Automatic epileptic seizure detection using multimodal biosignals recorded from a wearable device

Hamed Mohammadzadeh *[1, 2], Pouria Nazemi[1, 2], and Elaheh Imani †[1, 2]

[1]Neurosina
[2]Department of Computer Engineering, , Ferdowsi University of Mashhad

Abstract

This report proposes a method for detecting epileptic seizures using a simple wristband device, specifically the Empatica E4. The proposed method is designed over a publicly available dataset of physiological signals recorded from six patients with drug-resistant epilepsy. The dataset includes 3D accelerometry (ACC), blood volume pulse (BVP) measured by photoplethysmography (PPG), electrodermal activity (EDA), and temperature (TEMP) signals. The proposed method achieves a 74% sensitivity and a FAR of 35.8 d^{-1}, comparable to previous methods, considering it uses classic machine learning methods and only statistical and time-series features.

1 Introduction

Epilepsy is a chronic, noncommunicable brain disease that affects around 50 million people worldwide. It is characterized by recurrent seizures, which are brief episodes of involuntary movement that may involve a part of the body (partial) or the entire body (generalized). They are sometimes accompanied by loss of consciousness and control of bowel or bladder function [4]. The most common type (60%) of seizures is convulsive, which involves involuntary muscle contractions. More information about the distinct types of seizures can be found in Appendix A.

Seizure prediction for patients has been a challenging task, primarily due to the limitations of EEG monitoring. This traditional approach limits patients from engaging in everyday activities and proves to be uncomfortable with the constant presence of an implant or a head device. Consequently, a significant demand arises for more discreet wearable devices capable of seizure prediction or detection. One such promising solution lies in developing a simple wristband, which offers a minimalistic yet effective means to monitor and anticipate seizures without hindering the patients' daily lives. The Empatica E4 wristband is an example of such a device.

[6] conducted a study with patients with drug-resistant epilepsy and provided the patients with Empatica E4 devices to record various physiological signals, such as ACC, BVP (measured by PPG), EDA, TEMP, and HR, for 6 months. The researchers also collected iEEG (intracranial electroencephalogram) data during suspected seizure activity. Expert doctors reviewed these iEEG clips and identified the segments corresponding to seizure events. They made the data acquired by the Empatica E4 device publicly available as a challenge here.

2 Related Work

Extensive research has utilized multimodal devices, including simple wristbands, to detect and predict epileptic seizures. [2] used Empatica E4 to collect data, monitored 243 patients with

*hamedrq7@gmail.com, The work was done during an internship at Neurosina.
†elaheh.imani@gmail.com

epilepsy, and used Accelerometer and EDA modality to detect seizures. [2] did not analyze PPG (BVP) signals as the considerable movements during convulsive seizures render this signal too noisy for accurate ictal heart rate determination. Their study only included patients with focal to bilateral or generalized TCSs. Using a Gradient Tree Boosting Machine (GTBM) as the classifier, they achieved 100% sensitivity in cross-validation with 0.46 FAR per day.

[5] used the accelerometer modality to detect Convulsive Seizures and achieved 84.31% sensitivity with a mean FAR of 1.33 per day from 79 patients using the SVDD classifier. They collected 3-D accelerometer data from the patients using Apple iPod Touch.

3 Proposed Method

This project tackles a seizure detection task designed over the dataset from [6], which is publicly available here. The schematic view of Figure 1 depicts the pipeline used in the proposed detection process.

Figure 1: Pipeline overview of the experiments

3.1 Data acquisition

The data used in this project is collected from Empatica E4 and consists of raw data from 6 patients, which includes 3D accelerometry (ACC), blood volume pulse (BVP) measured by photoplethysmography (PPG), electrodermal activity (EDA), and Temperature (TEMP). The signals were recorded for a minimum of 6 months and were up-sampled to 128Hz. Each person also had an implanted brain device providing stimulation (NeuroPace RNS), which recorded their seizures using EEG, and these recordings were used to provide an accurate record of seizures. Table 1 shows the number of recorded seizures for each patient.

	1110	1869	1876	1904	1965	2002
# seizures	6	11	21	7	134	11
# valid seizures	4	10	12	6	x	8

Table 1: Number of recorded seizures for each patient. The second row shows the number of valid seizures, which is explained in section 3.3

3.2 Data Exploratory Analysis

Because of the sheer volume of the dataset, all the aspects of each modality can not be displayed at once, so several tools were developed to explore and plot each modality over a full day or right before/after seizures (Figure 2). Histograms for raw modalities were also calculated for each patient (Figure 3).

