Epileptic Seizure Classification Using Combined Labels and a Genetic Algorithm

Scot Davidson*‡, Niamh McCallan*, Kok Yew Ng*†, Pardis Biglarbeigi*, Dewar Finlay*,
Boon Leong Lan†, and James McLaughlin*

*NIBEC, Ulster University, Jordanstown Campus, Shore Road, Newtownabbey BT37 0QB, UK.
†Department of Electrical and Computer Systems Engineering and
Advanced Engineering Platform, School of Engineering, Monash University, Bandar Sunway, Malaysia.
‡Corresponding author. Email: davidson-s18@ulster.ac.uk

Abstract—**Epilepsy affects 50 million people worldwide and is one of the most common serious neurological disorders. Seizure detection and classification is a valuable tool for diagnosing and maintaining the condition. An automated classification algorithm will allow for accurate diagnosis. Utilising the Temple University Hospital (TUH) Seizure Corpus, six seizure types are compared; absence, complex partial, myoclonic, simple partial, tonic and tonic-clonic models. This study proposes a method that utilises unique features with a novel parallel classifier — Parallel Genetic Naive Bayes (NB) Seizure Classifier (PGNBSC). The PGNBSC algorithm searches through the features and by reclassifying the data each time, the algorithm will create a matrix for optimum search criteria. Ictal states from the EEGs are segmented into 1.8 s windows, where the epochs are then further decomposed into 13 different features from the first intrinsic mode function (IMF). The features are compared using an original NB classifier in the first model. This is improved upon in a second model by using a genetic algorithm (Binary Grey Wolf Optimisation, Option 1) with a NB classifier. The third model uses a combination of the simple partial and complex partial seizures to provide the highest classification accuracy for each of the six seizures amongst the three models (20%, 53%, and 85% for first, second, and third model, respectively).**

Index Terms—**Electroencephalography (EEG), Epileptic Seizure, Classification, Naive Bayes, Genetic Algorithm**

I. INTRODUCTION

Epilepsy is one of the most common neurological disorders in the world [1], affecting about 50 million people worldwide [2]. Epileptic seizures occur when millions of neurons are synchronously excited, resulting in a wave of electrical activity in the cerebral cortex [3]. Electroencephalography (EEG) is a noninvasive tool that measures cortical activity with millisecond temporal resolution. EEGs record the electrical potentials generated by the cerebral cortex nerve cells [4]. As a result, this tool is commonly used for the analysis and detection of seizures [5]. Epilepsy causes many difficulties in relation to the quality of life for the patient. The International League Against Epilepsy (ILAE) have outlined a list of identified seizures [6]. There are two responses on the patient, categorised as *motor* responses, where the seizures cause involuntary spasms, or *non-motor* responses, which affect the consciousness of the patient hence reducing their ability to pay attention.

Kukker et al. [7] used a fuzzy Q-learning genetic classifier to improve their original features. The signal is converted using empirical mode decomposition (EMD) where the intrinsic mode functions (IMFs) are converted using the Hilbert-Huang Transform (HHT). Nineteen features are extracted from the conversion with a distinct set of annotations created for each of the IMFs generated. The CHB-MIT dataset was used and a classification accuracy of 96.79 % was achieved. Ammar et al. [8] used a particle swarm genetic algorithm for a patient specific seizure classification system. This classification model used seizure and nonseizure as the annotations. The features extracted where kurtosis, skewness, and standard deviation. A smaller dataset (Bonn) was used with a support vector machine (SVM) classifier and an accuracy of 98.89 % was achieved. These state of the art designs focused primarily on the detection of seizure and non-seizure within the CHB-MIT and Bonn datasets, respectively. However, the classification of different seizure types was not investigated due to insufficient labelling as well as having limited patient-specific data found in these datasets.

This state-of-the-art work proposes a seizure classification system using multiclass parallel classifiers — Parallel Genetic Naive Bayes (NB) Seizure Classifier (PGNBSC) — to classify focal onset seizures against

other seizure types. As simple partial and complex partial seizures share morphology and exhibit similar responses on the patient, where the main factor being whether the patient remains conscious during the event, the proposed algorithm can discriminate these two seizure types against other seizures by combining them into a focal seizure. In addition, this algorithm is also tested against the Temple University Hospital (TUH) Seizure Corpus due to its extensive labelling of multiple seizure types as well as more patient-specific data compared to the CHB-MIT and Bonn datasets.

This manuscript is organised as follows: Section II explains the dataset used in this research; Section III describes the proposed PGNBSC classifier algorithm; Section IV presents and discusses the results obtained; and Section V provides the conclusions.

