

# 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 have been expertly annotated by a clinician, who labelled each 5-second timestep as 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 Residual Fusion layer, where the extracted features are combined into a fused representation and propagated through two densely connected blocks. The dataset was split by event, ensuring no overlap between events in the training and testing subsets. The model was trained using k-fold cross-validation, where k-1 folds were used for training and the remaining fold for validation. The results of these experiments highlight the effectiveness of Ictal-Phase Detection, with the model achieving an accuracy and f<sub>1</sub>-score of 0.9027 and 0.9035, respectively, on unseen test data. Notably, the model exhibited consistent generalisation, recording a True Positive Rate of 0.8342, 0.9485 and 0.9118 for the Normal, Pre-Ictal, and Ictal classes respectively, and an average False Positive Rate of 0.0502. In conclusion, this study introduces a new multimodal seizure detection technique and model that reduces the false alarm window and differentiates high and low-amplitude convulsive movements, laying the groundwork for further advancements in non-EEG-based 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 Epilepsy, epileptic seizures are divided into four sub-categories: focal, generalised, focal to bilateral tonic-clonic, and unknown [3]. Focal,

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).

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 the ictal state, particularly in focal seizures originating from the temporal lobe and generalising into a FBTC seizure. However, the absence of publicly available 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 novel seizure detection technique (Ictal-Phase Detection) capable of distinguishing between high-amplitude convulsive movement typical of the ictal phase and low-amplitude myoclonic movement observed in the pre-ictal phase.
- 2) AMBER is a new deep learning model designed for non-EEG, multimodal seizure detection. The model employs multiple branches (multi-branching) to independently process features, which are then fused into a single representation and propagated through a dense network, enabling effective multimodal detection for complex biomedical tasks.
- 3) The proposed model achieved accuracy,  $f_1$ , Cohen's Kappa, and MCC scores of 0.9027, 0.9035, 0.8498, and 0.8519, respectively. These results underscore the model's ability to effectively distinguish between positive and negative classes while accurately classifying the different phases of the ictal state.
- 4) We conducted a per-participant evaluation of the model's performance by grouping events in the test set by user and measuring the classification performance for each participant. The model achieved an average accuracy score of 0.8412 for participants exclusive to the testing subset, demonstrating the models ability to generalise effectively on unseen data from participants with different types of epilepsy.

## 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 generalised onset seizures. While conventional non-EEG seizure detection techniques primarily focus on the classification of seizure and non-seizure states, a significant number 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-amplitude and low-amplitude movements. We hypothesise that there are quantifiable variances in acceleration and heart rate between the normal (seizure-free), pre-ictal, and ictal phases. While accelerometers easily capture high-amplitude convulsive movements, detecting pre-ictal, low-amplitude movements is more challenging. However, subtle heart rate elevations during the pre-ictal phase, in conjunction with low amplitude movements could provide indications of seizure onset. Thus, heart rate data will serve as a secondary modality to complement the acceleration data, enhancing the model's ability to detect different seizure manifestation.

To address this, we propose the development of a new model that can independently leverage sensor data from two or more modalities. We posit that independent feature extraction will improve predictive reliability as a true representation of each modality is extracted. These representations can then be fused to form a 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 our hypothesis, Open Seizure Detector granted the University of West London early access to the Open Seizure Database [25] [26] [27].

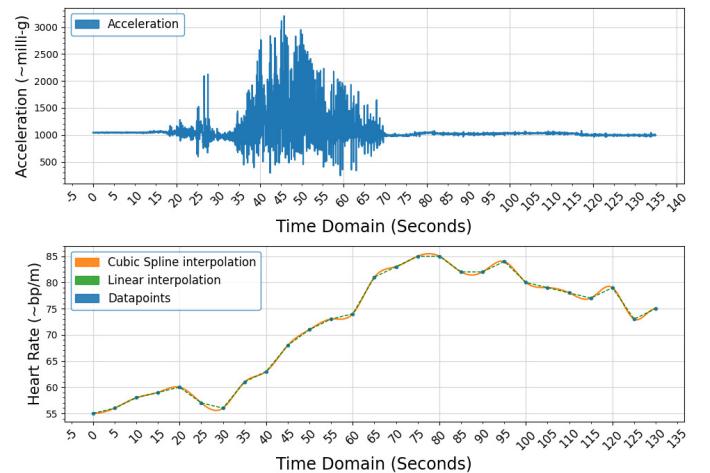
### A. The Open Seizure Database

This is the first study to use the newly developed Open Seizure Database (OSDB) - Version 3 [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.

### 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 dataset of 94 events, consisting 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 categorisation [28].

The final representative subset consists of non-EEG 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



**Fig. 1:** Event 15039 Time-series Plot: 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.

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 and can be expressed as  $A_t = \{a_1, a_2, \dots, a_{125}\}$  where  $a_n$  denotes the indexed acceleration datapoint within timestep  $t$ . The heart rate, data denoted as  $H$ , was recorded using a PPG sensor with a lower sample rate of 0.2 Hz, computing a single heart rate datapoint ( $h_1$ ) for each 5-second timestep, which can be expressed 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 the various phases of the ictal state, the clinical team annotated the timesteps for each event with a class label of 0, 1, or 2, corresponding to the normal, pre-ictal, and ictal phases respectively. It is important to note that, in a clinical or descriptive context, ictal phases are referred to in lowercase (normal, pre-ictal, and ictal), while the corresponding class labels are capitalised (Normal, Pre-Ictal, and Ictal).

The Normal class denotes movements outside of the convulsive scope, specifically during the inter-ictal and post-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 and clonic jerking 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 to guide the annotation process and can be expressed as:

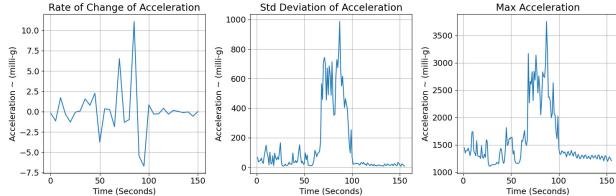
$$\text{Rate of Change of Acceleration} = \frac{A_t - A_{t-1}}{\Delta t} \quad (1)$$

$A_t - A_{t-1}$  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  $t$ .

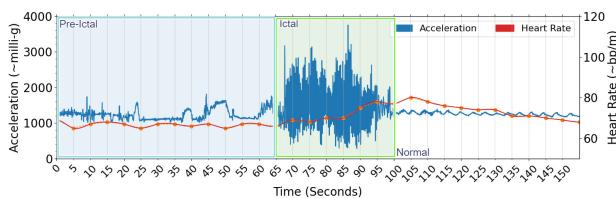
We conducted secondary computations to validate the results of Equation 1, calculating the average acceleration range, maximum acceleration, and standard deviation of acceleration in each timestep.

