Epileptic Seizure Detection in Clinical EEGs Using an XGboost-based Method

L. Wei[1,2] and C. Mooney[1,2]

1. School of Computer Science, University College Dublin, Dublin, Ireland
2. FutureNeuro SFI Research Centre, Royal College of Surgeons in Ireland, Dublin, Ireland
lan.wei@ucdconnect.ie, catherine.mooney@ucd.ie

Abstract— Epilepsy is one of the most common serious disorders of the brain, affecting about 50 million people worldwide. Electroencephalography (EEG) is an electrophysiological monitoring method which is used to measure tiny electrical changes of the brain, and it is frequently used to diagnose epilepsy. However, the visual annotation of EEG traces is time-consuming and typically requires experienced experts. Therefore, automatic seizure detection can help to reduce the time required to annotate EEGs. Automatic detection of seizures in clinical EEGs has been limited-to date. In this study, we present an XGBoost-based method to detect seizures in EEGs from the TUH-EEG Corpus. 4,597 EEG files were used to train the method, 1,013 EEGs were used as a validation set, and 1,026 EEG files were used to test the method. Sixty-four features were selected as the input to the training set, and Synthetic Minority Over-sampling Technique was used to balance the dataset. Our XGBoost-based method achieved sensitivity and false alarm/24 hours of 20.00% and 15.59, respectively, in the test set. The proposed XGBoost-based method has the potential to help researchers automatically analyse seizures in clinical EEG recordings.

keywords: **TUH-EEG Corpus, Seizure detection, Epilepsy, Machine learning**

I. Introduction

Epilepsy is one of the most common serious disorders of the brain, affecting about 50 million people worldwide [1]. Electroencephalography (EEG) is an electrophysiological monitoring method which can be used to measure tiny electrical changes of the brain. Epilepsy can lead to abnormal EEG readings [2], therefore, EEG is commonly used to diagnose epilepsy [3–5]. However, manually annotating EEG tracks is time-consuming, reproducibility between observers is low, and complicated by different types of seizures. Moreover, in a clinical setting, there is a shortage of experienced neurologists who can diagnose epilepsy through a visual examination. Automated detection systems are a powerful tool which can help to solve these problems by reducing expert annotation time and making annotations more reproducible.

Beginning in the 1970s, Viglione *et al.* used pattern recognition principles to develop an automatic epileptic seizure warning system. Their system transforms the signal to extract or detect the basic features from the processed signal, and determine the condition of the person from the detected features [6–8]. Since then, various approaches such as threshold-based [9, 10], machine learning-based [11, 12], and deep learning-based [13–16] approaches have been applied to this problem. However, some published seizure detection methods are trained on a small EEG data set with a small number of specific patients, resulting in these methods are not suitable for clinical use. In spite of the fact that thousands of EEGs are recorded each year in clinical settings around the world, relatively little data is publicly available in a useful form for the machine learning research community [17].

In 2020 Novela Neurotech joined forces with NeuroTechX to create an online crowdsourcing challenge using the TUH-EEG dataset to develop better seizure prediction algorithms [18]. The TUH-EEG Corpus is the world's largest publicly accessible archive of clinical EEG recordings [17]. Here we present results from our participation in this challenge.

II. Materials and methods

II-A. Dataset

The TUH-EEG data includes more than 40 unique channel configurations that have been manually annotated to represent different forms of seizure events. Most of the EEGs are recorded using at least 19 electrodes, corresponding to the international 10-20 EEG system. Shah *et al.* [19] have applied a combination of vertical and horizontal bipolar skin to create 22 different channels focusing on the focal area of the scalp. The 22 channels contain channel F7-T3, F8-T4, Fp1-F7, Fp2-F8, Fp1-F3, Fp2-F4, T5-O1, T6-O2, T3-T5, T4-T6, A1-T3, T4-A2, C3-CZ, CZ-C4, T3-C3, C4-T4, P3-O1, P4-O2, C3-P3, C4-P4, F3-C3 and F4-C4. In this study, we used signals from these 22 channels to estimate the features of interest. Figure 1 and Figure 2 show examples of non-seizure and seizure events in EEG recordings from channel Fp1-F3. 4,597 EEG recordings with 2,370 seizures were used to train, 1,013 EEG files with 673 seizure events were used to validate and 1,026 EEG files were used for independent testing of the method (Table 1).

