

Seizure Observation & Preemptive Detection with ML: Comparing EEG-Based Supervised Learning Models to Detect and Predict Seizures

Nikhil Krishnaswamy & Vishresh Deepak

Received March 10, 2024

Accepted April 10, 2024

Electronic access April 30, 2024

Over 50 million people worldwide suffer from epilepsy, an incurable and life-changing condition. Epilepsy is when excessive and abnormal brain cell activity occurs, resulting in a variety of symptoms, including unconsciousness and irregular behaviors. Epilepsy has a broad range of symptoms and severities, and the unpredictability of seizures makes it a tough condition to analyze. This study aims to utilize and compare supervised machine learning models for seizure detection and prediction algorithms to address the unpredictability and severity of seizures. By leveraging advanced computational analysis techniques and brain-wave data based on EEG readings, these models can potentially transform the landscape of epilepsy care, offering personalized and preemptive care to improve patient outcomes. Using real patient data, the research consists of processing data, identifying the appropriate models to be used, and training various models to analyze and detect seizures. The results of this paper provide insight into what types of machine learning models are best for EEG-based seizure detection and prediction. While each model has its benefits and drawbacks, the deciding factor is the accuracy of each model. After testing the models on seizure prediction and detection datasets, the overall best-performing training method was the Random Forest Classifier, with 98.59% accuracy in seizure detection and 90.8% accuracy for seizure prediction. These high accuracy rates show promising potential that this technology can transform epileptic care for over 50 million individuals who are suffering worldwide.

Introduction

Epilepsy is a significant neurological disorder affecting approximately 50 million people globally, characterized by its unpredictable seizures that can profoundly impact the quality of life of those affected (Ogoke et al., 2021)¹. The development of detective and predictive models based on electroencephalogram (EEG) data holds promise for mitigating the unpredictability of seizures and reducing the reliance on medications that often have potential side effects (Qi et al., 2022)². This paper aims to explore and compare different types of machine learning (ML) models that can predict an epileptic seizure before it occurs and detect one if and when it does happen. The models include Logistic Regression, Support Vector Machine (SVM), k-Nearest Neighbors (KNN), Random Forest Classification, Gradient Boosting, Gaussian Naive Bayes, Artificial Neural Networks (ANN), and Principal Component Analysis (PCA).

The recurrent yet unpredictable nature of seizures highlights the importance of using machine learning models for the detection and prediction of seizures (Ren et al., 2022)³. Recent advancements in AI have made it possible to apply modeling in care specifically in epilepsy research (Dang et al., 2021)⁴. These models have shown some promise in predicting seizure timing and providing insights for interventions (Goldenholz et

al., 2020)⁵. Researchers have focused on utilizing EEG data for modeling in epilepsy studies (Xu et al., 2022)⁶. Studies have demonstrated the effectiveness of features extracted from EEG signals in predicting seizures, emphasizing the potential of gradient boosting decision tree algorithms, in the field of seizure prediction (Xu et al., 2022)⁶. Moreover combining channel vision transformers has shown potential in improving seizure prediction accuracy (Hussein et al., 2022)⁷.

Furthermore, recent studies have demonstrated the feasibility of using artificial intelligence, particularly machine learning and convolutional neural networks, for the radiological identification of temporal lobe epilepsy based on magnetic resonance imaging (MRI) findings reflecting underlying mesial temporal sclerosis (Gleichgerricht et al., 2021)⁴. Additionally, the use of AI and machine learning has expanded in the field of epilepsy, showing potential for the classification of temporal lobe epilepsy with MRI data at the region of interest (ROI) level (Gleichgerricht et al., 2021)⁵.

Despite these advancements, there is a research gap in the comprehensive comparison and evaluation of various AI models for predicting epileptic seizures based on EEG data. This paper aims to explore and compare different types of machine learning (AI) models that can predict an epileptic seizure before it occurs and detect one if and when it does happen. The

models include Convolutional Neural Network (CNN), Logistic Regression, Support Vector Machine (SVM), k-Nearest Neighbors (KNN), Random Forest Classification, Gradient Boosting, Gaussian Naive Bayes, and Artificial Neural Networks (ANN). A Principal Component Analysis (PCA) algorithm was used for dimensionality reduction and tested for accuracy as well.

