DEPARTMENT OF
INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING

Spring Semester 2025

Exploration of Seizure Prediction Methods Tailored for Wearable Devices

Semester Project

Kyrill Toulev

toulevk@student.ethz.ch

4.6.2025

Advisors: Thorir Ingolfsson, OAT U 21, thoriri@iis.ee.ethz.ch
              Dr. Victor Kartsch, ETZ J68.2, victor.kartsch@iis.ee.ethz.ch
              Dr. Andrea Cossettini, OAT U 27, cossettini.andrea@iis.ee.ethz.ch

Professor: Prof. Dr. Luca Benini, lbenini@iis.ee.ethz.ch

Abstract

Electroencephalography (EEG)-based seizure prediction has the potential of transforming epilepsy care by providing timely warning of imminent seizures and enabling patients to take precautionary measures. However, most state-of-the-art algorithms rely on large, power-hungry neural networks and high-density electrode montages, both of which are impractical for continuous, real-world monitoring on wearable devices. This thesis investigates the feasibility of using a compact convolutional neural network (CNN) for real-time seizure prediction using less EEG channels.

The primary objective is to design a seizure-prediction pipeline that first maintains clinical performance (high sensitivity and lead time, low false-alarm rate per hour) on the MIT-CHB scalp-EEG database, with model complexity lower by by a factor of 10-100 relative to the SoA models, and operates on only four EEG channels, which can be integrated into glasses or other unobtrusive wearables. To achieve this, it requires overcoming two key challenges: Capturing subtle pre-ictal EEG signatures, and avoiding overfitting on the severely reduced parameter count. To address these challenges, we make use of EpiDeNet, a CNN with roughly 9k trainable parameters, specifically developed for Seizure related EEG signals, and compare it against conventional SoA Models with paramater counts ranging from 500k-12M. The model is evaluated under leave-one-seizure-out cross-validation, and the output fed trough a post-processing pipeline to enforce meaningfull per seizure metrics such as Event Sensitivity, FPR/h and Lead Time.

Our model achieves event-level sensitivity between 75.5 % and 94.4 % with FPR/h between 1.6 and 0.37, compared to state-of-the-art ranges of 81.7–91.5 % event-level sensitivity and 0.2–0.1 FPR/h. Even when reduced to four channels, EpiDeNet only loses about 7 % in event-level sensitivity. These results confirm that a compact, specialized CNN like EpiDeNet can still match top performance using few-channel EEG, which potentially opens the door to energy-efficient, wearable implementations. Future work could explore multimodal sensor fusion (e.g., adding PPG, ECG or accelerometry) and deployment of the entire pipeline on an embedded device.

Acknowledgments

I would like to express my deepest gratitude to my thesis supervisors, Thorir Ingolfsson, Dr. Victor Kartsch, Dr. Andrea Cossettini and Prof. Dr. Luca Benini, for their guidance and support throughout my semester thesis.

ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Declaration of originality

The signed declaration of originality is a component of every written paper or thesis authored during the course of studies. **In consultation with the supervisor**, one of the following two options must be selected:

☐ I hereby declare that I authored the work in question independently, i.e. that no one helped me to author it. Suggestions from the supervisor regarding language and content are excepted. I used no generative artificial intelligence technologies[1].

☒ I hereby declare that I authored the work in question independently. In doing so I only used the authorised aids, which included suggestions from the supervisor regarding language and content and generative artificial intelligence technologies. The use of the latter and the respective source declarations proceeded in consultation with the supervisor.

Title of paper or thesis:

Exploration of Seizure Prediction Methods Tailored for Wearable Devices

Authored by:
If the work was compiled in a group, the names of all authors are required.

Last name(s):

Toukev

First name(s):

Kyrill

With my signature I confirm the following:
- I have adhered to the rules set out in the Citation Guidelines.
- I have documented all methods, data and processes truthfully and fully.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for originality.

Place, date

Zürich, 1.6.25

Signature(s)

K. Tauber

If the work was compiled in a group, the names of all authors are required. Through their signatures they vouch jointly for the entire content of the written work.

---

[1] For further information please consult the ETH Zurich websites, e.g. https://ethz.ch/en/the-eth-zurich/education/ai-in-education.html and https://library.ethz.ch/en/researching-and-publishing/scientific-writing-at-eth-zurich.html (subject to change).

Contents

Contents

Chapter 1

Introduction

1.1 Epileptic Seizure

Epilepsy is a chronic neurological disorder characterized by recurrent seizures, affecting over 50 million people worldwide. Epileptic Seizures occur due to excessive electrical discharges in brain cells, leading to a range of symptoms including involuntary movements, altered sensations, emotional disturbances, and in some cases loss of consciousness. Patients often face significant challenges in their daily lives such as physical injuries, if the patient falls, sleep disturbances, since Epilepsy can disrupt sleep patterns or drowning risks and drive hazards. Hence, epilepsy profoundly affects quality of life, and the importance of accurate seizure prediction becomes evident. Especially for patients with drug-resistant epilepsy, which do not respond to medication and would benefit greatly from a wearable system, which enables continuous, real-time monitoring and seizure prediction.

The sudden electrical discharges that happen in the brain, which is the cause of epileptic seizures, can be measured with EEG. Electroencephalography (EEG) is a widely used non-invasive technique for measuring the brain's electrical activity reflecting cognitive and physiological processes, making it a valuable tool for neurological diagnostics. In particular, EEG is the gold standard for detecting and analyzing epilepsy-related brain activity. Given the unpredictable nature of seizures, continuous monitoring is essential to improve patient safety and optimize treatment strategies. Recent advancements in wearable EEG technology have enabled continuous monitoring with a reduced number of electrodes, providing a more practical and patient-friendly solution for real-world applications.[1, 2, 3, 4]

1.2 Seizure Prediction

Seizures can be addressed through either detection or prediction, with the primary distinction lying in their objectives: detection aims to determine if a seizure is currently

occurring, while prediction focuses on forecasting when a seizure is likely to happen. Beyond seizure detection, epileptic seizure prediction presents an opportunity to anticipate seizures before they occur, potentially improving patient quality of life and enabling timely interventions. As a consequence, individuals with epilepsy can adjust their activities, take precautionary measures or administer fast-acting medications, ultimately reducing the burden of epilepsy [5, 6].

Seizure prediction has significantly progressed in the last years [7], including methods that aim to detect pre-ictal/intra-ictal states through CNNs [8, 9] or to provide probabilistic long-term risk estimation taking into consideration cyclic patterns: circadian, multidien, and monthly rhythms rather than considering seizure as a randomly occurring event [10].

The goal of this project is to investigate seizure prediction techniques with a focus on wearable EEG devices.

1.3 Thesis Structure

The **Introduction** (Chapter 1) sets the stage for this thesis by presenting the topic, its significance, and the goals of the study.

In the **Background** (Chapter 2) the theory of seizures and seizure prediction methods with machine or deep learning including its possibility for deployment on wearable devices is discussed.

In the **Related Work** (Chapter 3) the state-of-the-art for seizure prediction is discussed, focusing on machine learning, deep learning and also the feasibility of seizure prediction models for wearable devices.

The **Implementation** (Chapter 4) provides a detailed explanation of the procedures and approaches used in the project. This chapter covers the dataset utilized, procedures for data extraction, preprocessing, the architecture of the deep learning models, and postprocessing methods.

The **Results** (Chapter 5) discusses the outcomes of the deep learning models by evaluating their performance and discusses the tradeoffs of the postprocessing.

The **Conclusions** (Chapter 6) summarizes the key findings of the project and reflects on their significance, discussing limitations of the methods used, proposing potential improvements for future work, and explores opportunities for applying the epileptic seizure prediction models in real-world scenarios.

Chapter 2

Background

2.1 Epileptic Seizures & EEG

Electroencephalography (EEG) is a non-invasive technique to record the brain's electrical activity of a subject with scalp electrodes. EEG records the electrical signals from the cerebral cortex (the most developed part of the human brain) by measuring the electrical activity of group's of neurons. By sampling voltage fluctuations EEG provides high temporal resolution and insight into neuronal activity but is lower in spatial resolution compared to an MRI, which could also be used for analyzing the brain activity. EEG waveforms (generally in the realm of 10–100 μV amplitude, 1–100 Hz bandwidth) are decomposed into five frequency bands, delta ($\leq$ 4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (12-26 Hz) and gamma (26-100 Hz). Epileptic preictal dynamics are most pronounced in the high-frequency band (32–128 Hz) [11] .This spectral content can be extracted by Fourier or wavelet transforms for feature-based machine learning. More recently, deep learning methods have shown that convolutional and recurrent architectures can learn the necessary features directly from raw EEG. However, there does not exist one or more clear features to characterize the EEG states, which is still an open research issue.

EEG analysis provides the primary insight into the brain's electrical dynamics before, during, and after seizures, which contain the characteristics that can be used to identify the different stages of an epileptic seizure. The different stages as seen in figure 2.1 are the following:

Pre-ictal State: A fixed time period before the occurrence of a seizure and does not occur at the rest of the times. Usually it is not visually apparent but changes in the underlying signal are often detectable due to a gradual rise in high-frequency power or changes in synchrony.

Ictal State: The change in EEG signal during a seizure, so the seizure itself. This is often visually apparent due to clear, sustained bursts of high-amplitude rhythmic activity

Interictal State: The State between two following seizure onsets. This is not directly seizure related.

Post-Ictal State: State after the occurrence of a seizure, after the Ictal. The EEG exhibits

slow waves and suppressed amplitude (post-ictal slowing), which can also be observed visually.
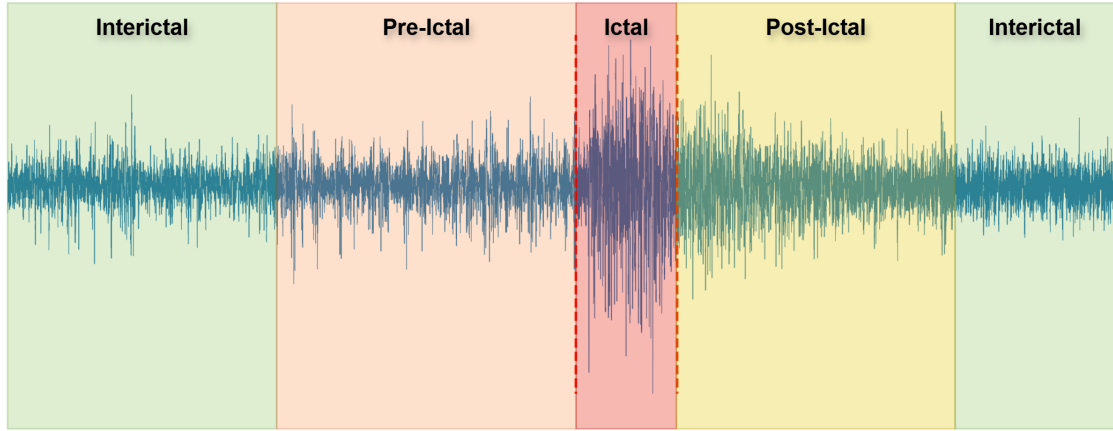

Figure 2.1: The different seizure phases, recorded with EEG

Visual EEG interpretation by experts can be time-consuming and subjective, so automated signal-processing and machine-learning pipelines, that are able to extract spectral features or learning them directly via deep networks are therefore essential for fast and objective identification of these seizure phases.
Hence EEG seems as a good fit for measuring, detecting and predicting epileptic seizures while its operational costs remain relatively low, hence enabling long-term monitoring for modern wearable EEG headsets. Also further reduction of the electrode counts and power draw without sacrificing key signal information is possible.

