ABSTRACT

Early detection of seizure onsets is a critical task in bio-informatics that can significantly help seizure patients with a controlled lifestyle. Recent research leverages various features from patients’ electroencephalograms (EEG) for their seizure detection model, some of which also focus on detecting seizures using a reduced number of EEG channels; however, they use longer time intervals of EEG and thus miss the early detection of seizure onsets. Imbalanced, noisy and diverse EEG data for patients makes early seizure onset detection further challenging. In my research, the first research contribution provides early detection of focal epileptic seizure onsets. In our upcoming research, we plan to explore focal seizure propagation and detect the onset of fully developed generalized seizures by reducing the number of channels. During this, we aim to leverage knowledge distillation to aid easy-to-use wearable devices for seizure detection and control. Lastly, we aim to equip our model with security features such that this model is robust against data poisoning and impersonation.

KEYWORDS

Multivariate EEG Time Series, Seizure Onset Detection, Knowledge Distillation, Security for AI Models

1 INTRODUCTION

Epileptic seizure patients often experience a disrupted lifestyle with risks of injury during seizures, and even the possibility of premature death. Researchers employ a continuous scalp electroencephalogram (EEG) data in their seizure detection models, but for better management and control, the challenge is to identify the onset of seizures as soon as they happen.

Miniscule seizure onset zone, which could be as low as 1 mm³ [15, 16], and limitations of focal seizure initiation to a hemisphere of the brain with probable spread across larger portions of the brain [1] over time make early detection of seizures a difficult task. Thus, only one of the many electrodes may show seizure patterns in the EEG data initially, while other channels continue to show unspecific patterns. Moreover, EEG data tends to be noisy, and seizure patterns vary significantly among patients; even within the same patient, these patterns can differ over time. Additionally, the data contains a class imbalance as seizures represent rare events, i.e., the minority class.

Furthermore, even when a model starts detecting focal epileptic seizure onsets precisely, it has been observed that the recall drops when the focal seizure propagates [8]. In other words, with time, when more electrodes start showing seizure-related patterns in the EEG, certain features which earlier were sufficient to precisely detect the seizure may not be enough to still help the model detect an underlying seizure. Also, the model that can precisely detect a focal epileptic seizure onset may not be suitable to detect a fully developed seizure or generalized seizure where all channels start showing the seizure patterns in the EEG at the same point in time.

Moreover, the models that can sufficiently address the above challenges may be prone to the risks of model poisoning or impersonation if they are made as per the patient’s need and ease to detect seizures. An intruder can use synthetic EEG data from a patient and destroy the trustworthiness of such models.

Based on the above-mentioned challenges, we formulate the following Research Questions (RQ) for epileptic seizure detection:

- **RQ1:** How can the onset of focal epileptic seizures be detected, and how quickly can this detection occur?
- **RQ2:** How can the propagation of focal epileptic seizures be detected to enhance the recall of early seizure onset detection? Would it be possible to achieve this using a reduced number of surface EEG electrodes?
- **RQ3:** How can seizure detection models be safeguarded against data poisoning and impersonation to ensure the reliability of detection devices for patients? Additionally, does incorporating security features into the model impact its performance in detecting seizure onset and type?

2 DEFINITIONS & PROBLEM FORMULATION

Here, we present the pertinent definitions and problem statement as done in our paper [8], that applies to all our research questions, despite variations in the scenarios and features for seizure detection.

**Definition 1 (Multivariate EEG Time Series).** A multivariate EEG time series is a collection of chronological observations \(X_t = [X_{1,t}, ..., X_{C,t}]\), where \(t\) is the time interval index, \(C\) is the set of the observed variables (EEG channels). Each EEG channel \(c \in C\) is observed at a discrete time interval \(t\), and \(X_{c,t}\) represents the value of the channel \(c\) at the time interval \(t\). The signal \(X_{c,t}\) is associated with the feature vector \(F_{c,t}\).

**Definition 2 (EEG Graph).** An EEG graph \(G = (V, E)\) is an undirected graph, where the nodes \(V\) represent the EEG channels...
movement-associated EEG features and use them in combination with earlier extracted features as a feature vector $F_{c,t}$, for each node (electrode) and give it as input to a graph-based model having the same number of nodes as the number in our EEG Graph and also experiment with Kolmogorov-Arnold Networks (KAN) [10] as a model. In our study of related literature [2, 3, 6, 7,], we observed that the features derived from the EEG of the electrodes C3 and C4 are responsible for muscle movements in the human body. To the best of our knowledge, it would be the first time that EEG features corresponding to muscle movements are being used to aid a seizure detection model. Additionally, through knowledge distillation, we aim to enable seizure onset detection using fewer electrodes than those initially used to collect the EEG data. So, based on the importance of each node in the classification performance, we later select an optimal subset of nodes that yields nearly the same result as the full electrode set.

