

# Enhanced Non-EEG Multimodal Seizure Detection: A Real-World Model for Identifying Generalised Seizures Across The Ictal State

J. Pordoy, G. Jones, N. Matoorian, M. Evans, N. Dadashiserej and M. Zolgharni

**Abstract**— Non-electroencephalogram seizure detection models hold promise for the early detection of generalised onset seizures. However, these models often experience high false alarm rates and difficulties in distinguishing normal movements from seizure manifestations. To address this, we were granted exclusive access to the newly developed Open Seizure Database, from which a representative dataset of 94 events was selected (42 generalised tonic-clonic seizures, 19 auras/focal seizures, and 33 seizures labelled as Other), with a combined duration of approximately 5 hours and 29 minutes. Each event contains acceleration and heart rate data which was expertly annotated by a clinician, who labelled every 5 second timestep with a class of Normal, Pre-Ictal, or Ictal. We then introduced the AMBER (Attention-guided Multi-Branching-pipeline with Enhanced Residual fusion) model. AMBER constructs multiple branches to form independent feature extraction pipelines for each sensing modality. The outputs of each branch are passed to our custom Enhanced Residual Fusion layer, where the extracted features are combined into a fused representation. The fused representation is then propagated through two densely connected blocks before being passed through a softmax activation function. The model was trained using k-fold cross validation, with k-1 fold used to train the model and the remaining fold was used to evaluate the model's performance. The results of these experiments underscore the efficacy of Ictal-Phase Detection, achieving an accuracy and  $f_1$ -score of 0.8995 and 0.8987. Notably, the model exhibited consistent generalisation, recording a True Positive Rate of 0.9564, 0.8325, and 0.9111 for the Normal, Pre-Ictal, and Ictal classes respectively. These findings were compounded by an average False Positive Rate, recording an overall score of 0.0502. In conclusion, this research introduces a new detection technique and model designed for multimodal seizure detection, with the potential to reduce the false alarm window and differentiate high and low amplitude convulsive movement. We believe the results of this study lay the groundwork for further advancements in non-electroencephalogram seizure detection research.

**Index Terms**— Seizure Detection, Multimodal, Attention, Residual Fusion, Multi-Branching, 1D CNN, Acceleration

## I. INTRODUCTION

Affecting approximately 1% of the global population, epilepsy is a chronic neurological disorder that is characterised by a lasting predisposition to generate recurring, unprovoked seizures [1]. Seizures are transient, paroxysmal alterations of the neurologic function resulting from an excessive, hypersynchronous electrical discharge of neurons [2]. The ictal state refers to the sequence of events and physiological changes encompassing an epileptic seizure and can be divided into the inter-ictal (normal state between seizures), pre-ictal (the state before a seizure), ictal (seizure activity) and post-ictal (the state following the seizure) phases. As defined by the International League Against

J. Pordoy, Dept of Computing, University of West London, St Mary's Rd, London W5 5RF, (e-mail: jamie.pordoy@uwl.ac.uk).

G. Jones, Open Seizure Detector, Hartlepool, UK, TS26 (e-mail: graham@openseizuredetector.org.uk).

N. Matoorian, Dept of Computing, University of West London, St Mary's Rd, London W5 5RF, (e-mail: Nasser.matoorian@uwl.ac.uk).

M. Evans, Pleotek, 46 Mount Prospect Park, Belfast, BT9 7BG (e-mail: molly.evans@live.com).

N. Dadashiserej, Dept of Computing, University of West London, St Mary's Rd, London W5 5RF (e-mail: nasim.dadashiserej@uwl.ac.uk).

M. Zolgharni, Dept of Computing, University of West London, St Mary's Rd, London W5 5RF (e-mail: massoud.Zolgharni@uwl.ac.uk).

Epilepsy, epileptic seizures are divided into four sub-categories: focal, generalised, focal to bilateral tonic-clonic, and unknown [3]. Focal seizures, also known as auras, are confined to one cerebral hemisphere and can affect a patient's muscle activity and cognitive awareness. Generalised seizures involve abnormal electrical fluctuations between communicating neuronal pathways and affect both hemispheres, resulting in muscle twitches, clonic jerking, a loss of consciousness, and compromised breathing [4]. A focal to bilateral tonic-clonic (FBTC) seizure begins as a focal event and subsequently generalises as the electrical discharge spreads across both hemispheres during the ictal phase into a generalised seizure. Despite pharmacological and surgical advancements, approximately 25% of patients diagnosed with epilepsy continue to experience spontaneous, unprovoked seizures. This condition is known as refractory or drug-resistant epilepsy. Epilepsy-related mortality remains a significant concern, with Sudden Unexpected Death in Epilepsy (SUDEP) documented as the predominant cause of premature epilepsy-related death with an incidence in adults recorded at 1.2 per 1000 patients [5].

Sveinsson *et al.* [6] conducted a clinical review to identify several leading risk factors for SUDEP, highlighting that patients with epilepsy (PWE) who experience uncontrolled generalised tonic-clonic (GTC) seizures are the primary at-risk group. Additionally, the risk of SUDEP is further increased if a PWE lives alone and experiences nocturnal GTC seizures, resulting in a 15-fold increased risk of early mortality. However, in 69% of cases, Sveinsson *et al.* also observed that premature death could have been prevented if there was an assistive form of in-home supervision during nocturnal hours. Due to the substantial proportion of unsupervised, in-home cases, there is a growing need for automated in-home seizure detection to mitigate the risk of SUDEP. As postulated by Friedman and Kazl [7], non-invasive detection systems could be used to augment in-home, nocturnal supervision and reduce the number of SUDEP cases through the early identification of generalised onset seizures.

## A. Related Work

Conventional electroencephalogram (EEG) monitoring is classed as the gold standard for the diagnosis, monitoring and detection of epileptic seizures. Clinicians can evaluate real-time neurological activity to diagnose the onset of a generalised event before it is clinically apparent. However, this method of detection is inadequate for real-world, in-home use as EEG monitoring is an expensive, time-consuming technique that requires a full clinical team with an electro-neurodiagnostic technician to affix a series of scalp-based electrodes for neurological observation [8].

Lamberts *et al.* [9] emphasised the need for preventive in-home measures, including nocturnal supervision to monitor high-risk patients susceptible to SUDEP and status epilepticus. To address the practical limitations associated with in-home EEG detection, researchers have explored the use of non-electroencephalogram (non-EEG) sensors to measure physiological changes in PWE. Thus, non-EEG detection has become a significant area of research, that employs non-invasive sensors to detect different types of epileptic seizure.

Most non-EEG systems are designed for real-world use outside of a hospital or epilepsy monitoring unit (EMU), and combine wireless

sensing modalities with state-of-the-art algorithms and commercial smart devices to form closed-loop detection systems [10]. Due to the large proportion of epilepsy-related deaths accounted for by SUDEP, there has been an increase in the number of studies devising novel closed-loop detection and ambulatory monitoring systems. These studies have examined the utility of accelerometers (ACMs), photoplethysmography (PPG), surface electromyography, and video motion sensors for non-EEG seizure detection. Among these, ACMs have demonstrated the greatest efficacy in detecting the initial convulsive muscle movements associated with GTC seizures. ACMs are non-invasive, lightweight, and cost-effective sensors that measure acceleration, making them a favourable modality for seizure detection. However, despite the success of ACM detection, a review of the literature found that ACM detection systems often have a high false alarm rate when detecting generalised onset seizures, ranging from several times a day [12] [13] to once every five days [11].

To reduce the false alarm rate, a 2020 study investigated how seizure and non-seizure movements could be distinguished using a three-dimensional spatial plane. The study measured 94.45% for classification accuracy, and successfully distinguished the convulsive movements of a GTC seizure from 12 common everyday movements [14]. However, this approach is only partially viable as on closer observation we can see that 80% of the data used in this study was simulated from healthy participants, thus it is hard to deduce whether this classification technique is feasible in the real-world.

Zia *et al.* [15], investigated real-time seizure detection using embedded wireless sensors and signal processing techniques to record 99% classification accuracy. However, this study only had access to data from three GTC seizures, which was recorded using a smartphone ACM. Whilst this study presented a series of novel ideas and contributes one of the first non-EEG datasets, its overall impact was hindered by the breadth and depth of the available data. Several ground-breaking studies have also used simulated patient data to pioneer non-EEG seizure detection research [13] [16].

Notably the seminal work of Conradsen *et al.* [17], whose research into non-invasive detection of seizure motor manifestations laid forth an early framework for multimodal classification still used today. However, a large proportion of the data used for research and commercial seizure detection originates in EMUs, often in the presence of a full clinical team with state-of-the-art technology, and a fully controlled environment [18]. Seizures in EMUs will be different to those in a residential environment, as PWE will have to handle the daily tasks and challenges of living with a neurological disorder without relying on a team of clinical experts assisted by video-EEG monitoring. Furthermore, EMUs are often affiliated with commercial collaborators and large consortiums, thus preventing the sharing of data and results due to intellectual property restrictions and conflicts of interest [19]. Most fields of neurological research have access to clinically annotated data, however as a research field non-EEG seizure detection is missing the core components required to facilitate rapid progression. As a result, we can see a trend across literature where studies have had to use simulated data to produce algorithms with excessively high accuracy scores and frequent false alarms.

Several studies have employed the EEG to extract neurological features to distinguish the different phases of the ictal state. Early studies encountered challenges in accurately distinguishing each ictal phase when analysing EEG recordings. A pioneering approach applied adaptive threshold modelling to detect convergence in short-term maximum Lyapunov exponent values from intracranial EEG recordings. However, this method yielded inconsistent results, with a p-value score of 0.5 and a high false prediction rate, rendering it unsuitable as a clinical diagnostic technique [20].

Similar research was conducted by Zeljković *et al.* [21] whose

optical flow extraction and band-pass temporal filtering techniques were used to analyse video-EEG recordings to differentiate the phases of the ictal state. However, this technique failed to predict 46% of video-EEG recordings, incorrectly classifying 4% and 42% of normal and ictal instances as pre-ictal respectively.

A more recent study by Eftekhari *et al.* [22], utilised N-gram-based pattern recognition with similarity metrics (Hamming distance and Needleman-Wunsch algorithm) to identify the epoch thresholds between ictal classes, achieving an average prediction sensitivity score of 93.81%, along with a false prediction rate of 0.06/h.

However, the seminal work by Sharif and Jafari [23] presents the most compelling EEG research. By employing optimised Poincaré plane analysis, each ictal phase was detected, recording an average sensitivity score between 91.8% and 96.6%, and an average false prediction rate of 0.05-0.08 per hour.

While previous studies have classified distinct ictal phases, the practical application of EEG-based seizure detection remains primarily confined to clinical settings due to operational costs. Conversely, non-EEG methodologies have demonstrated potential with real-world application. However, they face inherent challenges in distinguishing between different types of seizure-related movement, leading to recurrent false alarms and inaccuracies in seizure detection.

To improve the clinical utility and reliability of non-EEG seizure detection systems, it is imperative to address these limitations. While previous studies have explored the feasibility of detecting ictal classes using EEG techniques, to our knowledge, no prior research has investigated whether non-EEG modalities can distinguish between these phases. Thijss *et al.* [24] suggested that ictal autonomic changes could serve as diagnostic indicators, providing unique target variables for seizure detection mechanisms to identify. Autonomic manifestations often become more pronounced as a seizure traverses through the ictal state, particularly in focal seizures originating from the temporal lobe and generalising into a FBTC seizure. However, the absence of publicly available non-EEG data with clinical annotations has impeded the validation of this hypothesis and the development of a viable non-EEG detection methodology capable of discerning the different ictal phases.