**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.

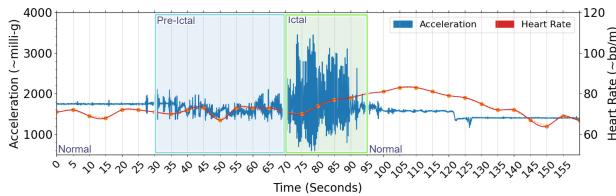
Class	Ictal Phase	Avg Acceleration Range [Milli-g]	Max Value [Milli-g]	STD [Milli-g]	Rate of Change [Milli-g]
0	Normal	800 - 1200	1500	0-50	$-1500 \leq A_t \leq 1000$
1	Pre-Ictal	1200 - 2000	2200	50-200	$-2000 < A_t \leq -1500$ or $1000 < A_t \leq 2000$
2	Ictal	2000 - 5000	7000	>200	$-2000 > A_t$ or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



**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 → 150s = Normal (Post-ictal)

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

The clinician then utilised these results along with partial video footage to annotate each 5-second timestep with a class label.

Table I presents the average thresholds calculated to guide the clinical team in annotating each timestep. Annotations were conducted using Oxford University's VGG Image Annotation software [29], 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 of 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 124 interpolated indices for  $H_t$ . For each instance of  $i$  within  $H_{t(i)}$ , a cubic polynomial can be calculated as:

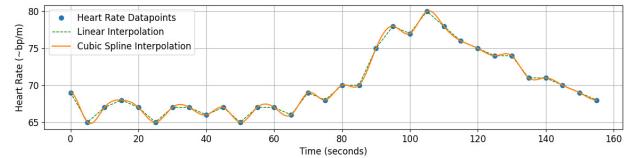
$$H_{t(i)} = a_i + b_i(t - t_i) + c_i(t - t_i)^2 + d_i(t - t_i)^3 \quad (2)$$

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  $H_t = \{h_1, h_2, \dots, h_{125}\}$ .

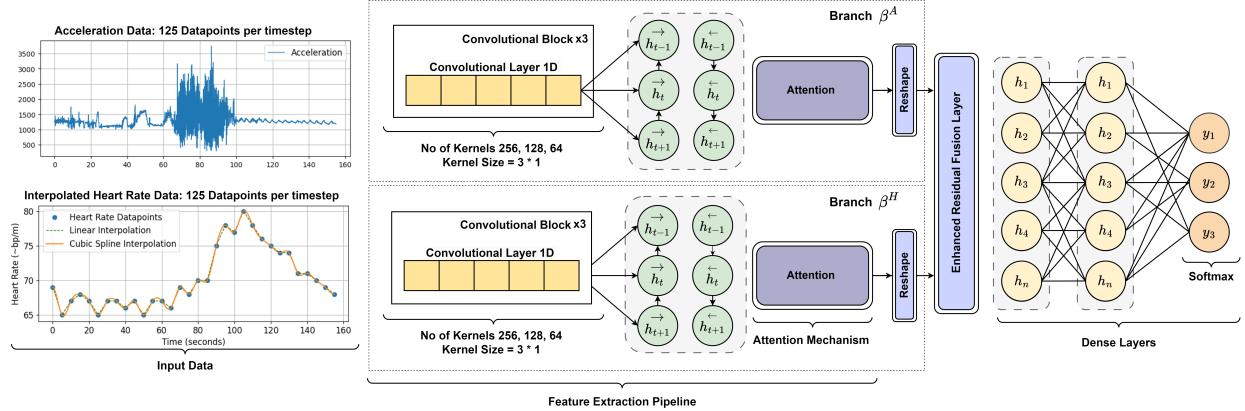
Since the heart rate datapoints are recorded using a PPG sensor at 5-second intervals, they serve as accurate cardiac biomarkers. While the original datapoints remain unchanged, interpolated values are used to align the lengths of  $A_t$  and  $H_t$ . It is important to note that the PPG signal cannot be substituted, as it directly reflects heart rate activity, while the interpolated values create a continuous signal between the PPG datapoints. This approach ensures the integrity of the heart rate signal is maintained, without negatively impacting the model's performance.

#### E. Preprocessing

In this subsection we describe the preprocessing techniques that were undertaken. Following the reshaping of our representative dataset, a class imbalance was observed in which the Normal class



**Fig. 4:** Cubic Spline Interpolation. Each timestep  $t$  corresponds to a single heart rate datapoint (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

accounted for 48.23% (1534 timesteps) of the overall dataset with the Pre-Ictal (901 timesteps) and Ictal (745 timesteps) classes accounting for 28.34% and 23.42% 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 the datapoints in each timestep. The result was a resampled dataset comprised of 4602 timesteps with a balanced distribution, where each class accounted for 33.3% ( $\approx 1534$  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 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 (4602, 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

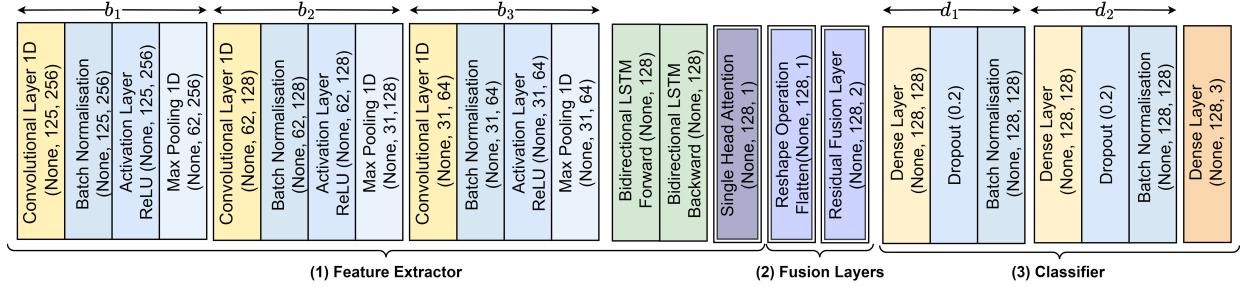
The following notation is a representation of input vector  $X_t$  as it traverses through the layers in the AMBER model (see Fig. 6). For each branch,  $X_t$  is passed as input through a dedicated feature extraction pipeline, comprised from three convolutional blocks denoted as  $b_1$ ,  $b_2$ , and  $b_3$ . 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:

