XGboost-based Method for Seizure Detection in Mouse Models of Epilepsy

L. Wei[1,2], R. Gerbatin[2,3], O. Mamad[2,3], H. Boutouil[2,3], C. Reschke[2,3], M. Lowery[4], D. Henshall[2,3], G. Morris[2,3] and C. Mooney[1,2]

1. School of Computer Science, University College Dublin, Dublin, Ireland
2. Department of Physiology & Medical Physics, Royal College of Surgeons in Ireland, Dublin, Ireland
3. FutureNeuro, SFI Research Centre for Chronic and Rare Neurological Diseases, RCSI, Dublin, Ireland
4. School of Electrical & Electronic Engineering, University College Dublin, Dublin, Ireland
lan.wei@ucdconnect.ie, catherine.mooney@ucd.ie

Epilepsy is a chronic neurological disease which affects over 50 million people worldwide [1], caused by the disruption of the finely tuned inhibitory and excitatory balance in brain networks, manifesting clinically as seizures. Electroencephalographic (EEG) monitoring in rodent disease models of epilepsy is critical in the understanding of disease mechanisms and the development of anti-seizure drugs. However, the visual annotation of EEG traces is time-consuming, and is complicated by different models and seizure types. Automated annotation systems can help to solve these problems by reducing expert annotation time and increasing the throughput and reliability of seizure quantification. As machine learning is becoming increasingly popular for modelling sequential signals such as EEG, several researchers have tried machine learning to detect seizures in EEG traces from mouse models of epilepsy. Most existing work [2, 3] can only detect seizures in single mouse models of epilepsy and research on multiple mouse models has been limited to-date.

In this work, we developed a Teager-Kaiser energy operator (TKEO)-based method that mimics how experts detect seizures in an intra-amygdala kainic acid (IAKA) mouse model of epilepsy. Furthermore, we propose a machine learning-based method which can review large volumes of data and discover specific trends and patterns that may not be apparent to humans. We compared the performance of the TKEO-based and machine learning-based method on the IAKA mouse model. We further tested these two methods on a Dravet syndrome (DS) mouse model of epilepsy to see whether these two methods could generalize to detect seizures in another mouse model of epilepsy.

The EEG data used in this study were recorded from nine mice with chemoconvulsant-induced epilepsy (IAKA; four mice) and genetically-induced epilepsy (DS; five mice). EEG recordings from the IAKA mice were sampled at a frequency of 500 Hz, while the DS mice were sampled at 512 Hz. Therefore, resampling was applied to down-sample the signal from 512 Hz to 500 Hz. A 50 Hz notch filter 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 5s with 2.5s overlap, with each epoch corresponding to seizure events or non-seizure events.

For the TKEO-based method, the mean TKEO of the first 20 minutes of EEG recording (the baseline period) of each mouse was calculated as the threshold. Then, the TKEO of each five-second epoch of the non-baseline period was compared to the threshold of each mouse. If the value in the epoch was two-fold higher than the threshold, the event was labelled as a potential seizure. The duration of the potential seizure was then checked, and if the duration of the event was greater than 10 seconds, it was defined as a seizure, if not, it was labelled as a regular event. Three IAKA mice (Seizures: 1,100 seconds; Non-seizures: 2,903,458 seconds) and four DS mice (Seizures: 340 seconds; Non-seizures: 224,465 seconds) were used for independent testing of the TKEO-based method.

XGBoost [4] is a boosting algorithm that has become increasingly popular recently. Boosting algorithms integrate many weak classifiers together to form a strong classifier. The XGBoost algorithm was implemented using the 'XGBClassifier' package of the sklearn library [5] within the Python 3 environment. Nineteen

features were extracted from the EEG signal of IAKA and DS mouse models in each five-second epoch. Features in frequency domain were estimated by Daubechies 4 wavelet. The estimated features are: TKEO; the absolute power of delta (0-4 Hz), theta (4-8 Hz), alpha (8-16 Hz), beta (16-32 Hz), and gamma (32-64 Hz) frequency band; relative power of delta, theta, alpha, beta, and gamma frequency band; total absolute power (0-250 Hz); mobility; mean; variance; kurtosis; skewness; signal envelope; and fractal dimension of the pre-processed signal. One IAKA mouse and one DS mouse were used to train (Seizures: 935 seconds; Non-seizures: 240,825 seconds) and validate (Seizures: 235 seconds; Non-seizures: 963,295 seconds) the algorithm. The other three IAKA mice (Seizures: 1,100 seconds; Non-seizures: 2,903,458 seconds) and four DS mice (Seizures: 340 seconds; Non-seizures: 224,465 seconds) were used for independent testing of the XGBoost-based method.