## B. Contributions

This section outlines the contributions of this study, which are summarised as follows.

- 1) This study introduces a new seizure detection technique that accurately differentiates between high-amplitude convulsive movement typical of the ictal phase and low-amplitude myoclonic movement observed pre-ictal.
- 2) A new deep learning architecture for non-EEG, multimodal seizure detection called AMBER was introduced. AMBER uses multi-branching to independently process acceleration and heart rate features. These branches are then fused and propagated through a dense network to classify the ictal state.
- 3) For our experiments, we recorded an overall accuracy,  $f_1$ , Cohen's Kappa and Matthews Correlation Coefficient (MCC) score of 0.8995, 0.8987, 0.8942 and 0.8499. These results underscore the model's reliability for multimodal detection and highlight its ability to generalise on unseen data.
- 4) An Enhanced Residual Fusion Class was introduced that combines inputs from two or more independent branches. The output of each branch is then passed through a multi-head attention mechanism and fused together using residual connections, leading to improved generalisation.

## II. METHODOLOGY

In this methodology, we introduce a non-EEG seizure detection technique called Ictal-Phase Detection. This detection technique is based on the hypothesis that it is feasible to distinguish pre-ictal movements from the convulsive jerking and clonic movements characteristic of the ictal phase in a generalised onset seizure.

While conventional non-EEG seizure detection techniques primarily focus on the classification of seizure and non-seizure states, a significant proportion of seizures exhibit a pre-ictal phase characterised by subtle, low-amplitude myoclonic movements that are often referred to as focal seizures or auras.

The aim of this study is to differentiate the phases of the ictal state by distinguishing between high and low-amplitude movement. We hypothesise that there are quantifiable variances in movement and heart rate between the normal (seizure free), pre-ictal and ictal phases which, if leveraged, could improve on existing non-EEG techniques.

To address this, we propose the development of a new model that can independently leverage sensor data from two or more modalities on an independent scale. We posit that independent feature extraction will improve predictive probability as a true representation of each modality is represented. These representations can then be fused to form single output vector of key features, thus enhancing the model's ability to generalise.

To our knowledge, this represents the first study undertaking non-EEG seizure detection research into the ictal state. As there are no publicly available datasets to validate this hypothesis, early access to the Open Seizure Database was granted to the University of West London by Open Seizure Detector for this study [25] [26] [27].

### A. The Open Seizure Database

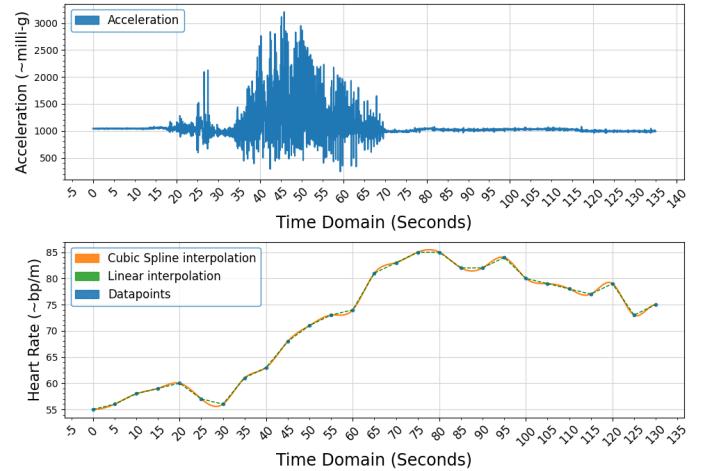
This is the first study to use the newly developed Open Seizure Database - Version 3 (OSDB) [28]. Designed by Open Seizure Detector (OSD) [26], to facilitate research into non-EEG seizure detection, the OSDB contains multimodal sensor data from 49 participants in real-world environments. The database is comprised from 494 events, encompassing 139 epileptic seizures and is collected over a duration of 453 days [28]. Continuous patient monitoring was facilitated for this period where patient data was recorded using wearable Garmin devices equipped with embedded ACM and PPG sensors.

The OSDB is publicly accessible, but exclusive access was granted to the University to West London to conduct this investigation into Ictal-Phase Detection.

### B. Dataset

For this study, we initially selected 139 events from the OSDB, categorised as GTC, Aura, Atonic/Fall, Other, and Simulated. For consistency, we removed 36 events that lacked simultaneous acceleration and heart rate data. We then filtered events labelled as Simulated and Atonic/Fall, as they were outside the scope of our experiments. This resulted in a final representative subset of 94 events. The subset consists of 42 GTC Seizures, 19 Auras/Focal Seizures, and 33 seizures categorised as Other. The OSDB labels a seizure as Other if the event lacks additional subtype classification [28].

The final representative subset consists of sensor data from 18 participants diagnosed with generalised epilepsy, with a combined duration of 5 hours, 29 minutes, and 5 seconds. Each event is assigned an identifier and structured using a one-to-many relationship, where one event is formed from multiple sub-events, each representing a 5-second timestep. Each timestep contains simultaneous acceleration and heart rate data as shown in Fig. 1. The acceleration data, denoted as  $A$ , was sampled at 25 Hz, resulting in 125 datapoints for each 5-second timestep. This can be expressed as  $A_t = \{a_1, a_2, \dots, a_{125}\}$



**Fig. 1:** Time series plot for Event 15039: The upper plot maps one-dimensional acceleration data as vector magnitude, while the lower plot shows interpolated heart rate data recorded in beats per minute.

where  $a_n$  denotes the indexed acceleration datapoints within timestep  $t$ . The heart rate, data denoted as  $H$ , was recorded at a lower sample rate of 0.2 Hz, computing a single heart rate datapoint ( $h_1$ ) for each 5 second timestep, which we can then express as  $H_t = \{h_1\}$ .

### C. Clinical Annotation

In this subsection, we detail the clinical annotation process for Ictal-Phase Detection. To test our hypothesis of differentiating between pre-ictal and ictal movements, our team clinician annotated the timesteps for each event with a class label of 0, 1, or 2, representing the Normal, Pre-Ictal, and Ictal classes respectively. It should be noted that, in a clinical or descriptive context, the normal, pre-ictal, and ictal phases of a seizure will be written in lowercase, while the Normal, Pre-Pre-Ictal, and Ictal classes will be capitalised.

The Normal class denotes movements outside of the convulsive scope, specifically during the post-ictal and inter-ictal phases. This approach was inspired by Abdulhay *et al.* [39], who also used the Normal class to express inter-ictal (seizure-free) states in EEG signal classification. The Pre-Ictal class represents focal seizures, auras, and low-amplitude myoclonic movement that typically precedes seizure onset. We then used the Ictal class to represent high-amplitude signals, indicative of the convulsive movements that occur between seizure onset and cessation. We conducted an analysis of each event to compute a set of averages which provided guidelines for the clinical annotation process (see Fig. 2). The rate of change of acceleration was the main technique used and can be expressed as:

$$\text{Rate of Change of Acceleration} = \frac{A_t - A_{t-1}}{\Delta t} \quad (1)$$

$A_t - A_{t-1}$  represents the difference in acceleration datapoints between two adjacent timesteps, where  $A_t$  represents acceleration at the end of timestep  $t$ ,  $A_{t-1}$  represents acceleration at the end of the previous timestep, and  $\Delta t$  denotes the duration of timestep  $t$ .

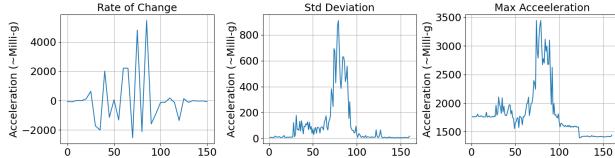
We conducted secondary computations to validate the results of Equation 1, calculating the average acceleration range, maximum acceleration, and standard deviation of acceleration in each timestep.

The clinician then utilised these results along with partial video footage to annotate each timestep with a class label.

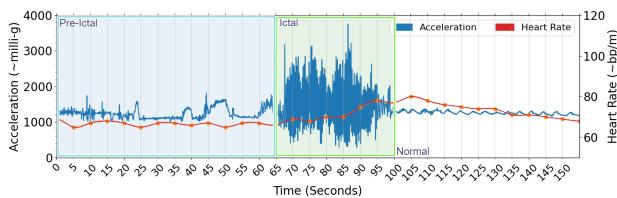
Table I presents the average thresholds calculated to guide the clinician in annotating each timestep. Annotations were conducted using Oxford University's VGG Image Annotation software [29],

**TABLE I:** Clinical Annotation Averages: These averages were employed to assign class labels (0 for Normal, 1 for Pre-Ictal, and 2 for Ictal) to each timestep. All measurements are in milli-g (1/1000 of a G). Notation: STD = standard deviation, Avg = average, and G = gravity.

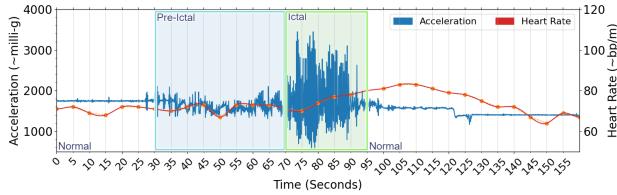
Class	Ictal Phase	Avg Acceleration Range [Milli-g]	Max Value [Milli-g]	STD [Milli-g]	Rate of Change [Milli-g]
0	Normal	800 - 1200	1500	0-50	$-1500 \leq A_t \leq 1000$
1	Pre-Ictal	1200 - 2000	2200	50-200	$-2000 < A_t \leq -1500$ or $1000 < A_t \leq 2000$
2	Ictal	2000 - 5000	7000	>200	$-2000 > A_t$ or $A_t > 2000$



**Fig. 2:** Event 7261 Analysis Plots: Rate of change of acceleration (left), standard deviation (centre) and maximum acceleration (right)



(a) Event 7261 Annotation Plot: 0s → 65s = Pre-Ictal, 65s → 100s = Ictal, 100s → 150s = Normal (Post-ictal)



(b) Event 9401 Annotation Plot: 0s → 30s = Normal, 30s → 65s = Pre-Ictal, 65s → 95s = Ictal, 95s → 155s = Normal (Post-ictal)

**Fig. 3:** Clinical Annotation Plots: White, blue and green annotations represent the Normal, Pre-Ictal and Ictal classes respectively.

where 2D boundary boxes were employed to visualise the segmented data into ictal phases (Fig. 3). The annotated timesteps were stored as a CSV file and grouped by event. All events were consolidated into a single dataframe, and then a new column called Outcome was appended to denote the ground truth annotations.

#### D. Cubic Spline Interpolation

This subsection details the interpolation techniques used to reshape our heart rate data. For each instance of  $t$ , acceleration data consists of 125 datapoints, whereas heart rate data is comprised from a single datapoint, resulting from the different sampling rates used when the OSDB was recorded.