Dataset Description

This study consists of two parts: identifying the best type of learning model for detecting when a seizure happens and experimenting with different machine learning models to test the plausibility of predicting a seizure before it occurs. The first dataset utilized in this study is from the UCI Machine Learning Repository and involves EEG recordings to identify epileptic seizures. The initial dataset is composed of one hundred files in five folders and each file represents a separate topic or individual, with brain activity recorded for 23.6 seconds. The time-series data is segmented into 4097 samples, each sample represents the EEG recording at a given occasion. 500 individuals in this dataset provide 4097 data points over 23.6 seconds. Initial pre-processing has been performed on the dataset. The 4097 data points for each subject were separated into 23 chunks, each with 178 data points, representing a 1-second interval.

The response variable, 'y,' takes values from the set 1, 2, 3, 4, 5, where:

- 5 indicates eyes open during EEG recording.
- 4 denotes eyes closed during EEG recording.
- 3 signifies recording EEG activity from the healthy brain area while identifying the tumor's location.
- 2 represents EEG recordings from the area where the tumor was located.
- 1 corresponds to the recording of seizure activity.

It is important to note that individuals in classes 2, 3, 4, and 5 are subjects without epileptic seizures, while only subjects in class 1 have experienced epileptic seizures. The reason behind creating this modified version of the dataset was to simplify accessibility through the creation of a .csv file. Although the original dataset encompasses five classes, many researchers opt for binary classification, specifically distinguishing class 1 (epileptic seizure) from the rest. However, by training the models to identify each specific state individually, as done in this paper, the models should be better trained to make distinctions between the first class and the rest of them.

The second part of the study, prediction, uses another public dataset, but one targeted and classified for prediction of seizures before they occur, from the American Epilepsy Society Seizure Prediction Challenge.

Intracranial EEG (iEEG) data clips are systematically organized into folders, encompassing training and testing data for each human or canine subject. The training data is structured into ten-minute EEG clips labeled as "Preictal" for pre-seizure data segments and "Interictal" for non-seizure data segments. Training data segments are sequentially numbered, while testing data segments are arranged in random order. Within the folders, data segments are stored in .mat files to be converted to a .csv and labeled according to the following convention:

preictal_segment_N.mat: The Nth preictal training data segment

interictal_segment_N.mat: The Nth non-seizure training data segment

test_segment_N.mat: The Nth testing data segment

Each .mat file contains a data structure with the following fields:

data: A matrix of EEG sample values arranged as rows x columns, representing electrode x time.

data_length_sec: The time duration of each data row.

sampling_frequency: The number of data samples representing 1 second of EEG data.

channels: A list of electrode names corresponding to the rows in the data field.

sequence: The index of the data segment within the one-hour series of clips. For instance, preictal_segment_6.mat has a sequence number of 6, representing the iEEG data from 50 to 60 minutes into the preictal data.

The interictal data were randomly selected from the complete data record, with the constraint that interictal segments be as distant as practically possible from any seizure to prevent contamination with preictal or postictal signals. In the extensive canine recordings, it was feasible to maintain a restriction of one week before or after a seizure. However, in the human recordings (which may have a total duration of less than a week), interictal data were limited to more than four hours before or after any seizure.

Methodology

Data Collection and Preprocessing

The analysis focuses on epileptic seizure detection using a dataset comprising interictal and preictal EEG segments from Patient 1. The data is organized into two categories: 'Patient_1_interictal_segment' and 'Patient_1_preictal_segment.' A total of 18 instances for each category are loaded and processed.

The preprocessing involves extracting relevant information from the specified file path. To facilitate binary classification, a class mapping is performed, transforming the multi-class problem into a binary one. The EEG data features are normalized to ensure consistency and comparability across different scales.

```

ESR.head()
[6] ✓ 0.1s Python
...
  Unnamed  X1  X2  X3  X4  X5  X6  X7  X8  X9  ...  X170  X171  X172  X173  X174  X175  X176  X177  X178  y
0  X21.V1.791  135  190  229  223  192  125  55  -9  -33  ...  -17  -15  -31  -77  -103  -127  -116  -83  -51  4
1  X15.V1.924  386  382  356  331  320  315  307  272  244  ...  164  150  146  152  157  156  154  143  129  1
2  X8.V1.1  -32  -39  -47  -37  -32  -36  -57  -73  -85  ...  57  64  48  19  -12  -30  -35  -35  -36  5
3  X16.V1.60  -105  -101  -96  -92  -89  -95  -102  -100  -87  ...  -82  -81  -80  -77  -85  -77  -72  -69  -65  5
4  X20.V1.54  -9  -65  -98  -102  -78  -48  -16  0  -21  ...  4  2  -12  -32  -41  -65  -83  -89  -73  5
5 rows x 180 columns