2.1.1 Seizure Detection vs Prediction

Seizure *detection* aims to identify the **Ictal** phase as it unfolds by detecting the high-amplitude, discharges that mark clinical onset of a seizure, whereas the *prediction* targets the **Pre-Ictal** phase, with more subtle spatiotemporal EEG changes (compared to Ictal) within a defined seizure-occurrence horizon (SOH) to issue an advance warning. Considering the different phases the main difference is that:
Seizure Detection: Binary Classification of Ictal vs non-Ictal (every other phase).
Seizure Prediction: Binary Classification of Pre-Ictal vs Interictal.
If the algorithm classifies the incoming data as Pre-Ictal, we know that a seizure is imminent, since Pre-Ictal precedes the Ictal phase. Thus we detect Pre-Ictal and predict the Ictal phase, hence the seizure. Detection methods typically perform per-window classification of ongoing EEG segments, while prediction methods employ temporal modeling (sliding-window analyses, with majority voting) raise an alarm & warn the patient before a seizure happens, thus allowing proactive interventions rather than retrospective seizure tracking. The most important seizure prediciton metrics are:
Preictal/Seizure Sensitivity: Also called *pre-ictal sensitivity or event sensitivity*, this is the

proportion of all seizures for which at least one alarm was raised during the designated Pre-Ictal time period. Formally,

$$\text{Sensitivity} = \frac{\text{\#Seizures with a correct pre-ictal alarm}}{\text{Total \# of Seizures}} \times 100\%$$

High sensitivity means that most seizures are successfully anticipated.

False-Alarm Rate (FPR/h) The average number of false positive alarms per hour of monitoring outside the Pre-Ictal window (typically measured during Inter-Ictal periods).

$$\text{FPR/h} = \frac{\text{\#False Alarms}}{\text{Total Interictal Hours}}$$

Lower FPR/h reduces alarm fatigue and improves patient trust, but there is a tradeoff between FPR/h and seizure sensitivity. Note that also other definitions of the FPR/h exist as in [12], where we would subtract the refractory period multiplied with the false positives from *Total Interictal Hours* since we suppress all alarms during that time.

Mean Lead Time Also known as the *mean warning time*, this is the average interval between the first correct pre-ictal alarm and the actual seizure onset. If t_alarm is the timestamp of the first true alert and t_onset the clinical onset, then

$$\text{Lead Time} = \frac{1}{N} \sum \left(t\_onset^{(i)} - t\_alarm^{(i)} \right)$$

Longer lead times allow more opportunity for preventive actions, but must be balanced against FPR/h.

Chapter 3

Related Work

3.1 Seizure Prediction with Machine Learning

Machine learning has become more and more important in modern Seizure Prediction research, driven by the need to replace time-consuming and often subjective EEG interpretation with fast, objective algorithms. Traditional EEG analysis relies on handcrafted features like spectral power, coherence, entropy measures or Lyapunov exponents that are extracted from preprocessed data after using band-pass and artifact filtering [13, 14]. Feature-selection methods such as principal-component analysis (PCA) or independent-component analysis (ICA) then can reduce dimensionality and highlight the most relevant seizure patterns [15]. These engineered features are then feed into classical classifiers, such as support-vector machines (SVM), k-nearest neighbors, decision trees or shallow neural networks, that label each window of EEG as either Pre-Ictal or Interictal [13]. Early studies using these techniques achieved sensitivities above 90 percent on benchmark datasets, demonstrating that even simple pipelines can anticipate seizures minutes before clinical onset [16]. However, ML-based studies heavily focused on the extraction of optimized features for prediction, which is a problem since relying exclusively on handcrafted features presents several limitations for seizure prediction algorithms and also there does not exist one or more definite features to characterize the EEG states [17].In contrast, end-to-end deep learning methods can discover and refine their own features directly from raw EEG, offering greater adaptability, with the potential to uncover previously unrecognized Pre-Ictal patterns hidden in the EEG signal [18].

3.2 Seizure Prediction with Deep Learning

Use of CNN for ES Prediction: One of the first **CNN-based** seizure prediction approaches that requires minimal EEG preprocessing was introduced by Truong et al. (2017) [19] They trained and tested their model on both the Freiburg iEEG and CHB-MIT

scalp EEG datasets by converting short EEG segments into 2D "images" via short-term Fourier transform (STFT) and feeding these into a convolutional neural network(CNN) to distinguish between Pre-ictal and Interictal states. With a 5-minute seizure prediction horizon (SPH) and a 30-minute seizure occurrence period (SOP), and using leave-one-out cross-validation, their model achieved 79.7 % sensitivity and a 0.24 false prediction rate (FPR/h) on raw EEG, improving to 89.8 % sensitivity and a 0.17 FPR/h when the data were standardized.

Use of RNN for ES Prediction: Tsiouris et al. were the first to apply **long short-term memory (LSTM)** networks for seizure prediction [20]. They evaluated three LSTM architectures, increasing in size and using feature vectors extracted from 5–50 s EEG segments. These feature vectors combined time-domain, frequency-domain, and graph-theoretic measures. Among them, the biggest, LSTM-3, achieved the best result. For LSTM-3, with feature vectors, average sensitivities were 99.28 % (15 min), 99.35 % (30 min), 99.63 % (60 min), and 99.84 % (120 min) pre-ictal, though this required **extensive** feature engineering. When the same LSTM-3 was fed raw EEG signals instead of feature vectors, performance dropped notably, indicating that end-to-end deep modeling with raw EEG remains an open challenge in this domain.

Use of Transformers and hybrid models for ES Prediction: Lee et al.[21] propose a hybrid ResNet-LSTM model for seizure prediction that first converts EEG signals into STFT-based spectrograms and then pre-trains a ResNet encoder using supervised contrastive learning on augmented data, before fine-tuning with an LSTM to capture temporal features. Evaluated via leave-one-out cross-validation on the CHB-MIT and SNUH datasets, the approach achieved 89.64 % sensitivity and a 0.058 false positive rate on CHB-MIT, and 79.89 % sensitivity with a 0.131 false positive rate on SNUH, outperforming conventional methods. Their results demonstrate that combining contrastive pre-training with a ResNet-LSTM architecture can improve seizure prediction accuracy while dealing with the data scarcity and EEG complexity.

Zhu et al. [22] propose a hybrid model that combines a multidimensional Transformer encoder with LSTM and GRU modules to capture both temporal and frequency features from STFT-derived EEG spectrograms. Evaluated on CHB-MIT, their approach achieved an average sensitivity of 98.24 % and specificity of 97.27 %, while on the Bonn dataset it attained around 99 % accuracy in both binary and tertiary classification tasks. By fusing Transformer and RNN components, their model outperformed comparable Transformer-only architectures

Unsupervised DL Method for ES Prediction: One major challenge in seizure prediction is that labeled EEG seizure data is quite scarce. To address this, Truong et al. [23] used a GAN trained on STFT spectrograms to extract features in an unsupervised manner, allowing real-time prediction without manual feature engineering.Their discriminator-based features yielded AUCs of 77.7 % (CHB-MIT), 75.5 % (Freiburg), and 65.1 % (EPILEPSIAE) with a 5 min SPH and 30 min SOP. Similarly, Ahmed et al.[24] trained a 2D convolutional autoencoder on unlabeled CHB-MIT EEG to learn features

automatically, then applied those encoder weights (transfer learning) to train patient-specific networks. They fed the learned features into a bidirectional LSTM and achieved 94.6 % sensitivity with only 0.04 false alarms/hour for a 1 h SPH.

While deep models often outperform traditional pipelines in sensitivity, robustness and performance, their large parameter counts (hundreds of thousands to tens of millions of learnable parameters) and high data demands present significant challenges for real-time, low-power deployment on wearable EEG platforms. The chapters that follow will build on these insights to design and evaluate compact CNN architectures, with parameter counts significantly lower (9k) but retain high prediction performance while enabling true wearable operation.

3.3 Seizure Prediction tailored for wearable devices

Zhao et al. [25] used neural architecture search to develop an compact CNN for seizure prediction, producing a baseline model of only 45 kB that achieved sensitivities above 99 % and AUCs up to 1.00 on multiple EEG/iEEG datasets while consuming under 10 mJ per inference . Through pruning, quantization, and depthwise-separable convolutions, they further shrank model size below 50 kB for scalp EEG and under 10 kB for intracranial EEG, performance levels competitive with larger architectures and meeting the energy constraints of wearable/implantable devices.
Samie et al. [26] present a seizure prediction algorithm for constrained IoT devices by reducing the EEG segment size to 5.12 s, using only 2–3 channels, and extracting simple FFT-based and time-domain statistical features to minimize memory and computation. They run a lightweight logistic regression model on an MSP432 microcontroller and offload an XGBoost classifier to a gateway, achieving a held-out AUC of 0.79, outperforming prior methods, while consuming just 13.13 mJ per segment. This demonstrates that careful feature simplification and channel selection can enable accurate, energy-efficient seizure prediction on wearable platforms.
Notably, EpiDeNet [27] (architecture can be seen in figure 4.5) has already been demonstrated as an energy-efficient seizure detection network on resource-constrained MCUs: Ingolfsson et al. introduced EpiDeNet-SSWCE and showed 91–92 % detection sensitivity on CHB-MIT and PEDESITE using only four temporal channels, with implementations on ARM Cortex M4F/M7 and GAP8/GAP9 achieving as little as 0.051 mJ per inference on GAP9 . Because the same core architecture was fully quantized and validated in real time on embedded hardware, EpiDeNet provides a proven, ultra-compact backbone. In this thesis, we therefore adopt EpiDeNet not only to leverage its low-power, low-latency profile but also to evaluate its capability for seizure prediction, which is a logical extension from detection, since prediction tasks require similar streaming, on-device processing constraints but impose distinct modeling challenges (e.g., capturing pre-ictal dynamics). By building on an already quantized, embedded-ready network, we can focus on adapting and validating its performance for the prediction setting without re-solving everything from the quantization/inference pipeline.

Chapter 4

Implementation

4.1 Overview Methodology

We will consider 2 slightly different methodologies, model 1 and 2, depicted in figure 4.1. Both pipelines share the same pre-processing steps (ictal/pre-ictal extraction, notch filtering, and z-normalization as described in Sections 4.3 and 4.4), but differ in how they split data, train/validate the network, and report final results.

This is done in a **subject-specific** manner: for each subject, we train and test the model exclusively on that subject's data, and we repeat this process across all subjects. By contrast, a **cross-subject** approach pools data from all individuals for training and testing and then evaluates performance on seizures from different subjects.

Model 1: 80:20 Random Split (Per-Window Evaluation)

- **Data Splitting:** All pre-ictal and interictal windows are pooled and randomly divided into 80 % training and 20 % testing.

- **Results:** Per-window predictions are aggregated to compute accuracy, sensitivity, and specificity on a window-level, per subject.

Model 2: Leave-One-Seizure-Out (LOSO) CV (Per-Seizure Evaluation)

- **Data Splitting:** Each seizure (its pre- and interictal windows) is held out once as the test fold, and the remaining $N-1$ seizures form the training pool.

- **Training/Validation:** Within each LOSO iteration, a small subset of the $N-1$ seizures serves as an internal validation set to prevent overfitting, then the trained model is tested on the held-out seizure.

- **Post-Processing & Results:** Per-window outputs on the held-out seizure are filtered via K-of-N majority voting, a refractory period, and predefined SPH/SOP explained

in section 4.6. Each seizure yields one binary alert (yes/no), FPR/h, and lead-time. We report averages on a seizure level, per subject.