We used linear interpolation to address the disparity in size between acceleration and heart rate features. This generated a straight-line connection of interpolated values between heart rate datapoints. However, this straight-line approximation did not accurately reflect the heart rates natural curvature and we felt linear interpolation could negatively impact our results. To further address this disparity and create a curvature that represents uniform dimensions for each feature, we employed cubic spline interpolation to mathematically construct a piecewise continuous curve that passes through each of the

heart rate datapoints  $(x_i, y_i)$  for  $i = 0, 1, \dots, N$ , where  $x_i$  represents the sequence index and  $y_i$  represents the heart rate value at that point. The resulting function  $y = f(x)$  represents the heart rate value  $y$  as a continuous function of  $x$ , providing a smooth transition through the datapoints. Since there are 125 datapoints in  $A_t$ , we need to generate 125 interpolated indices for  $H_t$ . For each instance of  $i$  within  $H_{t(i)}$ , a cubic polynomial can be calculated as:

$$H_{t(i)} = a_i + b_i(t - t_i) + c_i(t - t_i)^2 + d_i(t - t_i)^3 \quad (2)$$

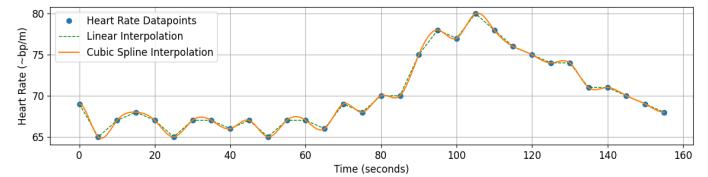
where  $H_{t(i)}$  denotes the computed polynomial for interval  $i$ . The coefficients  $a_i$ ,  $b_i$ ,  $c_i$ , and  $d_i$  are calculated to generate a smooth interpolation between datapoints. The linear coefficient  $b_i$  represents the rate of change between  $t_i$  and  $t_{i+1}$ . The quadratic coefficient  $c_i$  captures the quadratic behaviour, and the cubic coefficient  $d_i$  models the curvature between  $t_i$  and  $t_{i+1}$ . To calculate the polynomial values at  $t_i$ , coefficients  $a_i$ ,  $b_i$ ,  $c_i$ , and  $d_i$  are iteratively computed for each instance of  $i$  in  $t$ . This process yields a smooth cubic polynomial interpolation between datapoints at  $t_i$  and  $t_{i+1}$  (Fig. 4). As a result, the heart rate data for timestep  $t$  has been reshaped into a set of interpolated datapoints where  $H_t = \{h_1, h_2, \dots, h_{125}\}$ .

#### E. Preprocessing

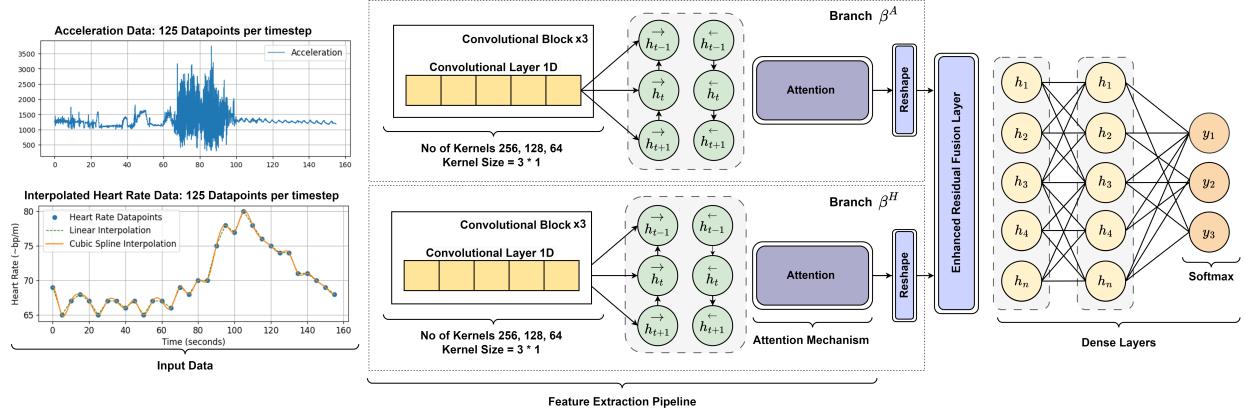
In this subsection we describe the preprocessing techniques that were undertaken. Following the reshaping of our representative subset, a class imbalance was observed in which the Normal class accounted for 40.54% (1601 timesteps) of the overall dataset with the Pre-Ictal (837 timesteps) and Ictal (1511 timesteps) classes accounting for 21.2% and 38.26% respectively. This imbalance was expected, given that the average duration of the ictal phase of a generalised seizure is  $\approx 39.3 \pm 17.7$  seconds, while the pre-ictal phase can range from seconds to minutes before generalisation [30].

To address this imbalance, time-series random oversampling was employed. This technique increased the representation of minority classes, making all classes proportionate to the majority class while preserving the sequential order of timesteps in each. The result was a resampled dataset comprised of 4803 timesteps with a balanced distribution, where each class accounted for 33.3% ( $\approx 1601$  timesteps).

We then reshaped the preprocessed dataset into a three-dimensional vector denoted as  $X_t$  to facilitate time series modelling. This input vector integrates the preprocessed acceleration data, represented as  $A_t$ , and heart rate data, represented as  $H_t$  into a single vector which



**Fig. 4:** Cubic Spline Interpolation. Each timestep  $t$  corresponds to a single heart rate data point (blue). Linear interpolation (green) connects adjacent timesteps, followed by cubic spline interpolation (orange), resulting in a smooth curve with 125 interpolated datapoints.



**Fig. 5:** Architecture of the Attention-guided Multi-Branching-pipeline with Enhanced Residual fusion (AMBER) Model

can be expressed as  $X_t = [a_1, h_1, a_2, h_2, \dots, a_{125}, h_{125}]$ . The reshaped vector  $X_t$  is formed from three dimensions: the number of timesteps, the timestep length, and the number of features, resulting in a final size of (4803, 125, 2).

#### F. Proposed Model

In this subsection, we introduced our model, AMBER (Attention-guided Multi-Branching-pipeline with Enhanced Residual fusion). The architecture of the proposed model is shown in Fig. 5. AMBER constructs multiple branches upon initialisation, where the number of branches is proportionate to the number of input features. Each branch establishes a dedicated feature extraction pipeline, incorporating attention mechanisms as the final layer to identify salient features. The output of each branch is then passed in parallel to the model's custom layer called, Enhanced Residual Fusion, which fuses the output of both branches into a single representation, combining the acceleration and heart rate features. The fused representation is then passed to the model's classifier. To introduce non-linearity, the output is propagated through two densely connected blocks before being passed to the final dense layer, which employs a softmax activation function for multi-class classification.

#### G. Model Overview

The AMBER model employs branch-specific vector partitioning to construct independent branches denoted as  $\beta$ , where the number of branches equals the number of input features. The model takes the reshaped multivariate array  $X_t$  as input, which is then partitioned and a permutation is applied to reshape the vectors' dimensions, forming two independent input vectors. The acceleration and heart rate input vectors now expressed as  $X^A$  and  $X^H$  are passed to branch  $\beta^A$  and  $\beta^H$  as input. We can express these vectors mathematically as  $X_t^A$  or  $X_t^H = [x_1, x_2, x_3, \dots, x_{125}] \in \mathbb{R}^{125}$ , where 125 real-valued datapoints ( $x$ ) are passed as input for each instance of  $t$ .

#### H. Feature Extraction Pipeline

The following notation is a representation of input vector  $X_t$  as it traverses through the layers in the AMBER model (see Fig. 6). For each branch,  $X_t$  is passed as input through a dedicated feature extraction pipeline, comprised from three convolutional blocks denoted as  $b_1$ ,  $b_2$ , and  $b_3$ . For each block, the 1D convolutional layer has a filter size 256, 128 and 64 respectively. We then configured the convolutional layer with a kernel size of 2, a stride of 1, followed by ReLU activation function to introduce non-linearity. The layer

then performs a convolutional operation which can be mathematically expressed as:

$$y_j^l = \sigma \left( \sum_{i=1}^{N_{l-1}} \text{Conv1D}(w_{i,j}^l, x_i^{l-1}) + b_j^l \right) \quad (3)$$

Let  $y_j^l$  represent the output of the  $j$ -th feature map in the  $l$ -th layer, computed by applying a sigmoid activation function to the sum of the convolutional operations between the filter of the  $l$ -th layer ( $w_{i,j}^l$ ) and the feature map from the previous layer ( $x_i^{l-1}$ ). Let  $b_j^l$  express the bias of the  $j$ -th feature map in layer  $l$ , which we then add to the result of the convolutional operation to introduce an offset. A summation is then computed for all neurons in the previous layer ( $N_{l-1}$ ), resulting in the activation of the  $j$ -th feature map in layer  $l$ .

The output of the convolutional operation undergoes standardisation through a batch normalisation layer to improve training stability and accelerate convergence. Proceeding, a ReLU activation layer is applied, introducing non-linearity to the block. We then conducted downsampling through a max pooling layer to extract features while reducing dimensionality. This process iterates through blocks  $b_1$  to  $b_3$ , where the features are flattened to form a 1D vector, denoted as  $X$ . The vector is then reshaped into a 3D tensor for instance of  $t$ . The reshaped 3D tensor,  $x_t$ , is then passed as input to the proceeding bidirectional Long Short-Term Memory (BiLSTM) layers.

The pipeline's bidirectional layers utilise a pair of hidden ( $h$ ) and cell ( $c$ ) states. For each time step  $t$ ,  $\vec{c}_t$  and  $\hat{h}_t$  represent these states in a forward direction, while  $\overleftarrow{c}_t$  and  $\overleftarrow{h}_t$  represent the states in a backward direction. To regulate the information flow to the cell state, three nonlinear gating mechanisms are employed, expressed as the input gate ( $i_t$ ), the output gate ( $o_t$ ), and the forget gate ( $f_t$ ).

For each timestep,  $f_t$  applies a sigmoid function to determine what information should be retained from the previous cell state ( $\vec{c}_{t-1}$ ) and  $x_t$ . This can be expressed by:

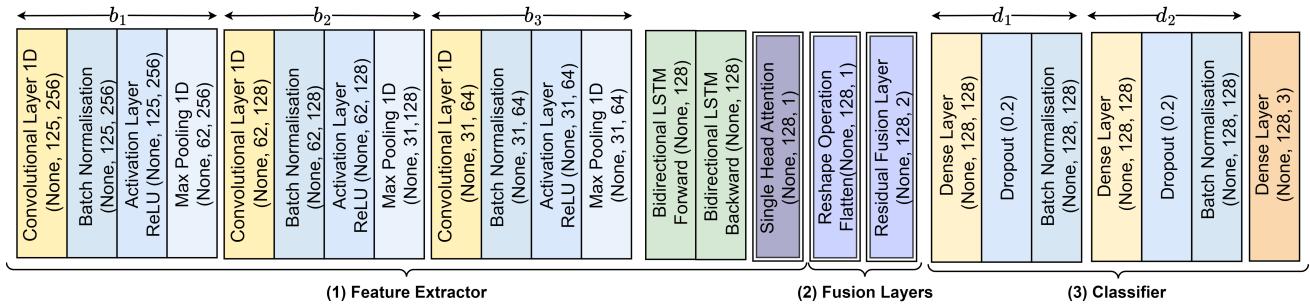
$$\vec{f}_t = \sigma(\vec{W}_f \cdot [\vec{h}_{t-1}, x_t] + \vec{b}_f) \quad (4)$$

We then use  $\vec{i}_t$  to calculate the information that should be added to  $\vec{c}_t$ , and then employ  $\vec{o}_t$  to calculate how much of  $\vec{c}_t$  is used as output for the current timestep. These gates are defined as:

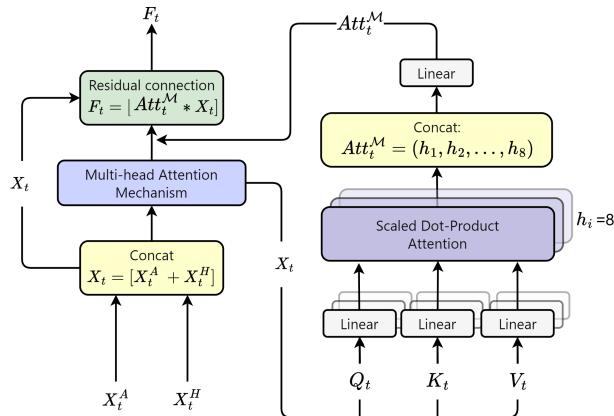
$$\vec{i}_t = \sigma(\vec{W}_i \cdot [\vec{h}_t, x_t] + \vec{b}_i) \quad (5)$$

$$\vec{o}_t = \sigma(\vec{W}_o \cdot [\vec{h}_{t-1}, x_t] + \vec{b}_o) \quad (6)$$

