Detection of Epileptic Seizures using Unsupervised Learning Techniques for Feature Extraction

Alexandra-Maria Tăuțan, Mihai Dogariu and Bogdan Ionescu
University Politehnica of Bucharest, Bucharest, Romania
alexandra.tautan@stud.etti.upb.ro

***Abstract*— Automatic epileptic seizure prediction from EEG (electroencephalogram) data is a challenging problem. This is due to the complex nature of the signal itself and of the generated abnormalities. In this paper, we investigate several deep network architectures i.e. stacked autoencoders and convolutional networks, for unsupervised EEG feature extraction. The proposed EEG features are used to solve the prediction of epileptic seizures via Support Vector Machines. This approach has many benefits: (i) it allows to achieve a high accuracy using small size sample data, e.g. 1 second EEG data; (ii) features are determined in an unsupervised manner, without the need for manual selection. Experimental validation is carried out on real-world data, i.e. the CHB-MIT dataset. We achieve an overall accuracy, sensitivity and specificity of up to 92%, 95% and 90% respectively.**

I. Introduction

Epilepsy is a neurological disease characterized by recurrent, involuntary seizure activity which leads to either observable clinical or subclinical symptoms. These seizures are caused by groups of abnormal functioning neurons spread out or localized in different regions of the brain. Placing the diagnostic of epilepsy requires both the observation of clinical symptoms and a more in-depth study of the brain's activity during suspected epileptic episodes.

The electroencephalogram (EEG), i.e., the recording of the electrical activity of the brain, provides such a diagnostic tool. All EEG based diagnostic methods require expert manual post-processing for selection of suspected epileptic segments. The procedure is cumbersome and time consuming. A robust algorithm for the automatic detection of epileptic seizures could reduce drastically the time needed for selecting suspected seizure segments from recorded data.

In this paper, we address the automatic classification of seizure and non-seizure EEG data. The automatic classification of seizure and non-seizure EEG data proves to be a challenge even today due to the complex nature of the phenomena and the underlying physiological signal. Seizure activity can be localized to specific regions of the brain or even specific subgroups of neurons. This would translate differently on the recording electrodes present all over the scalp or mounted invasively in different regions of the brain. Inter-subject variability is also high, both in the healthy and the epileptic EEG.

*This work was partially supported by the Ministry of Innovation and Research, UEFISCDI, project SPIA-VA, agreement 2SOL/2017, grant PN-III-P2-2.1-SOL-2016-02-0002

The remainder of the paper is organized as follows. Section II overviews the current state of the art and positions our contribution. Section III introduces the proposed approach for automatic prediction of epileptic seizures. Section IV provides the experimental validation and discusses the results. Section VI concludes the paper.

II. Previous Work

Automatic seizure detection has been extensively researched since the '80s. Early works focused on threshold-based non-patient specific algorithms for seizure detection [1]. Due to the variability observed in seizure data, the focus switched to patient-specific models.

The complexity of the pre-processing techniques increased and these methods were combined with more advanced, supervised classification algorithms. The reference work from [2] introduces the CHB-MIT database and presents a patient specific algorithm based on an Support Vector Machine (SVM) radial basis kernel classification. The overall sensitivity was 96%, tested in a leave-one-out cross-validation. One of the disadvantages of such a method is the necessity of *manual feature selection*. Unsupervised learning techniques provide more possibilities for extracting relevant features.

Due to their success in image processing tasks, Neural Networks (NN) have been adopted also in the field of EEG signal processing. Converting EEG data into frequency domain maps before using it as an input for a probabilistic neural network is proposed by the authors in [3]. A conversion to the frequency domain using short-time Fourier transform is used by the authors in [4] as input to a deep convolutional neural network. This approach achieves 81.2% sensitivity on the CHB-MIT database. The authors in [5] use all EEG channels to create a mutual information matrix of seizure and non-seizure signals which are then fed into a deep network with five convolutional layers and three fully connected layers. The approach was tested in a 5-fold cross-validation scenario on the CHB-MIT database, with a resulting accuracy, sensitivity and specificity of 98.13%, 98.85% and 97.47%, respectively. Another relevant approach is the one proposed in [6]. A stacked autoencoder with two hidden layers and one output layer as a logistic regression is built. Using a leave-one-record-out cross-validation method, a sensitivity of 100% was achieved on two patients from the CHB-MIT database. Similarly, a sparse autoencoder is used by the authors in [7] to create an EEG dictionary for a context-learning model. The method has an error rate of 22.93%. The authors in [8] reported a stacked autoencoder deep neural network architecture with two sparse encoders. The results obtained included an accuracy, sensitivity and precision of 94%, 93% and 95% respectively.