In summary, Model 1 tests the network's ability to classify individual windows drawn at random from the entire dataset, whereas Model 2 uses a more rigorous approach by evaluating how well the network generalizes to an unseen seizure, including a post-processing pipeline. By using these two methodologies, we can quantify both "ideal" per-window performance (Model 1) and a more general, clinically relevant per-seizure performance (Model 2).
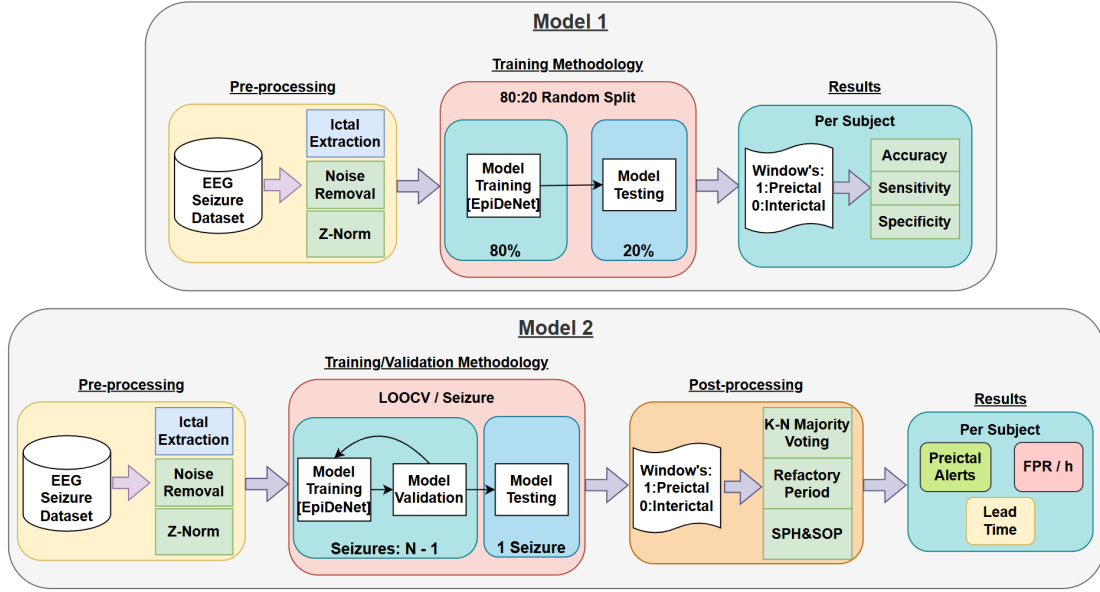

Figure 4.1: Model 1 and Model 2

4.2 Datasets

Before starting with the prediction task itself, the epileptic dataset used for the thesis should be defined. In the field, several publicly available EEG seizure datasets exist which became the benchmarks for developing and comparing seizure prediction algorithms. Among these, the **CHB-MIT** scalp EEG dataset [28] is one of the most widely used: it contains multi-hour recordings from pediatric patients with annotated seizure onsets and is especially popular. Other notable datasets include the **Freiburg iEEG dataset**, which provides high-resolution intracranial recordings from patients during presurgical evaluation, the **EPILEPSIAE** database, offering a large collection of continuous scalp-EEG (and associated video/clinical metadata) from adult subjects, the **Bonn University dataset**, which consists of short single-channel segments , and the **TUH EEG Seizure Corpus (TUSZ)**, a comprehensive clinical scalp-EEG collection primarily used
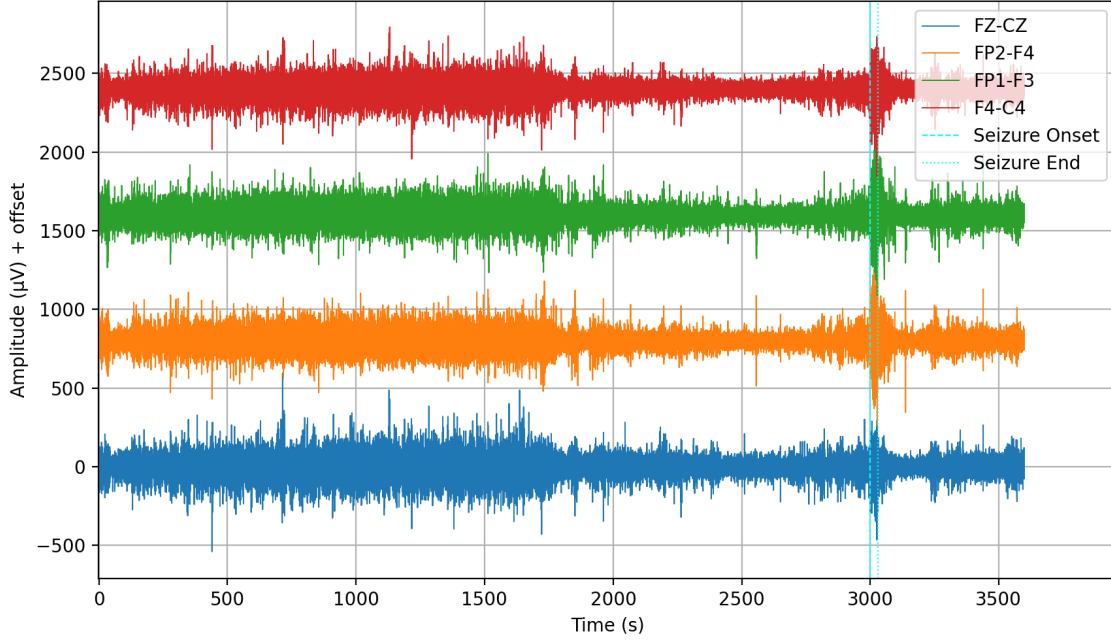
Figure 4.2: 4 EEG Channels of the CHB-MIT Dataset

for detection tasks.

In this thesis, we focus on the CHB-MIT dataset because it is the most widely used and has been employed in embedded implementations. The dataset includes 24 subjects (aged roughly 1.5 to 22 years), each recorded for approximately 20-160 hours, during which one or more seizures occurred. EEG signals were acquired using a 10–20 electrode montage with 23 scalp channels (plus two auxiliary channels for ECG for some subjects) all sampled at 256 Hz (some of which can be seen in Figure 4.2). The recordings are stored in European Data Format (EDF), with seizure events hand-marked by expert reviewers, where each annotation includes the exact start and end times of clinical seizures. In total, the database contains 198 clinically verified seizures across the 24 individuals. In the next sections, we describe how we extract fixed-length preictal and interictal segments from these EDF files, and how they are preprocessed to train our compact CNN model, EpiDeNet.

4.3 Data Extraction

As introduced in the Background (Chapter 2), seizure prediction is a binary classification task distinguishing two EEG states: interictal (the period between seizures, not directly related to an imminent seizure) and pre-ictal (the minutes leading up to seizure onset). Therefore, our first task is to correctly extract interictal and pre-ictal segments from the CHB-MIT dataset. The CHB-MIT data are stored in EDF format, which includes metadata such as recording start/end times, seizure start/end times (if any), the number

of channels, and any changes in channel configuration. We read this information and store it in a dictionary for each subject, so that each recording's relevant details are readily available for subsequent extraction.

Next, the **pre-ictal extraction** follows. We define "pre-ictal" as the 30 minutes immediately before seizure onset, since this is the most widely used pre-ictal duration [29]. We also experiment with other pre-ictal durations (e.g., 5, 10, and 60 minutes), and those results are presented in Chapter 5. When extracting a 30-minute pre-ictal segment, two complications can arise:

1. **Seizure-to-seizure proximity:** If two seizures occur less than 30 minutes apart, then the second seizure's nominal 30 minute pre-ictal window would overlap with the actual seizure or its post-ictal period. We therefore define a post-ictal period of 10 minutes (a commonly used range is 5–30 minutes [30]) and exclude any pre-ictal data that falls within that post-ictal window of a preceding seizure. We only accept pre-ictal windows of at least 10 minutes. If two seizures occur so closely that the available pre-ictal period is shorter than 10 minutes, we assign the earlier seizure's pre-ictal segment to both events.

2. **Recording gaps:** Between successive EDF recordings, there may be a gap of several seconds to hours during which no EEG was recorded. If a seizure begins less than 30 minutes after the start of one recording, then the pre-ictal window must draw from the tail end of the previous recording. For example, if seizure onset is 10 minutes into recording n, we first take those 10 minutes from recording n and then retrieve the remaining 20 minutes from the last portion of recording $n-1$, adjusting for any gap between the two files. If the gap is greater than 20 minutes in this case, we would only consider the 10 minutes from recording n, since we cannot reach recording n-1 due to the gap (This means that the first 20 min of the pre-ictal fall into the unrecorded time).

The next thing to consider is the **interictal extraction**. For that we flatten all the recordings into one continuous timeline and consider all the data as interictal that is away from the unsafe zone. Therefore we consider a buffer of 180 min between each pre-ictal start an post-ictal end, meaning that the interictal data is at least 3h away from any seizure related event/activity. Again this is not clear, but most studies use betwenn 60 min to 4h. The Data extraction is visualized in Figure 4.3

EEG channel configurations can vary between recordings for the same subject: most sessions have 23 channels (some have 24 or 26), but in some cases the montage changes partway through (after a couple of recordings), adding or removing electrodes. To handle these changing dimensions, we identify the smallest common channel set that appears in every recording for a given subject and use only those channels for all of that subject's sessions. In subjects with the same montage used across all recordings, the full channel set is used. The minimal channel set is 18 and for most patients the full, 23 channels, set is utilized.
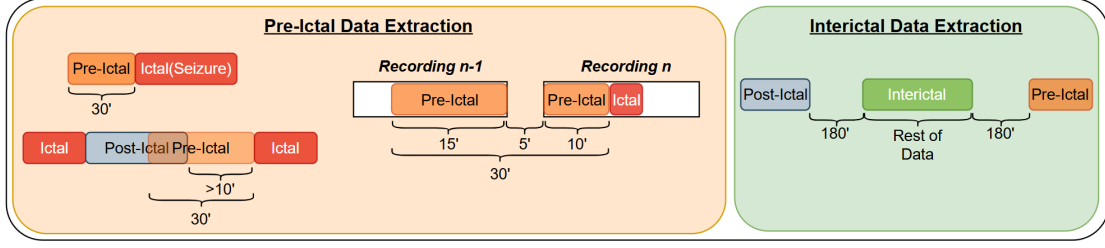
Figure 4.3: The Pre-Ictal and Interictal Extraciton

4.4 Data Preprocessing

To prepare the data for training we apply a 60 Hz notch filter (and its first harmonic at 120 Hz) to each channel in every EDF recording to remove power-line interference. Once the noise is suppressed, we perform z-score normalization on the EEG signals [31].Critically, all normalization statistics (namely a per-channel mean and standard deviation) are computed solely from the training folds. Those same values are then used to normalize the left-out fold and this way, no information from the validation set leaks into the training procedure and thus we have a more realistic estimate of how the model will perform on truly unseen data. Since we do not perform any feature extraction, filtering and normalization are the only preprocessing on the raw EEG inputs, which then can be fed to our classifier.