II. Dataset

The dataset from the TUH Seizure Corpus v1.5.2 (27 May 2020) was used in this study. This dataset is a subset of a much larger EEG corpus [9], [10]. This dataset is divided by the TUH into two sets, defined as training and testing with demographics being made equal across both sets. The dataset was annotated by a team established by the university based on the signals and the neurologists' reports. It is annotated by giving the seizure start and end time in a text file. This is converted to a sample number and each window of the seizure is extracted between the start and the end samples. The configurations for the training and testing are shown in Table I. The TUH includes only two pieces of information on the individual patients, which are the age and the gender of each patient. These demographics are negated because age is only a contributing factor in the absence seizure, which already has a unique waveform [6].

III. Feature Extraction and Selection

A. Preprocessing of Raw EEGs

To maintain consistency in the dataset, only the 19 channels common to all seizures defined by the International 10–20 System were used [11]. By doing so, the unwanted channels such as the electrocardiogram (ECG), electromyography (EMG), as well as photo stimulus channels were omitted. The signals were resampled to 250 Hz using an antialiasing low-pass finite impulse response filter as this is the lowest sampling frequency common to all of the seizures. The 60 Hz line noise was removed with a bandstop infinite impulse response (IIR) filter. The EMD was then performed and the first IMF was taken for the analysis of the signals [7]. The signals

were then divided into 1.8 s windows, which was chosen because it reflects on the shortest ictal window of all the different seizure types. If a seizure contains only one window, then the window is taken from the centre of the seizure. Each feature was calculated for each channel of the windowed signals.

B. Time Domain Features

1) Standard Deviation: Ictal episodes have a higher energy compared to the non-ictal episodes. However, some seizures have a lower energy response such as absence, complex partial, and simple partial [12]–[14]. The standard deviation σ measures how much energy is created away from the mean, where a higher energy level would naturally produce a higher σ.

2) Shannon Entropy: The Shannon entropy method is used to measure the chaotic nature of EEGs.

3) Kurtosis: Kurtosis γ_2 is the degree of "peakedness" of a real valued random variable, which is used to determine if σ is created from small constant deviations or if large inconsistent deviations are present in the signal [15].

4) Hjorth Criteria: Hjorth parameters measures a signal via three criteria: activity, mobility, and complexity. Activity uses variance as its basis and is therefore not used as a feature. However, mobility h_m and complexity h_c can be used to measure the mean frequency of a signal and deviation from a pure sine wave, respectively [16].

5) Skewness: Skewness measures the symmetry of the time series data.

6) Energy and Nonlinear Energy: The total energy of the time series data is used because different seizures can have a muscle artefact which typically has a higher energy content. The mean nonlinear energy measures more than the total energy by measuring the fluctuations in energy from the time series data.

C. Fractal Analysis

Fractals are used to measure the self-similarity of the given time series values. Absence, tonic, and tonic-clonic seizures have rhythms generating self similar waveforms [14], [17].

1) Higuchi Fractal Dimension: The time series data have to be firstly decomposed into a set of subseries [18]. A curve is then created to measure the similarity of the subseries. The length of the curve can be computed using

$$L_m = \frac{1}{k}\left[\left(\sum_{i=1}^{r}|x(m+ik)-x(m+(i-1)k)|\right)\frac{N-1}{rk}\right], \quad (1)$$

where $r = \left\lceil\frac{N-m}{k}\right\rceil$ and $x(i)$ is the subseries.

Seizure Type	Training	Testing	wTrain	wTest	Duration_Tr (s)	Duration_Tt (s)	wDuration_Tr (s)	wDuration_Tt (s)
Absence	13	7	253	173	495	357	455	311
Complex partial	134	35	15,373	4629	27,922	840	27,671	1512
Myoclonic	2	1	717	10	1293	19	1290	18
Simple Partial	5	3	1055	112	1941	204	1899	202
Tonic	19	9	243	396	455	749	437	819
Tonic-Clonic	17	11	1080	1981	1967	3582	1944	3566
TOTAL	**190**	**66**	**18,721**	**7301**	**34,073**	**5751**	**33696**	**6428**

*wTrain: The amount of labels in training that are equal to the corresponding seizure.
*wTest: The amount of labels in testing that are equal to the corresponding seizure.
*wDuration: The windowed duration of training (Tr) and testing (Tt) sets.