Secure Modeling for Early Seizure Onset Detection (RP3): To address the third research question, we first aim to train the model obtained in RC2 per patient with their respective synthetic EEG data generated using a Generative Adversarial Network (GAN), marked as negative samples, so that during inference, the model can correctly classify only the real EEG to have seizure onsets, thereby handling the model poisoning issue. We also mix some other patients’ EEG’s and mark them as negative samples for the patient under consideration for the model. This helps the model learn to avoid impersonation. By hyperparameter tuning, we choose the optimal model for each patient such that the security-equipped model performs nearly the same as what it could have done without the poisoning and impersonation-based training.

Preliminary Results: We conduct all experimental evaluations on the publicly available TUH EEG dataset [12]. Table 1 presents the performance of model - Co-ReaSON [8] as our first research contribution (RC1). It outperforms the stated baselines on macro-Avg (mA) F1 by at least 5 p.p. and detects the focal seizure onset in a maximum of 9.19 seconds. However, as compared to the precision, the recall is relatively low. Since, we observed that in cases where the focal seizure propagates with time, our model misses detecting its seizure onset. To this end, we base our next research on RP2.

Acknowledgements: This work was partially funded by the Ministry of Culture and Science of the State of North Rhine-Westphalia, Germany (“iBehave”).

<table>
<thead>
<tr>
<th>Model</th>
<th>P</th>
<th>R</th>
<th>F1</th>
<th>mA P</th>
<th>mA R</th>
<th>mA F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chrononet [11]</td>
<td>0.62</td>
<td>0.11</td>
<td>0.18</td>
<td>0.66</td>
<td>0.37</td>
<td>0.47</td>
</tr>
<tr>
<td>RUSBoost [13]</td>
<td>0.15</td>
<td>0.58</td>
<td>0.23</td>
<td>0.53</td>
<td>0.63</td>
<td>0.51</td>
</tr>
<tr>
<td>TDNN+LSTM [14]</td>
<td>0.62</td>
<td>0.28</td>
<td>0.39</td>
<td>0.69</td>
<td>0.41</td>
<td>0.51</td>
</tr>
<tr>
<td>ResNet+LSTM [9]</td>
<td>0.65</td>
<td>0.47</td>
<td>0.55</td>
<td>0.73</td>
<td>0.54</td>
<td>0.62</td>
</tr>
<tr>
<td>GCN Teacher [17]</td>
<td>0.94</td>
<td>0.48</td>
<td>0.63</td>
<td>0.96</td>
<td>0.74</td>
<td>0.81</td>
</tr>
<tr>
<td>Co-ReaSON [8]</td>
<td>0.99</td>
<td>0.58</td>
<td>0.73</td>
<td>0.98</td>
<td>0.79</td>
<td>0.86</td>
</tr>
</tbody>
</table>

Figure 1: Our Approach: Features extracted from brain surface EEG are served as input to models in different Research Contribution/Plan (RC/RP) for secure epileptic seizure detection; Features extracted with green arrows are used in RC1; Seizure Propagation (S.P.) scenario with moving window over EEG is considered in RP2; Model security compliance is considered in RP3, which then gives our final integrated model.

(sensors) and the edges $E$ represent the connections between the sensors within the international 10-20 system of electrode placement on the scalp for recording EEG signals [5].

Problem Definition: Given a multivariate EEG time series $X_t$, and the sequence of observations at previous $k$ time intervals $X_{t-k}, \ldots, X_{t-1}$, the task of seizure onset detection is to identify if $X_t$ represents a seizure onset such that the seizure is observed in any channel $c$ at time interval $t$ for a patient.

3 RESEARCH CONTRIBUTION/PLANS AND PRELIMINARY RESULTS

In this section, we describe how we address the above-mentioned research questions (RQ1, RQ2, RQ3) using the respective three research contribution/plan (RC1, RP2, RP3). Eventually, we aim to integrate them together to come up with a single robust system for early seizure detection, as illustrated in Figure 1.

Focal Epileptic Seizure Onset Detection (RC1): As illustrated in Fig. 1, for RC1, our recently published research [8] builds EEG Graph edges and leverages them to extract novel correlation-based image features. These correlation features are combined with other extracted features like DWT-based statistical and image features and raw EEG image features and feed together as input to a pre-trained ResNet18 [4] model with a fully connected layer (Co-ReaSON [8]).

Knowledge Distillation for Early Detection of Propagating Seizures (RP2): To improve the recall for cases where the focal epileptic seizure propagates, we aim to have a time interval window that shifts 1 second on the EEG while extracting the same features as in RC1. This ensures capturing the temporal dependencies during seizure propagation. Additionally, we plan to leverage muscle
REFERENCES