For each timestep, parallel computations are conducted in both the forward and backward directions. Hidden states ( $\vec{h}_t$ ,  $\overleftarrow{h}_t$ ) from both



**Fig. 6:** Layered Representation of AMBER: (1) = Feature Extraction Pipeline with  $3 * \text{convolutional blocks}$  ( $b_1, b_2, b_3$ ),  $1 * \text{BiLSTM}$  and  $1 * \text{Attention}$ , (2) = Fusion Layer (Enhanced Residual Fusion), (3) = Classifier with  $2 * \text{dense blocks}$  ( $d_1, d_2$ ) and a softmax layer



**Fig. 7:** Architecture of the Enhanced Residual Fusion Layer illustrating how attention weights  $X^A$  and  $X^H$  are outputted from branch  $\beta^A$  and  $\beta^H$ , concatenated, passed to a multi-head attention mechanism, and then fused with the original inputs via a residual connection

directions are concatenated to form  $h_t$ . The cell states  $(\tilde{c}_t^f, \tilde{c}_t^b)$  are aggregated using  $f_t$  and  $i_t$ , forming the output  $c_t$ . The output gate  $o_t$  then refines the cell state by applying a sigmoid function, regulating the gating mechanism with a hyperbolic tangent function to control the information flow. The results from the forward and backward layers are concatenated to form the bidirectional output of  $t$ , which we can express as  $X_t = [\tilde{h}_t^f, \tilde{h}_t^b]$ .

We then pass  $X_t$  through a single-headed attention mechanism, denoted as  $Att_t$ , positioned as the final layer in each pipeline to extract relevant features. Vector  $X_t$  undergoes a linear transformation to construct three attention matrices: Query ( $Q_t$ ), Key ( $K_t$ ), and Value ( $V_t$ ). This transformation is expressed as:

$$Q_t = X_t \cdot W_Q, \quad K_t = X_t \cdot W_K, \quad V_t = X_t \cdot W_V \quad (7)$$

In these equations,  $W_Q$ ,  $W_K$ , and  $W_V$  are pre-set weighted matrices with the dimensions  $d \times d_k$ , where  $d$  is the dimensionality of the input feature. The attention scores for timestep  $t$  are then computed by taking the dot product of  $Q_t$  and transposed matrix  $K_t^T$ .

$$\text{Attention Scores}_t = Q_t \cdot K_t^T \quad (8)$$

These scores are then scaled by the square root of  $d_k$  to maintain numerical stability, which we can express as:

$$\text{Scaled Scores}_t = \frac{\text{Attention Scores}_t}{\sqrt{d_k}} \quad (9)$$

The Scaled Scores are then passed through a softmax function and multiplied by  $V_t$  to compute a final set of attention weights for  $Att_t$ .

$$X_t = \text{Softmax}(\text{Scaled Scores}_t) \cdot V_t \quad (10)$$

### *I. Enhanced Residual Fusion Layer*

In this subsection, we introduce our custom Enhanced Residual Fusion Layer, designed to merge the outputs of two or more branches into a single fused representation. Attention weights  $X_t^A$  and  $X_t^H$  from branches  $\beta^A$  and  $\beta^H$  are passed in parallel to the Enhanced Residual Fusion Layer, where the weights from both branches are concatenated to create vector  $X_t$ . Mathematically, this operation is expressed as  $X_t = \text{concat}[X_t^A, X_t^H]$ , with  $X_t$  formed from the dimensions  $\mathbb{R}^{(|A|+|H|) \times d}$ .

Vector  $X_t$  is then passed through a multi-head attention mechanism, denoted as  $\text{Att}_t^M$ , which is comprised from 8 attention heads denoted as  $h_i$ , each designed to operate on a projection space of 32 dimensions.

Attention heads  $h_1$  -  $h_8$ , independently compute  $Q_t$ ,  $K_t$  and  $V_t$  matrices using learned weights. Since these equations have already been explained in Equation 7, they will not be repeated here.

The attention scores for each head are then calculated using a scaled dot-product, which is expressed by the following equation:

$$h_i = \text{Softmax} \left( \frac{Q_t \cdot K_t^\top}{\sqrt{d_k}} \right) \cdot V_t \quad (11)$$

where  $K_t$  and  $V_t$  are the key and value matrices to compute the  $i$ -th attention head. The dimensions of  $K_t$  and  $V_t$  matrices can be expressed as  $\mathbb{R}^{(|A|+|H|) \times d_k}$  and  $\mathbb{R}^{(|A|+|H|) \times d_v}$ , respectively.

The output for each head is then concatenated and linearly projected to form the output of  $\text{Att}_t^M$ :

$$Att_t^{\mathcal{M}} \equiv \text{concat}(h_1, h_2, \dots, h_n) : W_O \quad (12)$$

To complete the residual fusion layer,  $Att_t^M$  is then multiplied to the original concatenated input vector  $X_t$  through a residual connection, outputting a fused vector denoted as  $F_t$ :

$$F_t \equiv [X_t + Att_t^{\mathcal{M}}] \quad (13)$$

The residual connection was used to retain key features from both branches while integrating the new information derived from  $Att_t^M$ . Thus, the Enhanced Residual Fusion Layer combines inputs from  $X^A$  and  $X^H$  to form an attention guided fused representation.

## 1. Classifier and Loss Function

Vector  $F_t$  is then propagated through a fully connected network, passing through dense blocks  $d_1$  and  $d_2$ . Each dense block consists of

**TABLE II:** K-fold Cross Validation Results. Notation: Val = Validation

k-Fold	Accuracy	Loss	Val Accuracy	Val Loss	Test Accuracy
Fold 1	0.8849	0.0583	0.8930	0.0505	0.9050
Fold 2	0.8897	0.0537	0.8906	0.0506	0.8976
Fold 3	0.8834	0.0561	0.9048	0.0446	0.8960
Fold 4	0.8914	0.0537	0.8738	0.0554	0.9013
Fold 5	0.8746	0.0587	0.9000	0.0461	0.9077

a single dense layer consisting of 128 neurons, followed by a dropout layer with a dropout rate of 0.2, followed by a batch normalisation layer. The output of  $d_2$  is then passed to the final dense layer, which employs a softmax activation function to calculate the prediction probability of  $t$ . We can calculate this as:

$$\hat{y}_t = \text{Softmax}(F_t \cdot w + b) \quad (14)$$

where  $w$  is the layers weights,  $b$  is the networks bias and  $\hat{y}_t$  is the predicted ictal class for timestep  $t$ .

### III. EXPERIMENTAL SETUP

In this section, we detail the experimental setup employed for this study. A shared Google Drive was established, facilitating collaboration between technical and clinical team members for annotating data. For each event, clinical annotations were stored using a CSV file, which was accompanied by a visualisation plot depicting the annotated phases (see Fig. 3). Experiments were then conducted using Python 3.7 and TensorFlow 2.4.

To improve generalisation, several regularisation techniques were applied. An  $l_2$  regulariser was added to the dense layers in  $d_1$  and  $d_2$  to control the model's complexity and prevent overfitting, with the regularisation strength adjusted to 0.001. Additionally, the ReduceLROnPlateau callback function was implemented to decrease the model's learning rate by 50% if loss did not improve for 20 consecutive epochs. Furthermore, we employed early stopping to halt training after 20 consecutive epochs without improvement.

To train our proposed model, we employed stratified k-Fold Cross Validation. This approach partitioned our dataset into equal-sized instances of  $k$ , where  $k = 5$ . Folds were then partitioned by event Id, with  $k - 1$  (80%) folds used to train the model, while the remaining instance of  $k$  (20%) were used for testing. We partitioned our data by event Id to minimise biases and prevent overlap, ensuring that the events used for training were excluded from the testing subsets, thereby providing an unbiased evaluation of the model's performance.

For each set of experiments in  $k$ , the proposed model was trained for 100 epochs. Through iterative experimentation and hyperparameter optimisation, we selected an RMSprop optimiser with a reduced learning rate of 0.0001 and an augmented epsilon score of  $1e - 09$ . We then chose a batch size of 16 and employed a categorical cross-entropy loss function to quantify the model's performance.

Experiments were conducted to evaluate the models performance for each  $k$ -fold. This involved calculating the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) predictions made by the model. These metrics then served as predictive indicators of the model's performance where:

- 1) TP represents the number of predictions where our model correctly classified a positive class as positive.
- 2) FP represents the number of predictions where our model incorrectly classified a negative class as positive.
- 3) FN represents the number of predictions where our model correctly classified a negative class as negative.
- 4) TN represents the number of predictions where our model incorrectly classified a positive class as negative.

### IV. RESULTS

This section introduces the results of the proposed model and demonstrates the application of the Ictal-Phase Detection. The results presented in this section are comprised from several sets of experiments aimed at evaluating the performance of Ictal-Phase Detection on unseen test data.

Table II presents the training and validation results for each  $k$ -fold. Accuracy scores range from 0.8746 to 0.8914 across all  $k$ -folds. The model's loss metric exhibited slight variability, with loss values ranging from around 0.0537 to 0.0587, indicating similar degrees of predictive error for each instance of  $k$ . The model's validation accuracy produced a similar curvature, recording between 0.8738 to 0.9000, while validation loss fluctuated between 0.0446 to 0.0554. These observations underscore slight performance fluctuations across each instance of  $k$ , as consistent results were recorded between the training and validation subsets, indicating stable generalisation.

Notably, in Fold 3, the proposed model achieved the highest validation accuracy at approximately 0.9048, coupled with the lowest validation loss recorded at 0.0446. Conversely, in Fold 5 the model recorded slightly lower scores, recording 0.9000 and 0.0461, for validation accuracy and loss respectively. These validation results demonstrate strong generalisation, indicating the model's ability to perform well across all folds. The test accuracy for each fold showed consistent performance, ranging from 0.8960 to 0.9077. Fold 5 achieved the highest test accuracy scores recording 0.9077, closely followed by Fold 1 at 0.9050.

In summary, the model demonstrated consistent performance across all  $k$ -folds. The validation metrics showed slight variability, with an average accuracy and loss score of 0.8848 and 0.0561, respectively. Furthermore, our test results showcase consistent performance for all instances of  $k$ , highlighted by strong generalisation and an average test accuracy score of 0.899.