2) Katz Fractal Dimension: Katz utilises the same subseries defined above. It is another fractal estimation algorithm where the successive points are measured to compute the self-similarity such that the fractal dimension of the subseries is used as another feature.

3) Spectral Entropy: To calculate the spectral entropy of a time series data, it is initially converted to the frequency domain using fast Fourier transform, which is then further converted into power. The mean, maximum, and minimum are then taken as the features to be fed into the PGNBSC. To balance the dataset, an upsampling method is used where the least represented labels are repeated by the nearest integer factor of the largest label. In the first model, the largest label is the complex partial seizure. In the second model, the largest is the complex partial seizure. In the third model, the largest label is the focal-onset.

D. Genetic Algorithms

Genetic algorithms are optimisation methods that mimic Darwin theory. They are exploratory procedures that seek to find the near optimal solutions to complex problems [19]. Emary et al. [20] adapted a previous genetic algorithm called Binary Grey Wolf Optimisation (BGWO) to be applied for feature selection. The proposed work in this research uses the BGWO Algorithm 1 from Emary et al. [20] because it provides the shortest compilation time and is shown in Algorithm 1.

To use a genetic algorithm, the key feature is the fitness function. The fitness function is initialised using a randomiser. This selects genes for the starting lineage of the subsequent pairs, which are defined as W_i in Algorithm 1. As a fail-safe, should the algorithms selects no features, then the cost is set to ∞. The same classifier defined in Section III-E is used to evaluate the features

and the F1-Score is used for the comparison performance of the selected features. As parallel classifiers are used, the F1-score can be used without having to use the global F1-score formula. The F1-score is perfect if the score is 1 however genetic algorithms are perfect if they score zero. Therefore, we use $1 - F1$ to manipulate the F1-score towards 0. To ensure the gene-line and the subsequent gene-lines are improving, a bias is created using the the iteration number. As the algorithm is improving the bias becomes higher meaning the gene-lines must be improving or the new gene-line is entered into the gene-pool.

An early stopping criterion is also set in the algorithm that measures the six previous values. This is measured by comparing the deviation in the entire signal with the deviation in the final six values. If the deviation is <

Algorithm 1 Binary Grey Wolf Optimisation

I iterations
X_i Population Initalisation
W_1 = first search agent
W_2 = second search agent
W_3 = third search agent
$t_i = 0$
    for $i = 1 : 3$ **do**
        $W_i \leftarrow$ New Global position
    end for
    while $t_i < I$ **do**
        Apply fitness function to W_1, W_2, W_3
        Evaluate W_1, W_2, W_3
        Compare W_1, W_2, W_3
        $t_i = t_i + 1$
    end while

5% of the total deviation, then the stopping criterion is finished and the features are extracted.

E. Classification

The classifier chosen is the Naive Bayes classifier. It is a probabilistic classifier based on the Bayes theorem under the assumption that any feature of a particular class is independent of any other feature. Error estimation is computed based on the maximum likelihood [21]. The prior probabilities are set to uniform to not give any any class advantage especially when using *one vs all*. The kernel used is a Gaussian distribution and is set to be unbounded. To compare the features and their ability to detect the individual seizures, a parallel classification system is needed. Each of the classifiers is trained in a *one vs all* set up. To remove the need for six different confusion matrices, a heatmap was generated for each of the classifiers with the False Positive (FP) and the False Negative (FN) placed at the bottom of the heatmap. The TP value is placed in the predicted and true label of the heatmap the same as a confusion matrix. This makes comparison between each of the individual classifiers easier. To evaluate the performance of the genetic algorithm on each iteration a NB classifier is used. The NB classifier is faster at classification in comparison to the other machine learning algorithms. When the data is being iterated many times over the same dataset processing speed is critical.

IV. RESULTS AND DISCUSSION

The performance metrics used in this state of the art work are shown below

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \qquad (2)$$

$$F1 = \frac{TP}{TP + \frac{FP + FN}{2}}, \qquad (3)$$

where TP, TN, FP, and FN are the true positive, true negative, false positive, and false negative, respectively. The first model chosen is the baseline model. This model does not use the genetic algorithm and does not combine the complex partial seizure and simple partial seizure into one label. This model is a single multiclass classifier. The performance of this model is shown in Figure 1, where there is a high misclassification with the complex partial seizure being classified as simple partial seizures. It is unexpected that the simple partial and the complex partial are being misdiagnosed as a tonic-clonic seizure. A tonic-clonic seizure can have a focal onset, which can propagate into a bilateral tonic-clonic seizure [6].