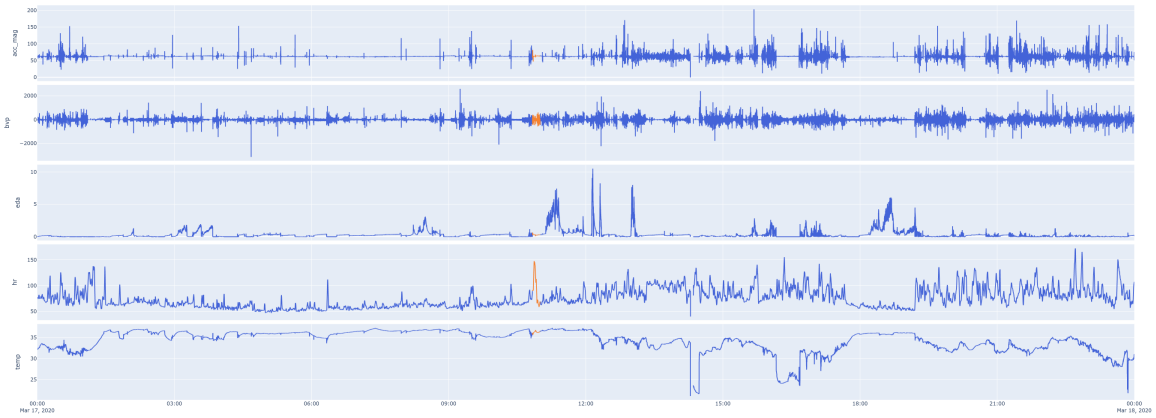

Figure 2: Raw modalities of patient 1110 in a whole day, where The red lines indicate seizure segments. These plots were developed using Plotly

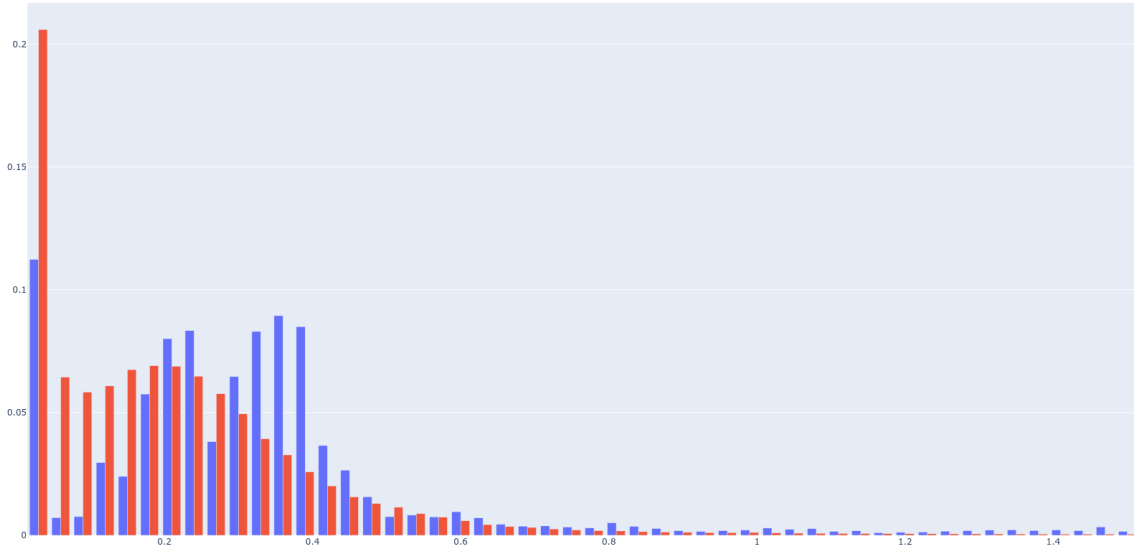

Figure 3: Normalized Histograms of EDA modality for patient 1110, for Seizure (Red bars) and Non-Seizure (Blue bars) segments

Upon visual inspection of the modalities, it is evident that some data segments are invalid. This is mainly because the patient sometimes removes the wristband during or after a seizure (See Figure 8 in Appendix B). To address this issue, [1] developed a signal quality algorithm that produces five additional signals to determine whether or not the wristband is on the patient's body and if the collected data is valid.