II-A1. Channel Selection

EEG recording is an extremely complex process, and the unique channel configuration of each EEG or clinical site needs to be adapted. As can be seen from Figure 3, when a seizure occurs, the signal of channel Fp1-F3 changes significantly, but the signal of channel P3-O1

Table 1. Number and duration of seizures used to train, validate and test the method. The test set was used by the Neureka 2020 Epilepsy Challenge to independently evaluate the method, and the details of this set are unknown.

	Train	Validation	Test
Number of EEGs	4,597	1,013	1,026
Number of patients	592	50	Unknown
Number of seizures	2,370	673	Unknown
Seizure duration(s)	168,139.23	58,445.11	Unknown
Non-seizure duration(s)	2,540,144.77	554,786.89	Unknown

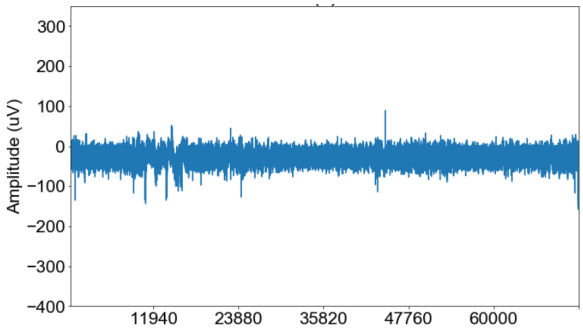

Figure 1. Example of a non-seizure region in an EEG recording from channel Fp1-F3.

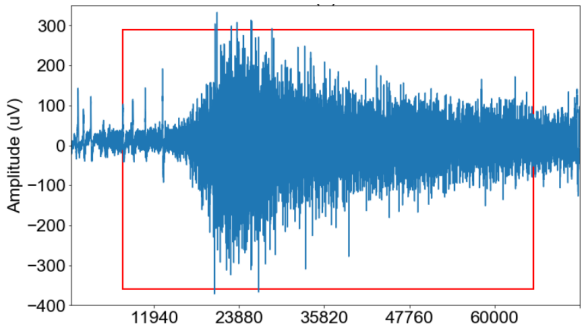

Figure 2. Example of a seizure in an EEG recording from channel Fp1-F3 (the red block indicates the seizure region).

does not. Therefore, not every channel's signal changes significantly when a seizure occurs. If we can reduce the number of channels, it is especially beneficial to make the channel universal across all EEGs and to provide a reasonable level of performance. In addition, it will reduce the dimensionality of the problem.

The TUH-EEG data contains multiple types of epilepsy, such as non-specific focal seizure (FNSZ), generalised seizure (GNSZ), and complex partial seizure (CPSZ). For different types of epilepsy, seizures may occur in different locations of the brain. For example, even focal epilepsy can occur in any lobe on either side, but the most commonly observed in the temporal or frontal lobe epilepsy [20]. Therefore, in order to minimise the loss of information, we selected four channels from different brain locations. These are channel F8-T4 to get information for the temporal lobe, T5-O1 for the occipital lobe, and Fp1-F3 for the frontal lobe. Moreover, channel T5-O1 and Fp1-F3 can get information from the left cerebral hemisphere and T5-O1 for the right cerebral hemisphere. We also selected channel F3-C3 to get the information from the central cerebral hemisphere.

II-B. Data Pre-processing

EEG recordings from the TUH-EEG dataset were sampled at different sample frequencies of 250Hz, 256Hz, 400Hz, and 1000Hz. Therefore, resampling was applied to down-sample the signal to 250 Hz. A notch filter (60 Hz) was applied to remove power line interference, and the DC offset was removed from the EEG recordings. The EEG signal was divided into epochs of 1s with 0.5s overlap, with each epoch corresponding to seizure events or non-seizure events.