The second model uses the genetic algorithm to find the features that provide the most optimum F1-score using the labels in the TUH Seizure Corpus. This is a parallel classifier and contains six NB classifiers. Figure 2 shows the performance of this model. The tonic seizure had the highest F1-score with 0.96 followed by the absence seizure with 0.93. The tonic-clonic and simple partial seizure performed poorly by having higher FPs and FNs than TPs in the system. The heatmap does highlight a large reduction in FPs. There is also an increase in the classification of myoclonic seizures by 400% compared to Figure 1.

The third model uses the genetic algorithm to find the features that provide the most optimum F1-score with the combined label of both complex partial and simple partial seizures. This is a parallel classifier and contains five NB classifiers. The performance of the third model is shown in Figure 3. Combining the complex partial seizure and the simple partial seizure label has increased the accuracy of the entire system; tonic-clonic seizures had an increase in TPs by 100% and FPs decreased by 391% compared to Figure 2. However, FNs of tonic-clonic seizures had increased by 13%. Also, combining the labels for complex partial and simple partial seizures has had a positive effect on their FNs and FPs, with a combined reduction of 88%. In Figure 2, the combined TP value for complex partial and simple partial seizures is 4667. However, in Figure 3 the combined value is 4654, which is only a decrease of $< 1\%$.

In comparison with Kukker et al. [7] the models did not achieve the same accuracy of 96.79%. However, the proposed models have improved classification performance by using smaller time windows in the design and the multiclass classification system adopted.

Depending on the chosen classifier, an implementation of a genetic algorithm should be considered for two reasons; firstly to increase the accuracy of classification, and secondly to reduce the size of the feature array.

However, using parallel classifiers for multiclass seizure classification increases the storage space required for the final implementation in comparison with having a single classifier performing multiclass detection. The processing power is also increased because of the multiple iterations performed. Having said that, this increase in resources requirements is compensated by a higher classification accuracy and an overall more robust classification system. The combination of the focal onset seizures is required to increase the systems accuracy with no real-life effects on the likelihood of the patient.
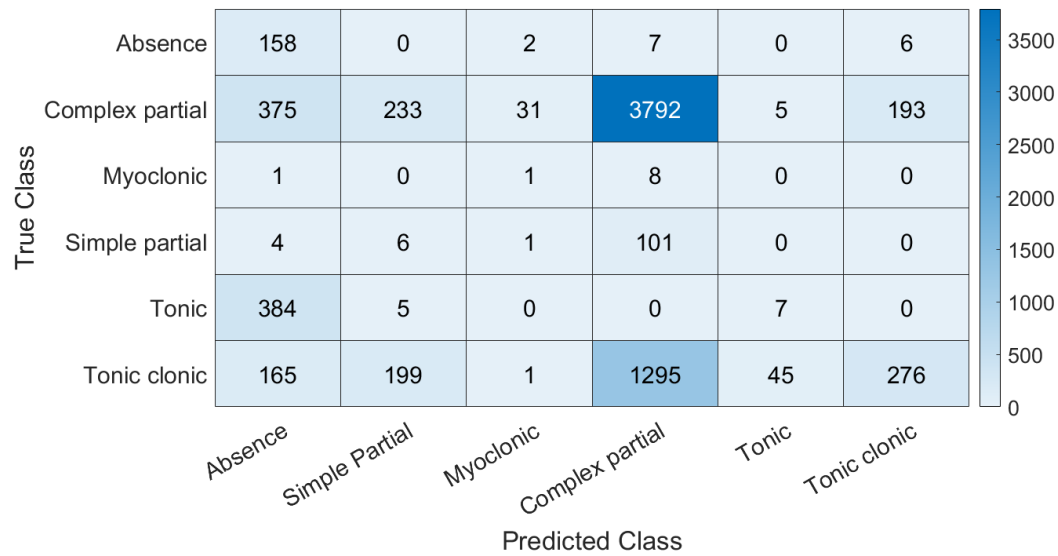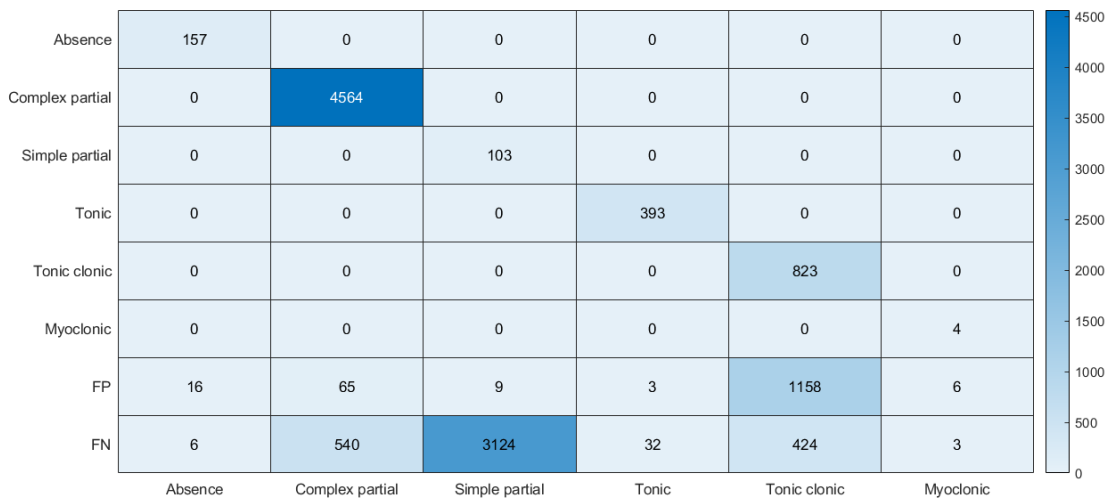
Fig. 1. Confusion matrix of the first model.