#### A. Comparative Analysis with State-of-the-Art Models

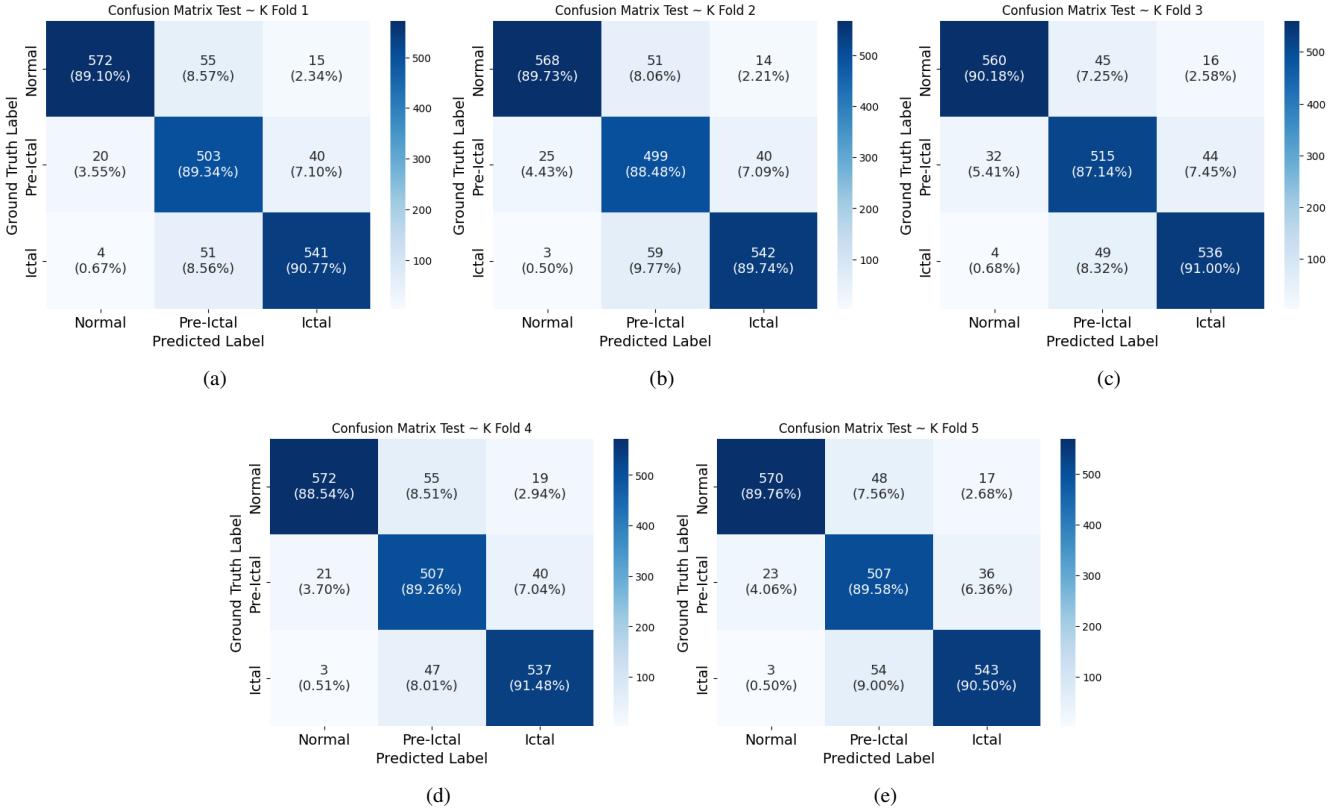
In this subsection we compare the performance of AMBER with several state-of-the-art models using the same Ictal-Phase dataset and hyperparameter configurations (see Table III). The AMBER model demonstrated strong performance across several classification metrics, recording an average Test Accuracy of 0.8995 and an  $f_1$ -score of 0.8987. The model also recorded a Cohen's Kappa score of 0.8492 and a MCC of 0.8499. Furthermore, we recorded a Positive Predictive Value (PPV) and a True Positive Rate (TPR) of 0.8994 and 0.8989 respectively, highlighting the AMBER model's proficiency in accurate classification while minimising the frequency of false positives.

In comparison, models such as the 1D CNN, Stacked BiLSTM, and Stacked LSTM recorded lower performance scores, while models employing attention mechanisms (as seen in  $m_2$  and  $m_4$ ) showed increased performance. Notably, the models that employed multi-branching techniques, specifically models  $m_2$ ,  $m_3$ , and  $m_5$ , recorded high accuracy and  $f_1$  scores ranging from 0.8613 to 0.8698 and 0.8759 to 0.8952 respectively.

Notably, the early experiments we conducted employed multi-branching, although without a residual fusion layer. While the results of these experiments produce promising results, they exhibited lower performance scores when compared to AMBER with a custom residual fusion layer. These findings validated the use of the Enhanced Residual Fusion layer and demonstrated how it improved AMBER's ability to capture complex patterns from independent features. This improvement is evident when comparing the results of models  $m_1$  and  $m_2$ , as both models are constructed from the same architectural components, although  $m_1$  incorporates the Residual Fusion layer.

**TABLE III:** Comparative Analysis: Evaluation of performance scores obtained by the AMBER model compared to several state-of-the-art classification models and configuration builds. All models are trained using the same dataset, preprocessing techniques and hyperparameters.

<b>Id</b>	<b>Model</b>	<b>Accuracy</b>	<b>f<sub>1</sub>-Score</b>	<b>Cohen's Kappa</b>	<b>MCC</b>	<b>PPV</b>	<b>TPR</b>
<i>m<sub>1</sub></i>	AMBER	0.8995	0.8987	0.8492	0.8499	0.8994	0.8989
<i>m<sub>2</sub></i>	Multi-Branche CNN-BiLSTM + Attention	0.8698	0.8805	0.8102	0.8204	0.8743	0.8999
<i>m<sub>3</sub></i>	Multi-Branche CNN-BiLSTM	0.8613	0.8952	0.8397	0.8416	0.8845	0.8628
<i>m<sub>4</sub></i>	1D CNN + Attention	0.8653	0.8754	0.8007	0.8106	0.8778	0.8832
<i>m<sub>5</sub></i>	Multi-Branche CNN + Adam	0.8647	0.8759	0.8004	0.8102	0.8731	0.8886
<i>m<sub>6</sub></i>	1D CNN Stacked BiLSTM	0.8631	0.8534	0.8552	0.8287	0.9079	0.8702
<i>m<sub>7</sub></i>	Stacked BiLSTM	0.8454	0.8725	0.8012	0.8101	0.8532	0.8641
<i>m<sub>8</sub></i>	1D CNN + ADAM	0.8449	0.8722	0.8009	0.8104	0.8530	0.8676
<i>m<sub>9</sub></i>	Stacked LSTM	0.8306	0.8617	0.7899	0.7996	0.8432	0.8515
<i>m<sub>10</sub></i>	1D CNN	0.8149	0.8455	0.7695	0.7802	0.8278	0.8415



**Fig. 8:** AMBER Confusion Matrix Results: Each matrix represents a  $k$ -fold testing subset from a 5-fold cross-validation. Predicted labels are compared against annotated ground truth for Normal, Pre-Ictal, and Ictal classes.

### B. Confusion Matrix Analysis

In this subsection, we present the confusion matrix results of the AMBER model. Each experiment represents a 20% fold in the dataset used to test the model, while the remaining 80% is used for training. This iterative process is repeated for each fold in  $k$ , with the selected test set traversing sequentially through each fold. This ensures that each fold uses a different set of partitions to train and test the model. The results of these experiments are shown in Fig. 8, where each plot represents the confusion matrix experiments conducted for each instance of  $k$ , mapping the annotated ground truth against the models predicted labels.

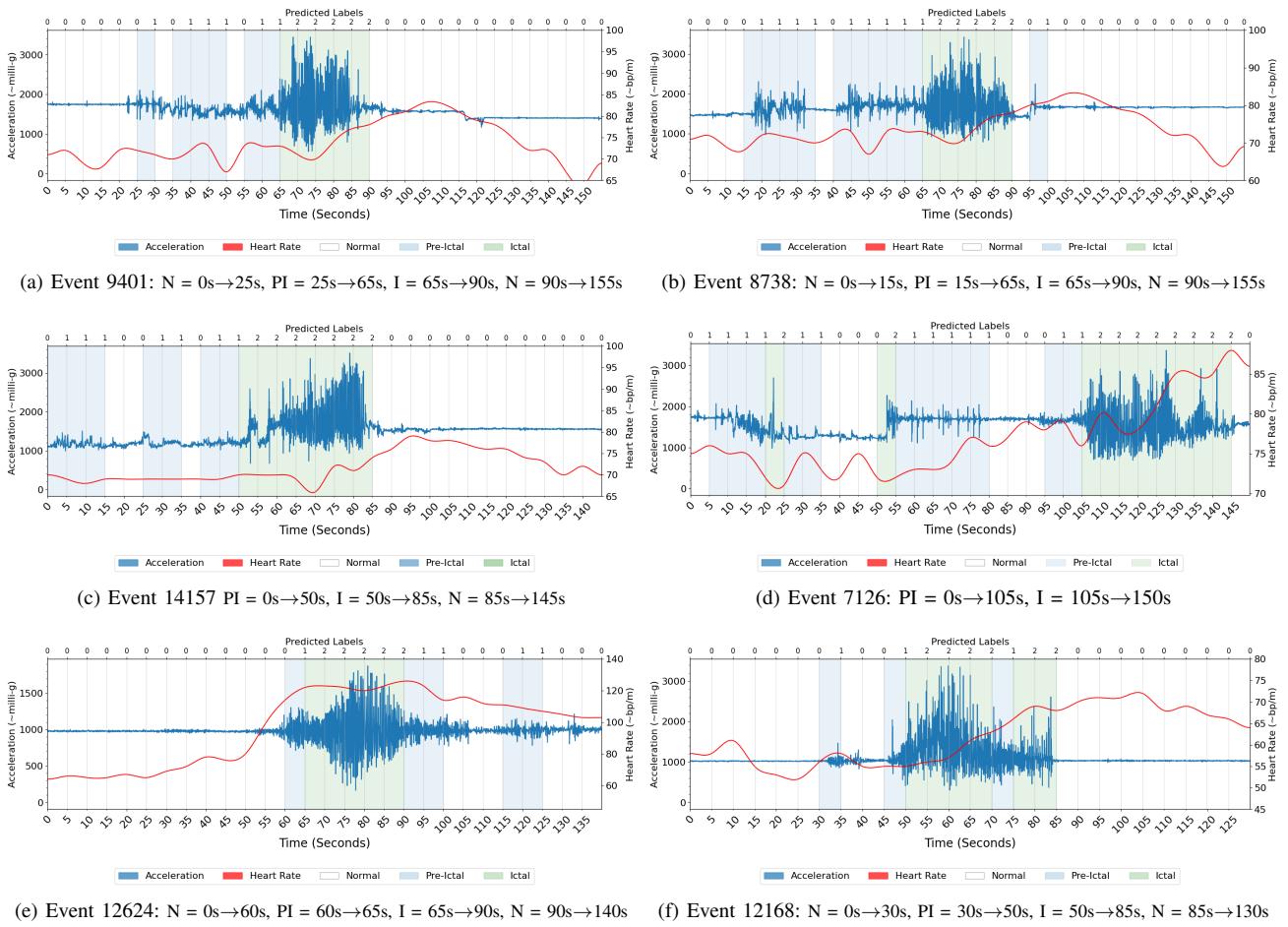
Fig. 8(a) shows the results of Fold 1, where the model correctly identified 572 and 541 timesteps from the Normal and Ictal classes, recording 89.10% and 90.77% of TPs. However, challenges were encountered when distinguishing the Pre-Ictal class, recording 89.34% of timesteps correctly. However 60 Pre-Ictal timesteps were

incorrectly labelled, with 3.55% as Normal, and 7.10% as Ictal.

In Fold 2 (see Fig. 8(b)), a similar pattern was observed. The model achieved high accuracy in detecting 542 (89.74%) Ictal timesteps. Furthermore, consistency was observed when classifying Normal and Pre-Ictal classes, recording a TP score of 88.48% and 89.77%, respectively. Notably, the model only misclassified 3 (0.50%) timesteps as Normal, indicating a clear distinction between Ictal-based movements and human movements represented by the Normal class.

In Fold 3 (see Fig. 8(c)), further consistency was observed, correctly detecting 90.18%, 87.14% and 91.00% of TP timesteps for Normal, Pre-Ictal and Ictal respectively. Nevertheless, challenges were encountered when distinguishing the Pre-Ictal class, with 76 FPs incorrectly labelled.