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

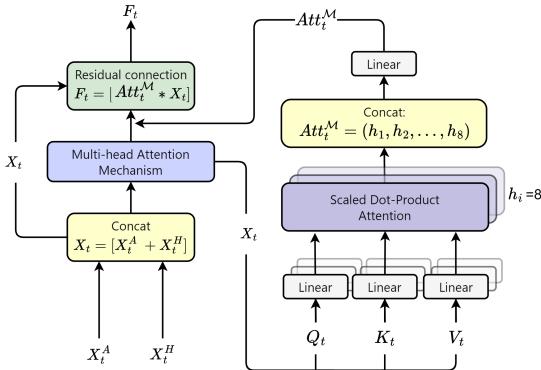
Let  $y_j^l$  represent the output of the  $j$ -th feature map in the  $l$ -th layer, computed by applying a sigmoid activation function to the sum of the convolutional operations between the filter of the  $l$ -th layer ( $w_{i,j}^l$ ) and the feature map from the previous layer ( $x_i^{l-1}$ ). Let  $b_j^l$  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  $t$ ,  $\overrightarrow{c}_t$  and  $\overrightarrow{h}_t$  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,



**Fig. 6:** Layered Representation of AMBER: (1) = Feature Extraction Pipeline with 3 \* convolutional blocks ( $b_1, b_2, b_3$ ), 1 \* BiLSTM and 1 \* Attention, (2) = Fusion Layer (Enhanced Residual Fusion), (3) = Classifier with 2 \* 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

three non-linear 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 ( $\vec{h}_t, \vec{h}_t$ ) from both directions are concatenated to form  $h_t$ . The cell states ( $\vec{c}_t, \vec{c}_t$ ) 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  $X_t = [\vec{h}_t, \vec{h}_t]$ .

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 enhanced Residual Fusion layer, designed to merge the outputs of two or more branches into a single fused representation (see Fig. 7). Attention weights  $X_t^A$  and  $X_t^H$  from branches  $\beta^A$  and  $\beta^H$  are passed in parallel to the 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  $Att_t^M$ , which consists of 8 attention heads, denoted as  $h_1$  through  $h_8$ . Each attention head operates on a low-dimensional (32-dimensions) projection space derived from the original concatenated inputs.

The attention heads,  $h_1$  through  $h_8$ , independently compute the  $Q_t$ ,  $K_t$ , and  $V_t$  matrices using learned weights. Since the computation of attention weights has already been explained in Equation 7, we do not repeat it here. The attention scores for each head are then calculated using a scaled dot-product, as shown in the following equation:

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

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

**TABLE II:** Comparative Analysis: Evaluation of performance scores obtained by the AMBER model compared to several state-of-the-art classification models and configuration. 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</i> <sub>1</sub>	AMBER	0.9027	0.9035	0.8498	0.8519	0.8914	0.8981
<i>m</i> <sub>2</sub>	Multi-Branched CNN-BiLSTM + Attention	0.8698	0.8805	0.8102	0.8204	0.8743	0.8999
<i>m</i> <sub>3</sub>	Multi-Branched CNN-BiLSTM	0.8613	0.8952	0.8397	0.8416	0.8845	0.8628
<i>m</i> <sub>4</sub>	1D-CNN + Attention	0.8653	0.8754	0.8007	0.8106	0.8778	0.8832
<i>m</i> <sub>5</sub>	Multi-Branched CNN + Adam	0.8647	0.8759	0.8004	0.8102	0.8731	0.8886
<i>m</i> <sub>6</sub>	1D-CNN Stacked BiLSTM	0.8631	0.8534	0.8552	0.8287	0.9079	0.8702
<i>m</i> <sub>7</sub>	Stacked BiLSTM	0.8454	0.8725	0.8012	0.8101	0.8532	0.8641
<i>m</i> <sub>8</sub>	1D-CNN + ADAM	0.8449	0.8722	0.8009	0.8104	0.8530	0.8676
<i>m</i> <sub>9</sub>	Stacked LSTM	0.8306	0.8617	0.7899	0.7996	0.8432	0.8515
<i>m</i> <sub>10</sub>	1D-CNN	0.8149	0.8455	0.7695	0.7802	0.8278	0.8415

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

$$\text{Att}_t^M = \text{concat}(h_1, h_2, \dots, h_8) \cdot W_O \quad (12)$$

To complete the Residual Fusion layer,  $\text{Att}_t^M$  is then multiplied with the original concatenated input vector  $X_t$  through a residual connection, outputting a fused vector denoted as  $F_t$ :

$$F_t = [X_t * \text{Att}_t^M] \quad (13)$$

The residual connection is used to retain key features from both branches while integrating the new information derived from  $\text{Att}_t^M$ . Thus, the Residual Fusion layer combines inputs from  $X^A$  and  $X^H$  to form an attention guided, fused representation of extracted features.

### J. 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 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 timestep  $t$ . This can be calculated as:

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

where  $w$  represents the layers weights,  $b$  is the networks bias and  $\hat{y}_t$  is the predicted class label 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 that represents the annotated phases (see Fig. 3). Experiments were then conducted using Python 3.7 and TensorFlow 2.4.

The dataset was divided into two subsets: 70% was allocated for training, while the remaining 30%, consisting of 30 randomly selected events, was partitioned for independent testing. The testing subset was comprised of events from 10 participants, five of whom were independent and did not contribute to the training data, while the remaining five provided events for both subsets. However, the subsets were partitioned at the event level to ensure that data from the same event could not be in both the training and the testing subsets. The training subset was further partitioned using stratified  $k$ -Fold Cross-Validation, dividing the training data into 5 ( $k=5$ ) equal folds. We then used  $k-1$  folds (80%) to train the model, while the remaining instance of  $k$  (20%) was utilised for our validation experiments.

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  $1e - 09$ . To improve generalisation and prevent overfitting, we applied an  $l_2$  regulariser to the dense layers in  $d_1$  and  $d_2$ , with the regularisation penalty set to 0.001. Furthermore, if there were no improvements in the loss function for 10 consecutive epochs, the ReduceLROnPlateau callback function was used to reduce the learning rate by 50%. 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 by calculating the number of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) predictions.