In this work, we investigate the use of Stacked Autoencoders (SAE) and Convolutional Neural Networks (CNN) for unsupervised EEG feature extraction. We propose a network architecture composed of three convolutional and four fully connected layers which are fed to an SVM to achieve the final data classification. The proposed architecture is able to provide high accuracy prediction using samples of only one second length. Previous techniques that used frequency representations of the data would not have been possible with such a narrow time window. In addition, the proposed system does not require any pre-processing or conversion of the input data. Unsupervised learning techniques provide the great advantage of extracting adequate features from the data. Unlike classical methods where the experimenter manually selects the adequate features, the proposed methods learn and adapt from the data itself. This makes the approach more reliable and ready to use in real-world scenarios, where the data is not previously known.

III. Proposed approach

A. Autoencoders

An autoencoder is a type of unsupervised neural network that can be used for feature extraction. A sparse autoencoder has one hidden layer in between the input and the output. By forcing the size of the output to be the same as the input, the network is forced to extract features in the hidden layer that enable it to reconstruct the input at the output. A stacked autoencoder is created by placing several sparse autoencoders on top of each other [9].

After experimenting with several architectures, we investigate the use of a six layer stacked autoencoder. The first three layers encode the input according to the size of each layer. Unsupervised feature representations are extracted at each phase. Layer 3 represents the coded input. The next two layers represent the decoding of the coded input, while the last layer is the reconstructed input. All layers are fully connected. A graphical representation of the network model is presented in Fig. 1. A ReLU (Rectified Linear Unit) activation was used for all autoencoder layers, except for the last which used a sigmoid activation. ReLU activation improves the network performance [10], while the sigmoid activation allows the output layer to take both positive and negative values.

The features extracted after each layer of the encoder are fed into an SVM classifier for classification into seizure and non-seizure EEG segments.

B. Proposed Network Architecture

Based on the results obtained with the unsupervised feature extraction from three fully connected layers, we propose a new network architecture that also integrates convolutional layers. The additional convolutional layers added before the fully connected layers provide supplemental filtering to the EEG data. After the first three convolutional layers, the data is flattened. An additional layer of size 1 is added to the output of the network to reduce the dimensionality of the calculated features. Its reduced output is fed to an SVM classifier for seizure/non-seizure segment prediction. Fig. 2 shows the proposed network architecture.

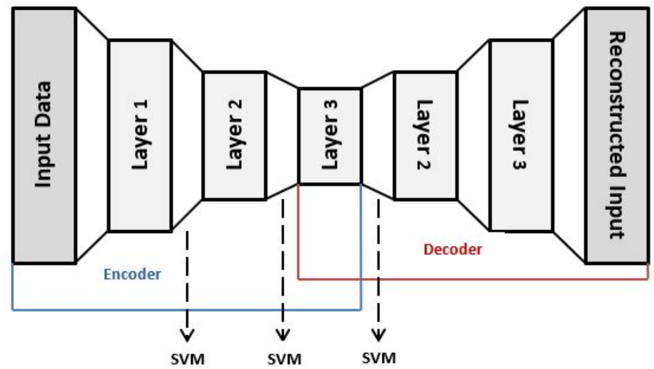

Fig. 1. Architecture of a Stacked Autoencoder for feature extraction combined with SVM for classification.

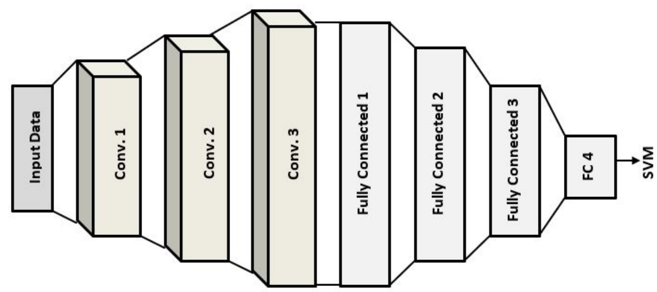

Fig. 2. Proposed network architecture combining convolutional (Conv.) and fully connected (FC) layers.

Unlike the work in [5] and [4], where authors use similar network architectures with 2D inputs, the first three layers of our proposed network are 1D convolutional layers and require no data reconfiguration. For the convolutional layers, the filter numbers increase per layer (32, 64, 128) with the kernel size remaining 2.