Fig. 8(d) shows the results of Fold 4, where the model exhibited consistent performance scores, identifying 537 (91.48%) Ictal timesteps. Further consistency was observed, detecting 88.54% of Normal timesteps and 89.26% of Pre-Ictal timesteps respectively.



**Fig. 9:** GTC Experiments: The lower *x*-axis represents time in 5-second intervals (e.g., Timesteps 0s → 10s represents the timesteps from 0 to 10 seconds). The *y*-axis denotes the acceleration (left) and heart rate (right) signals which are measured in milli-g and beats per minute respectively. The upper *x*-axis shows the predicted class labels for each timestep, where 0, 1, and 2 correspond to Normal (white), Pre-Ictal (blue), and Ictal (green) phases. The clinically annotated ground truth labels are provided as sub-captions for each figure, indicating the ground truth labels, defining the timesteps that start and end each phase where s = seconds, N = Normal, PI = Pre-Ictal, and I = Ictal.

**TABLE IV:** Classification Analysis Distributed by Ictal Class – Notation = TPR = True Positive Rate, TNR = True Negative Rate, PPV = Positive Predicted Value, NPV = Negative Predicted Value

Metric	Normal	Pre-Ictal	Ictal	Weighted Average
TPR	0.9564	0.8325	0.9111	0.8999
TNR	0.9461	0.9505	0.9527	0.9494
PPV	0.8976	0.8957	0.9050	0.8994
NPV	0.9777	0.9174	0.9559	0.9496
FPR	0.0539	0.0495	0.0473	0.0502
FNR	0.0436	0.1675	0.0889	0.1136

Finally in Fold 5 (see Fig. 8(e)), the model identified 543 instances of the Ictal class, recording a TP score of 90.50%. The model also detected 570 and 507 TPs, accounting for 89.76% and 89.58% for Normal and Pre-Ictal classes. For Pre-Ictal we observed 23 (4.06%) FPs for Normal, and 36 (6.36%) FPs for Ictal. The model also recorded a FP score of 0.50% and 9.00% for Normal and Pre-Ictal respectively.

Overall, the analysis of each fold reveals consistently high TP scores for the Ictal class, with an average TP score of approximately 90.69% across all instances of *k*.

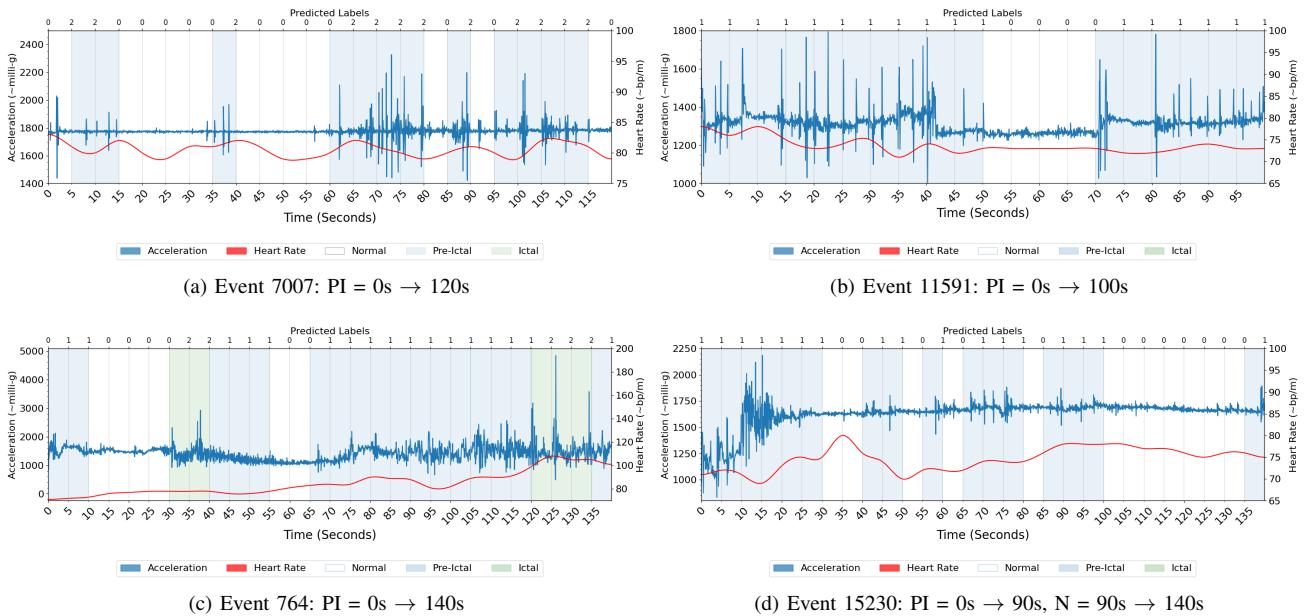
These results indicate the model’s proficiency in detecting the

convulsive movements associated with high-amplitude seizure manifestations. However, challenges were observed in distinguishing between Normal and Pre-Ictal classes, as evidenced by the number of FPs observed for the Pre-Ictal results for each fold. The number of FPs predicted for the Pre-Ictal class suggest a an overlap in feature space with the Normal and Ictal classes. The results of these experiments highlight the model’s strong and consistent performance in identifying Ictal instances, performing well on each interdependent test set with the lowest TP score detected across all folds at 87.14%.

To address these challenges, we evaluated each class using a comprehensive set of classification metrics. Table 2 summarises the performance for each ictal class, highlighting the model’s strengths and areas for improvement in classifying the ictal phase.

The model recorded a TPR for all classes, exceeding 0.83. However, a degree of uncertainty was observed between the Normal and Pre-Ictal classes, as evidenced by the low TPR for Pre-Ictal (0.8325) compared to the Normal (0.9564) and Ictal (0.9111) classes. Additionally, a consistently high True Negative Rate (TNR) across all classes was recorded, underscoring the model’s ability to identifying non-seizure movements, recording an average TNR score of 0.9494.

Despite consistently high TPRs and TNRs, our model exhibited a trade-off with PPV, which has an average score of 0.8994 across all classes, reflecting the proportion of FPs that were recorded. However,



**Fig. 10:** Aura/Focal Seizure Experiments: For a full description of the experiments, annotations and class labels see the caption in Fig. 9

the False Positive Rate (FPR) scores for each class are relatively low ( $< 0.053$ ) and demonstrated consistent generalisation across all classes. These findings emphasise the model's performance across both positive and negative assessment metrics and identify areas where the model encountered uncertainty.

### C. Ictal Phase Detection

This subsection demonstrates the Ictal Phase Detection technique on a representative subset of events. Fig. 9 shows the results of six experiments where each event is labelled as GTC seizure. The upper  $x$ -axis denotes the predicted class labels, where 0 represents Normal (White), 1 represents Pre-Ictal (Blue), and 2 represents Ictal (Green) respectively. Fig. 9(a) to Fig. 9(d) are show signs of a FBTC seizure as there is a clear pre-ictal phase indicative of an aura or myoclonic movements. In contrast, Fig. 9(e) and Fig. 9(f) showcase the variability of different events labelled as GTC seizure, as these seizures have no predefined pre-ictal phase; instead, the first sign of pre-ictal movement occurs seconds before onset.

From these results, it is evident that the proposed technique can distinguish the convulsive movements of the Ictal phase from the movements observed in the Normal and Pre-Ictal phases. The Ictal phase was correctly classified across all 6 plots, accurately detecting  $\approx 92.1\%$  (35/38) of Ictal timesteps. A degree of uncertainty was observed in Fig. 9(d), where 2 Pre-Ictal classes were incorrectly labelled as Ictal. However, the onset and cessation timesteps of each event was clearly detected, demonstrating this technique's potential in detecting the convulsive movements of the Ictal phase.

Fig. 10 show the results of several experiments that detect the Pre-Ictal class. These events were labelled as Aura/Focal seizure, and were indicative of low-amplitude movements. The movement and heart rate signals in these events are more challenging to distinguish than high-amplitude movement compounded by an elevated heart rate as observed in the Ictal class. Convulsive movements are indicative of rapid fluctuations, high amplitude movement and traversing a significant distance in terms of vector magnitude. However, Pre-Ictal events are often compounded by sequential timesteps with minimal or near-zero movement (Fig. 10(a)), where the recorded acceleration

signal remains flat, indicating negligible variations in movement and heart rate.

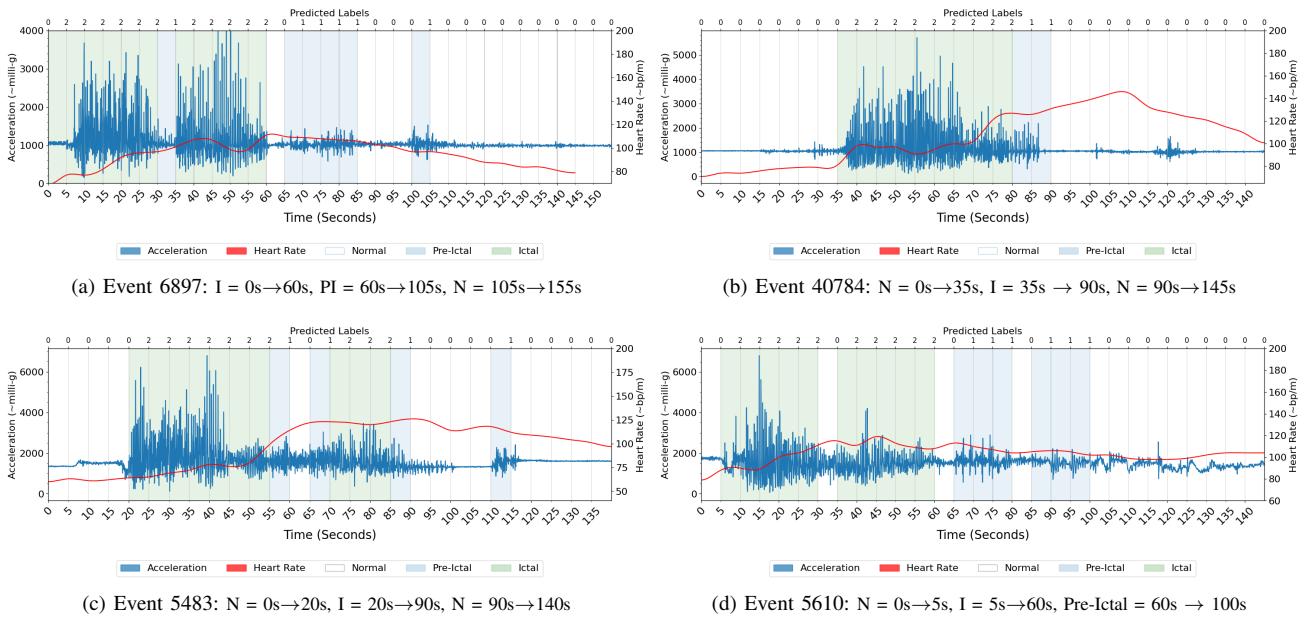
In Fig. 10(a), the proposed model distinguished the movements as Pre-Ictal, accurately labelling each timestep that showed waveform fluctuations in acceleration, compounded by elevated heart rate. However, the model did not label timesteps 15s - 35s and 40s - 60s correctly, due to the minimal variation in movements and reoccurring decline in heart rate between 84 - 80 bpm.

In Fig. 10(b), the model detected 80% (16/20) of the timesteps. However, for timestep 50s - 70s, we can see that the acceleration and heart rate waveforms exhibited minimal fluctuations, indicating a period of relative stability. The signals observed in these timesteps are indicative of an overlap in feature space between the Normal and Pre-Ictal classes.

