 2004 IEEE International Joint Conference on Neural Networks*, 2004.

[17] [ONLINE] Andrew L. Maas, Awni Y. Hannun, Andrew Y. Ng, "Rectifier Nonlinearities Improve Neural Network Acoustic Models ", Computer Science Department, Stanford University, CA 94305 USA - http://ai.stanford.edu/~amaas/papers/relu_hybrid_icml2013_final.pdf .

[18] Alexander Rosenberg Johansen, Jing Jin, Tomasz Maszczyk, Justin Dauwels, Sydney S. Cash, M. Brandon Westover, "Epileptiform spike detection via convolutional neural networks", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016.

[19] L. Vidyaratne, A. Glandon, M. Alam, K. M. Iftekharuddin, "Deep recurrent neural network for seizure detection", *International Joint Conference on Neural Networks (IJCNN)*, 2016.