3.3 Annotating Seizures Onset

Since the original dataset was collected for prediction purposes, the exact time of the seizures is not marked for each seizure, and it is limited to a 10-minute interval. A heuristic algorithm was created to annotate the Onset in a seizure segment to address this issue. A seizure segment is the second segment after the last preictal segment, as determined by the original dataset labels. The heuristic algorithm determines the Onset by applying a threshold to a final score, a weighted combination of the moving standard deviation of ACC_MAG, the absolute value of the gradient of the moving average of EDA, and the gradient of heart rate. These weights are patient-specific and are determined manually. Figure 4 shows some examples of the Onset annotated by the heuristic algorithm. For more details, refer to the "find_onset()" method.

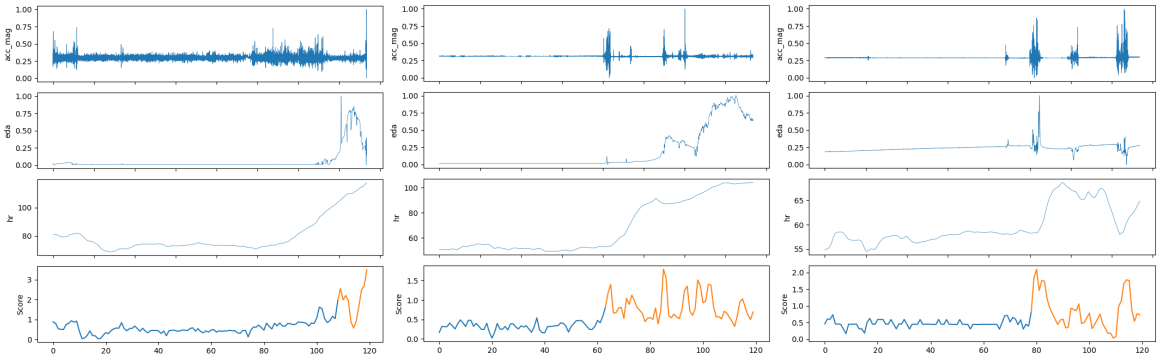

Figure 4: Some examples of the Onset found by the heuristic algorithm. Raw modalities of ACC_MAG, EDA, Heart Rate, and the final score are plotted in each row, and The red lines in the final score subplot are annotated as ictal.

It is important to note that some segments, especially seizure segments, are invalid. In annotating the seizure Onset, these invalid seizure segments (determined by the signal quality control algorithm [1] and visual inspection) are removed from the entire pipeline and not used in the training and testing processes (Figure 8 in Appendix B shows an example of an invalid seizure segment). Patient 1965 had 134 seizures, but many of them were invalid. Additionally, when selecting weights for this patient in the heuristic algorithm, no specific seizure patterns were found, and the ictals were usually visually ambiguous. As a result, this patient was not used or evaluated in the rest of the pipeline.

3.4 Selecting Training Data

Traditional classifiers such as SVM, Random Forest, and Decision Tree are not designed to handle temporal data. Therefore, when working with temporal data, it is necessary to segment the data into Epochs and feed each Epoch individually to a classifier. In our case, each seizure is a segment of the dataset. Since there are only a few seizures for each patient (except patient 1965), using all dataset segments for training is impossible. Using all segments creates an unbalanced dataset with an uneven distribution of positive (ictal) and negative (non-ictal) epochs. Only 30 valid nighttime segments, 30 valid daytime segments, all valid seizure segments, and 4 pre-ictal segments for each seizure were selected for training to address this issue. For more information, refer to the "generate_meta_data.py" file.

3.5 Feature Extraction

A list of features extracted from each modality and the candidate window sizes are shown in Table 2. Various configurations of window size, the amount of overlapping, and three Epoch lengths (60s, 120s, 240s) were tested to find their effectiveness and optimal value. (see "generate_clf_data.py" for more details).

Modality	Features	Window Sizes
Acc_mag	Standard deviation, Gradient, Kurtosis, Spectral Entropy, Power, Zero crossing rate (X, Y, Z)	2, 5*, 10*
EDA Tonic	Mean, Gradient, Kurtosis, Spectral Entropy, Power	10*, 20*
EDA phasic	Standard deviation, Gradient, Kurtosis, Spectral Entropy, Power	2, 5*, 10*
Heart Rate	Mean, Gradient, Kurtosis, Spectral Entropy, Power	10*, 20*
BVP	Standard deviation, Gradient, Kurtosis, Spectral Entropy, Power	2, 5*, 10*
Temperature	Mean, Gradient, Kurtosis, Spectral Entropy	10*, 20*

Table 2: Configurations tested for different Epoch lengths to find optimal features and window sizes.