4.5 Deep Learning Architecture and Training

After the data extraction and preprocessing we have 2 types of data, the interictal and the preictal. The two models then differ in how they form training and test sets. In **Model 1** all interictal and preictal windows are pooled and randomly permuted. We then split this combined set into a training fold containing 80% of the windows and a test fold containing the remaining 20%. No seizure specific grouping is preserved resulting in each window being treated independently

Model 2 uses a different approach. Each subject has a number of seizures n, between one and 40, stretching over its recordings. After the data extraction, we have n preictal data chunks since each of them comes from one seizure (except for the case, when 2 seizures are in close proximity to each other). The fully extracted interictal data has length t. Then t/n is the interictal time divided by the number of preictal chunks,(i.e seizures). The interictal timeline, partitioned into n segments, looks like this:

$$\left[0, \frac{T}{n}\right), \quad \left[\frac{T}{n}, 2\frac{T}{n}\right), \quad \ldots, \quad \left[(n-1)\frac{T}{n}, T\right).$$

Each preictal chunk is then randomly combined with a continuous chunk of interictal data with size t/n. In the end we have n chunks, each containing t/n interictal data and one preictal chunk. Since model 2 uses Leave-One-Seizure-Out-Cross-Validation, we

train the model on $(n - 1)$ chunks and one is left out for testing. This is done for each chunk once, resulting in n-fold cross validation, where we test on the held out preictal (i.e seizure). In each fold, we also set aside 10% - 20% of the training pairs as an internal validation set to monitor overfitting and to decide which epoch's weights to use for the final evaluation. The whole process is visualized in figure 4.4.
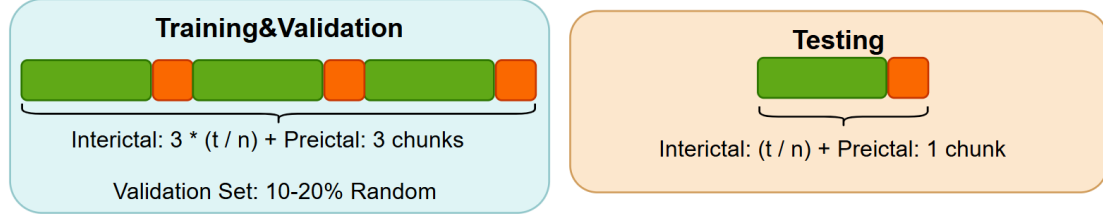

Figure 4.4: Here we have 4 preictal chunks(seizures, n=4) and t represents the amount of interictal data. This would be 4-fold CV, since each left out fold/chunk is used once for testing.

Before dealing with the network itself, we first have to look into the data and the loss function we want to use. Most of a subject's EEG recording (approximately 90% or higher) does not contain seizure-related activity. Hence, when we extract preictal (positive) and interictal (negative) segments from a given patient's data, we encounter a severe class imbalance: for some subjects, the ratio of interictal to preictal samples can range from 5:1 up to 50:1, depending on the number of seizures and the total recording duration. Under such conditions, using the standard Cross-Entropy (CE) loss during model training causes the classifier to overfit to the majority class (interictal), yielding poor sensitivity to the preictal state.
There are several techniques for dealing with class imbalance, including:

- **Focal Loss** [32], which under-weights well-classified examples and focuses learning on hard, misclassified samples.

- **Weighted CE Loss** [33], where different weights are assigned to positive and negative classes.

- **Oversampling/Undersampling** methods, in which one can undersample the majority class (interictal) and/or oversample the minority class (preictal). A variety of such sampling strategies are implemented in the scikit-imbalanced-learn library [34].

In this thesis, we adopt the **Subject-Specific Weighted Cross-Entropy (SSWCE)** loss [27], defined as:

$$\text{SSWCE}(y, p) = \text{CE}(y, p) + \alpha \left(1 - \text{SP}\right) + \beta \left(1 - \text{SN}\right),$$

where:

4 Implementation

- y is the true label and p is the predicted probability,

- SP denotes specificity (true-negative rate),

- SN denotes sensitivity (true-positive rate), and

- α and β are subject-specific hyperparameters.

By choosing α and β appropriately for each patient, SSWCE allows us to weight specificity and sensitivity independently, thereby directly addressing the imbalance between interictal and preictal examples.

The Data is then loaded into PyTorch tensors which have shape (B=32,1,C=17-23,T) where B is the batch size, C the EEG channels (the hight dimension) and T the Window/segment size (the width dimension, corresponding to time steps/samples). C differs per subject and T depends on the window size we use. Tested were window sizes from 5 to 60s, anf considering the sample size of 256, that would be T=1280 for 5s and T=15'300 for 60s. The additional dimension is added since we consider the original signal as a single "image channel" (i.e., a grayscale image if you draw an analogy to image processing). This is done because the tensors fed into the 2D Convolutional Layer in Pytorch are of the form (batchSize,inChannels,height,width).

Block	Type	#Filters	Kernel	Output
ϕ^1	Conv2D	4	(1,4)	$(4, C, T)$
	MaxPool		(1,8)	$(4, C, T/8)$
ϕ^2	Conv2D	16	(1,16)	$(16, C, T/8)$
	MaxPool		(1,4)	$(16, C, T/32)$
ϕ^3	Conv2D	16	(1,8)	$(16, C, T/32)$
	MaxPool		(1,4)	$(16, C, T/128)$
ϕ^4	Conv2D	16	(16,1)	$(16, C, T/128)$
	MaxPool		(4,1)	$(16, C/4, T/128)$
ϕ^5	Conv2D	16	(8,1)	$(16, C/4, T/128)$
	AdaptiveAvgPool		(1,1)	$(16, 1, 1)$
	Flatten			(16)
	Dropout			(16)
ϕ^6	Dense			(2)

Table 4.5 demonstrates the CNN architecture used. The model takes the tensors described above as input followed by several convolution, pooling, and batch normalization layers. Finally, the features are aggregated and fed through a fully connected layer that outputs two logits corresponding to the two classes (i.e., Interictal or Pre-Ictal).

In the initial convolutional blocks ϕ^1 - ϕ^3 the network primarily learns frequency-related representations. By applying convolutions with small kernels along the time dimension, the model can detect characteristic spectral signatures, such as alpha waves(8–12Hz) or

other critical EEG frequency ranges. As the network proceeds to intermediate layers ϕ^4 - ϕ^5 it begins to combine these frequency filters with spatial filters. Because each EEG channel can carry a slightly different brain signal, these blocks identify where (i.e., which channels or sets of channels) certain frequency patterns occur. Ultimately, in the final layers (ϕ^6, the fully connected layer), the model fuses together the frequency-specific and spatially-aware features to produce class logits. This final step is where the network decides, "Given the frequency–spatial signature of this EEG window, does it most likely belong to interictal or preictal?"

4.6 Post-Processing

This stage is only relevant for Model 2, as Model 1 does not employ any post-processing. Once the network has been trained, it outputs a binary label for each window, either interictal (class 0) or preictal (class 1). However, window-level predictions alone do not directly translate into meaningful, per-seizure predictions. To derive per-seizure alerts and compute metrics such as event sensitivity, false-alarm rate per hour (FPR/h), and lead time, we apply a post-processing pipeline consisting of:

1. **K-of-N Majority Voting.** Each window produces a prediction, either 0 or 1. We slide a fixed-length buffer of N consecutive windows and declare the n-th window as 1 ("positive") if at least K of those N preceeding windows are classified as preictal. This reduces spuriosity from isolated "preictal" votes and helps smooth the output of the classifier.

2. **Refractory Period (R).** After a positive prediction (preictal), we want to suppress the following positive windows in the next R minutes. This prevents multiple alerts from clustering around the same seizure event.

3. **Alarm-Raising Logic with SPH and SOP.** We iterate through all post-processed predictions in chronological order. Whenever a positive prediction occurs (indicating we are in the pre-ictal phase), we immediately raise an alarm, since a seizure is presumed imminent. To decide whether that alarm is correct, we apply the SOP/SPH framework:

   - **Seizure Prediction Horizon (SPH):** the final interval of length SPH immediately before seizure onset during which any alarm is too late to be acted upon because we want to give the patient some time to be able to react. Alarms whose timestamps fall within

   $$\left(t\_\text{onset} - \text{SPH}, t\_\text{onset} \right]$$

   are simply ignored (they are neither counted as true positives nor as false alarms).

   - **Seizure Occurrence Period (SOP):** the interval of length SOP immediately preceding the SPH. In other words, the "true-positive window" is

   $$\left[t\_\text{onset} - \text{SOP}, t\_\text{onset} - \text{SPH} \right].$$

Any alarm timestamp in this window is counted as a correct prediction (true positive), because it guarantees at least SPH seconds of warning and occurs no more than SOP + SPH seconds before onset and it occurs during the preictal.

- **False alarms:** Any raised alarm that does not lie in the TP window and also does not fall within the SPH interval is counted as a false positive.

In code, we compute:

```
valid_start = t_onset - SOP
valid_end   = t_onset - SPH
```

and then classify each alarm as follows:

a) If valid_start $\leq$ alarm_time $\leq$ valid_end, it is a true positive.

b) If $t\_onset - \text{SPH} <$ alarm_time $\leq t\_onset$, it is ignored (too late).

c) Otherwise, it is a false positive.

This ensures that alarms are only counted as correct if they appear early enough (outside the SPH) yet still within the pre-ictal window (the SOP).

After passing through these three steps, each held-out seizure yields exactly one binary outcome ("alert" vs. "no alert"), a false-alarm rate per hour (FPR/h), and a lead time (the interval between the first valid alert and the actual seizure onset). We then average these metrics across all seizures in a subject to obtain a per-subject event sensitivity, FPR/h, and mean lead time.

Chapter 5

Results

In this chapter, we present quantitative results demonstrating the performance of our seizure-prediction pipeline. We report window-level metrics from model 1 and event-level (per-seizure) metrics from model 2 and check for the tradeoff in post-processing. We then examine how reducing the number of EEG channels (from full montage down to four electrodes) affects both AUC and event-level metrics, and compare our best configurations against prior state-of-the-art (SoA) methods and go into the current limitations and difficulties faced during the whole process.

5.1 Window-Level Performance & Comparison to SoA

As shown in Figure 4.1, Model 1 is trained on an 80:20 split and reports only window-level metrics. In contrast, Model 2 uses a leave-one-out cross-validation (LOOCV) procedure so that each seizure (its preictal data) is held out during training, afterwards its window-level predictions are then aggregated to produce event-level results.

In this section we will focus on Model 1 and compare it against the state of the art. Many seizure-prediction studies report only per-window metrics, using a random train/test split to see if a classifier can distinguish interictal from preictal windows, regardless of which seizure they came from. Further they also only include "lead" seizures, which are seizures at least four hours away from any preceding seizure. We follow the same approach in order to have a basis for comparison with those works. Regarding the train/test split, preictal segments from every seizure end up in both training and testing. However, this means that the model has seen preictal data from every seizure during training and is tested on different windows of the same seizures. Although this method demonstrates that preictal and interictal windows are distinguishable, it isn't practical in real life, because we can't train on data from a seizure before it actually happens.

We implemented Model 1 primarily to compare with existing state-of-the-art methods (Table 5.1), where EpideNet even outperforms a much larger CNN. Note that EpiDeNet was trained on perfectly balanced preictal-and-interictal data, whereas the CNN relied on over- and under-sampling to achieve balance. Also, the CNN's reported scores are the

average of 10 independent runs, while EpiDeNet was run just once (although repeated trials showed negligible variance). Despite using **fifty times** fewer parameters, EpiDeNet achieves essentially the same ability to distinguish preictal from interictal windows as the much larger CNN, when sizure prediction itself is not considered directly. Our main focus, however, is Model 2, which uses a leave-one-out scheme so that each seizure (and its preictal segment) is completely held out during training which is closer to a real-life scenario and reports performance only on unseen seizures.