Fig. 2. A heatmap describing the performance of the second model.

V. CONCLUSION

The implementation of a genetic algorithm can be crucial when building a machine learning platform to increase classification of seizure types. It requires more training time but further investigations can highlight which features are comparative. The value of spending increased time at the training stage will ensure that the model increases in accuracy along with combining focal onset seizures to further increase the classification rate.

To further improve the models accuracy, a new set of features that are more tailored to each seizure type can be investigated. More studies can also be carried out to balance the classification performance and the resources requirements factor.

REFERENCES

[1] M.J. Brodie, S.C. Schachter, and P. Kwan. *Fast Facts: Epilepsy*. Karger Medical and Scientific Publishers, 2012.
[2] World Health Organisation. Key facts: epilepsy, 2019. accessed on: 11/06/2021.
[3] A.T. Tzallas, M.G. Tsipouras, and D.I. Fotiadis. Epileptic Seizure Detection in EEGs Using Time–Frequency Analysis.
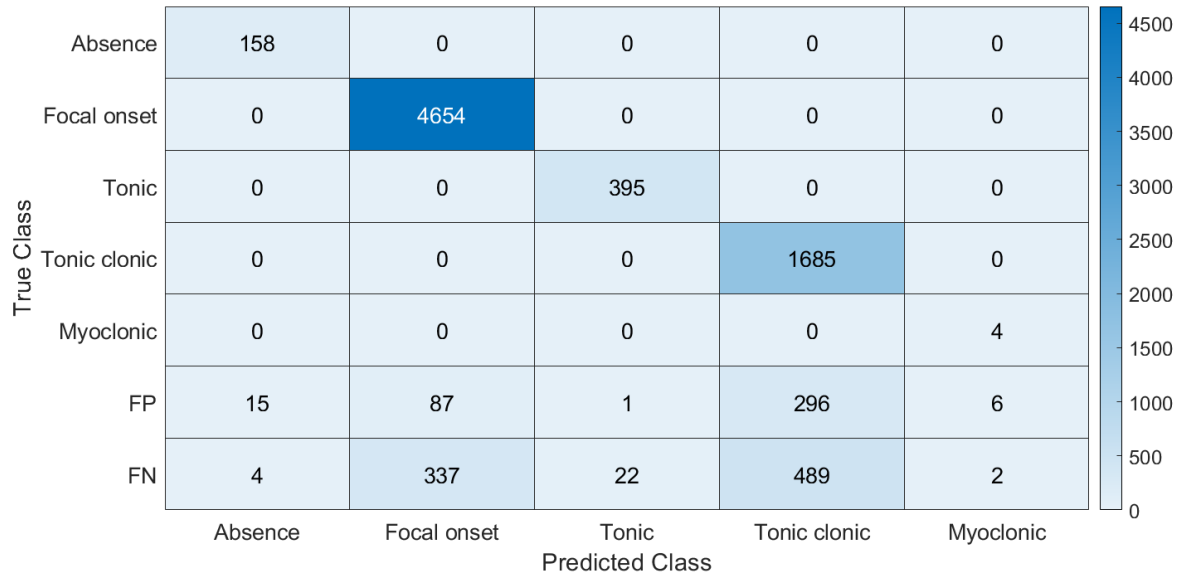
Fig. 3. Heatmap describing the proposed PGNBSC classifier with the simple partial and the complex partial seizures combined as focal-onset.