IV. Experimental validation

A. Test Dataset

For the evaluation, we use the CHB-MIT dataset [2], [11]. Given its open source availability, the dataset is one of the most used in the evaluation of algorithms for automatic seizure detection. The data was collected at the Boston Children's Hospital and contains scalp EEG data from 24 patients. The EEG signals were collected at a sampling rate of 256 Hz with a 16-bit resolution. In most cases, 23 EEG channels were recorded based on the 10-20 International Electrode Positioning System. The data was provided with annotations of seizure events.

TABLE I

VARIATION IN THE LAYER SIZE OF THE DEEP STACKED AUTOENCODER

Network	Layer 1	Layer 2	Layer 3
Deep Stacked Autoencoder 1 (SAE1)	128	64	32
Deep Stacked Autoencoder 2 (SAE2)	500	1,000	1,500

Seizure and non-seizure EEG segments of equal length are selected from previously annotated data. From each marked seizure onset an equal number of samples are selected prior to the annotation point (non-seizure segment) and after the annotation point (seizure segment). By selecting the same number of seizure and non-seizure EEG segments, the dataset can be considered balanced. No filtering and no pre-processing is applied. All available channels are considered as part of the database for classification.

B. Evaluation method

Evaluation of the proposed algorithm was performed per subject (for all 24 available recordings) using a leave-one-record-out cross-validation approach. Assuming that N EEG segments (seizure and non-seizure data) are available, $N-1$ segments are used for the training of the model, while the remaining segment is used for its testing. The data is permuted into N different such train and test sets. The number of segments N depends on the total number of seizures annotated for each recording. Each result is attributed a true positive, true negative, false positive and false negative value. The metrics for evaluation are globally computed per patient. These are accuracy, sensitivity and specificity [12]. Sensitivity measures the amount of seizures that are predicted as seizures, while specificity provides information on the true negative rate.

C. Parameter tuning

1) Autoencoder: In testing the proposed deep stacked autoencoder architecture, several parameters were varied: (i) the input window size, (ii) the size of the autoencoding layers, (iii) the number of dropout layers after each encoding layer. Variations in the proposed layer architecture were evaluated through the metrics described in section IV-B. These were computed per subject and the mean across all subjects is reported.

The input window size was varied from 1s, 3s, 5s and 20s for each architecture for testing purposes. In all cases, the performance of the models increased with an increase of the window size. A 20 second window results in the best performance, but depending on the network architecture, a 1 second input window can also be acceptable.

The layer sizes were changed according to Table I. The first network - Stacked Autoencoder 1 (SAE1) - decreases the size of the encoding layers, concentrating the coded input into only 32 points. The second network - Stacked Autoencoder 2 (SAE2) - expands the layer nodes, layer 3 being three times as large as layer 1. The decrease and increase in the layer size is the same for each window size

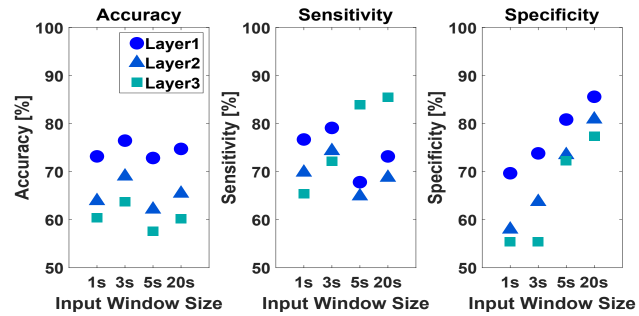

Fig. 3. Achieved performance varying the input window size for SAE1.

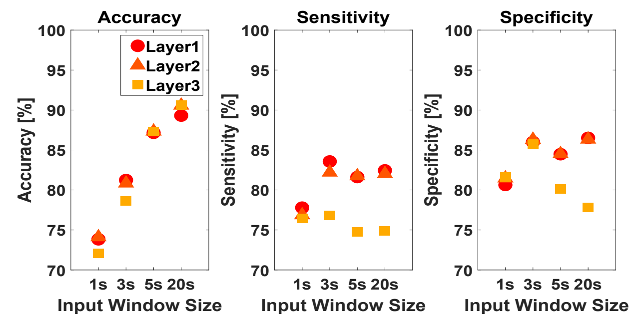

Fig. 4. Achieved performance varying the input window size for SAE2.

used as input. Performance measures for the two network models are presented in Fig. 3 and 4. The expanding SAE2 provides overall better results when compared to SAE1 for all input window size variations. When using the first layer for feature extraction, the results are better than for the second and third layer.