Fig 10(c) showcases acceleration waveforms similar to those observed during the Ictal phase of GTC seizure, although with a much lower magnitude of acceleration. The waveforms are generally hover below the 2000 milli-g threshold and shown sudden spikes which are densely clustered compared to the Pre-Ictal movements in Fig. 10(a) and Fig. 10(b). However, in Fig. 10(c), a degree of overlap is observed between Pre-Ictal and Ictal class (30s - 40s and 55s - 65s) with acceleration spikes exceeding  $\approx 2000$  milli-g in each instance. Furthermore, an overlap with the Normal class (10s- 30s and 120s - 135s) was also observed, indicative of the varying fluctuations that are encountered in the Pre-Ictal class.

In Fig. 10(d), the proposed model accurately identified high-amplitude Pre-Ictal movement, even when exceeding 2000 milli-g. Between timesteps 0s and 20s, a burst of Pre-Ictal movement is evident, accompanied by a 25% increase in heart rate. Subsequently, movement intensity diminishes, with heart rate returning to 70 - 75 bpm baseline. Acceleration waveforms then oscillate between Normal and Pre-Ictal phases, stabilising over time. Instances of low-amplitude movement occur at timesteps 30s - 40s, 50s - 55s, and 60s - 65s were undetected due to minimal fluctuations in acceleration and heart rate.

These plots underscore the challenges we encountered when classifying the Pre-Ictal class. This is because the model must account for the broad spectrum of normal, non-epileptic movements, while



**Fig. 11:** Other Seizure Experiments: For a full description of the experiments, annotations and class labels see the caption in Fig. 9

addressing the variability inherent in the different types of Auras and subtle myoclonic movements that precede seizure onset.

Fig. 11 illustrates several experiments conducted on events labelled as Other Seizure. The structure of these acceleration waveforms differs from those shown in Fig. 9. In most instances of the Other Seizure class, there is minimal low-amplitude movement indicative of the Pre-Ictal phase seen in other generalised events. Furthermore, these seizures exhibit higher intensity compared to those analysed in Fig. 9 where the maximum acceleration peaked at approximately 3000 milli-g. However, in Fig. 11(a) and Fig. 11(b) acceleration exceeds 4000 milli-g, while in Fig. 11(c) and Fig. 11(d) the acceleration waveforms exceed 6000 milli-g, representing a doubling of intensity compared to all plots in Fig. 9. There are also notable heart rate fluctuations in Fig. 11 where each plot exceeds 100 beats per minute, with a notable increase occurring at the time of onset and continuing until cessation. Overall the proposed model accurately detected the Ictal Phase in Fig. 11, approximately detecting 91.6% (11/12), 81.81% (9/11), 73.33% (11/15) and 90.91% (10/11) for Fig. 11(a) - Fig. 11(d) respectively.

## V. DISCUSSION

In this section, we discuss our investigation into Ictal Phase Detection, a non-EEG technique designed to distinguish the different phases of the ictal state. For this study, we hypothesised that it was possible to differentiate high-amplitude convulsive movement characteristic of the ictal phase from low-amplitude myoclonic movement indicative of the pre-ictal phase, or aura.

To address challenges in multimodal detection, we introduced AMBER, an attention-guided deep learning architecture. AMBER constructs independent branches for each modality, generating attention weights that are subsequently fused into a single branch. The fusion process is facilitated by our custom fusion layer, called Enhanced Residual Fusion. This layer combines inputs from the individual branches, leveraging a multi-head attention mechanism to extract emphasised features and incorporating residual connections to enhance feature propagation. The novelty of AMBER lies in its ability to extract distinct features from independent modalities and fuse them effectively, emphasising key inputs while minimising redundancy.

Our experiments demonstrated significant performance improvements using the Enhanced Residual Fusion layer, compared to instances of the same architecture without this feature as shown in Table III.

### A. Summary of Results

The results of our experiments underscore the effectiveness of the AMBER model combined with the Ictal Phase Detection technique. We recorded an overall accuracy of 0.8995, an  $f1$ -score of 0.8987, and recorded Cohen's Kappa and MCC scores of 0.8492 and 0.8499 respectively. Continuous levels of performance access each fold, particularly the Cohen's Kappa and MCC scores surpassing 0.84, underscore the model's consistent generalisation when classifying positive and negative classes.

From the experiments conducted, we observed a trend in the models ability to correctly identify non-seizure movements from the Normal class, reflected in the models TNR, which is consistently above above 94% across all folds (see Table IV). These results indicate that the model has learned to distinguish non-seizure movements. Additionally, our analysis of the models TPR across all folds revealed insights into the model's performance. The model achieved a TPR of approximately 0.9564 for Normal, 0.8325 for Pre-Ictal and 0.9111 for Ictal. These TPR scores are compounded by the models PPV, ranging between 0.8976 to 0.9050, indicating a high level of confidence in the model's ability to correctly identify positive classes. Amber also recorded a low FPR of 0.0539, 0.0495, and 0.0473, coupled with a FNR of 0.0436, 0.1675, and 0.0889, respectively. These results show that the model can effectively distinguish between TP and FP predictions, indicating the models ability to differentiate between Pre-Ictal and Ictal timesteps while minimising the occurrence of false alarms in each event.

From our analysis of GTC events, the model demonstrated an approximate 91% accuracy in detecting the exact timesteps where seizure onset and cessation occurred (see Fig. 9(a), Fig. 9(b), Fig 9(d)). Moreover, in events categorised as Other, the model achieved an 89% detection rate, identifying the timestep where onset (Ictal class) began. These results highlight the model's capability to discern consistent low-amplitude movement followed by a sudden spike

indicative of seizure onset and generalised activity (see Fig. 11(b) - Fig. 11(c)).

### B. Observations and Findings

An analysis of our results show that there was an observable increase in cardiac manifestations during the pre-ictal and ictal phases, especially in seizures labelled as Other. Consistent with existing literature [36] [37], ictal-tachycardia manifestations were observed prior to seizure onset in a significant portion of the events. Of the 94 events employed for this study, 39% ( $N=36/94$ ) of seizures exhibited ictal cardiac manifestations, characterised by a heart rate increase that exceeded 100 beats per minute, demonstrating ictal tachycardia manifestations. Subsequently, when events are categorised by type, we can see that only one Aura event exceeded the ictal tachycardia threshold, however this increased to 24% in GTC events and 70% in seizures labelled as Other.

Secondary observations show that a heart rate increase of  $\approx 10\%$  was observed from the initial timestep where seizure onset was detected. In most cases, cardiac activity increased after the initial 15-20 seconds of convulsive activity following onset. Once clonic activity ceased, the heart rate signal stabilised, returning to the pre-ictal baseline. Similar patterns relating to cardiac activity and acceleration were consistently observed across different events, which enabled our model to leverage pattern recognition techniques to accurately classify the ictal phases.

### C. Summary

In summary, the observed uncertainty in the Pre-Ictal phase can be attributed to the broad spectrum of movements this class encapsulates. Challenges arise in differentiating high-amplitude Normal movements from low-amplitude pre-Ictal movements, and high-amplitude pre-Ictal movement from low-amplitude Ictal movements. This differentiation of class types introduces a degree of uncertainty when classifying the pre-Ictal class, further compounded by the Normal class accounting for Post-ictal and everyday (Inter-Ictal) human movement. However, it should be noted that despite these challenges, the model demonstrates proficiency in detecting generalised events preceding their onset through Pre-Ictal detection. Rather than predicting a seizure at the first signs of high-amplitude movement, our model assessed whether the analysed signal is indicative of high amplitude or low amplitude movement. This, combined with a multimodal approach that leverages heart rate readings, presents a robust seizure detection technique that has real-world application.

Comparatively, traditional studies in existing literature often measure accuracy based on whether the model detected the event. In contrast, our approach classifies each 5-second timestep of an event, producing an overall performance score based on the number of correctly distinguished timesteps as Normal, Pre-Ictal, or Ictal. This approach allows for a more comprehensive evaluation of the model's performance, providing insights into its ability to distinguish between different phases of seizure activity accurately.

### D. Limitations

In this study several limitations were encountered with the main limitation being the use of a single class label (Normal) to represent both common human movements during the inter-ictal phase (the period between seizures) and the post-ictal phase (the recovery period following a seizure). The decision to use a single class label was influenced by Abdulhay *et al.* [39], who classified EEG signals characterised by disturbances in temporal and spectral content using identical class labels. However, in the context of non-EEG detection,

this approach is constrained by the diversity of potential movements during the inter-ictal phase. While the use of these class labels has been established in literature, in a non-EEG context, a single class label does not account for the diverse range of movements observed during the pre and post-seizure periods.

Another limitation was encountered when annotating non-EEG data into ictal phases. Due to the absence of established methodologies, developing a suitable technique for quantifying the different timesteps across each event proved challenging. This process necessitated extensive collaboration between the technical and clinical team members, given the lack of existing benchmarks. Thus, achieving consensus and refining our approach required considerable time and effort.

### E. Future Work

We hypothesise that augmenting the number of input classes could improve the degree of uncertainty encountered when distinguishing between Normal and Pre-Ictal classes. Furthermore, applying several classes to account for common human movements could further refine classification of the pre and post-ictal phases. This would improve the model's ability to distinguish different types of normal movement (lying down, walking, brushing teeth), while allowing the model to isolate specific high and low amplitude movements characteristic of the Pre-Ictal and Ictal class.

## VI. CONCLUSION

In summary, this study has made advancements in non-EEG seizure detection through the introduction of the Ictal Phase Detection technique. This technique has shown how high-amplitude and low-amplitude movements indicative of the pre-ictal and ictal phases can be distinguished. We feel this has the potential to reduce the number of false alarms and enable future detection methodologies to identify the convulsive movements that occur during generalisation, rather than triggering a false alarm at the first indication of random, low-amplitude movement.

In conclusion, this study contributes a new detection technique and model designed for multimodal seizure detection. We feel this lays the groundwork for further advancements in non-EEG based seizure detection methodologies.

We are committed to advancing this research and ensuring that the AMBER model is accessible to the public to foster collaboration and innovation in the field. To facilitate progress in non-EEG seizure detection, we are dedicated to sharing the findings, models, and codebase from this study. We believe AMBER holds potential applications in other fields related to one-dimensional time series research. Therefore, AMBER will be made publicly available under the MIT license, inviting collaboration to drive further advancements.

## ACKNOWLEDGEMENTS

We extend our sincere gratitude to Open Seizure Detector for providing early access to the Open Seizure Database. Their steadfast dedication to open research will greatly enhance the progress of non-EEG seizure detection research. We wholeheartedly acknowledge and commend their unwavering support and commitment to advancing scientific exploration and knowledge within the epilepsy community.

Additionally, we express our appreciation to the participants of the Open Seizure Database. Without their dedication and commitment, this research would not be feasible.

## AVAILABILITY OF DATA AND MATERIALS

We are committed to advancing non-EEG detection research and are pleased to announce the availability of our study's results, models, and code base. The AMBER model is now distributed under the MIT license as open-source [41]. We invite any technical contributions aimed at enhancing the model's performance and extending its application in non-EEG seizure detection. The full code base and results for this project are now accessible on GitHub [42].