II-B1. Feature Estimation

From Figure 4, we can see that compared with a non-seizure event, seizure events contain higher amplitude and higher frequency. The frequency of a non-seizure event is usually below 10Hz, but the frequency of the seizure event is much higher than 10Hz. Teager-kaiser energy operator (TKEO) calculates the instantaneous amplitude and frequency of the signal. Therefore, we choose TKEO and relative high-frequency power which is the sum of the relative power of beta (16-32Hz) and gamma (32-64Hz) as features of our method. We also measure the other features from the time domain and frequency domain to improve our method to analyze EEG activity. In this study, we used Butterworth filter to get signals in the band of interest, 1s epochs with 0.5s overlap were used to develop 16 features in each channel. We used four channels (channel F8-T4, Fp1-F3, F3-C3, and T5-O1) to develop our algorithm (64 features in total). These features from each channel are as follows:

Time domain (9): The mean, standard deviation, signal envelope, kurtosis, skewness, complexity, mobility, TKEO and fractal dimension of the pre-processed absolute amplitude of EEG recordings.

Frequency domain (7): A Butterworth band-pass filter with order 6 was used to get the relative band power of delta (0.1-4Hz), theta (4-8Hz), alpha (8-16Hz), beta (16-32Hz) and gamma (32-64Hz). The absolute band power of the EEG amplitude and the sum of relative beta and gamma were also extracted as features for our seizure detection method.

II-B2. Dataset Balancing

The overall duration of seizure periods is much shorter than that of the non-seizure periods. This dataset contains 336,278 seizure events and 5,080,290 non-seizure events in the training set (the duration of each event is 0.5s). This leads to class imbalance, which makes it difficult to train machine learning algorithms. Therefore, Synthetic Minority Over-sampling Technique (SMOTE) [21] was used to balance the data in the training set. The working method of SMOTE is to select the close sample in the feature space, draw a line between the

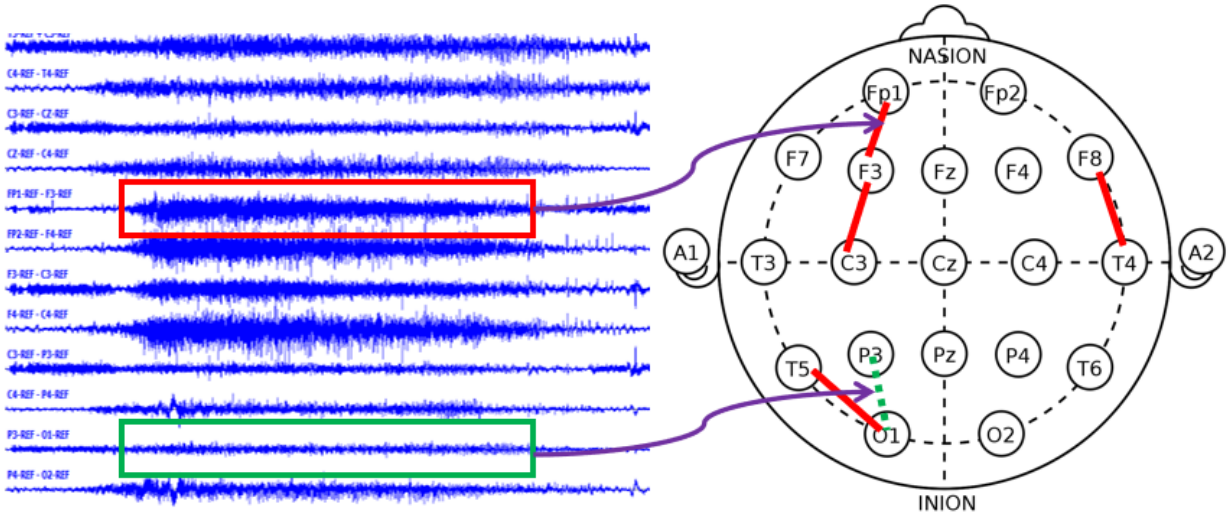

Figure 3. An example of tonic-clonic seizures (TCSZ) in EEG signals. The abnormal events are easily observed in channels CZ-C4, Fp1-F3, Fp2-F4, F3-C3 and F4-C4. However, in the same period, the abnormal events can not be observed in channel C3-CZ, C3-P3, C4-P4 and P3-O1. The signal in the red block indicates the presence of abnormal EEG; the signal in the green block is a normal EEG signal. The red lines located in different parts of the brain indicate the channels selected in this study. Channel F8-T4 at the temporal lobe, T5-O1 located at the occipital lobe, Fp1-F3 at the frontal lobe, and F3-C3 in the central cerebral hemisphere.