```

Fig. 1 A preview of the dataset used for detection models

```

/kaggle/input/seizure-prediction/Patient_1/Patient_1/Patient_1_test_segment_0112.mat
/kaggle/input/seizure-prediction/Patient_1/Patient_1/Patient_1_test_segment_0030.mat
/kaggle/input/seizure-prediction/Patient_1/Patient_1/Patient_1_test_segment_0109.mat
/kaggle/input/seizure-prediction/Patient_1/Patient_1/Patient_1_test_segment_0111.mat
/kaggle/input/seizure-prediction/Patient_1/Patient_1/Patient_1_test_segment_0090.mat
/kaggle/input/seizure-prediction/Patient_1/Patient_1/Patient_1_test_segment_0191.mat

```

Fig. 2 A preview of the dataset used for prediction models

Subsequently, the dataset is split into training and testing sets (90% training, 10% testing) to facilitate model evaluation.

To address the class imbalance, random oversampling is applied, ensuring a balanced representation of both interictal and preictal segments in the training data. The dataset is shuffled to enhance the robustness of the models.

Model Selection:

A diverse set of machine learning models were chosen to explore their effectiveness in epileptic seizure detection:

- **Logistic Regression:** A baseline linear model suitable for binary classification tasks.
- **Support Vector Machines (SVM):** Known for their ability to handle complex decision boundaries.
- **k-Nearest Neighbors (KNN):** A non-parametric algorithm relying on proximity to neighbors for classification.
- **Random Forest Classification:** An ensemble learning method utilizing multiple decision trees for improved accuracy.
- **Gradient Boosting:** A boosting technique that builds models sequentially, correcting errors of the previous ones.
- **Gaussian Naive Bayes:** A probabilistic model based on Bayes' theorem, suitable for classification tasks.

- **Artificial Neural Networks (ANN):** Deep learning approach with multiple layers mimicking the human brain's neural network.

Training and Evaluation

- **Data Splitting:** The dataset was divided into training and testing sets (80% training, 20% testing) to assess model generalization.
- **Feature Scaling:** Standard scaling was applied to ensure all features have the same scale, preventing the dominance of certain features.
- **Model Training:**
 - Each selected model was trained using the training dataset.
 - Relevant hyperparameters were fine-tuned to optimize model performance.
- **Model Evaluation:**
 - The trained models were evaluated using the testing dataset.
 - Metrics such as accuracy, precision, recall, and F1-score were computed for comprehensive performance assessment.
- **Comparative Analysis:**

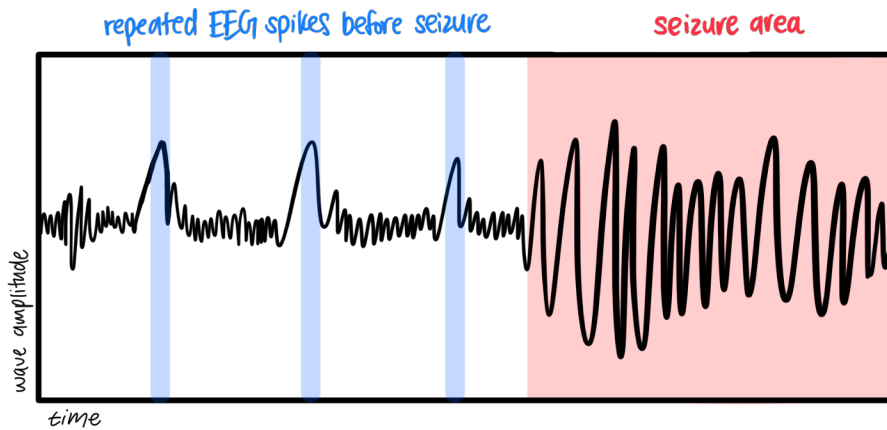


Fig. 3 By detecting the preictal spikes before the seizure, a model can predict a seizure before it occurs

```
# Example: Training Logistic Regression
from sklearn.linear_model import LogisticRegression
log_reg_model = LogisticRegression()
log_reg_model.fit(X_train, y_train)
```

Fig. 4 Code example for logistic regression model training

```
# Example: Evaluating Logistic Regression
from sklearn.metrics import accuracy_score, classification_report
y_pred_log_reg = log_reg_model.predict(X_test)
accuracy = accuracy_score(y_test, y_pred_log_reg)
classification_rep = classification_report(y_test, y_pred_log_reg)
print(f"Accuracy: {accuracy:.2f}")
print("Classification Report:\n", classification_rep)
```