The high number of output parameters obtained from the deep layer might slow down the computation and might introduce overfitting. To avoid these problems, dropout layers can be used [13]. For 1 second of input data the performance of SAE1 and SAE2 are tested by adding dropout layers with a dropout rate of 0.6 after each encoding layer. When adding dropout layers to SAE2 the performance slightly increases, while for SAE1 it decreases. This is to be expected as SAE1 compresses the input information and dropping some of these values would mean eliminating useful information. Whereas for SAE2, the input information is expanded in the encoding layers, creating irrelevant parameters. The dropout layer trims some of the additional information (see Fig. 5).

Our experiments show that optimal autoencoder design for seizure and non-seizure segment classification is that of SAE2 with dropout layers after each encoding layer. The best performance is obtained on a 20s input window, but smaller input window sizes provide higher performances than in the case of SAE1.

2) Proposed network architecture: The proposed deep convolutional neural network architecture for feature extraction is also tested with respect to the input window size variation and the size of the fully connected layers. The input window size is varied similarly on 1s, 3s, 5s and 20s time windows. Fully connected layer sizes are either decreased

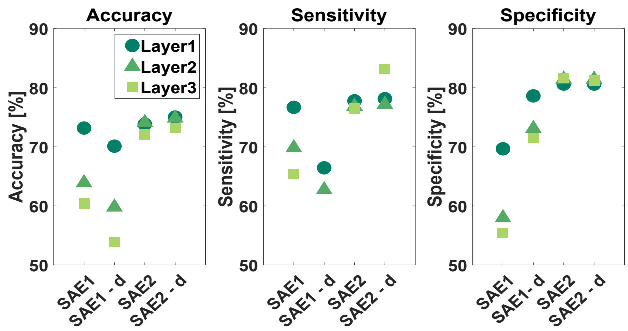

Fig. 5. Achieved performance when dropout layers are added after each autoencoder's layer for SAE1 and SAE2 (input data is of 1 second).

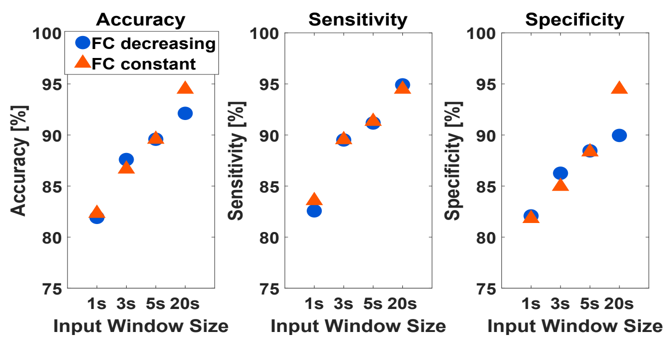

Fig. 6. Achieved performance for different fully connected (FC) layer sizes: FC decreasing (128, 64, 32) and FC constant (128, 128, 128).

TABLE II

SEIZURE AND NON-SEIZURE SEGMENT CLASSIFICATION

Network	Accuracy %	Sensitivity %	Specificity %
SAE2 - mean	86.21	90.10	82.31
CNN + FC - mean	92.12	94.90	89.75
SAE2 - S19	97.82	100	95.65
CNN + FC - S7	97.10	98.50	95.77

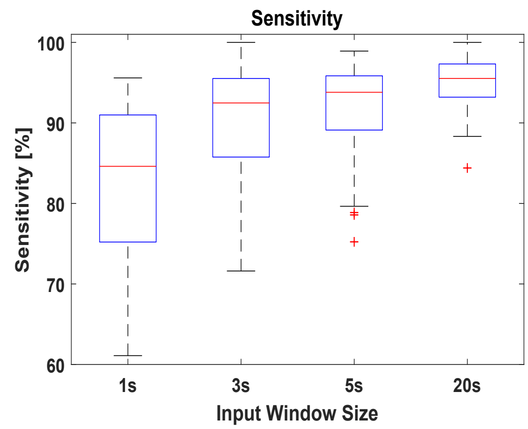

Fig. 7. Variation of the sensitivity of the individualized CNN + FC network model across different input window sizes. Each boxplot represents the variation in sensitivity obtained from the entire database for individualized models with different input time windows.