The TKEO-based seizure detection method performed well on the IAKA mouse model, with sensitivity and specificity of 72.1% and 97.5%, respectively. However, when we tested this method using the EEGs from the DS mouse model the sensitivity was 73.5%, but the specificity dropped to 62.4%. In contrast, the XGBoost-based method performed well on EEGs from both mouse models and achieved sensitivity and specificity of 93.0% and 99.3%, respectively, for the IAKA mouse model, and sensitivity and specificity of 99.5% and 98.0%, respectively, for the DS mouse model. In summary, our XGBoost-based method has the potential to assist researchers to detect seizures more quickly and reliably in EEG recordings of single-channel, multi-type seizures in long mouse EEG recordings. In future work, we will validate these methods using larger numbers of mice from different mouse models of epilepsy.

ACKNOWLEDGEMENTS

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, and research supported by the European Union Seventh Framework (FP7) 'EpimiRNA' project under grant agreement 602130. GM is supported by a Marie Skłodowska-Curie Actions Individual Fellowship 'EpimiRTherapy', H2020-MSCA-IF-2018 840262. 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] M. J. Brodie, "Road to refractory epilepsy: The Glasgow story," *Epilepsia*, vol. 54, pp. 5–8, 2013.

[2] Y. Pan, S. S. Ge, A. Al Mamun, and F. R. Tang, "Detection of seizures in EEG signal using weighted locally linear embedding and SVM classifier," in *2008 IEEE Conference on Cybernetics and Intelligent Systems*. IEEE, 2008, pp. 358–363.

[3] H.-J. Jang and K.-O. Cho, "Dual deep neural network-based classifiers to detect experimental seizures," *The Korean Journal of Physiology & Pharmacology*, vol. 23, no. 2, pp. 131–139, 2019.

[4] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794.

[5] 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," *Journal of machine learning research*, vol. 12, no. Oct, pp. 2825–2830, 2011.

XGBoost-based Method for Seizure Detection in Mouse Models of Epilepsy

L. Wei[1,2], R. Gerbatin[2,3], O. Mamad[2,3], H. Boutouil[2,3], C. R Reschke[2,3], M. Lowery[4], D. Henshall[2,3], G. Morris[2,3] and C. Mooney[1,2]

1. School of Computer Science, University College Dublin, Dublin, Ireland 2. Department of Physiology & Medical Physics, Royal College of Surgeons in Ireland, Dublin, Ireland

3. FutureNeuro, The SFI Research Centre for Chronic and Rare Neurological Diseases, RCSI, Dublin, Ireland 4. School of Electrical & Electronic Engineering, University College Dublin, Dublin, Ireland

Background

- Epilepsy is the second most common neurological disorder behind stroke, according to the World Health Organization (WHO).
- Epilepsy is caused by the disruption of the finely tuned inhibitory and excitatory balance in brain networks, manifesting clinically as seizures.
- Electroencephalographic (EEG) monitoring in rodent disease models of epilepsy is critical in the understanding of disease mechanisms and the development of anti-seizure drugs.
- Seizure detection with EEG requires a direct examination by a physician and substantial amount of time and effort.
- Automated detection is a powerful method to devote to this task which can reduce the annotation time of experts.
- Research on seizure detection methods applicable to multiple mouse models has been limited to date.
- In this study, an automated method for seizure detection in EEGs from different mouse models of epilepsy is proposed.

Data

Mouse model I : Intra-amygdala kainic acid (IAKA) Adult male SV129

EEG signals were recorded in each mouse for 14 days, with 20 minutes of baseline on the first day. 40 mins later, the intraperitoneal lorazepam (8 mg/kg) was injected to reduce morbidity and mortality. Following a latent period of 3-5 days, spontaneous recurrent seizures started to develop.