Fig. 5 Example code for how accuracy is determined on test data

- The performance of each model was compared based on accuracy and other relevant metrics
- Visualizations, such as confusion matrices, were employed to facilitate comparative analysis.

```
# Example: Confusion Matrix for Logistic Regression
from sklearn.metrics import confusion_matrix
conf_matrix = confusion_matrix(y_test, y_pred_log_reg)
print("Confusion Matrix:\n", conf_matrix)
```

Fig. 6 Example code for how the confusion matrix is displayed

This comprehensive methodology ensures a systematic ex-

ploration of various machine learning models and a thorough evaluation of their effectiveness in epileptic seizure detection. The choice of diverse models and meticulous evaluation process aims to provide valuable insights into the comparative performance of each algorithm.

Machine Learning Process

The code was compiled in stages to compare the impact on accuracy at each stage. In Stage 1, data loading, preprocessing, model selection, training, and evaluation were performed. In Stage 2, additional preprocessing steps and a new set of models were introduced for a more comprehensive analysis.

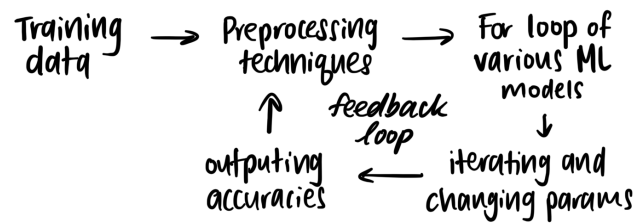


Fig. 7 Flowchart demonstrating feedback loop for training and testing data

In the initial stage of the analysis, the data undergoes loading and preprocessing, involving the extraction of information from the specified path. Class mapping is executed to transform the multi-class problem into a binary classification task (see dataset description above), streamlining the subsequent modeling process. The normalization of EEG data features ensures consistency and comparability across different scales. Following this, the dataset is partitioned into training and testing sets to facilitate model evaluation. Addressing class imbalance is

crucial, and random oversampling is applied to mitigate this issue. A diverse set of models is then selected for evaluation, encompassing various machine learning techniques. Subsequently, each chosen model is trained on the resampled training data, and predictions are made on the test set. The Random Forest Classifier undergoes hyperparameter optimization through grid search to enhance its predictive performance. The best model, identified through this process, is then trained and evaluated on the test set.

In the second stage, the analysis progresses with additional data loading and preprocessing steps. The dataset is loaded, and class mapping is applied similarly to the first stage. To simplify the dataset structure, the 'Unnamed' column is dropped. Further preprocessing involves the normalization of EEG data features and the application of the Interquartile Range (IQR) method to manage outliers. Random oversampling is employed to address class imbalance, and the dataset is once again split into training and testing sets. An expanded set of models, including Logistic Regression, Support Vector Machines, K-Nearest Neighbors, Random Forest, Gradient Boosting, Gaussian Naive Bayes, and Artificial Neural Networks, are introduced for evaluation. Each model undergoes training on the oversampled training data, and predictions are made on the test set. Additionally, Principal Component Analysis (PCA) is implemented to assess its impact on accuracy. The results are summarized in a comparison table, providing a concise overview of model names and their corresponding accuracy scores for easy interpretation and comparison.

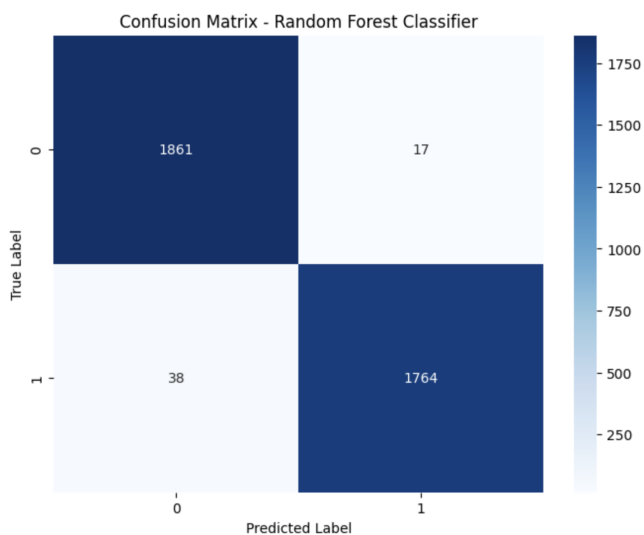


Fig. 8 Random forest Classifier Confusion Matrix that shows the true vs, predicted values of the testing data

Similar techniques are used for the prediction part of this paper, however, this part was completed in one stage and the notebook was run on Kaggle with a GPU accelerator because of

the amount of data required.