When extracting features from each modality, using a variable window size is desirable. For instance, a modality such as temperature changes slowly, and thus, it is best to use bigger window sizes. On the other hand, a modality like the accelerometer changes at a high frequency and requires smaller window sizes to capture meaningful information. This variable length of window size for each modality can result in an unequal number of samples generated from each feature. For instance, consider Epochs of 120-second and 10-second windows with no overlap to extract the mean of the EDA signal and 30-second windows to extract the mean of the Temperature signal; this generates 12 samples from the accelerometer and 4 samples from the Temperature, which is a mismatch and needs to be addressed. To manage this mismatch, values of samples from bigger windows are copied into samples of smaller windows, such that in the previous example, we would have 12 two-dimensional samples from each Epoch. One dimension is the mean of EDA, which is different for each sample, and the other is the mean of Temperature, which is repeated for every three samples (see the "resample_datapointbased_test()" function for more details).

Boxplots of features for categories ictal, nighttime, daytime, and pre-ictal were calculated to Anlyse extracted features, and some are shown in Figure 5.

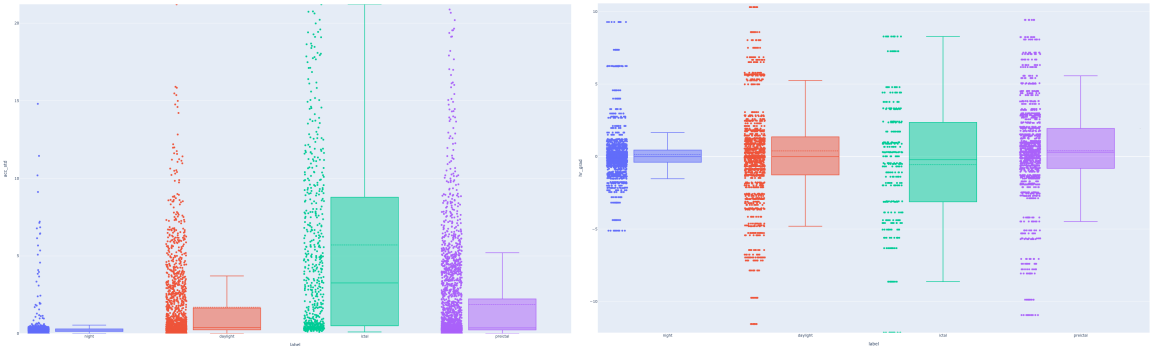

Figure 5: Boxplot of ACC_MAG s.t.d for night, daylight, ictal, and pre-ictal is shown on the left. Boxplot of the Heart Rate gradient is also shown on the right (both plots are for patient 2002).

3.6 Selecting the best window size and best features

A test bed was created to determine the optimal feature configuration and epoch length. The test was conducted for each patient's seizure separately, in a leave-one-seizure-out manner[1]. A random forest classifier was used since it can handle imbalanced data[2]. The best hyperparameters of random forest for each configuration were selected based on the recall of the test ictals.

Epoch lengths of 120 and the best window sizes marked by * in Table 2 were selected as the best configurations based on the mean recall of the classifier. The ROC curve of the results of test data of different configurations for each patient is shown in Figure 6. Performance on 4 out of 5 patients is above chance consistently along all configurations.

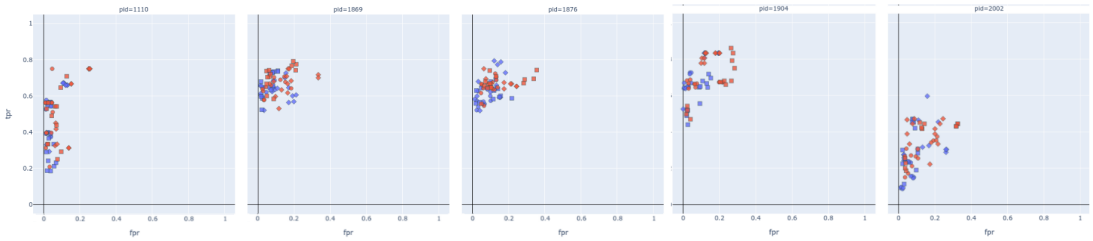

Figure 6: ROC curve of the results of test data of different configurations for patients 1110, 1869, 1876, 1904, and 2002 from left to right.

3.7 Train and test classifier

Random Forest classifiers were developed for each patient seizure using a leave-one-seizure-out approach for the final model. The best feature and epoch-length configurations from section 3.6 were utilized in this process. In order to test the classifier for each seizure, all segments of the seizure day were tested to report the sensitivity and the number of false alarms. A total of 720 epochs were generated from the whole day using 120-second epochs, and a post-processing algorithm was used to refine these results.