samples in the feature space, and plot a new sample at a point on the line. It starts by randomly selecting an example from the minority class and then finds the nearest neighbour k to this example (k=5 in this study). It then randomly selects a neighbourhood in the feature space and creates a composite example at a randomly selected point between the two samples.

II-B3. Classification Algorithms

XGBoost [22] is a decision-tree-based ensemble machine learning algorithm that uses a gradient boosting framework, which integrates many weak classifiers to form a robust classifier. In this study, the XGBoost algorithm was implemented within the Python 3 environment, scikit-learn [23]. Scikit-learn is an efficient tool for predictive data analysis, it is open source, accessible to everybody, and reusable in various contexts.

Three parameters were optimised in this study: n-estimators: the number of sequential trees to be modelled; gamma: gamma specifies the minimum loss reduction required to make a split (a node is split when the resulting split gives a positive reduction in the loss function); and max-depth: maximum tree depth for base learners. We optimised the parameters based on the performance of the validation set to improve the performance of the algorithm in the seizure detection. The n-estimators values were tested from 50 to 1000, the gamma values were tested from 0.010 to 0.020, and max-depth were tested from 8 to 13. When n-estimators = 100, gamma = 0.015, and max-depth = 10, the performance on the validation set was best. In addition, the learning rate of XGBoost classifier in this work is 0.01, the seed is set to 22, and the random state is 2.

II-B4. Data Post-processing

In this study, the EEG data were collected in a clinical settings, where there is greater variability and noise compared to a tightly controlled research setting. Therefore, the seizures detected by the XGBoost-based method with an interval less than 2s were grouped together, and their duration was defined as the first component's start time to the end time of the last component. The duration of the seizure is typically greater than 15s; therefore, if the duration of seizure detected by the XGBoost-based method was not greater than 15s, they were relabeled as a non-seizure. Otherwise, the seizure was defined as the final seizure event.

II-C. Performance Evaluation

Seizure events can vary in duration from a few seconds to many hours. In some applications, correctly detecting the number of events is as crucial as their duration [24]. In this study, the sensitivity (Sens), accuracy (Acc), and F1 score (F1) of the XGBoost-based method in estimating the seizure events were evaluated. The area under the receiver operating characteristics curve (AUROC) and false alarm rate were also used to evaluate the XGBoost-based method. Since the 1970s, the AUROC has been widely studied and applied in the field of medical diagnosis [25], and AUROC is an unbiased measurement for imbalanced data [26]. The false alarm rate is defined as the number of false alarms (giving an alarm or warning in the case of a non-event) per the total number of 'non-events' (the number of times the event did not occur). These evaluation metrics are calculated as follows:

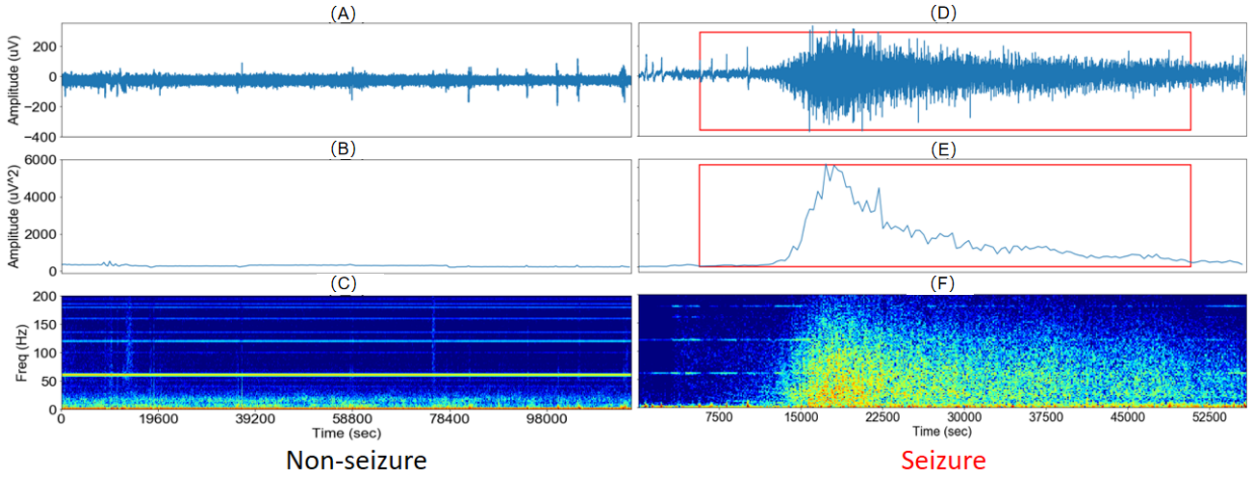

Figure 4. Figures A and D show examples of the seizure pattern in channel Fp1-F3 (the signal in red block indicates the presence of seizure); Figures B and E show TKEO of the corresponding EEG signal; and Figures C and F show the spectrogram of the signal

$$\begin{aligned} Sens &= \frac{TP}{TP+FN} \\ Acc &= \frac{TP+TN}{TP+TN+FP+FN} \\ F1 &= \frac{TP}{TP+1/2*(FP+FN)} \end{aligned} \quad (1)$$

where: True Positive (TP): Annotated as a seizure and predicted as a seizure
True Negative (TN): Not annotated as a seizure and not predicted as a seizure
False Negative (FN): Annotated as a seizure and not predicted as a seizure
False Positive (FP): Not annotated as a seizure and predicted as a seizure.

II-D. Neureka 2020 Epilepsy Challenge

In 2020 Novela Neurotech joined forces with NeuroTechX to create an online crowdsourcing challenge using the TUH-EEG dataset to develop better seizure prediction algorithms [18]. For this competition, the training set and validation set were released to build the method. An independent test set was used for testing the methods. Sensitivity and false alarms were measured, and the score was then combined into a single scalar scoring metric as follow:

$$Score = Sens - 2.5 * FA/24hr - 7.5 * (Channels(N)/19)$$

This measure has been shown to correlate well with a system's ability to segment a signal into seizure and background events accurately.

III. Results

Table 2 presents the results of the XGBoost-based seizure detection method on the training set, validation set and test set EEG recordings. From Table 2, we can see that the proposed method in this study achieved a sensitivity of 59.80% and 51.90% on the training and validation set, respectively. In the test set, the method achieved lower sensitivity, compared with the validation set, at 20.00%, but obtained lower FA/24hr at 15.59. Figure 5 shows the ROC curve of the validation set.

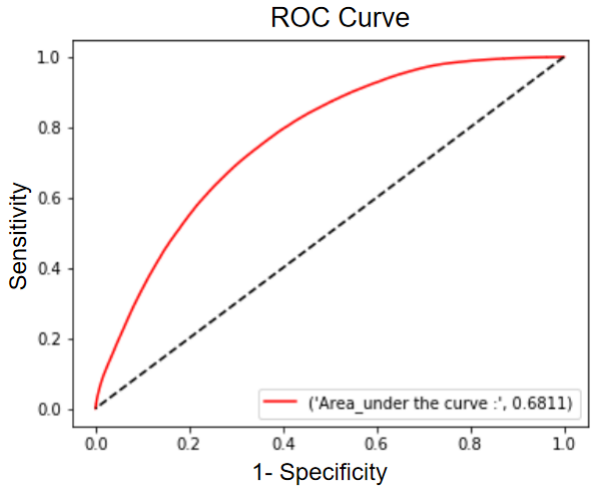

Figure 5. ROC curve showing the sensitivity plotted against 1 - specificity on the validation set.

Table 2. Performance of the XGBoost-based seizure detection method.