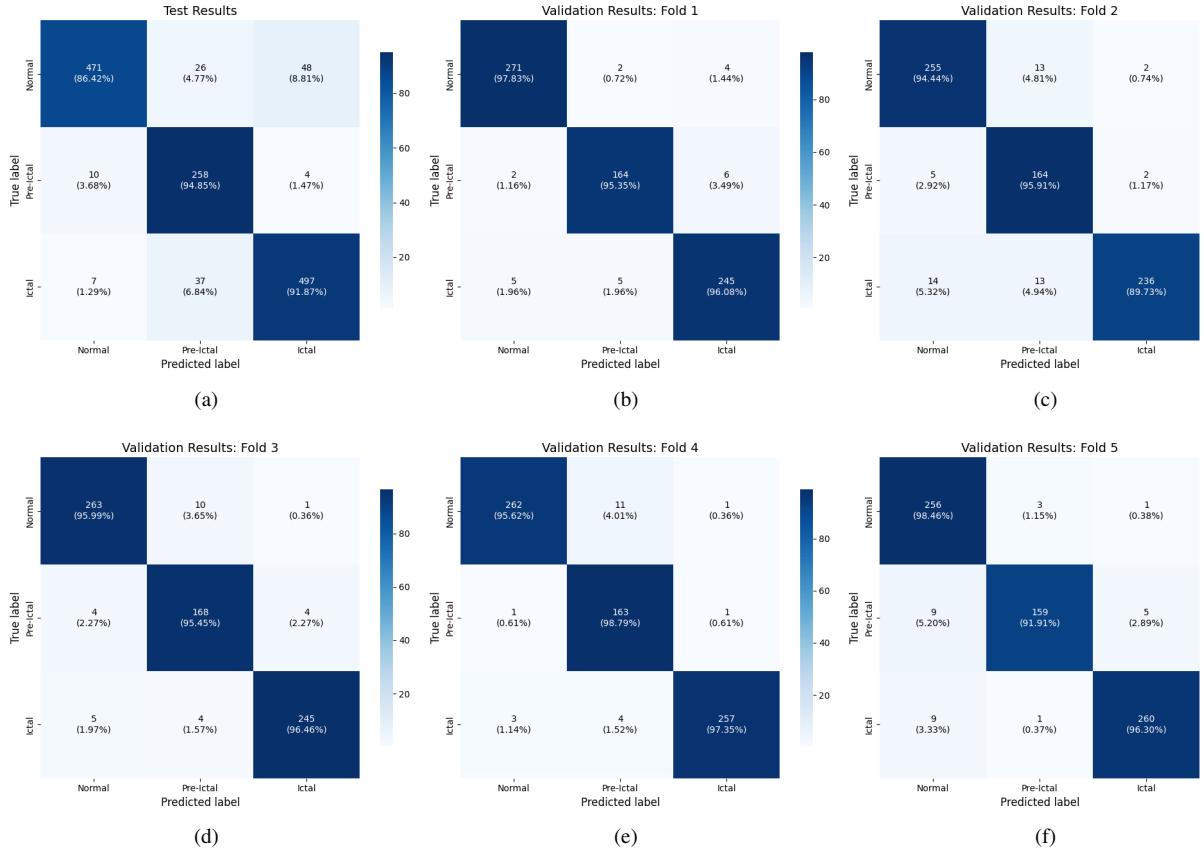
- 1) TP represents the number of predictions where our model correctly classified a positive class as positive.
- 2) FP represents the number of predictions where our model incorrectly classified a negative class as positive.
- 3) FN represents the number of predictions where our model correctly classified 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 presents the results of this study and demonstrates the application of the Ictal-Phase Detection technique. The results are derived from several sets of experiments aimed at evaluating overall performance. The model achieved an average training accuracy of 0.9187, ranging from 0.9090 to 0.9242 across 5 folds (see supplementary material). Validation metrics indicated robust generalisation, with accuracy scores slightly higher, between 0.9304 and 0.9701, and a steady validation loss between 0.0317 and 0.0445. These results underscore the model's ability to maintain high accuracy without overfitting, performing consistently across all folds.

### 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 data and hyperparameter configurations (Table II). Our model ( $m_1$ ) achieved a balance between generalisation and predictive precision, recording an overall accuracy of 0.9027 and an  $f_1$ -score of 0.9035 on unseen test data. Additional experiments measured a Cohen Kappa score of 0.8498 and a MCC of 0.8519, indicating consistent performance across both positive and negative classes. Furthermore, the Positive Predictive Value (PPV) of 0.8914 and True Positive Rate (TPR) of 0.8981 further demonstrate the model's precision in identifying TPs while effectively minimising FPs.



**Fig. 8:** Confusion Matrix Results: Each matrix maps the models predicted results to the ground truth labels for Normal, Pre-Ictal, and Ictal classes. Fig. 8(a) presents the results of our model on the testing subset. Fig. 8(b) - Fig. 8(f) shows the validation results across each fold

In comparison, models such as the 1D-CNN ( $m_{10}$ ), Stacked LSTM ( $m_9$ ) and Stacked BiLSTM ( $m_7$ ) recorded lower performance scores, while those incorporating attention mechanisms (specifically,  $m_2$  and  $m_4$ ) demonstrated enhanced performance. Models utilising multi-branching ( $m_2$ ,  $m_3$ , and  $m_5$ ) recorded notably higher accuracy and  $f_1$  scores, ranging from 0.8613 to 0.8698 and 0.8759 to 0.8952, respectively. Models  $m_2$  and  $m_3$  employed multi-branching architectures, recording accuracy scores of 0.8696 and 0.8613, respectively, outperforming 80% of other architectures. However,  $m_2$  and  $m_3$  did not employ a Residual Fusion layer, and the extracted features were concatenated, resulting in lower performance. These results underscore the importance of independent feature extraction processes and highlight the advantages of fusing features to maximise performance.

### B. Per-Class Classification Analysis

In this subsection, we present the results of our model across several confusion matrix experiments (Fig. 8), plotting the annotated ground truth against the models predicted labels.

Fig. 8(a) shows the confusion matrix results for our model on the unseen testing subset. For the Normal class, 471 timesteps were correctly classified, achieving a TP score of 86.42%. However, 13.58% were misclassified, with 4.77% incorrectly classified as Pre-Ictal and 8.81% as Ictal, suggesting an overlap in feature space. For the Pre-Ictal class, the model classified 94.91% of TPs, correctly classifying 258 out of 272 timesteps, with only 3.68% misclassified as Normal and 1.47% as Ictal. Similarly, the model recorded a TP score of 91.87% for the Ictal class, with 6.84% misclassified as Pre-Ictal and 1.29% as Normal, indicating some overlap with Pre-Ictal markers.

Similar results were observed in our validation experiments as shown in Fig. 8(b) - Fig. 8(f). Each experiment represents a  $k$ -1 partition of the training data used to validate each fold. The model exhibited strong performance, with the Normal class recording an average TP score of 96.38% across all folds, with scores ranging from 94.44% to 98.46%. The model showed a low FP score of 2.88% for the Normal class, indicating minimal misclassifications when averaged across the validation data. Similar results were observed for the Pre-Ictal and Ictal classes, recording an average TP score of 96.88% and 96.43% respectively. These results demonstrate the model's effectiveness in identifying Ictal timesteps, with only 2.96% misclassified as FP and 2.48% as FN.