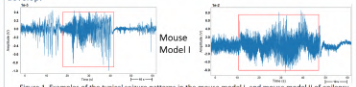

Figure 1. Examples of the typical seizure patterns in the mouse model I and mouse model II of epilepsy (the signal in the red block indicates the presence of seizure event).

Mouse model II : Dravet Syndrome (DS) F1.*Scn1a(+/-)tm1Kea*

Mice were bred with a mutation which mimics DS, a rare and severe epileptic encephalopathy. DS, F1.*Scn1a(+/-)tm1Kea* experience recurrent spontaneous seizures. These were recorded using tethered EEG monitoring from 12:30 to 6:30 pm between postnatal days p21-p28.

Table 1. Number and duration of seizures labelled by experts

	IAKA	DS
Number of mice	4	5
Average number of seizures	14	2
Total seizures number	54	11
Seizure duration (seconds)	2,212	398
Non-seizure duration (seconds)	4,053,284	278,758

Method

- Teager-Kaiser energy operator (TKEO)-based method

The discrete TKEO is defined

$$\varphi[x_t] = x^2 - x_{t-1} * x_{t+1}$$

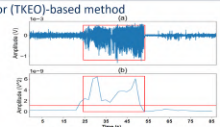

Figure 2. TKEO applied to the IAKA mouse EEG signal (a) the original signal (b) the signal after the TKEO approach wasapplied (the red line is the threshold)

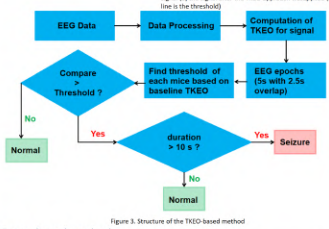

Figure 3. Structure of the TKEO-based method

- XGBoost-based method

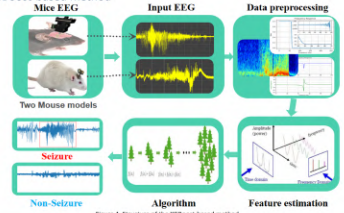

Figure 4. Structure of the XGBoost-based method

Results

Table 2. Performance of the TKEO-based seizure detection method on the independent test set.

	Sens (+/- std) (%)	Spec (+/- std) (%)	Acc (+/- std) (%)
IAKA Test (N=3)	72.1 (+/-12.0)	97.5 (+/- 0.1)	98.0 (+/- 0.7)
DS Test (N=4)	73.5 (+/-24.9)	62.4(+/-25.8)	62.3(+/-25.7)

Table 3. Performance of the XGBoost-based seizure detection method on the independent test set.

	Sens (+/- std) (%)	Spec (+/- std) (%)	Acc (+/- std) (%)
IAKA Test (N=3)	93.0 (+/-20.6)	99.3 (+/- 1.4)	99.3 (+/- 1.4)
DS Test (N=4)	99.5 (+/- 1.1)	98.0 (+/- 1.4)	98.0 (+/- 1.4)

std: standard deviation

Conclusion

- ✓ XGBoost-based method performed better than TKEO-based method in both mouse models of epilepsy
- ✓ A novel XGBoost-based method to detect seizures
- ✓ Assist researchers in the automated analysis of seizures in mouse model of epilepsy
- ✓ Single-channel, multi-type seizures in long mice EEG recordings
- ✓ Removes user bias when detecting seizures
- ✓ Fast, reliable, reproducible

Future work

The method need to be validated in data sets from :

- ☐ Larger numbers of mice
- ☐ Different mouse models of epilepsy

Acknowledgements

This publication has emanated from research supported in part by a research grant from Science FoundationIreland (SFI) under Grant Number 16/RC/3948 and co-funded under the European Regional DevelopmentFund and by FutureNeuro industry partners, and research supported by the European Union Seventh Frame-work (FP7) 'EpimiRNA' project under grant agreement 602130. GM is supported by a Marie Skłodowska-Curie Actions Individual Fellowship 'EpimiRTherapy', H2020-MSCA-IF-2018 840262. We acknowledgethe Research IT HPC Service at University College Dublin for providing computational facilities and sup-port that contributed to the research results reported in this paper.