In Kaggle, a sequential model is constructed using TensorFlow and Keras. The CNN architecture comprises convolutional layers, max-pooling layers, dropout layers for regularization, and densely connected layers for classification. The model is trained on the normalized and oversampled training data, and its performance is evaluated on the test set. Additional classical machine learning models, including Logistic Regression, SVM, KNN, Random Forest, Gradient Boosting, and Naive Bayes, are trained and evaluated. These models were applied after flattening the spectrogram representations of EEG data, removing outliers, and employing other preprocessing techniques.

Evaluation and Comparative Analysis

Model evaluation is conducted using relevant metrics, such as accuracy, precision, recall, and F1-score, focusing on the CNN and classical machine learning models. The CNN's test loss and accuracy were reported. A comparison table (see table 1) is generated to summarize the accuracy of each model, facilitating an easy interpretation of their relative performances.

Model Persistence

The trained CNN model is saved with joblib while additional classical machine learning models are saved using the pickle library, providing a means for future use without retraining.

Data Visualizations

Visualizations, such as ROC curves and confusion matrices, were also used to enhance the understanding of model performance.

Results

Stage One: Baseline Model Evaluation

The initial evaluation of machine learning models yielded the following accuracies:

These baseline accuracies provide a foundation for further investigation into model enhancements in the subsequent stage.

Stage Two: Model Refinement and Feature Reduction

In the second stage, techniques like outlier removal and PCA were incorporated to improve model performance and reduce dataset dimensionality. The PCA was used with default parameters and had an accuracy of 42.5%, it did not improve the overall performance of the models suggesting that seizure data is not functional when reduced in dimensionality.

```

# Build CNN model
model = Sequential()
model.add(Conv2D(32, kernel_size=(3, 3), activation='relu', input_shape=input_shape))
model.add(Conv2D(32, kernel_size=(3, 3), activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Dropout(0.25))
model.add(Flatten())
model.add(Dense(32, activation='relu'))
model.add(Dropout(0.5))
model.add(Dense(1, activation='sigmoid'))
model.compile(loss=tf.keras.losses.binary_crossentropy,
              optimizer=tf.keras.optimizers.RMSprop(),
              metrics=['accuracy'])

```

Fig. 9 Random forest Classifier Confusion Matrix that shows the true vs, predicted values of the testing data

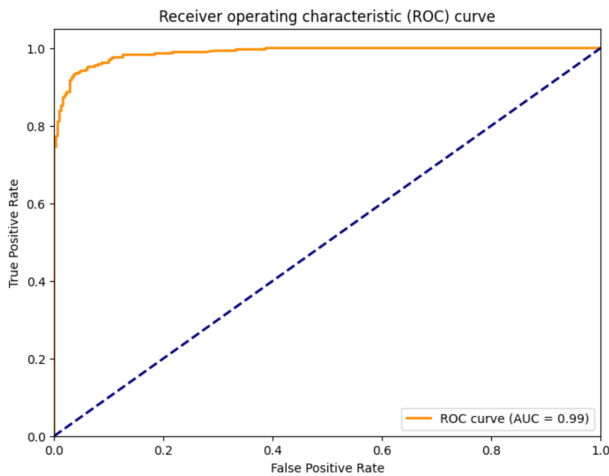


Fig. 10 ROC curve used to determine the accuracy of the CNN model, the greater the area underneath the graph, the better the model performance.

The Random Forest Classifier demonstrated notable improvement, achieving an accuracy of 98.59%, proving its effectiveness in this context.

Support Vector Machines and Multilayer Perceptron Classifiers also exhibited considerable accuracy, showcasing their robustness in handling complex patterns.

Gaussian Naive Bayes and Logistic Regression, while providing insights, demonstrated comparatively lower accuracies, suggesting limitations in capturing intricate relationships within the data.

In terms of the prediction models, the accuracies were as follows:

The CNN was trained separately due to high parameters.