Although the model recorded 96.38% TPs for the Normal class during validation, performance on the test set dropped to 86.42%, suggesting significant variability. While there was noticeable overlap between the Normal and Pre-Ictal classes, it is important to note that the model demonstrated strong generalisation, clearly distinguishing between the Pre-Ictal and Ictal classes.

In Table III, we present further classification results, distributed by class, based on predicted results in Fig. 8(a). These results show the model recorded a high TPR across all classes, with scores between 0.8342 and 0.9485, averaging 0.8981. A high TNR was observed across all classes (average 0.9497), highlighting the model's effectiveness in correctly classifying non-seizure timesteps. We observed variability in the PPV, with the Normal class achieving a high score of 0.9652, while the Pre-Ictal class recorded a lower score of 0.8037. This disparity indicates a higher incidence of FPs for Pre-Ictal events compared to the Normal class. The model's False Positive Rate (FPR) remained consistently low across all classes, averaging 0.0502,

**TABLE III:** Classification Analysis Distributed by Ictal Class - Notation = TPR = True Positive Rate, TNR = True Negative Rate, PPV = Positive Predicted Value, NPV = Negative Predicted Value, FPR = False Positive Rate, FNR = False Negative Rate

Metric	Normal	Pre-Ictal	Ictal	Average
TPR	0.8342	0.9485	0.9118	0.8981
TNR	0.9768	0.9421	0.9304	0.9497
PPV	0.9652	0.8037	0.9053	0.8914
NPV	0.9149	0.9865	0.9456	0.9497
FPR	0.0232	0.0583	0.0696	0.0502
FNR	0.1658	0.0515	0.0881	0.1018

indicating effective generalisation and inference. Furthermore, the model achieved an average False Negative Rate (FNR) of 0.1018, with the highest FNR of 0.1658 observed in the Normal class. While the model performed well in detecting Ictal events, we can see there is room for improvement when distinguishing between Normal and Pre-Ictal classes.

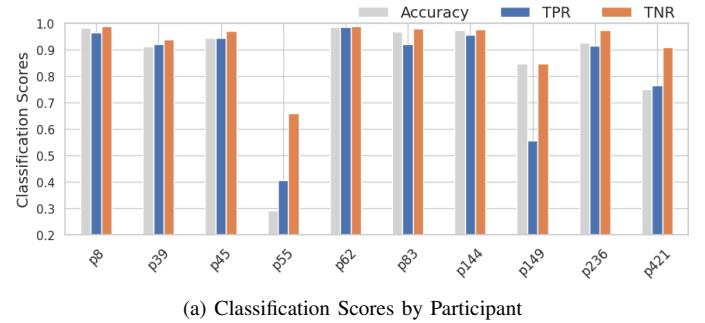
### C. Per-Participant Classification Analysis

This subsection provides an analysis of the model's performance on the unseen testing subset, focusing on individual participants rather than events. As shown in Fig. 9, further experiments were conducted to evaluate the model's effectiveness across different participants, allowing for a more granular understanding of the overall performance. The test set includes events from 15 tonic-clonic seizures, 8 auras, and 8 additional seizures denoted as Other. To ensure an independent evaluation, events from Participants *p83*, *p55*, *p62*, *p236*, and *p421* were excluded from the training data and used exclusively in the testing subset. In contrast, Participants *p8*, *p39*, *p45*, *p144*, and *p149* contributed events to both the training and testing subsets.

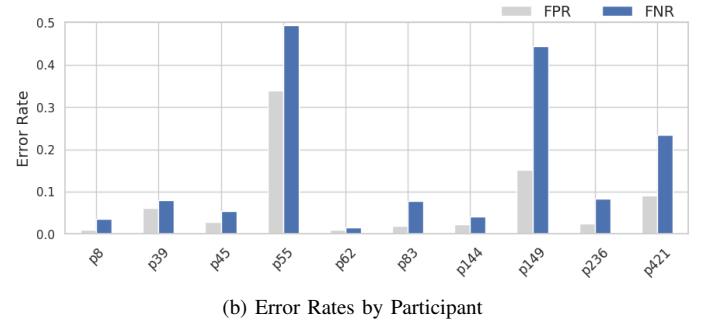
We recorded an overall accuracy score of 0.8874 across all participants, although performance exhibited notable variability (see Fig. 9(a)). Participants *p8*, *p39*, *p45*, and *p144* demonstrated high accuracy scores (above 0.9139), with an average TPR of 0.9477, indicating effective generalisation across these subjects. However, for Participant *p149*, the model showed lower performance, recording an accuracy of 0.8484 and a low TPR of 0.5555, suggesting challenges in accurately classifying the Pre-Ictal and Ictal classes for this participant.

The participants exclusively in the test set exhibited varying performance, with accuracy scores ranging from 0.5714 to 0.9870. Notably, Participants *p83* and *p209* achieved high accuracy scores of 0.9695 and 0.9870, respectively, with TPRs above 0.9215. Conversely, Participants *p55* and *p421* exhibited lower performance, particularly in terms of accuracy and FNR. For Participant *p55*, the model recorded an accuracy of 0.5714 and an FNR of 0.4313 (see Fig. 9(b)), highlighting difficulties in classifying the pre-ictal phase, likely due to variability in unseen seizure manifestations.

Participants whose events were exclusively in the test set had a lower average accuracy of 0.8412. In contrast, participants who contributed events to both the training and test sets exhibited higher accuracy, ranging from 0.9139 to 0.9870, suggesting that including data from the same participants in both sets improved the model's ability to generalise across varying seizure types. These results underscore the inherent challenges in generalising to different events, as each participant's signals are unique, but common fluctuations can be observed within the same participant over time (see Fig. 10(a) – Fig. 10(b)). Despite slightly lower performance scores for test-only participants, it is important to note that all seizures in the test set were successfully detected, demonstrating the model's robustness even when evaluated on unseen data.