## REFERENCES

- [1] E. Beghi, G. Giussani, and J. W. Sander, "The natural history and prognosis of epilepsy", *Epileptic Disorders: International Epilepsy Journal with Videotape*, vol. 17, no. 3, pp. 243–253, Sep. 2015.
- [2] H. Anwar, Q. U. Khan, N. Nadeem, I. Pervaiz, M. Ali, and F. F. Cheema, "Epileptic seizures", *Discoveries*, vol. 8, no. 2, pp. e110, Jun. 2020.
- [3] S. T. Sarmast, A. M. Abdullahi, and N. Jahan, "Current Classification of Seizures and Epilepsies: Scope, Limitations and Recommendations for Future Action", *Cureus*, vol. 12, no. 9, pp. e10549, Sep. 2020.
- [4] A. Kumar, K. Maini, K. Arya, S. Sharma, "Simple Partial Seizure", *StatPearls*, [Online]. Available: <https://www.ncbi.nlm.nih.gov/books/NBK500005/>, [Accessed: Sep. 15, 2022].
- [5] C. DeGiorgio, A. Curtis, D. Hertling, and B. Moseley, "Sudden Unexpected Death in Epilepsy: Risk Factors, Biomarkers and Prevention", *Acta Neurologica Scandinavica*, vol. 139, no. 11, Dec. 2018, doi: 10.1111/ane.13049.
- [6] O. Sveinsson, T. Andersson, P. Mattsson, S. Carlsson, and T. Tomson, "Clinical risk factors in SUDEP", *Neurology*, vol. 94, no. 4, Jan. 2020, doi: 10.1212/WNL.0000000000008741.
- [7] R. Friedman, and J. Kazl, "Seizure Detection and SUDEP Prevention", *Bryn Mawr Communications*, [Online]. Available: <https://practicalneurology.com>, [Accessed: Dec. 19, 2022].
- [8] R. M. J. Cook, T. J. O'Brien, S. F. Berkovic, M. Murphy, A. Morokoff, G. Fabinyi *et al.*, "Prediction of seizure likelihood with a long-term, implanted seizure advisory system in patients with drug-resistant epilepsy: a first-in-man study", *The Lancet Neurology*, vol. 12, no. 6, pp. 563–571, Jun. 2013, doi: 10.1016/S1474-4422(13)70075-9.
- [9] R. J. Lamberts, R. D. Thijss, A. Laffan, Y. Langan, and J. W. Sander, "Sudden unexpected death in epilepsy: people with nocturnal seizures may be at highest risk", *Epilepsia*, vol. 53, no. 2, pp. 253–257, Feb. 2012, doi:10.1111/j.1528-1167.2011.03360.x
- [10] M. Dümpelmann, "Early seizure detection for closed loop direct neurostimulation devices in epilepsy", *Journal of Neural Engineering*, vol. 16, no. 4, p. 041001, Aug. 2019, doi: 10.1088/1741-2552/ab094a.
- [11] S. Beniczky, T. Polster, T. W. Kjaer, and H. Hjalgrim, "Detection of generalised tonic-clonic seizures by a wireless wrist accelerometer: a prospective, multicenter study", *Epilepsia*, vol. 54, no. 4, p. e58–61, Apr. 2013, doi: 10.1111/epi.12120.
- [12] K. Cuppens, P. Karsmakers, A. Van de Vel, B. Bonroy, M. Milosevic, S. Luca *et al.*, "Accelerometry-based home monitoring for detection of nocturnal hypermotor seizures based on novelty detection", *IEEE journal of biomedical and health informatics*, vol. 18, no. 3, p. 1026–1033, May. 2014, doi: 10.1109/JBHI.2013.2285015.
- [13] J. Lockman, R. S. Fisher, and D. M. Olson, "Detection of seizure-like movements using a wrist accelerometer", *Epilepsy & Behavior: E&B*, vol. 20, no. 4, p. 638–641, Apr. 2011, doi: 10.1016/j.yebeh.2011.01.019.
- [14] J. Pordoy, Y. Zhang, N. Matoorian, and M. Zolgharni, "Predicting epileptic seizures with a stacked long short-term memory network", *International Journal of Automation, Artificial Intelligence and Machine Learning*, vol. 1, no. 1, p. 93–108, Oct. 2020.
- [15] S. Zia, A. N. Khan, M. Mukhtar, S. E. Ali, J. Shahid, and M. Sohail, "Detection of Motor Seizures and Falls in Mobile Application using Machine Learning Classifiers", *Artificial Intelligence, and Communications Technology*, , p. 62–68,Jul. 2020, doi: 10.1109/IAICT50021.2020.9172028.
- [16] P. Vergara, J. Villar, E. Marín, M. Menéndez González, and J. Sedano, "Pre-Clinical Study on the Detection of Simulated Epileptic Seizures," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 24, no. 6, pp. 33–46, Dec. 2016, doi: 10.1142/S0218488516400092.
- [17] I. Conradien, S. Beniczky, P. Wolf, T. W. Kjaer, T. Sams, and H. B. D. Sorensen, "Automatic multi-modal intelligent seizure acquisition (MISA) system for detection of motor seizures from electromyographic data and motion data, *Computer Methods and Programs in Biomedicine*, 107(2), 97–110, Aug. 2012, doi:10.1016/j.cmpb.2011.06.005
- [18] G.-Q. Zhang, L. Cui, S. Lhatoo, S. U. Schuele, and S. S. Sahoo, "MEDCIS: Multi-Modality Epilepsy Data Capture and Integration System", *AMIA Annu Symp Proc*, vol. 2014, pp. 1248–1257, Nov. 2014.
- [19] J. Wagenaar, G. Worrell, Z. Ives, M. Dümpelmann, B. Litt, and A. Schulze-Bonhage, "Collaborating and sharing data in Epilepsy Research," *J. Clin. Neurophysiol. Off. Publ. Am. Electroencephalogr. Soc.*, vol. 32, no. 3, pp. 235–239, Jun. 2015.
- [20] J. C. Sackellares, D. Shiu, J. C. Principe, M. C. K. Yang, L. K. Dance, W. Suharitdamrong *et al.*, "Predictability Analysis for an Automated Seizure Prediction Algorithm", *Journal of Clinical Neurophysiology*, vol. 23, no. 6, pp. 509, Dec. 2006.
- [21] V. Željković, V. Valev, C. Tameze, and M. Bojic, "Pre-Ictal-phase detection algorithm based on one-dimensional EEG signals and two-dimensional formed images analysis," *2013 International Conference on High Performance Computing & Simulation (HPCS)*, pp. 607–614, Jul. 2013, doi: 10.1109/HPCSim.2013.6641477.
- [22] A. Eftekhar, W. Juffali, J. El-Imad, T. G. Constantinou, and C. Toumazou, "Ngram-Derived Pattern Recognition for the Detection and Prediction of Epileptic Seizures", *PLOS ONE*, vol. 9, no. 6, pp. e96235, Jun. 2014, doi: 10.1371/journal.pone.0096235.
- [23] B. Sharif and A. H. Jafari, "Prediction of epileptic seizures from EEG using analysis of ictal rules on Poincaré plane," *Computer Methods and Programs in Biomedicine*, vol. 145, pp. 11–22, Jul. 2017, doi: 10.1016/j.cmpb.2017.04.001.
- [24] R. D. Thijss, P. Ryvlin, and R. Surges, "Autonomic manifestations of epilepsy: emerging pathways to sudden death?", *Nature Reviews Neurology*, vol. 17, no. 12, pp. 774–788, Dec. 2021, doi: 10.1038/s41582-021-00574-w.
- [25] The Open Seizure Database (OSDB), Open Seizure Detector, Hartlepool, UK, 2022. [Online] Available: <https://github.com/OpenSeizureDetector/OpenSeizureDatabase>
- [26] Open Seizure Detector, Open Seizure Detector, Hartlepool, UK, 2013. [Online] Available: <https://www.openseizuredetector.org.uk>
- [27] University of West London, University of West London, London, UK, 2022. [Online] Available: <https://www.uwl.ac.uk>
- [28] J. Pordoy, The Open Seizure Database Facilitating Research Into Non-EEG Seizure Detection. (2023,8), [https://www.techrxiv.org/articles/preprint/The\\_Open\\_Seizure\\_Database\\_Facilitating\\_Research\\_Into\\_Non-EEG\\_Seizure\\_Detection/23957625](https://www.techrxiv.org/articles/preprint/The_Open_Seizure_Database_Facilitating_Research_Into_Non-EEG_Seizure_Detection/23957625)
- [29] E. Coto and A. Zissermann, VGG Image Classification (VIC) Engine. (<http://www.robots.ox.ac.uk/vgg/software/vic/>, 2017), Version: X.Y.Z, Accessed: 31st Aug 2023.
- [30] L. Vilella, N. Lacuey, J. P. Hampson, L. Zhu, S. Omidi, M. Ochoa-Urrea *et al.*, "Association of Peri-ictal Brainstem Posturing With Seizure Severity and Breathing Compromise in Patients With Generalized Convulsive Seizures", *Neurology*, vol. 96, no. 3, pp. e352 – e365, Jan. 2021, doi:10.1212/WNL.000000000000112.
- [31] A. Ullah, S. U. Rehman, S. Tu, R. M. Mahmood, Fawad and M. Ehatisham-ul-haq, "A Hybrid Deep CNN Model for Abnormal Arrhythmia Detection Based on Cardiac ECG Signal", *Sensors (Basel, Switzerland)*, vol. 21, no. 3, pp. 951, Feb. 2021, doi: 10.3390/s21030951.
- [32] G. Jana, R. Sharma, and A. Agrawal, "A 1D-CNN-Spectrogram Based Approach for Seizure Detection from EEG Signal", *Procedia Computer Science*, vol. 167, pp. 403–412, 2020, doi:10.1016/j.procs.2020.03.248.
- [33] J. Zhang, Y. Zeng, and B. Starly, "Recurrent neural networks with long term temporal dependencies in machine tool wear diagnosis and prognosis", *SN Applied Sciences*, vol. 3, no. 4, pp. 442, Mar. 2021, doi.org/10.1007/s42452-021-04427-5.
- [34] J. Raitoharju, "Convolutional neural networks", in *Deep Learning for Robot Perception and Cognition* 1<sup>st</sup>ed., vol. 3, J. Academic Press, 2022, pp. 35–69.
- [35] A. Ghatak, "Optimization", in *Deep Learning with R* 1<sup>st</sup>ed., vol. 1, Singapore, Springer, 2022, pp. 103–147, doi:10.1007/978-981-13-5850-0-5
- [36] W. Chen, C. L. Guo, PS. Zhang, "Heart rate changes in partial seizures: analysis of influencing factors among refractory patients", *BMC Neurol*, vol. 14, no. 153, Jun. 2014, doi.org/10.1186/1471-2377-14-135.

- [37] S. S. Allana, H. N. Ahmed, K. Shah, A. F. Kelly, "Ictal bradycardia and atrioventricular block: a cardiac manifestation of epilepsy.", *Oxford medical case reports*, vol. 2014, no. 2, pp. 33–35, May 2014, doi.org/10.1093/omcr/omu015.
- [38] K. S. Pedersen, W. R. Bamlet, A. L. Oberg, M. de Andrade, M. E. Matsumoto, H. Tang *et al.*, "Leukocyte DNA methylation signature differentiates pancreatic cancer patients from healthy controls", *PLoS One*, vol. 6, no. 3, e18223, Mar 2011, doi: 10.1371/journal.pone.0018223.
- [39] E. Abdulhay, M. Alafeef, A. Abdelhay, A. Al-Bashir, Classification of Normal, Ictal and Inter-ictal EEG via Direct Quadrature and Random Forest Tree, *J Med Biol Eng*, vol. 37(6), pp. 843–857, Jun 2017.
- [40] A. Mostafa Khalil *et al.*, "Real-Time System Prediction for Heart Rate Using Deep Learning and Stream Processing Platforms," Published on February 23, 2021.
- [41] AMBER" GitHub, (2024), [Online]. [Online] Available: <https://github.com/jpordoy/AMBER>
- [42] Ictal Phase Detection, GitHub, (2024),[Online]. [Online] Available: <https://github.com/jpordoy/Ictal-Phase-Detection>