(128, 64, 32) or kept constant at 128 nodes. Fig. 6 presents the overall mean metrics for all subjects on the two network architectures. Increasing the number of nodes to 128 for all fully connected layers does not bring improvements to the results. Thus maintaining the first fully connected network configuration provides sufficient performance with a smaller number of computations.

V. RESULTS AND DISCUSSION

All testing and evaluation was performed on the CHB-MIT database. The summary of the results of both the stacked autoencoder and the proposed convolutional neural network architecture (with an input window of 20 seconds) are presented in Table II. The mean metrics over all individualized models are presented along with the results of the best performing individualized patient models. The SAE2 algorithm showed best results on patient 19, whereas the CNN-FC on patient 7. Although on specific individualized models SAE2 has slightly higher accuracy and sensitivity values, the convolutional neural network based architecture performs better over the entire patient data set.

Fig. 7 provides more information on the variation of the CNN-FC method performance over the entire data set. Sensitivity is the most relevant metrics of the three used as it indicates the ability of the classifier to detect the true positives. Overall, less variation in sensitivity and performance is observed when compared to our previous study [14]. Our previous study used manually extracted wavelet-based features on 20s input windows. This was fed to an SVM classifier, in a standard supervised learning approach. In our previous work, a high variation in classification accuracy was obtained for individual patient models. These accuracies ranged from 50% for some of the patients to above 90% for others. Using the CNN-FC method, the inter-subject variation of the performance metrics is reduced. When using 20s as input window, the sensitivity over the entire database for individualized models is between 85-100%.

Fig. 8 shows the features created by the model when using a 1s and 20s input. When increasing the input window size, the features are better separated between seizure and non-seizure EEG segment which is also reflected in the performance measures.

The variation in sensitivity of the model across the database decreases with an increase in input window size. Although the mean sensitivity for a 1s input window is lower than that obtained with a 20s input window, it is still sufficiently high (85%) and comparable to other values found in literature ([4] - 81.2%, [5] - 98.85% , [6] - 100%). A trade off exists between selecting the input window size and the obtained performance.

The biggest advantage of using unsupervised feature learning methods such as the proposed stacked autoencoder or the 1D based convolutional neural network is the small amount of input data required for classification. For smaller input window sizes, higher time resolutions are possible in applications requiring automatic annotation of data. Epileptic seizures can last for several minutes or seconds. Epileptiform discharges with duration smaller than 10s can be in some

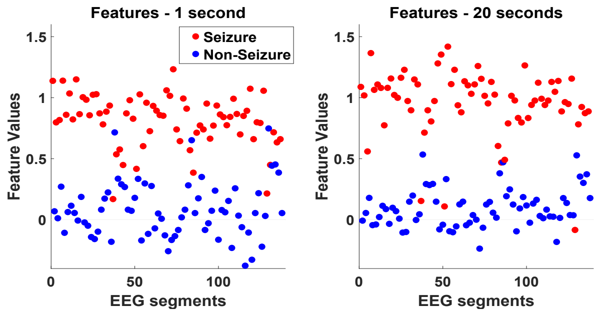

Fig. 8. Features obtained as the output of the CNN-FC network for 1s and 20s input seizure and non-seizure EEG segments.

cases observed on the recorded EEGs. Using shorter time windows can be beneficial in a more precise identification of such events in large amounts of data. When compared to classical manual frequency feature extraction techniques [3] or even CNN based methods that require 2D inputs [4], our method provides equivalent performance with smaller input windows. A 30s window is required for computing the Short Time Fourier Transform used as input for the CNN based method proposed in [4]. The mutual information based image creation described by the authors in [5] requires 8s EEG segments as input. The autoencoder used in [7] also reduced the amount of input data to 3s samples. The algorithm proposed by the authors in [6] also uses a stacked autoencoder on 1s of input EEG data. The reported sensitivity is of 100%. The algorithm was tested only on two patients from the CHB-MIT database and furthermore requires manual threshold adjustments. The choice of the network architecture and of the input time window used should be a balance between the required time resolution and performance. Several other deep neural network algorithms have been developed and tested for the detection of epileptic seizures. However, these are evaluated also on different available data sets. The recent work presented in [15] uses a 13 layer CNN with resulting in a sensitivity of 95%. These values cannot be compared to our results as the input data was different.