(a) Classification Scores by Participant



(b) Error Rates by Participant

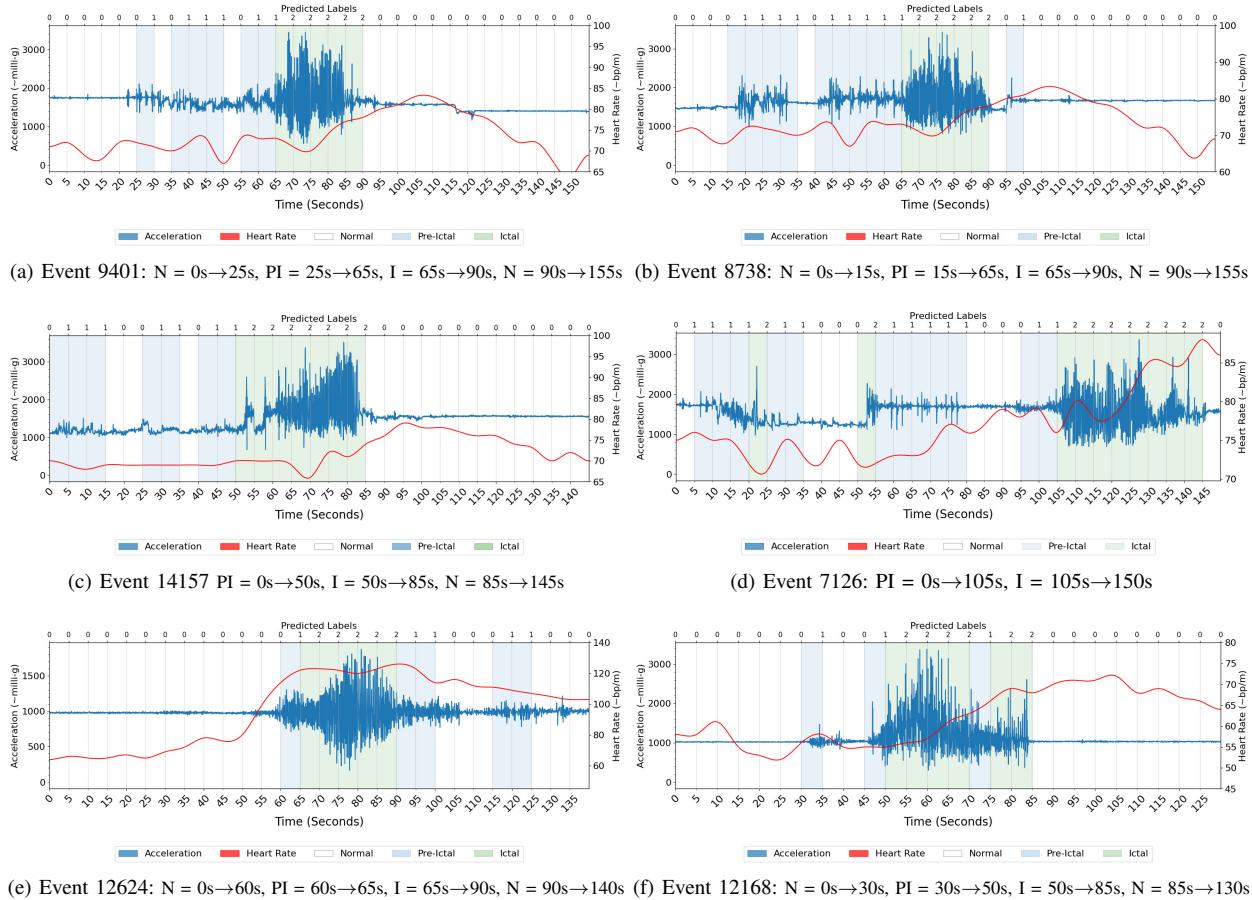
**Fig. 9:** Classification results distributed by participant. Fig 9(a) maps accuracy, TPR and TNR, presenting a comprehensive view of classification success and the model's reliability. Fig. 9(b) illustrates error-related metrics, measuring the FPR and FNR for each participant

### D. Ictal-Phase Detection

This subsection demonstrates the Ictal-Phase Detection technique on a representative subset of events. Fig. 10 shows the results of six experiments where each event is labelled as GTC seizure. The upper *x*-axis displays the predicted labels, with 0, 1, and 2 representing the Normal (white), Pre-Ictal (blue), and Ictal (green) classes respectively.

In Fig. 10(a) to Fig. 10(d), the presence of a clear pre-ictal phase is evident, as indicated by both the acceleration and heart rate signals. The acceleration data displays gradual increases in amplitude, corresponding with higher intensity movements associated with pre-ictal manifestations. Simultaneously, heart rate trends show a steady increase of approximately 10% from the point of onset to cessation. This pattern can be observed in both plots, where both modalities present a clear transition into the ictal phase, marked by an elevated increase in beats per minute and repetitive convulsive movements exceeding 2000 milli-g. In contrast, Fig. 10(e) and Fig. 10(f) show GTC events with no clear pre-ictal phase, where the first signs of convulsive movements emerge seconds before the onset of the ictal phase. Despite the absence of a well-defined pre-ictal phase, both modalities exhibit trends that align with the observations seen in Fig. 10(a) - Fig. 10(d). Heart rate increases of approximately 10% were observed prior to repetitive fluctuations in acceleration, reinforcing the notion that even without a pre-ictal phase, there are clear physiological changes that indicate seizure onset.

From these results, it is evident that the proposed model can distinguish the convulsive movements of the Ictal class from the movements observed in the Normal and Pre-Ictal classes. The ictal phase was correctly classified across all six plots, accurately detecting approximately 92.1% (35/38) of Ictal timesteps. A degree of uncertainty was observed in Fig. 10(d), where two Pre-Ictal classes were incorrectly labelled as Ictal. However, the onset and cessation timesteps of each event was clearly detected, demonstrating the models ability to detect convulsive movements during the ictal phase.



**Fig. 10:** 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, defining the timesteps that start and end each phase where s = seconds, N = Normal, PI = Pre-Ictal, and I = Ictal.

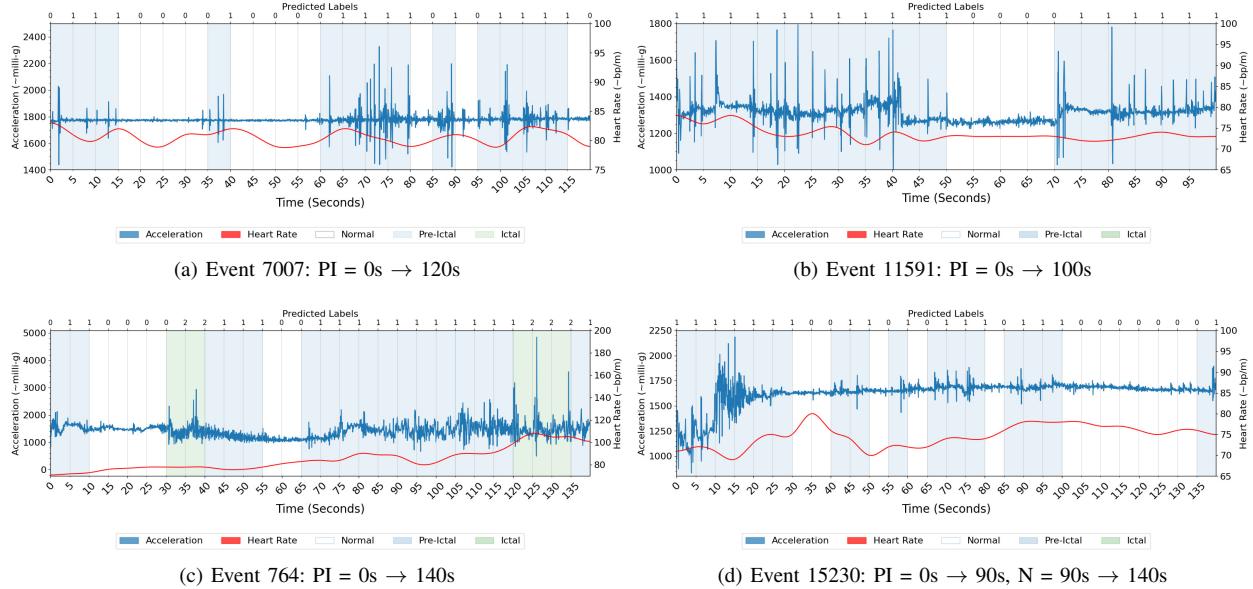
Fig. 11 presents the results of our experiments classifying low-amplitude signals indicative of an Aura or Focal seizure. The signal data for these events are more challenging to distinguish compared to the high-amplitude movements compounded by the elevated heart rate observed in the ictal phase. Convulsive movements are indicative of rapid fluctuations that traverse a significant distance in terms of vector magnitude, whereas Pre-Ictal events are often characterised by sequential timesteps with minimal or near-zero movement.