	Sens(%)	Acc(%)	F1	AUROC	FA/24hr
Train (N=4,597)	59.80	67.01	0.5881	0.8030	55.69
Val (N=1,013)	51.90	58.85	0.4935	0.6811	81.35
Test (N=1,026)	20.00	-	-	-	15.59

Train: training set; Val: validation set; Test: testing set.

Table 3 shows previously published seizure detection methods which were developed using the TUH-EEG data. In our method, we used four channels with 64 features as the input of the XGBoost-based method.

In contrast, the work of Shah *et al.* [19] used various EEG channels and different CNN layers to construct a 2D CNN method. When four channels were used as input, the sensitivity of 3 CNN layers is 33.11%, and FA/24hr is 325.54. When one layer of CNN was used, the sensitivity is 34.09%, and FA/24hr is 332.15. For our XGBoost-based method, we achieved lower FA/24hr (81.35) and higher sensitivity (51.90%) on the validation set. In the work of Ziyabari *et al.* [24], 22 channels of EEG signal were used as input, and various methods were developed to detect the seizure in TUH-EEG data. Among these methods, CNN/MLP reached the highest sensitivity of 31.58%, with FA/24hr of 91 in the validation set. CNN/LSTM achieved the lowest FA/24hr at 8, with the sensitivity of 31.58% in the validation set. Compared with our XGBoost-based method, higher sensitivity in the validation set was obtained by using fewer channels (N=4).

Table 3. The performance of the proposed seizure detection method compared to other published method

Ref	Method	Channel	Sens (%)	FA/24hr
[19]	2D CNN-L3 (Val)	22	39.15	22.83
	2D CNN-L3 (Val)	20	34.54	49.25
	2D CNN-L3 (Val)	16	36.54	53.99
	2D CNN-L3 (Val)	8	33.44	38.19
	2D CNN-L3 (Val)	4	33.11	325.54
	2D CNN-L2 (Val)	8	30.66	28.57
	2D CNN-L1 (Val)	4	34.09	332.15
	2D CNN-L3 (Val)	2	31.15	308.74
[24]	HMM/Sda (Val)	22	17.29	82.00
	HMM/LSTM (Val)	22	22.84	68.00
	IPCA/LSTM (Val)	22	22.12	83.00
	CNN/MLP (Val)	22	31.58	91.00
	CNN/LSTM (Val)	22	12.48	8.00
This work	XGBoost (Val)	4	51.90	81.35

L: layer (eg: L2 means two layers); Tra: training; Val: validation;

Table 4 presents the performance of other participants in the Neureka 2020 Epilepsy Challenge. From the Table 4 we can see that even without using all 22 EEG channels, some relatively precise results can still be obtained. Compared with other work proposed in this Challenge, we (position 5) obtained higher sensitivity with fewer channels, but the FAs/24hr is higher than the other four teams (position 1-4).

Table 4. Comparison of the performance of other work in the Neureka 2020 Epilepsy Challenge

Po	Team/Ind.	Sen (%)	FAs/24hr	Chan	Score
1	Biomed Irregulars	12.37	1.44	16	2.46
2	NeuroSyd	2.04	0.17	2	0.82
3	USTC-EEG	8.93	0.71	17	0.45
4	RocketShoes	5.98	3.36	3	-3.60
5	**Lan Wei (Ind.)**	**20.00**	**15.59**	**4**	**-20.56**
6	EEG Miners	16.00	16.54	9	-28.89
7	Anonymous (Ind.)	21.65	28.05	4	-50.05
8	James Msonda (Ind.)	11.33	29.27	10	-65.79
9	TABS	9.03	31.21	19	-76.50
10	cpl team	5.66	94.34	1	-230.59
11	DeepAlert	9.86	172.92	10	-426.40
12	Interfaces	26.53	186.63	1	-440.44
13	Neurocomputacion	0.22	758.48	11	-1,900.32
14	TeamPT2	34.75	927.12	19	-2,290.53
15	Last Dance	10.13	1,385.03	1	-3,452.83

Po: position; Chan: channels

IV. Discussion

In this study, an XGBoost-based seizure detection method were presented to detect seizures in TUH-EEG data. Some published algorithms [27–30] report accurate detection of seizures in long EEG recordings. However, most of these works used data which is not publicly available and is therefore hard to replicate. Moreover, differences in data sets and performance evaluation methods make direct comparisons with other published seizure detection methods difficult. Therefore, we compared our XGBoost-based method with the same data set (TUH-EEG data) and the same evaluation criteria in this study.