The post-processing algorithm works by using a moving window of 120 seconds. If the number of samples classified as ictal exceeds a certain threshold (threshold is optimized as a hyperparameter) within a window, the entire window is classified as a seizure. After each seizure prediction, the classifier's outputs are ignored for 10 minutes. More information on post-processing can be found in the "post_processing()" function.

Various feature configurations, thresholds, and random forest hyperparameters were tested, and the best result for each patient is listed in Table 3.

	Threshold	Window Sizes	Sensitivity	Mean FAR (d^{-1})
1110	0.83	10s, 20s, 20s	**75%**	**25.0**
1869	0.66	10s, 10s, 10s	80%	40.7
1876	0.5	5s, 10s, 10s	66%	50.8
1904	0.83	10s, 20s, 20s	66%	25.0
2002	0.5	5s, 10s, 20s	50%	37.8

Table 3: Final results and configuration for each patient. Window sizes from left to right correspond to [ACC, EDA Phasic, BVP], [EDA Tonic], [Temperature, Heart Rate].

After averaging the sensitivity and FAR of the patients, the model achieved a final performance of sensitivity of 74% and a FAR of 35.8 d^{-1}. Note that a random classifier would have a sensitivity

[1] in leave-one-seizure-out validation a seizure segment and its pre-ictal segments were kept as test data, and the classifier was trained using the nighttime and daytime segments along with other seizures of the patient (and their preictals).

[2] see `$class\_weight$` parameter for `$RandomForestClassifier$` in sklearn documents.

of 50% and FAR of 144 d^{-1} using the same post-processing algorithm. Figure 7 compares the performance of our model to other studies conducted on a mixture of seizures, reported by [3].

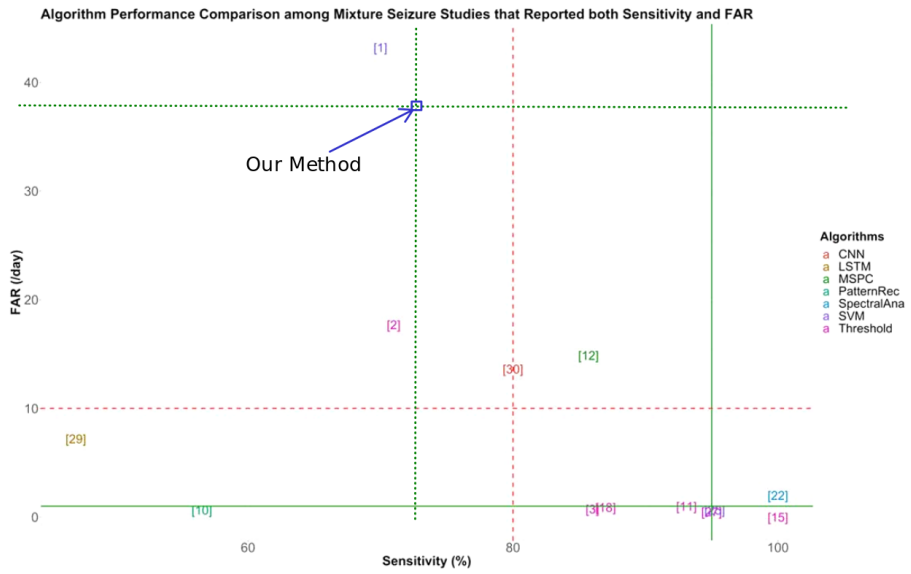

Figure 7: Performance Comparison of different studies of mixture seizure detection reported by [3], and the performance of the proposed method is outlined by the blue arrow.

A Seizure Classification

Seizures are first classified by their onset. Using seizure onset (how and where in the brain the seizure started), seizures can be divided into Focal onset Seizures, where the seizure starts on one side of the brain, generalized onset seizures, where the seizure affects both hemispheres of the brain and Unknown onset seizures, where the beginning of seizure is not entirely known. each type of seizure can be further classified by looking into additional features that occurred with the seizure:

- **Physical Component** For focal, generalized, and unknown onset seizures, are there **motor** behaviors (i.e., was there movement involved at the start of the seizure) or **nonmotor** characteristics (i.e., sensory experiences, such as smelling or tasting something)
- **Awareness** For focal seizures, is awareness impaired (i.e., a loss of consciousness) at any time during the seizure (*focal aware* and *focal unaware*)
- Does the seizure originate as a focal seizure but then spread to both hemispheres of the brain (formerly known as secondarily generalized)

PSYCHOGENIC NON-EPILEPTIC SEIZURES

Psychogenic non-epileptic seizures (PNES) happen when a person experiences symptoms similar to those of epilepsy but without any abnormal activity in the brain. The causes of non-epileptic seizures are not fully understood, but non-epileptic seizures may be associated with anxiety, depression, and personality disorders. There is also a link between post-traumatic stress disorder (PTSD) and non-epileptic seizures.