It's worth noting that in [35], authors reports "FPR/h," but since they only compute per-window metrics (and never describe any post-processing), this is almost certainly just the per-window false-positive rate (i.e., 1 – specificity), not actual false alarms per hour. When applying true FPR/h, which is calculated as false alarms divided by total interictal hours, then even a single false alarm in 20 h of recording (as in Subject 1) would yield an FPR/h of 0.05, not 0.001. And also when running the script for 10 runs(as they did), and only having the one false positive, this would still result in more than 0.001 when taking the average .In other words, when refactoring the FPR calculation, the landscape of the work done in [35] changes significantly.

Subject	LS / Seizure	CNN (Params: 0.5M)			EpiDeNet (Params: 9k)		
		Sens. (%)	FPR (1-Spec)	AUC (%)	Sens. (%)	FPR (1-Spec)	Acc. (%)
chb01	3 / 7	100.0	0.001	100.0	100.0	0.00	100.00
chb05	3 / 5	99.7	0.072	99.3	100.0	0.00	100.00
chb06	6 / 10	96.0	0.138	96.7	98.23	0.08	95.37
chb08	3 / 5	99.9	0.027	99.8	100.0	0.00	100.00
chb10	5 / 7	97.9	0.095	98.5	100.0	0.01	99.44
chb14	4 / 8	98.9	0.109	98.3	94.20	0.03	95.83
chb22	3 / 3	99.5	0.078	99.2	100.0	0.04	97.96
Overall	**27 / 45**	98.8	0.074	98.8	98.92	0.02	98.37

Table 5.1: Comparison of CNN[35] vs. EpiDeNet (Random 80:20 split).

5.2 Event-Level Performance & Comparison to SoA

In Section 4.5, we described how Model 2 is trained and tested using LOOCV, and in Section 4.6 we explained how per-window predictions are converted into per-seizure events. Here, we report the event-level results on truly unseen seizures, including seizure sensitivity, false-positive rate per hour (FPR/h). Table 5.2 compares EpiDeNet's performance (after post-processing) against several state-of-the-art models.

EpiDeNet achieves a comparable event-level sensitivity i.e., the fraction of seizures correctly predicted per subject, but its FPR/h is about 0.15 higher than the best benchmarks. The window-level AUC (measured before post-processing) is similar to SeizureNet's and only slightly below ResNet or TGCNN. Note that event-level sensitivity and FPR/h depend strongly on the chosen post-processing parameters ($K = 8$, $N = 10$ here), and

Figure 5.1 shows how changing *K* and *N* affects these metrics.

It is worth noting that our comparison models also reserve a validation set during training to prevent overfitting, but they do so by taking the last 25 % of both preictal and interictal training data. In contrast, we found that using a randomly chosen 10 % validation split, rather than the final 25 % did improved event sensitivity by about 10–15 %. Although randomly sampling time-series windows could be problematic, our final evaluation still only uses completely unseen preictal data since the randomly sampled data is only used for the internal validation set. Some of the trade-offs and benefits of this choice are discussed in Section 5.6. The other 3 models used 30 s windows and STFT preprocessing, whereas we processed almost raw EEG in 5 s windows, applying only normalization and noise filtering.

We also evaluated a version with no internal validation set (as model 2 4.1 but in the training loop, the green square, there is no internal model validation, and we directly proceed to the testing), using 100% of each subject's data for training. In that case, we must manually decide when to stop training for each subject, which is an approach related to tuning α and β, since it has to be done for each subject specifically. While this could yield better performance, it means the stopping epoch (and other hyperparameter) must be chosen separately for every subject to prevent overfitting.

Unlike the three SoA models in Table 5.2, which balance preictal/interictal data via over/undersampling and use standard cross-entropy loss, EpiDeNet is trained with SS-WCE (Subject-Specific Weighted CE). We set $\alpha = 0.5$ and $\beta = 1.0$ to compensate for class imbalance, but these parameters could be tuned per subject (e.g., via grid search). In fact, for Subject 19 the default α, β resulted in zero event sensitivity. Increasing the interictal buffer time and adjusting α, β lead to the current result. Other than that, altering *K* and *N* improved its performance slightly, but it remained lower than for other subjects and lower than the competing models.

Patient	SeizureCNN (Params: 200k)			ResNet (Params: 11.7M)			TGCNN (Params: 100k)			EpiDeNet (Params: 9k)		
	AUC (%)	Event (%)	FPR (/h)	AUC (%)	Event (%)	FPR (/h)	AUC (%)	Event (%)	FPR (/h)	AUC (%)	Event (%)	FPR (/h)
Pt 1	91.9	85.7	0.24	93.4	100.0	0.22	99.9	100.0	0.17	99.7	100.0	0.504
Pt 2	33.5	33.3	0.00	63.1	33.3	0.00	73.1	66.7	0.09	49.1	33.3	0.508
Pt 3	96.8	100.0	0.18	94.5	100.0	0.18	98.8	100.0	0.09	88.5	83.3	0.106
Pt 5	87.1	80.0	0.19	93.8	100.0	0.07	99.9	100.0	0.00	91.5	100.0	0.201
Pt 8	85.6	80.0	0.29	89.3	80.0	0.29	95.8	100.0	0.00	86.8	80.0	0.575
Pt 9	74.2	50.0	0.12	64.7	50.0	0.08	83.1	50.0	0.06	64.2	50.0	0.412
Pt 10	55.6	33.3	0.00	78.5	66.7	0.27	87.8	83.3	0.46	90.5	85.7	0.233
Pt 13	96.8	80.0	0.14	94.7	80.0	0.13	99.7	100.0	0.13	90.5	83.3	0.457
Pt 14	66.2	80.0	0.40	72.4	80.0	0.71	81.7	80.0	0.61	95.0	100.0	0.384
Pt 16	89.8	90.0	0.36	85.4	90.0	0.18	94.7	90.0	0.18	87.1	100.0	0.389
Pt 17	76.4	66.7	0.50	63.3	66.7	0.40	87.4	66.7	0.30	66.6	66.7	0.706
Pt 18	93.5	100.0	0.28	92.1	100.0	0.16	98.8	100.0	0.16	56.3	75.0	0.302
Pt 19	99.9	100.0	0.00	98.4	100.0	0.00	99.5	100.0	0.00	84.6	80.0	0.104
Pt 20	98.4	100.0	0.25	98.8	100.0	0.14	99.9	100.0	0.19	97.3	83.3	0.215
Pt 21	90.3	100.0	0.23	94.8	100.0	0.23	95.5	100.0	0.14	83.7	100.0	0.419
Pt 23	99.9	100.0	0.33	99.9	100.0	0.07	99.9	100.0	0.00	99.5	100.0	0.197
Aver	**83.5**	**81.7**	**0.175**	**86.1**	**86.6**	**0.152**	**93.5**	**91.5**	**0.145**	**83.2**	**82.5**	**0.357**

Table 5.2: Per-Patient Performance Comparison: SeizureNet[19] , ResNet[36], TGCNN[36], and EpiDeNet[27]

5.3 Tuning of Post-Processing and Risk Levels

After training Model 2, we obtain per-window predictions, which we then convert into event-level metrics using post-processing. As described in Section 4.6, the most important post-processing parameters are K and N in the majority-voting scheme that smooths the window-level outputs.

Figure 5.1 illustrates the trade-off between event sensitivity and FPR/h for different values of K (with $N = 10$ fixed). We used the same subjects as in Table 5.2, except for Subject 19, because its model outputs were generated earlier using the default $K = 8$, $N = 10$ and were not saved for re-tuning.

From the plot, we see that a smaller K raises event sensitivity but also increases FPR/h, while a larger K reduces FPR/h at the expense of sensitivity. The choice of K therefore depends on whether maximizing sensitivity or minimizing false alarms is more important for a given application. High event sensitivity and low FPR/h are desirable nevertheless.
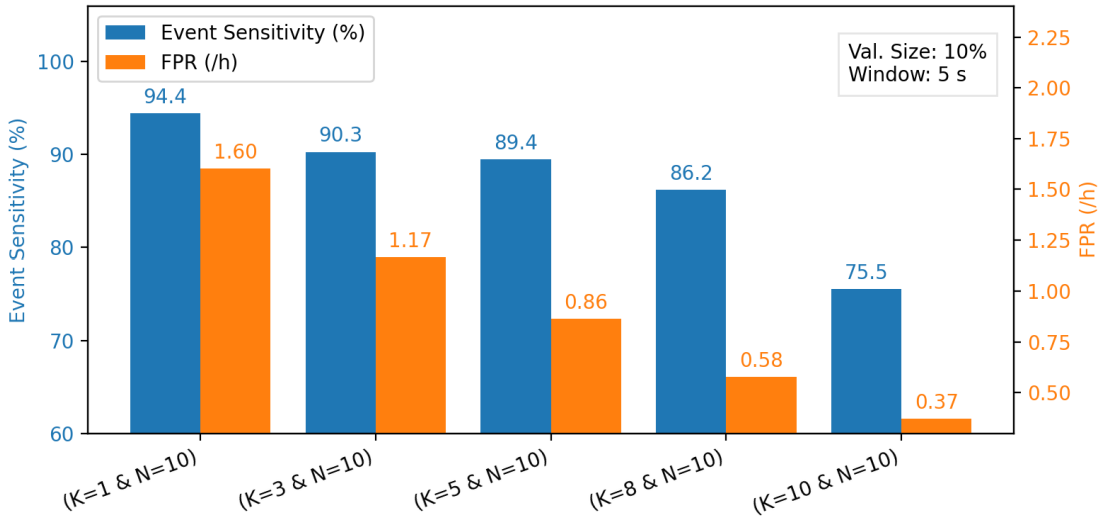

Figure 5.1: This shows the Postprocessing Tradeoff for the K-N Majority voting, with different values for K

Instead of issuing discrete alarms, we can output a continuous "risk level" by computing a probability score for each window based on the previous N predictions. In other words, rather than requiring K out of N windows to be classified as preictal before raising an alarm, we assign each window a risk score equal to the fraction of the last N windows predicted as preictal. Figure 5.2 illustrates this approach. Because we randomly mix interictal and preictal segments, the actual preictal period (marked by the blue dashed line) may begin earlier or later than expected. After that line, we observe how the risk level steadily climbs toward seizure onset.

In this example, we set the "low→moderate" threshold at $T\_1 = 0.3$ and the "moderate→high" threshold at $T\_2 = 0.7$. An alarm is effectively raised whenever the risk score

exceeds 0.7 (i.e., $K = 7$ out of $N = 10$ windows are preictal). Notice in Figure 5.2 that the first spike around 18:43 is a false alarm, since it occurs before the dashed line (i.e., during interictal data). A second red peak, which is above 0.7, does not produce a new alarm because it falls within a refractory period. Once the true preictal segment begins (after the dashed line), the model correctly predicts every window, and the risk level remains at its maximum thereafter.

Because each subject and seizure behave differently, some risk-level plots show almost no "moderate risk" during interictal periods and remain at "high risk" throughout the entire preictal phase. Others display more noise, with red peaks appearing even before the dashed line. This visualization makes it easy to see how risk grows over time and how our chosen thresholds influence false alarms and timely alerts.