We compared the results of the proposed seizure detection method with other work (see Table 3). Sensitivity analysis plays a critical role in assessing the robustness of results or conclusions based on preliminary analyses of clinical trial data [31]. Compared with the published work presented in Table 3, our XGBoost-based method present in this study achieved the highest sensitivity (51.90%) in the validation set.

The limitation of this study is that the FAs/24hr of our seizure detection method is not satisfactory. In future work, we would like to develop a method to detect seizures in the whole TUH-EEG data set with higher sensitivity and lower FA/24hr. We will analyse the difference between the features of the seizure period and the normal part of the EEG, and estimate some features that can distinguish the two events better to develop the model. Features will be fed into other machine learning algorithms to obtain a more accurate seizure detection method for TUH-EEG data.

V. Conclusions

In this study, we described an XGBoost-based method to detect seizures in the world's largest publicly accessible archive of clinical EEG data set (TUH-EEG). The XGBoost-based method shows a sensitivity of 20.00% and FA/24hr of 15.29 by using four channels in the test set of TUH-EEG data, which may assist researchers use fewer EEG channels in these long EEG recordings to automatically analyze seizures.

ACKNOWLEDGMENTS

This publication has emanated from research supported in part by a research grant from Science Foundation Ireland (SFI) under Grant Number 16/RC/3948 and co-funded under the European Regional Development Fund and by FutureNeuro industry partners. We acknowledge the Research IT HPC Service at University College Dublin for providing computational facilities and support that contributed to the research results reported in this paper.

REFERENCES

[1] W. H. Organization, G. C. against Epilepsy, P. for Neurological Diseases, N. W. H. Organization), I. B. for Epilepsy, W. H. O. D. of Mental Health, S. Abuse, I. B. of Epilepsy, and I. L. against Epilepsy, *Atlas: epilepsy care in the world*. World Health Organization, 2005.

[2] B. Semiology and M. K. Abou-Khalil, "Atlas of eeg & seizure," 2006.

[3] M. A. Mikati, E. Trevathan, K. S. Krishnamoorthy, and C. T. Lombroso, "Pyridoxine-dependent epilepsy: Eeg investigations and long-term follow-up," *Electroencephalography and clinical neurophysiology*, vol. 78, no. 3, pp. 215–221, 1991.

[4] S. Smith, "Eeg in the diagnosis, classification, and management of patients with epilepsy," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 76, no. suppl 2, pp. ii2–ii7, 2005.

[5] F. Mormann, K. Lehnertz, P. David, and C. E. Elger, "Mean phase coherence as a measure for phase synchronization and its application to the eeg of epilepsy patients," *Physica D: Nonlinear Phenomena*, vol. 144, no. 3-4, pp. 358–369, 2000.

[6] S. Viglione, V. Ordon, and F. Risch, "A methodology for detecting ongoing changes in the EEG prior to clinical seizures," *21st western institute on epilepsy*, pp. 27–8, 1970.

[7] S. S. Viglione, V. A. Ordon, W. B. Martin, and J. C. C. Kesler, "Epileptic seizure warning system," Feb. 4 1975, uS Patent 3,863,625.

[8] S. Viglione and G. Walsh, "Proceedings: Epileptic seizure prediction." *Electroencephalography and clinical neurophysiology*, vol. 39, no. 4, pp. 435–436, 1975.

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

[10] L. Ayoubian, H. Lacoma, and J. Gotman, "Automatic seizure detection in seeg using high frequency activities in wavelet domain," *Medical engineering & physics*, vol. 35, no. 3, pp. 319–328, 2013.