In Fig. 11(a) and Fig. 11(b), the pre-ictal phase is marked by low-intensity signals with sudden spikes in acceleration. In Fig. 11(a), the proposed model classified the event, accurately predicting 87% (13/15) of the timesteps that showed acceleration waveform fluctuations in acceleration, compounded by elevated heart rate readings. However, the model did not label any seizure activity between 15s and 35s and 40s - 60s, due to the minimal variation in movements and reoccurring decline in heart rate between 84 - 78./m. In Fig. 11(b), the model detected 80% (16/20) of the timesteps. However, at 50s to 70s, we can see that the acceleration and heart rate waveforms exhibited minimal fluctuations, indicating an overlap in feature space between the Normal and Pre-Ictal classes.

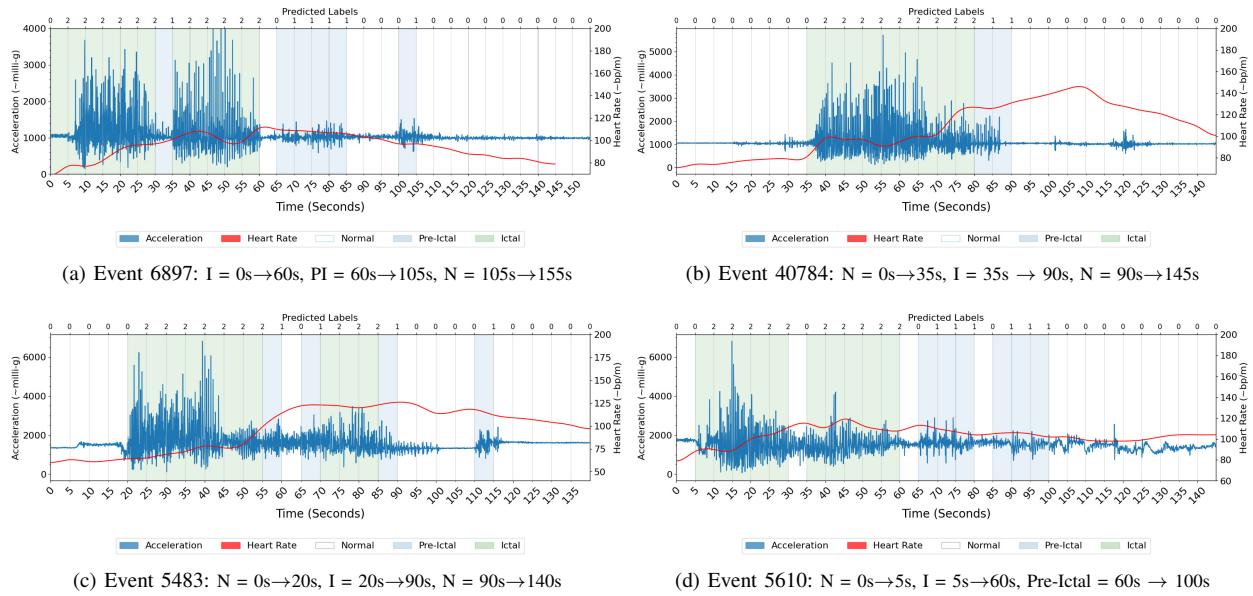
Fig. 11(c) highlights the complexities of classifying Pre-Ictal signals, with the model misclassifying five timesteps as Ictal due to sudden acceleration spikes. Despite these challenges, the data reveals a substantial overlap with the Normal class at specific intervals (10s

- 30s and 55s - 65s). Subtle heart rate elevations further complicate the classification process; however, the model accurately identified 22 out of 28 timesteps as either Pre-Ictal or Ictal, highlighting the models ability to detect seizure-related timesteps. In Fig. 11(d), high-amplitude Pre-Ictal movements exceeding 2000 milli-g coincide with a 25% heart rate increase that gradually returns to a baseline. However, some low-amplitude signals in later phases went undetected due to minimal fluctuations. These plots highlight 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 addressing the variability inherent in the different types of Auras and myoclonic movements and heart rate fluctuations that precede the ictal phase.

Fig. 12 illustrates several experiments conducted on events labelled as Other Seizure. The structure of these acceleration waveforms differs from those shown in Fig. 10. 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. 10, where the maximum acceleration peaked at approximately 3000 milli-g. However, in Fig. 12(a) and Fig. 12(b) acceleration exceeds 4000 milli-g, while in Fig. 12(c) and Fig. 12(d) the acceleration waveforms exceed 6000 milli-g, representing a doubling of intensity compared to Fig. 10. There are also notable heart rate



**Fig. 11:** Aura/Focal Seizure Experiments: For a full description of the experiments, annotations and class labels see the caption in Fig. 10



**Fig. 12:** Other Seizure Experiments: For a full description of the experiments, annotations and class labels see the caption in Fig. 10

fluctuations in Fig. 12 where each plot exceeds 100 beats per minute, with a notable increase occurring at the time of onset, continuing until cessation. Overall the proposed model accurately detected the ictal phase in Fig. 12, approximately detecting 91.6% (11/12), 81.81% (9/11), 73.33% (10/14) and 90.91% (10/11) of Ictal timesteps for Fig. 12(a) - Fig. 12(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 when detecting epileptic seizures. We hypothesised that high-amplitude convulsive movements, characteristic of the ictal phase, can be differentiated from low-amplitude myoclonic movements indicative of the pre-ictal phase. To the best of our knowledge, this study is the first to investigate the viability of non-

EEG multimodal data to detect these distinct phases, offering a novel approach for seizure monitoring and detection. To address the challenges in distinguishing between Normal, Pre-Ictal, and Ictal classes, we introduced a new model (AMBER) that constructs independent branches for each modality, generating attention weights that fuse extracted features into a single representation to enhance feature extraction and propagation.