The refractory period prevents an alarm from being raised for every positive window. Without it, every window in the preictal phase (after the initial buildup) would trigger a new alarm. In Figure 5.2, the refractory period also suppresses what would have been a second false alarm, which in this case is rather lucky.
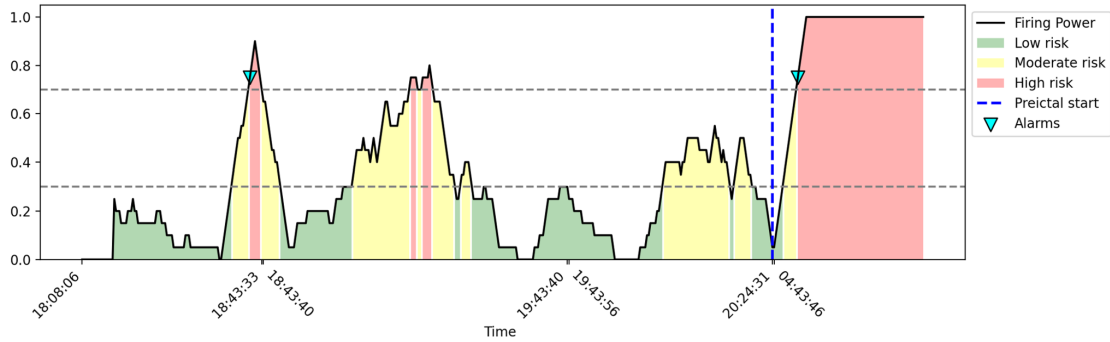

Figure 5.2: Risk Levels over time. The blue dashed line indicates where the preictal starts.

5.4 Impact of Different Preical Lengths

Another consideration is how varying the fixed preictal window affects performance, as shown in Figure 5.3. When we increase the preictal length, the average lead time before seizure onset also increases (since "preictal" now covers more time) and the FPR/h generally decreases. However, changing the preictal duration alters the class imbalance: for example, using a 60 min preictal window instead of 30 min doubles the amount of preictal data, halving the interictal-to-preictal ratio. We must account for this imbalance in our loss function, which may explain some of the observed sensitivity and FPR/h trends.

These findings align with studies that fixed different preictal lengths per patient (e.g., [20]), where longer preictal windows typically led to lower FPR/h and higher event sensitivity. Finally, recent work suggests that the optimal preictal interval can vary

by subject, and even by individual seizure, rather than being a single fixed duration. For example, Segal et al. [37] determine each seizure's preictal length algorithmically using multiple EEG features, rather than assuming a uniform preictal duration across all events.
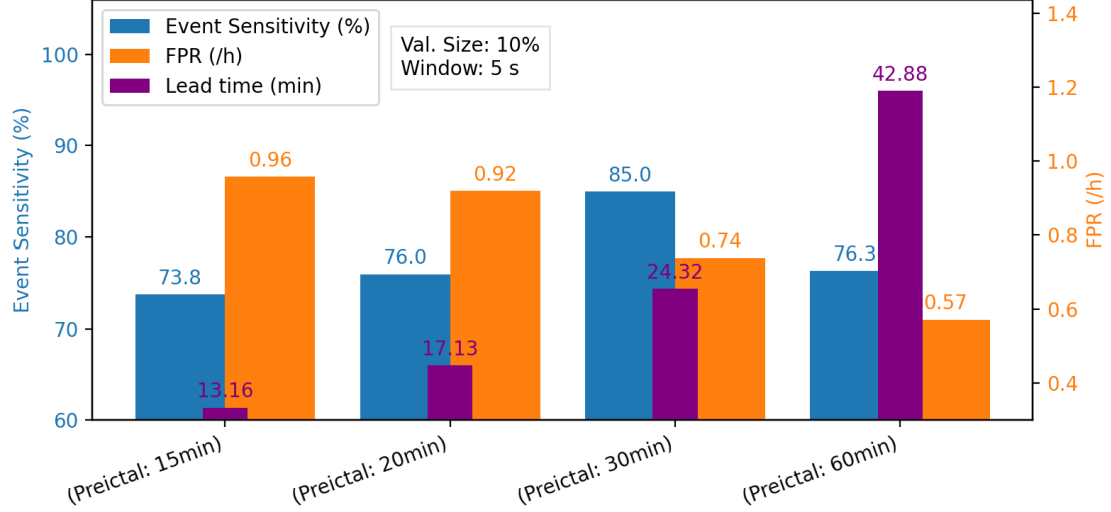

Figure 5.3: The effect when choosing different Preictal Lengths .

5.5 Impact of Channel Reduction

Another important experiment is reducing the number of EEG channels to find a configuration suitable for a wearable device. We evaluated five channel sets: 4, 8, 12, 16, and the full montage, and plotted the results for **all** subjects (except 12 & 24) in Figure 5.4. (Note: for some subjects, "full" contains only 18 electrodes, so results could be similar to the 16-channel case.)

The exact channels used for each configuration, also considering their suitability for wearable's were:

- **4 Channels (wearable glasses)**: `FP1-F7, F7-T7, FP2-F8, F8-T8`
  Wearability: There exist wearable devices, that use 4 out of the 6 electrodes [38] (Note T3 and T4 have been renamed to T7 and T8 but are the same), or [39] that also makes use of Fp1 & Fp2.

- **8 Channels**: `FP1-F7, F7-T7, T7-P7, FP1-F3, FP2-F8, F8-T8, FP2-F4, F4-C4`
  Wearability: Including slightly more electrodes and channels, to capture more brain activity, similar to [40].

- **12 Channels**: `FP1-F7, F7-T7, T7-P7, P7-01, FP1-F3, F3-C3, FP2-F8, F8-T8, T8-P8, P8-02, FP2-F4, F4-C4`

- **16 Channels**: `FP1-F7, F7-T7, T7-P7, P7-01, FP1-F3, F3-C3, C3-P3, P3-01, FP2-F8, F8-T8, T8-P8, P8-02, FP2-F4, F4-C4, C4-P4, P4-02, FZ-CZ, CZ-PZ` **Wearability:** Probably to much electrodes/channels for a every day wearable [41].

- **Full Montage**: all available scalp channels (up to 23, but sometimes only 18 for certain subjects)

Fewer EEG electrodes improve wearability, power consumption, and patient comfort but come at the cost of reduced spatial coverage, which can delay or even miss seizures originating outside the remaining channels' monitoring areas [42]. By comparing these five configurations, we see the effects of having less channels which is crucial for the design of a practical, everyday wearable system. Other works implementing a reduced channel set for seizure prediction include [43, 44, 45, 46, 47, 48, 49]. Note that the 18 common channels across every subject and recordings are: 'FP1-F7', 'F7-T7', 'T7-P7', 'P7-O1', 'FP1-F3', 'F3-C3', 'C3-P3', 'P3-O1', 'FP2-F4', 'F4-C4', 'C4-P4', 'P4-O2', 'FP2-F8', 'F8-T8', 'T8-P8', 'P8-O2', 'FZ-CZ', 'CZ-PZ'.
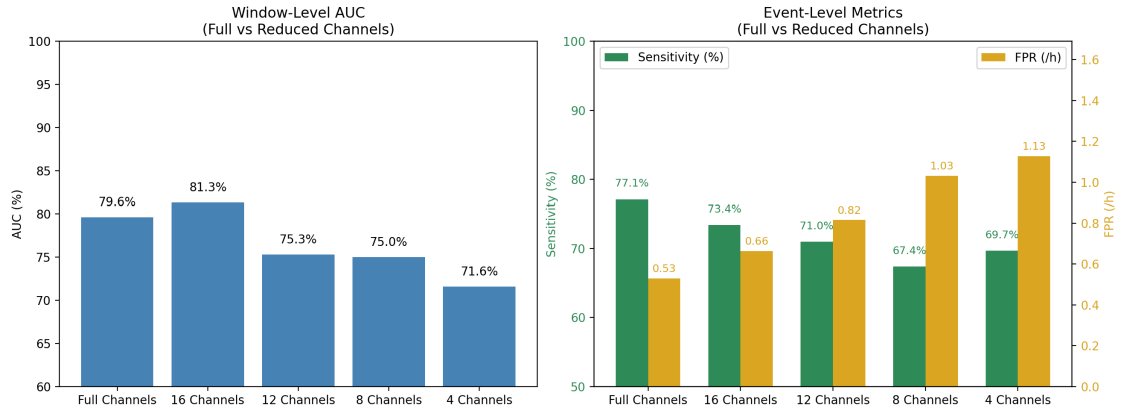

Figure 5.4: Left Plot: Per-window AUC scores for the different channel counts. Right Plot: The Post-Processing metrics event-sensitivity and FPR/h.

5.6 Current limitations and Difficulties

During our developement we faced two key issues. One was the **class imbalance** between preictal and interictal, which was tackled via a custom tunable Cross Entropy loss function. Another issue was also the difficulty using the last 25% form the training set as the internal validation set, used to guard against overfitting. In our experiments, withholding the last 25% for validation sometimes gave poor validation stability (high validation loss, erratic metrics) even though the model later performed well on a truly held-out seizure (LOOCV). Conversely, a seemingly stable validation loss sometimes gave poor generalization on the unseen seizure. To avoid these inconsistencies, we switched to randomly sampling 10 % of the training windows (preictal/interictal) for

validation. This ensured the validation set always contained a representative mix of preictal and interictal periods, even if those "preictal" windows came from multiple seizures, hence the training-stop criterion (early stopping) became more reliable. Note that, although random sampling solves the internal validation stability issue, it does not guarantee that the network will generalize to a held-out seizure, since each seizure's onset patterns can differ significantly, leading to some preictal patterns of seizures being hardly detectable for the model.

Conserning the **imbalance**: Figure 5.6 summarizes each subject's metrics, averaged over 4 runs. *Interictal (h)* denotes the total hours labeled as interictal for each subject, depending on the number of seizures and the interictal–preictal buffer. When the interictal–preictal imbalance exceeds around 15:1 (seen in Subjects 2, 4, 7, 9, 11, and 19), the model often overfits to the majority (interictal) class, predicting almost no preictal windows and yielding very low overall event sensitivity and high FPR/h. Also, very low per-window sensitivity can be observed in those cases. Subject 11 is the only exception: despite an imbalance >15, it still achieved reasonable sensitivity, suggesting easier preictal signatures. To improve performance for these highly imbalanced cases, we must more aggressively adjust SSWCE's weights (α, β) or explore additional strategies (e.g., focal loss or oversampling, etc.) as discussed in 4.5, since SSWCE has primarily been validated on detection rather than prediction tasks. Furthermore, some subjects (e.g., 2 and 4) appear inherently more difficult: even reducing interictal duration (increasing the buffer) did not significantly help because their preictal patterns seem less pronounced and harder to learn. Also note, subject 12 was excluded since 40 seizures occur for this subject and with our currently defined buffers we get no interictal data and for subject 24 the data extraction has to be changed slightly, in comparison with the other subjects. All results in Table 5.6 use a 5s window size. We also tested longer windows (e.g., 10s, 30s) and found lower FPR/h in some cases, but overall event sensitivity did not uniformly improve. Note that these metrics can shift when varying postprocessing parameters; our default smoothing uses $K = 8$, $N = 10$.

5.7 Implementation Details

The models were implemented in `PyTorch` and trained on the Integrated Systems Laboratory's (ETH Zürich) `sassauna` GPU servers. Code development was performed using Visual Studio Code, with real-time monitoring of training progress and metrics via TensorBoard. The input data set was the CHB-MIT scalp EEG database [28], which contains multi-channel recordings from 23 pediatric subjects, including 198 seizures across 844 hours of continuous EEG.