B Example of invalid seizure segment

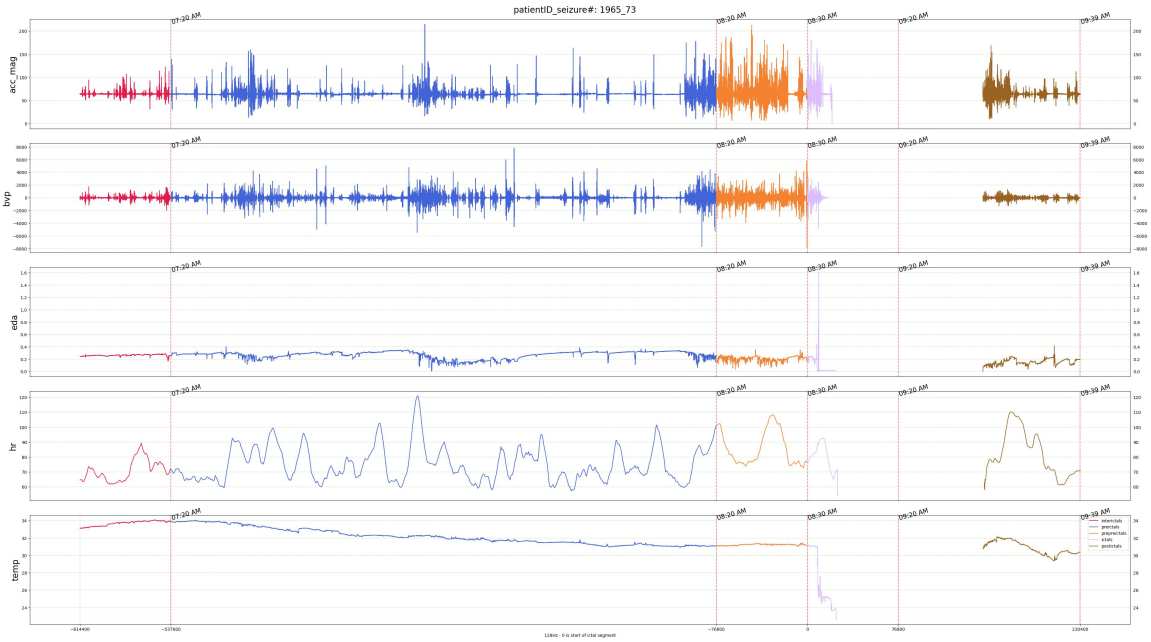

Figure 8: An example of an invalid seizure segment, where the patient has taken off the wristband after seizure (The Cyan lines)

References

[1] Sebastian Böttcher et al. "Data quality evaluation in wearable monitoring". In: *Scientific Reports* 12 (Dec. 2022), p. 21412. DOI: 10.1038/s41598-022-25949-x.

[2] Sebastian Böttcher et al. "Detecting Tonic-Clonic Seizures in Multimodal Biosignal Data From Wearables: Methodology Design and Validation (Preprint)". In: (Feb. 2021). DOI: 10.2196/preprints.27674.

[3] Fangyi Chen et al. "Seizures detection using multimodal signals: a scoping review". In: *Physiological Measurement* 43.7 (July 2022), 07TR01. DOI: 10.1088/1361-6579/ac7a8d. URL: https://dx.doi.org/10.1088/1361-6579/ac7a8d.

[4] *Epilepsy, World health Organization*. https://www.who.int/news-room/fact-sheets/detail/epilepsy. Accessed: 2010-09-30.

[5] Shitanshu Kusmakar et al. "Automated Detection of Convulsive Seizures Using a Wearable Accelerometer Device". In: *IEEE Transactions on Biomedical Engineering* 66.2 (2019), pp. 421–432. DOI: 10.1109/TBME.2018.2845865.

[6] Mona et al. Nasseri. "Ambulatory seizure forecasting with a wrist-worn device using long-short term memory deep learning." In: (2021).