IEEE Transactions on Information Technology in Biomedicine, 13(5):703–710, 2009.

[4] A. Subasi and E. Erçelebi. Classification of EEG Signals Using Neural Network and Logistic Regression. *Computer Methods and Programs in Biomedicine*, 78(2):87–99, 2005.

[5] H. Adeli, Z. Zhou, and N. Dadmehr. Analysis of EEG Records in an Epileptic Patient Using Wavelet Transform. *Journal of Neuroscience Methods*, 123(1):69–87, 2003.

[6] R.S. Fisher, J.H. Cross, J.A. French, N. Higurashi, E. Hirsch, F.E. Jansen, L. Lagae, S.L. Moshé, J. Peltola, E. Roulet Perez, I.E. Scheffer, and S.M. Zuberi. Operational Classification of Seizure Types by the International League Against Epilepsy: Position Paper of the ILAE Commission for Classification and Terminology. *Epilepsia*, 58(4):522–530, 2017.

[7] A. Kukker and R. Sharma. A Genetic Algorithm Assisted Fuzzy Q-Learning Epileptic Seizure Classifier. *Computers & Electrical Engineering*, 92:107154, 2021.

[8] S. Ammar, O. Trigui, and M. Senouci. Genetic and Practical Swarm Optimisation Algorithms for Patient-Specific Seizure Detection Systems. In *2018 4th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, pages 1–4, 2018.

[9] V. Shah, E. von Weltin, S. Lopez, J.R. McHugh, L. Veloso, M. Golmohammadi, I. Obeid, and J. Picone. The Temple University Hospital Seizure Detection Corpus. *Frontiers in Neuroinformatics*, 12:83, 2018.

[10] A. Harati, S. López, I. Obeid, J. Picone, M.P. Jacobson, and S. Tobochnik. The TUH EEG CORPUS: A Big Data Resource for Automated EEG Interpretation. In *2014 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, pages 1–5, 2014.

[11] R. Oostenveld and P. Praamstra. The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology*, 112(4):713–719, 2001.

[12] B. Weir. The Morphology of the Spike-wave Complex. *Electroencephalography and Clinical Neurophysiology*, 19(3):284–290, 1965.

[13] O. Devinsky, K. Kelley, R.J. Porter, and W.H. Theodore. Clinical and Electroencephalographic Features of Simple Partial Seizures. *Neurology*, 38(9):1347–1347, 1988.

[14] F. Gilliam, B.J. Steinhoff, H.-J. Bittermann, R. Kuzniecky, E. Faught, and B. Abou–Khalil. Adult Myoclonic Epilepsy: A Distinct Syndrome of Idiopathic Generalized Epilepsy. *Neurology*, 55(7):1030–1033, 2000.

[15] L.R. Quitadamo, R. Mai, F. Gozzo, V. Pelliccia, F. Cardinale, and S. Seri. Kurtosis-based Detection of Intracranial High-frequency Oscillations for the Identification of the Seizure Onset Zone. *International Journal of Neural Systems*, 28(07):1850001, 2018.

[16] Y. Höller, E. Trinka, G. Kalss, G. Schiepek, and R. Michaelis. Correlation of EEG Spectra, Connectivity, and Information Theoretical Biomarkers with Psychological States in the Epilepsy Monitoring Unit — A pilot study. *Epilepsy & Behavior*, 99:106485, 2019.

[17] K. Kobayashi, T. Inoue, Y. Watanabe, M. Oka, F. Endoh, H. Yoshinaga, and Y. Ohtsuka. Spectral Analysis of EEG Gamma Rhythms Associated with Tonic Seizures in Lennox–Gastaut Syndrome. *Epilepsy Research*, 86(1):15–22, 2009.

[18] T.Q.D. Khoa, V.Q. Ha, and V.V. Toi. Higuchi Fractal Properties of Onset Epilepsy Electroencephalogram. *Computational and Mathematical Methods in Medicine*, page 461426, 2012.

[19] J.F. Frenzel. Genetic algorithms. *IEEE Potentials*, 12(3):21–24, 1993.

[20] E. Emary, H.M. Zawbaa, and A.E. Hassanien. Binary Grey Wolf Optimization Approaches for Feature Selection. *Neurocomputing*, 172:371–381, 2016.

[21] A.H. Fielding. *Cluster and Classification Techniques for the Biosciences*. Cambridge University Press, 2006.