5.7.1 Hyperparameter Settings and Configuration

All experiments share the following base hyperparameters unless otherwise noted:

- **Batch size:** 32

Subject	AUC	Sensitivity	Specificity	FPR	Accuracy	Overall Sens.	False Alarms	Interictal (h)	FPR/h	Lead (min)	Imbalance
1	0.989 ± 0.013	0.871 ± 0.005	0.838 ± 0.018	0.162 ± 0.018	0.842 ± 0.013	0.917	9.5	15.87	0.599	28.66 ± 1.01	4.53
2	0.510 ± 0.258	0.242 ± 0.021	0.809 ± 0.110	0.191 ± 0.110	0.778 ± 0.103	0.500	16.5	25.61	0.644	19.62 ± 7.72	17.07
3	0.889 ± 0.011	0.640 ± 0.098	0.968 ± 0.003	0.032 ± 0.003	0.937 ± 0.012	0.917	4.0	28.39	0.141	27.18 ± 0.73	8.11
4	0.424 ± 0.119	0.005 ± 0.002	0.887 ± 0.104	0.113 ± 0.104	0.877 ± 0.103	0.000	43.0	133.35	0.322	—	66.67
5	0.867 ± 0.029	0.563 ± 0.057	0.899 ± 0.005	0.101 ± 0.005	0.851 ± 0.013	0.900	5.0	14.96	0.334	27.50 ± 1.82	5.98
6	0.713 ± 0.030	0.460 ± 0.032	0.794 ± 0.044	0.206 ± 0.044	0.740 ± 0.032	0.750	18.5	26.15	0.707	20.72 ± 3.96	5.23
7	0.438 ± 0.121	0.044 ± 0.038	0.911 ± 0.003	0.089 ± 0.003	0.886 ± 0.004	0.167	23.0	51.32	0.448	$21.67 \pm nan$	34.21
8	0.867 ± 0.047	0.782 ± 0.017	0.848 ± 0.029	0.152 ± 0.029	0.809 ± 0.022	0.800	1.5	1.74	0.863	26.44 ± 1.92	0.70
9	0.670 ± 0.066	0.442 ± 0.052	0.805 ± 0.049	0.195 ± 0.049	0.791 ± 0.049	0.625	31.0	48.59	0.638	24.71 ± 5.72	24.30
10	0.836 ± 0.067	0.584 ± 0.055	0.874 ± 0.078	0.126 ± 0.078	0.841 ± 0.075	0.857	11.0	25.78	0.427	22.14 ± 1.95	7.37
11	0.970 ± 0.022	0.680 ± 0.069	0.958 ± 0.003	0.042 ± 0.003	0.949 ± 0.005	1.000	8.5	32.00	0.266	19.26 ± 6.10	32.00
13	0.899 ± 0.024	0.847 ± 0.003	0.834 ± 0.007	0.166 ± 0.007	0.831 ± 0.013	0.917	8.0	15.30	0.523	21.60 ± 1.73	3.06
14	0.864 ± 0.003	0.811 ± 0.031	0.696 ± 0.020	0.304 ± 0.020	0.738 ± 0.001	1.000	6.0	5.21	1.151	21.79 ± 3.74	1.30
15	0.917 ± 0.000	0.835 ± 0.001	0.834 ± 0.037	0.166 ± 0.037	0.831 ± 0.014	0.733	3.0	1.33	2.251	24.02 ± 0.05	0.16
16	0.755 ± 0.001	0.680 ± 0.001	0.802 ± 0.017	0.198 ± 0.017	0.757 ± 0.011	0.750	3.0	5.14	0.583	26.68 ± 3.60	1.14
17	0.640 ± 0.003	0.510 ± 0.022	0.815 ± 0.001	0.185 ± 0.001	0.786 ± 0.001	0.667	7.5	14.16	0.530	28.94 ± 0.32	9.44
18	0.616 ± 0.010	0.345 ± 0.121	0.907 ± 0.002	0.093 ± 0.002	0.868 ± 0.007	0.625	10.5	26.48	0.397	23.59 ± 1.23	8.83
19	0.896 ± 0.035	0.053 ± 0.050	0.997 ± 0.001	0.003 ± 0.001	0.963 ± 0.001	0.250	0.0	27.00	0.000	$29.17 \pm nan$	18.00
20	0.956 ± 0.018	0.801 ± 0.012	0.969 ± 0.041	0.031 ± 0.041	0.951 ± 0.037	0.833	2.5	18.57	0.135	27.35 ± 2.31	4.64
21	0.824 ± 0.048	0.604 ± 0.010	0.815 ± 0.067	0.185 ± 0.067	0.799 ± 0.063	1.000	14.0	23.87	0.586	25.41 ± 1.22	11.94
22	0.790 ± 0.049	0.439 ± 0.009	0.907 ± 0.008	0.093 ± 0.008	0.872 ± 0.007	1.000	7.5	17.56	0.427	22.53 ± 1.20	11.71
23	0.998 ± 0.002	0.957 ± 0.009	0.958 ± 0.010	0.042 ± 0.010	0.958 ± 0.010	1.000	3.0	15.22	0.197	28.47 ± 1.23	4.35
ALL	0.788 ± 0.172	0.554 ± 0.279	0.869 ± 0.074	0.131 ± 0.074	0.848 ± 0.071	0.737	10.8	26.07	0.553	24.64 ± 3.22	– – –

Table 5.3: Per-Subject Performance Metrics. Excluded is Subject 12 and 24. Metrics: AUC, Sensitivity, Specificity, FPR, Accuracy are per-window, where as Overall Sens. False Alarms, Interictal, FPR/h, Lead are per-event metrics.

- **Number of epochs:** 100 (with early stopping)

- **Window Size:** 5s (How may samples were considerd to create one labled window (preictal or interictal) used for training.

- **Runs:** 2-4, this are the number of runs each subject specific model was ran for each subject, then taking the average from the runs to get the final metrics.

- **Optimizer:** Adam with learning rate $= 1 \times 10^{-4}$ and weight decay $= 1 \times 10^{-4}$

- **Dropout:** 0.5 applied immediately before the final fully-connected layer

- **Validation split:** 10% of training data randomly held out for validation

- **Early stopping:** Training halted if validation loss did not improve for patience $= 30$ consecutive epochs

- **Loss function:** the SSWCE loss function with parameters $\alpha = 0.5$, $\beta = 1.0$ (to put more enphasis on the sensitivity: the correct prediction of preictal samples) for all subjects (This can be tuned for each subject individually, depending on the data imbalance for that subject

5.7.2 Post-Processing and Evaluation Parameters

During inference, the windows from the model output were post-processed using a K-of-N majority-voting filter to reduce spurious alarms:

- **Majority voting:** $k = 8$, $n = 10$ (i.e. at least 8 positive windows out of the most recent 10 must be positive to issue an alarm)

- **Refractory period:** $R = 30$ minutes after any raised alarm (to avoid repeated alarms during the same preictal phase)

- **Seizure Prediction Horizon (SPH):** 5 minutes

- **Seizure Occurrence Period (SOP):** 30 minutes

Chapter 6

Conclusion and Future Work

This work investigated whether seizure-prediction algorithms can be scaled down to run on wearable EEG devices without sacrificing performance too much. Our core problem was that most state-of-the-art predictors rely on large deep learning networks and high-density electrode arrays, which are constraints that make continuous, real-world monitoring impractical. To address this, we make use of EpiDeNet, a compact convolutional neural network of roughly 9 k parameters, trained and validated on the CHB-MIT database under leave-one-seizure-out cross-validation and with standard post-processing for seizure prediction and report per-seizure metrics such as event sensitivity, FPR/h and lead time.

Our evaluation shows that EpiDeNet matches state-of-the-art models (Table 5.2) in event-level sensitivity and false-alarm rate, while using far fewer parameters. When we reduce from the full montage to just four channels (a 6× reduction), we lose about 7 % in seizure sensitivity and roughly double FPR/h. However, as Figure 5.1 demonstrates, post-processing parameters (e.g., K and N) can be tuned to recover FPR/h at the cost of sensitivity, or vice versa. Earlier experiments with different window sizes or smoothing settings produced smaller relative changes in sensitivity and FPR/h, though their absolute sensitivity was lower.

Future work could focus on subject-specific tuning, once for the SSWCE loss (adjusting α and β) and for other hyperparameters, in order to improve performance due to the imbalance and also on "harder" seizures, and/or subjects. Architectural refinements may also help. In particular, combining EEG with ECG, PPG, or accelerometry has shown promise for seizure detection [50], and a similar fusion could boost prediction accuracy. Finally, the deployment of the full pipeline on an embedded platform (e.g., BioGAP [51]), and preliminary studies of power consumption and inference latency are discussed for EpiDeNet in [27].

List of Figures

List of Tables

Bibliography

[1] T. M. Ingolfsson, A. Cossettini, X. Wang, E. Tabanelli, G. Tagliavini, P. Ryvlin, L. Benini, and S. Benatti, "Towards long-term non-invasive monitoring for epilepsy via wearable eeg devices," in *2021 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 2021, pp. 01–04.

[2] P. Busia, A. Cossettini, T. M. Ingolfsson, S. Benatti, A. Burrello, M. Scherer, M. A. Scrugli, P. Meloni, and L. Benini, "Eegformer: Transformer-based epilepsy detection on raw eeg traces for low-channel-count wearable continuous monitoring devices," in *2022 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 2022, pp. 640–644.

[3] D. Sopic, A. Aminifar, and D. Atienza, "e-glass: A wearable system for real-time detection of epileptic seizures," in *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2018, pp. 1–5.

[4] J. Munch Nielsen, I. Zibrandtsen, P. Masulli, T. Sørensen, T. Andersen, and T. Wesenberg Kjær, "Towards a wearable multi-modal seizure detection system in epilepsy: A pilot study," *Clinical Neurophysiology*, vol. 136, pp. 40–48, 2022.

[5] R. Andrzejak, H. Zaveri, A. Schulze-Bonhage, M. Leguia, W. Stacey, M. Richardson, L. Kuhlmann, and K. Lehnertz, "Seizure forecasting: Where do we stand?" *Epilepsia*, vol. 64, 03 2023.

[6] B. H. Brinkmann, P. J. Karoly, E. S. Nurse, S. B. Dumanis, M. Nasseri, P. F. Viana, A. Schulze-Bonhage, D. R. Freestone, G. Worrell, M. P. Richardson, and M. J. Cook, "Seizure diaries and forecasting with wearables: Epilepsy monitoring outside the clinic," *Frontiers in Neurology*, vol. Volume 12 - 2021, 2021. [Online]. Available: https://www.frontiersin.org/journals/neurology/articles/10.3389/fneur.2021.690404

[7] K. Rasheed, A. Qayyum, J. Qadir, S. Sivathamboo, P. Kwan, L. Kuhlmann, T. O'Brien, and A. Razi, "Machine learning for predicting epileptic seizures using eeg signals: A review," *IEEE Reviews in Biomedical Engineering*, vol. 14, pp. 139–155, 2021.

[8] "A ResNet-LSTM hybrid model for predicting epileptic seizures using a pretrained model with supervised contrastive learning," vol. 14. [Online]. Available: https://www.nature.com/articles/s41598-023-43328-y

[9] D. Liang, A. Liu, Y. Gao, C. Li, R. Qian, and X. Chen, "Semi-supervised domain-adaptive seizure prediction via feature alignment and consistency regularization," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–12, 2023.

[10] L. Kuhlmann, P. Karoly, D. Freestone, B. Brinkmann, A. Temko, A. Barachant, F. Li, G. Titericz, B. Lang, D. Lavery, K. Roman, D. Broadhead, S. Dobson, G. Jones, Q. Tang, I. Ivanenko, O. Panichev, T. Proix, M. Náhlík, and M. Cook, "Epilepsyecosystem.org: crowd-sourcing reproducible seizure prediction with long-term human intracranial eeg," *Brain : a journal of neurology*, vol. 141, 08 2018.