	Model	Accuracy
0	LogisticRegression	0.564783
1	SVC	0.881304
2	DecisionTreeClassifier	0.776957
3	RandomForestClassifier	0.876522
4	GradientBoostingClassifier	0.797391
5	GaussianNB	0.558261
6	MLPClassifier	0.851739

Fig. 11 Stage one accuracies for different detection models

	Model	Accuracy
0	LogisticRegression	0.575272
1	SVC	0.922554
2	KNeighborsClassifier	0.897011
3	RandomForestClassifier	0.985870
4	GradientBoostingClassifier	0.823641
5	GaussianNB	0.585326
6	MLPClassifier	0.944293

Fig. 12 Stage two accuracies for different detection models

Discussion

High-accuracy machine learning models for seizure detection and prediction deployment can have a revolutionary impact on

	Model	Accuracy
0	LogisticRegression	0.840278
1	KNeighborsClassifier	0.758333
2	RandomForestClassifier	0.908333
3	GaussianNB	0.664352

Fig. 13 Accuracies for different prediction models

```
1114 - val_accuracy: 0.9433
Test loss: 0.1113804504275322
Test accuracy: 0.9433333277702332
```

Fig. 14 Accuracy of CNN model for seizure prediction

epilepsy care. The potential to correctly anticipate approaching seizures creates opportunities for effective and personalized intervention at the right time, and the proposed machine learning models, paired with the right hardware, has the potential to save lives. These models may be used by clinicians as decision support tools, helping them adapt interventions to suit individual patient profiles.

- **Early Intervention:** The high-accuracy models allows healthcare practitioners to take preventive measures at an early stage which, in turn, may reduce the severity and effects of seizures. Such early warnings may trigger changes in medication, lifestyle adjustments, or the initiation of various personalized interventions.
- **Reduced False Alarms:** Especially, the Random Forest Classifier revealed a strong robustness and thus implies that false alarms are less likely to occur. This is critical in reducing unnecessary interventions or alerts, thereby ensuring that the interventions are targeted and efficacious.
- **Individualized Treatment Plans:** These models, being accurate, make it possible to develop personalized treatment plans. Customizing interventions based on the specific patterns presented by each patient can increase the effectiveness of epilepsy therapy overall.
- **Improved Quality of Life:** By making correct forecasts, these models help to enhance the standard of life among people suffering from epilepsy. Early anticipatory measures and prompt interventions can contribute to enhanced seizure control and minimize the negative impact on daily routines.
- **Remote Monitoring and Telemedicine:** High-precision seizure detection models allow for remote monitoring of patients. They can support telemedicine programs by helping

healthcare providers to evaluate patient conditions without necessarily visiting them often.

These models could potentially change how seizures are managed, moving towards more proactive, personalized, and patient-focused approaches instead of existing universal ones.

Conclusion

The implementation of high-accuracy, low-runtime, machine learning models to predict epileptic seizures in real time provides promise for revolutionizing epileptic care. Through prediction, this algorithm can tackle the individuality and unpredictability of seizures, two of the largest issues with seizures. By equipping doctors and medical professionals with the necessary tools to enable personalized, accurate, and timely diagnosis methods, our algorithm can effectively simplify the diagnosis and care processes. Additionally, the research provides insight into software related to data prediction with continuous monitoring which can be implemented with advancing in-ear EEG devices. This feature allows both patients and health care professionals to understand the specific condition with more specificity and precision and also provides comprehensive information, creating a better understanding. The utilization of high-precision detection models translates to a proactive and patient-specific approach to epileptic care. The results of this paper have shown that high-accuracy models are possible and promising for the future of epileptic care.

References

- 1 C. Ogoke, W. Igwe and E. Umeadi, *Annals of Clinical and Biomedical Research*, **2**, year.
- 2 X. Qi, S. Hu, S. Liu, L. Zhao and Y. Zhang, *Ieee Transactions on Neural Systems and Rehabilitation Engineering*, **30**, 957–966.
- 3 Z. Ren, X. Han and B. Wang, *Frontiers in Neurology*, **13**, year.
- 4 E. Gleichgerrcht, B. Munsell, S. Alhusaini, M. Alvim, N. Bargalló, B. Bender and L. Bonilha, *Neuroimage Clinical*, **31**, 102765.
- 5 E. Gleichgerrcht, B. Munsell, S. Keller, D. Drane, J. Jensen, M. Spampinato and L. Bonilha, *Brain Communications*, **4**, year.
- 6 X. Xu, M. Lin and T. Xu, *International Journal of Environmental Research and Public Health*, **19**, 11326.
- 7 R. Hussein, S. Lee and R. Ward, *Biomedicines*, **10**, 1551.