VI. Conclusions and Future Work

In this article we investigated the use of stacked autoencoders and CNN based network architectures for unsupervised feature extraction, with the objective of automatically detecting seizures from EEG data. We proposed a network architecture composed of three convolutional layers followed by four fully connected layers for feature extraction. The output is fed to an SVM classifier for the final prediction. Experimental validation was carried out using the CHB-MIT dataset. Results show the advantage of this approach leading to an accuracy of 92.12% as a mean over all the individual models, with the highest individual model accuracy of 97.82%. A high accuracy was maintained even with a small amount of input data.

As a point of improvement for the current work, our algorithm should be tested on multiple EEG databases for a better evaluation of performance. At the moment, all available EEG channels are used as input for the models. Channel selection might improve the performance of the classification. Moreover, unsupervised features can be used to improve channel selection methods. The proposed CNN-FC network architecture might be further improved through adding additional network layers. It is possible that similar performances can be obtained with smaller network layers. Our future work will focus on the development of a generalized, non-patient specific model with sufficiently high performance and low input data requirements.

References

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

[2] A. Shoeb and J. Guttag, "Application of Machine Learning To Epileptic Seizure Detection," *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, pp. 975–982, 2010. [Online]. Available: http://www.icml2010.org/papers/493.pdf

[3] S. B. Wilson, "Algorithm architectures for patient dependent seizure detection," *Clinical Neurophysiology*, vol. 117, no. 6, pp. 1204–1216, 2006.

[4] N. D. Truong, L. Kuhlmann, M. R. Bonyadi, and O. Kavehei, "Semi-supervised Seizure Prediction with Generative Adversarial Networks," pp. 1–6, 2018. [Online]. Available: http://arxiv.org/abs/1806.08235

[5] H. Ke, D. Chen, X. Li, Y. Tang, T. Shah, and R. Ranjan, "Towards Brain Big Data Classification: Epileptic EEG Identification With a Lightweight VGGNet on Global MIC," *IEEE Access - Special section on cyber-physical-social computing and networking*, vol. 6, pp. 14 772–14 733, 2018.

[6] A. Supratak, L. Li, and Y. Guo, "Feature Extraction with Stacked Autoencoders for Epileptic Seizure Detection," in *Conf. Proc IEEE Eng Med Biol Soc*, 2014, pp. 4184–4187.

[7] G. Xun, X. Jia, and A. Zhang, "Detecting epileptic seizures with electroencephalogram via a context-learning model," *BMC Medical Informatics and Decision Making*, vol. 16, pp. 97–109, 2015.

[8] M. P. Hosseini, H. Soltanian-Zadeh, K. Elisevich, and D. Pompili, "Cloud-based deep learning of big EEG data for epileptic seizure prediction," *2016 IEEE Global Conference on Signal and Information Processing, GlobalSIP 2016 - Proceedings*, pp. 1151–1155, 2017.

[9] Y. Bengio, "Learning Deep Architectures for AI Trends in Machine Learning," *Foundations and Trends in Machine Learning*, vol. 2, no. 1, pp. 1–127, 2009. [Online]. Available: http://www.iro.umontreal.ca/bengioy

[10] V. Nair and G. Hinton, "Rectified Linear Units Improve Restricted Boltzmann Machines," in *Proc. 27th International Conference on Machine Learning*, no. 3, 2010, p. 6421113. [Online]. Available: https://www.cs.toronto.edu/ hinton/absps/reluICML.pdf

[11] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-k. Peng, and H. E. Stanley, "PhysioBanl, PhysioToolkit, and PhysioNet, Components of a New Research Resource for Complex Physiologic Signals," *Circulation*, vol. 101, no. 23, pp. E215–20, 2000.

[12] M. Sokolova, N. Japkowicz, and S. Szpakowicz, "Beyond Accuracy, F-score and ROC: A Family of Discriminant Measures for Performance Evaluation," in *Australasian Joint Conference on Artificial Intelligence*, 2006, pp. 1015–1021.

[13] S. Nitish, H. Geoffrey, K. Alex, S. Ilya, and S. Ruslan, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *Journal of Machine Learning Research*, vol. 15, pp. 1929–1958, 2014.

[14] A. M. Tautan, A. I. Munteanu, D. D. Taralunga, R. Strungaru, and G. M. Ungureanu, "Automated Classification of Epileptiform Discharges in EEG Signals Using the Wavelet Transform," in *2018 International Conference and Exposition on Electrical And Power Engineering (EPE)*. IEEE, 2018, pp. 877–882.

[15] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli, "Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals," *Computers in Biology and Medicine*, vol. 100, no. July 2017, pp. 270–278, 2018. [Online]. Available: https://doi.org/10.1016/j.compbiomed.2017.09.017