[11] A. Alkan, E. Koklukaya, and A. Subasi, "Automatic seizure detection in eeg using logistic regression and artificial neural network," *Journal of Neuroscience Methods*, vol. 148, no. 2, pp. 167–176, 2005.

[12] Y. Liu, W. Zhou, Q. Yuan, and S. Chen, "Automatic seizure detection using wavelet transform and svm in long-term intracranial eeg," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 20, no. 6, pp. 749–755, 2012.

[13] Y. Yuan, G. Xun, K. Jia, and A. Zhang, "A multi-view deep learning method for epileptic seizure detection using short-time fourier transform," *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, 2017, pp. 213–222.

[14] R. Hussein, H. Palangi, R. Ward, and Z. J. Wang, "Epileptic seizure detection: a deep learning approach," *arXiv preprint arXiv:1803.09848*, 2018.

[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, pp. 270–278, 2018.

[16] K. M. Tsiouris, V. C. Pezoulas, M. Zervakis, S. Konitsiotis, D. D. Koutsouris, and D. I. Fotiadis, "A long short-term memory deep learning network for the prediction of epileptic seizures using eeg signals," *Computers in biology and medicine*, vol. 99, pp. 24–37, 2018.

[17] I. Obeid and J. Picone, "The temple university hospital eeg data corpus," *Frontiers in neuroscience*, vol. 10, p. 196, 2016.

[18] Neureka 2020 epilepsy challenge. (available at: *https://neureka-challenge.com/*).

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

[20] J. W. Britton, L. C. Frey, J. L. Hopp, P. Korb, M. Koubeissi, W. Lievens, E. M. Pestana-Knight, and E. L. St, "Electroencephalography (eeg): An introductory text and atlas of normal and abnormal findings in adults, children, and infants," 2016.

[21] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of artificial intelligence research*, vol. 16, pp. 321–357, 2002.

[22] D. Nielsen, "Tree Boosting With XGBoost Why Does XGBoost Win "Every" Machine Learning Competition? ," Master's thesis, NTNU, 2016.

[23] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, "Scikit-learn: Machine learning in python," *the Journal of machine Learning research*, vol. 12, pp. 2825–2830, 2011.

[24] S. Ziyabari, V. Shah, M. Golmohammadi, I. Obeid, and J. Picone, "Objective evaluation metrics for automatic classification of eeg events," *arXiv preprint arXiv:1712.10107*, 2017.

[25] J. Huang and C. X. Ling, "Using auc and accuracy in evaluating learning algorithms," *IEEE Transactions on knowledge and Data Engineering*, vol. 17, no. 3, pp. 299–310, 2005.

[26] S. Wang, S. Sun, and J. Xu, "Auc-maximized deep convolutional neural fields for protein sequence labeling," *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 2016, pp. 1–16.

[27] N. Sadati, H. R. Mohseni, and A. Maghsoudi, "Epileptic seizure detection using neural fuzzy networks," *2006 IEEE international conference on fuzzy systems*. IEEE, 2006, pp. 596–600.

[28] R. Meier, H. Dittrich, A. Schulze-Bonhage, and A. Aertsen, "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, no. 3, pp. 119–131, 2008.

[29] W. Deburchgraeve, P. Cherian, M. De Vos, R. Swarte, J. Blok, G. H. Visser, P. Govaert, and S. Van Huffel, "Automated neonatal seizure detection mimicking a human observer reading eeg," *Clinical Neurophysiology*, vol. 119, no. 11, pp. 2447–2454, 2008.

[30] Y. Song, J. Crowcroft, and J. Zhang, "Automatic epileptic seizure detection in eegs based on optimized sample entropy and extreme learning machine," *Journal of neuroscience methods*, vol. 210, no. 2, pp. 132–146, 2012.

[31] L. Thabane, L. Mbuagbaw, S. Zhang, Z. Samaan, M. Marcucci, C. Ye, M. Thabane, L. Giangregorio, B. Dennis, D. Kosa *et al.*, "A tutorial on sensitivity analyses in clinical trials: the what, why, when and how," *BMC medical research methodology*, vol. 13, no. 1, pp. 1–12, 2013.