The results of our experiments underscore the effectiveness of our model, achieving an accuracy score of 0.9027 and effectively distinguishing between different phases of the ictal state. With Cohen's Kappa and MCC scores of 0.8498 and 0.8519, respectively, demonstrating strong agreement between the predicted and ground-truth labels, highlighting consistent generalisation across positive and negative classes. The high TNR of 0.9497 reflects the model's ability to accurately detect non-seizure timesteps, while an average TPR of

0.8981 confirms its effectiveness in detecting seizures. A slight bias toward specificity is observed, with a marginally lower TPR for the Normal class compared to its TNR. The PPV range of 0.8037 to 0.9652 indicates high confidence in positive classifications, although a FNR of 0.1018 indicates some seizure timesteps are occasionally missed. The model's trade-off between a high TNR and moderate FNR indicates a preference for minimising FPs over FNs.

From our analysis of GTC events, we recorded approximately 91% of timesteps where seizure onset and cessation occurred (see Fig. 10(a), Fig. 10(b) and Fig. 10(d)). Moreover, in events classified as Other, the model accurately identified 89% of the specific timesteps marking the onset of seizure activity. These results highlight the model's ability to detect low-amplitude movement followed by a sudden spikes indicative of the generalised activity seen in the ictal phase (see Fig. 12(b) - Fig. 12(c)).

To assess the model's performance on individual participants, we conducted a participant-level analysis. The model achieved an average accuracy score of 0.8874 across the 10 participants, with notable variability in performance. Participants *p8*, *p39*, and *p144* demonstrated high accuracy scores of 0.9828, 0.9139, and 0.9762, respectively, indicating effective generalisation for these subjects. In contrast, participants *p55* and *p421* showed lower accuracy, particularly with higher FNRs in Pre-Ictal class. For *p55*, the accuracy dropped to 0.5714, highlighting the challenges in capturing unique seizure manifestations not observed in the training data. Notably, participants who contributed events to both the training and testing subsets, such as *p8*, *p39*, *p45* and *p209*, exhibited higher accuracy scores ranging from 0.9139 to 0.9870. This suggests that while seizure events are unique to each participant, they share common characteristics across different manifestations, facilitating better generalisation when the model is exposed to a variety of events from the same participant.

An important observation from our analysis was the noticeable increase in cardiac manifestations during the pre-ictal and ictal phases, particularly in seizures categorised as Other. This trend highlights the potential link between cardiac activity and seizure progression. 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 used in this study, 36.18% ( $N=34/94$ ) exhibited ictal tachycardia manifestations, characterised by a heart rate increase exceeding 100 beats per minute. While only one event labelled as Aura exceeded the ictal tachycardia threshold, this number rose to 10 for GTC and 23 for seizures denoted as Other. A consistent heart rate increase of approximately 10% was observed starting from seizure onset, lasting roughly 15-20 seconds, and returning to baseline once clonic activity ceased. This pattern, evident across event types, underscores the potential of heart rate as a reliable biomarker for Ictal-Phase Detection. In contrast, Aura events, characterised by low-amplitude pre-ictal activity, showed this response less frequently. However, the widespread occurrence of ictal tachycardia in seizures with more distinct during the ictal phase for events categorised as GTC and Other.

#### A. 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 detection of the Pre-Ictal class. Rather than predicting a seizure at the first sign 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.

Traditional studies in existing literature measure performance based on whether an event is detected. In contrast, our approach focuses on classifying each 5-second timestep, calculating the overall performance score based on the models ability to classify three distinct classes: Normal, Pre-Ictal, and Ictal. We feel this technique enhances diagnostic accuracy by improving the detection of generalised onset seizures and enabling the early identification of pre-seizure states, offering a more reliable method for seizure detection in real-world environments.

#### B. Limitations and Future Work

While our study provides valuable insights, several limitations should be considered. A key challenge was using a single class label (Normal) to represent common human movements during the inter-ictal and post-ictal phases. This decision, influenced by previous work such as Abdulhay *et al.* [39], proved restrictive in the context of non-EEG detection. The variability of movements in these phases is considerable, making a single label insufficient to capture this diversity. Future work will expand the number of input classes to reduce uncertainty between Normal and Pre-Ictal classes. Incorporating a range of human activities, such as lying down, walking, or brushing teeth, could improve phase-specific classification by capturing subtle variations in baseline movements and isolating seizure-related signals. This approach aims to enhance both the specificity and sensitivity of the model in identifying distinct movement patterns characteristic of the ictal state.

A further limitation was encountered during the annotation of the non-EEG data due to the lack of established methodologies for segmenting and quantifying timesteps across events. Developing an appropriate technique required close collaboration between our technical and clinical teams, as no clear benchmarks were available for this process. Reaching a consensus on the annotation proved time-consuming, highlighting the complexity of the task. Future work could focus on developing standardised annotation techniques for labelling non-EEG data, which would streamline the process and improve the reliability of event classification.

## VI. CONCLUSION

In summary, this study has made advancements in non-EEG seizure detection through the introduction of the Ictal-Phase Detection technique. By applying a multimodal approach that combined acceleration and heart rate data, we highlighted the potential of cardiac manifestations as biomarkers and demonstrated how high-amplitude and low-amplitude movements, indicative of the pre-ictal and ictal phases, can be distinguished. Our proposed model (AMBER) demonstrated high performance, with an overall accuracy of 0.9027, clearly distinguishing between seizure and non-seizure states. We believe this approach has the potential to reduce the number of FPs and enable future detection methodologies to identify different types of convulsive movement. In conclusion, this study contributes a new detection technique and model designed for multimodal seizure detection and lays the groundwork for further advancements in non-EEG seizure detection.

We are committed to advancing this research and promoting collaboration within the field by ensuring the accessibility of our findings.

To further support the development of non-EEG seizure detection, we are dedicated to sharing the results, model, and codebase from this study.

## ACKNOWLEDGEMENTS

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

We also wish to express our appreciation to the participants of the Open Seizure Database. Without their dedication and commitment, this research would not have been possible.

## 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, dataset, and code base. The AMBER model is now distributed under the MIT license as open-source [41]. We welcome 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 available 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. Zeljković, 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, "N-gram-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 Generalised 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. Mehmood, 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>