[11] S. Blanco, A. Garay, and D. Coulombie, "Comparison of frequency bands using spectral entropy for epileptic seizure prediction," *ISRN Neurology*, vol. 2013, p. 287327, 2013. [Online]. Available: https://doi.org/10.1155/2013/287327

[12] J. Batista, M. F. Pinto, M. Tavares, F. Lopes, A. Oliveira, and C. Teixeira, "Eeg epilepsy seizure prediction: the post-processing stage as a chronology," *Scientific Reports*, vol. 14, p. 407, 2024.

[13] S. M. Usman, A. Hassan, F. Riaz, and Q. Chaudry, "Efficient prediction and classification of epileptic seizures using eeg data based on univariate linear features," *Journal of Computers*, vol. 13, no. 6, pp. 616–621, jan 2018.

[14] F. Mormann, T. Kreuz, C. Rieke, R. G. Andrzejak, A. Kraskov, P. David, C. E. Elger, and K. Lehnertz, "On the predictability of epileptic seizures," *Clinical Neurophysiology*, vol. 116, no. 3, pp. 569–587, mar 2005.

[15] A. Subasi and M. Ismail Gursoy, "Eeg signal classification using pca, ica, lda and support vector machines," *Expert Systems with Applications*, vol. 37, no. 12, pp. 8659–8666, 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417410005695

[16] M. K. Siddiqui, R. Morales-Menendez, X. Huang, and N. Hussain, "A review of epileptic seizure detection using machine learning classifiers," *Brain Informatics*, vol. 7, no. 1, pp. 1–18, May 2020.

[17] A. Shoeibi, N. Ghassemi, R. Alizadehsani, M. Rouhani, H. Hosseini-Nejad, A. Khosravi, M. Panahiazar, and S. Nahavandi, "A comprehensive comparison of handcrafted features and convolutional autoencoders for epileptic seizures detection in eeg signals," *Expert Systems with Applications*, vol. 163, p. 113788, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417420306114

[18] I. Kiral-Kornek, S. Roy, E. Nurse, B. Mashford, P. Karoly, T. Carroll, D. Payne, S. Saha, S. Baldassano, T. O'Brien, D. Grayden, M. Cook, D. Freestone, and S. Harrer, "Epileptic seizure prediction using big data and deep learning: Toward a mobile system," *EBioMedicine*, vol. 27, pp. 103–111, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S235239641730470X

[19] N. D. Truong, A. D. Nguyen, L. Kuhlmann, M. R. Bonyadi, J. Yang, and O. Kavehei, "A generalised seizure prediction with convolutional neural networks for intracranial and scalp electroencephalogram data analysis," *CoRR*, vol. abs/1707.01976, 2017. [Online]. Available: http://arxiv.org/abs/1707.01976

[20] ostas . Tsiouris, V. C. Pezoulas, M. Zervakis, S. Konitsiotis, D. D. Koutsouris, and D. I. Fotiadis, "A long short-term memory deep learning network for the prediction of epileptic seizures using eeg signals," *Computers in Biology and Medicine*, vol. 99, pp. 24–37, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S001048251830132X

[21] D. Lee, B. Kim, T. Kim, I. Joe, J. Chong, K. Min, and K. Jung, "A resnet-lstm hybrid model for predicting epileptic seizures using a pretrained model with supervised contrastive learning," *Scientific Reports*, vol. 14, p. 1319, jan 2024.

[22] R. Zhu, W.-X. Pan, J.-X. Liu, and J.-L. Shang, "Epileptic seizure prediction via multidimensional transformer and recurrent neural network fusion," *Journal of Translational Medicine*, vol. 22, no. 1, p. 895, Oct 2024.

[23] N. D. Truong, L. Kuhlmann, M. R. Bonyadi, D. Querlioz, L. Zhou, and O. Kavehei, "Epileptic seizure forecasting with generative adversarial networks," *IEEE Access*, vol. 7, pp. 143 999–144 009, 2019.

[24] A. M. Abdelhameed and M. Bayoumi, "Semi-supervised deep learning system for epileptic seizures onset prediction," in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2018, pp. 1186–1191.

[25] S. Zhao, J. Yang, and M. Sawan, "Energy-efficient neural network for epileptic seizure prediction," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 1, pp. 401–411, Jan 2022.

[26] F. Samie, S. Paul, L. Bauer, and J. Henkel, "Highly efficient and accurate seizure prediction on constrained iot devices," in *2018 Design, Automation Test in Europe Conference Exhibition (DATE)*, 2018, pp. 955–960.

[27] T. M. Ingolfsson, U. Chakraborty, X. Wang, S. Beniczky, P. Ducouret, S. Benatti, P. Ryvlin, A. Cossettini, and L. Benini, "Epidenet: An energy-efficient approach to seizure detection for embedded systems," in *2023 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 2023, pp. 1–5.

[28] A. H. Shoeb and J. V. Guttag, "Chb-mit scalp eeg database," https://physionet.org/content/chbmit/1.0.0/, 2009, physioNet.

[29] M. Bandarabadi, J. Rasekhi, C. A. Teixeira, M. R. Karami, and A. Dourado, "On the proper selection of preictal period for seizure prediction," *Epilepsy & Behavior*, vol. 46, pp. 158–166, May 2015.

[30] W. Abood and S. Bandyopadhyay, "Postictal seizure state," in *StatPearls [Internet]*. Treasure Island (FL): StatPearls Publishing, 2021, updated 2021 Jul 16. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK526004/

[31] D. Singh and B. Singh, "Investigating the impact of data normalization on classification performance," *Applied Soft Computing*, vol. 97, p. 105524, May 2020.

[32] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 2980–2988.

[33] S. Wang, W. Liu, J. Wu, L. Cao, Q. Meng, and P. J. Kennedy, "Training deep neural networks on imbalanced data sets," in *2016 International Joint Conference on Neural Networks (IJCNN)*, 2016, pp. 4368–4374.

[34] G. Lemaître, F. Nogueira, and C. K. Aridas, "Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning," *Journal of Machine Learning Research*, vol. 18, no. 17, p. 1–5, 2017. [Online]. Available: http://jmlr.org/papers/v18/16-365.html

[35] Y. Xu, J. Yang, S. Zhao, H. Wu, and M. Sawan, "An end-to-end deep learning approach for epileptic seizure prediction," in *2020 2nd IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS)*, 2020, pp. 266–270.

[36] C. Li, X. Huang, R. Song, R. Qian, X. Liu, and X. Chen, "Eeg-based seizure prediction via transformer guided cnn," *Measurement*, vol. 203, p. 111948, 2022.

[37] G. Segal, N. Keidar, M. Herskovitz, and Y. Yaniv, "Personalized preictal eeg pattern characterization: do timing and localization matter?" *Frontiers in Neuroscience*, vol. 19, p. 1526963, 2025.

[38] D. Sopić, A. Aminifar, and D. Atienza, "e-glass: A wearable system for real-time detection of epileptic seizures," in *Proceedings of the 2018 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2018, pp. 1–5.

[39] Z. Gao, X. Cui, W. Wan, Z. Qin, and Z. Gu, "Signal quality investigation of a new wearable frontal lobe eeg device," *Sensors*, vol. 22, no. 5, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/5/1898

[40] M. Jain and C. M. Markan, "Calibration of off-the-shelf low-cost wearable eeg headset for application in field studies," *arXiv preprint arXiv:2209.12633*, 2022. [Online]. Available: https://arxiv.org/abs/2209.12633

[41] W. Sensing, "Dsi-24 - wearable sensing | dry eeg," https://wearablesensing.com/dsi-24/, 2025, accessed: 2025-06-03.

[42] H. Gelbard-Sagiv, S. Pardo, N. Getter, M. Guendelman, F. Benninger, D. Kraus, O. Shriki, and S. Ben-Sasson, "Optimizing electrode configurations for wearable

eeg seizure detection using machine learning," *Sensors*, vol. 23, no. 13, p. 5805, 2023. [Online]. Available: https://www.mdpi.com/1424-8220/23/13/5805

[43] P. Afonso, T. Sagué, D. F. Pedraza, J. L. Vázquez, J. R. López, and V. Fernández, "Wearable epileptic seizure prediction system based on ear eeg, ecg, and ppg," *Sensors*, vol. 22, no. 23, p. 9372, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/23/9372

[44] M. J. Cook, T. J. O'Brien, S. F. Berkovic, M. Murphy, A. Morokoff, G. Fabinyi, W. D'Souza, R. Yerra, J. Archer, R. Litewka, and et al., "Seizure forecasting with wearable devices and subcutaneous eeg," *Neurology*, vol. 99, no. 11, pp. e1068–e1076, 2022. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/40411751/

[45] E. Grigorash, X. Li, H. A. Faucett, L. A. Kaper, C. A. Flask, M. A. Luciano, M. Naulaerts, S. Rampp, R. Fiebrink, G. A. Worrell, and et al., "Non-invasive seizure forecasting using ultra-long-term wearable eeg," *epilepsia*, vol. 65, no. 2, pp. 387–398, 2024. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/40411751/

[46] C. Meisel, R. El Atrache, M. Jackson, S. Schubach, C. Ufongene, and T. Loddenkemper, "Deep learning from wristband sensor data: Towards wearable, non-invasive seizure forecasting," *arXiv preprint arXiv:1906.00511*, 2019. [Online]. Available: https://arxiv.org/abs/1906.00511

[47] F. Tian, J. Yang, S. Zhao, and M. Sawan, "Neuromorphic spiking-cnn for low-power seizure prediction," *arXiv preprint arXiv:2102.12773*, 2021. [Online]. Available: https://arxiv.org/abs/2102.12773

[48] S. Gupta, V. Ranga, and P. Agrawal, "Epilnet: Iot-based 1d cnn for on-site seizure prediction," *arXiv preprint arXiv:2111.03265*, 2021. [Online]. Available: https://arxiv.org/abs/2111.03265

[49] H. Gelbard-Sagiv, S. Pardo, N. Getter, M. Guendelman, F. Benninger, D. Kraus, O. Shriki, and S. Ben-Sasson, "Optimizing electrode configurations for wearable eeg seizure detection using machine learning," *Sensors*, vol. 23, no. 13, p. 5805, 2023. [Online]. Available: https://www.mdpi.com/1424-8220/23/13/5805

[50] T. M. Ingolfsson, X. Wang, U. Chakraborty, S. Benatti, A. Bernini, P. Ducouret, P. Ryvlin, S. Beniczky, L. Benini, and A. Cossettini, "Brainfusenet: Enhancing wearable seizure detection through eeg-ppg-accelerometer sensor fusion and efficient edge deployment," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 18, no. 4, pp. 720–733, 2024.

[51] S. Frey, M. Guermandi, S. Benatti, V. Kartsch, A. Cossettini, and L. Benini, "Biogap: a 10-core fp-capable ultra-low power iot processor, with medical-grade afe and ble connectivity for wearable biosignal processing," 2023. [Online]. Available: https://arxiv.org/abs/2307.01619

[52] K. Lee, K.-M. Choi, S. Park, S.-H. Lee, and C.-H. Im, "Selection of the optimal channel configuration for implementing wearable eeg devices for the diagnosis of mild cognitive impairment," *Alzheimer's Research & Therapy*, vol. 14, no. 1, p. 170, 2022. [Online]. Available: https://alzres.biomedcentral.com/articles/10.1186/s13195-022-01115-3