Efficient Edge-AI: Towards the Future of Implantable and Smart Medical Devices

Luis Fernando Herbozo Contreras

Principal Supervisor

Prof. Omid Kavehei

Auxiliary Supervisors

A/Prof. Armin Nikpour, Prof. Alistair McEwan

A thesis submitted in fulfillment
of the requirements of the degree of
Doctor of Philosophy

THE UNIVERSITY OF
SYDNEY

Faculty of Engineering
School of Biomedical Engineering
The University of Sydney

May 2026

Statement of Originality

I hereby declare that this submission is my original work and, to the best of my knowledge and belief, contains no material previously published or authored by another individual, nor any content that has been substantially used to obtain another degree or diploma from this university or any other institution of higher learning, except where duly acknowledged in the text.

This research involving human data was conducted in accordance with the relevant ethical guidelines and regulations. Ethics approval for the use of the Royal Prince Alfred Hospital (RPAH) dataset was granted by the NSW Local Health District (LHD) Human Research Ethics Committee (Reference: X19-0323-2019/STE16040). The approved project covers collaboration between The University of Sydney and the Comprehensive Epilepsy Services, Department of Neurology, Royal Prince Alfred Hospital, Australia.

Luis Fernando Herbozo Contreras

28 May 2026

Data Availability

The datasets used in this thesis are described as follows:

- The Temple University Hospital EEG Corpus (TUH) EEG Corpus is publicly available at: https://isip.piconepress.com/projects/tuh_eeg/.
- The EPILEPSIAE dataset for Electroencephalogram (EEG) and Intracranial EEG (iEEG) is available upon purchase and registration via: https://www.epilepsy-database.eu/.
- The Freiburg dataset (FB) dataset used to be available openly at: https://epilepsy.uni-freiburg.de/freiburg-seizure-prediction-project/, but it seems it no longer accepts registration.
- The CHB-MIT Scalp EEG Database is publicly available via PhysioNet: https://physionet.org/content/chbmit/1.0.0/.
- The Royal Prince Alfred Hospital (RPAH) dataset was used with exclusive ethical approval from the Local Health District's Human Ethics Review Board for using and processing de-identified human data from RPAH. This data is not publicly available.

Abstract

Epilepsy remains a major global neurological disorder, affecting over 1% of the population worldwide and imposing substantial clinical, social, and economic burdens. Pharmacological therapy is the first line of treatment; however, approximately 30–40% of patients remain drug-resistant, leaving neuromodulation as one of the few viable therapeutic options. Despite advances in diagnostic and therapeutic technologies, existing neuromodulation systems are predominantly open-loop or rely on cloud-based Artificial Intelligence (AI), which is constrained by latency, power consumption, data privacy, and long-term scalability. These limitations hinder the deployment of fully autonomous, personalised, and implantable closed-loop neurostimulation systems.

This thesis explores neuromorphic computing as a foundational paradigm for next-generation closed-loop neuromodulation, with a focus on seizure detection and prediction in epilepsy. The thesis begins with a perspective article that introduces the concept of neuromorphic neuromodulation and outlines how neuromorphic principles can enable self-responsive, on-device intelligence for personalised therapies. By developing algorithms and testing learning mechanisms inspired by biological neural dynamics, this work enables low-latency and memory-efficient intelligence to operate directly at the edge. Building on this foundation, a seizure detection system is developed using bio-inspired neural network models incorporating liquid-time constant neurons and dendritic spiking mechanisms with heterogeneous temporal dynamics. These models allow efficient temporal representation of neural signals without reliance on computationally expensive feature extraction and demonstrate robust out-of-sample generalisation across patients and recording conditions on large-scale clinical EEG datasets. In addition, the thesis focuses on edge learning, proposing a neuromorphic learning rule that enables on-device adaptation for patient-specific treatments under strict power and memory constraints and can be adopted by the developed models. Finally, learnable activation functions are further explored within artificial neural networks, improving training efficiency and interpretability, highlighting a pathway for integration with neuromorphic technology.

Together, these contributions advance neuromorphic neurotechnology toward autonomous, personalised, and continuously learning closed-loop systems, with implications extending beyond epilepsy to a broad range of neurological disorders.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisors, Prof. Omid Kavehei, A/Prof. Armin Nikpour and Prof. Alistair McEwan, for their invaluable guidance, encouragement, and continuous support throughout my Ph.D. journey. Their expertise, patience, and insightful feedback have been instrumental in shaping both this research and my development as a researcher.

I also gratefully acknowledge the support of Mr. Satendra Pratap, IT technologist at the Comprehensive Epilepsy Service, Royal Prince Alfred Hospital (RPAH), Sydney, Australia, for his technical assistance and collaboration. I further thank Dr. Jason K. Eshraghian from the Department of Electrical and Computer Engineering at the University of California, Santa Cruz, and Dr. Nhan Duy Truong for their continued mentorship and constructive discussions, which greatly enriched this work.

I acknowledge the partial financial support provided by the Faculty of Engineering Research Scholarship from The University of Sydney. I also gratefully acknowledge the financial support from the School Research Funding Scheme (SRFS) at The University of Sydney.

I would like to thank my friends, especially Zhaojing Huang and Leping Yu, with whom I have co-authored extensively and had many meaningful technical discussions. I am also grateful to my colleagues for their encouragement, collaboration, and support throughout my Ph.D. studies.

Lastly, and most importantly, I am deeply grateful to my fiancée, Rachel Shao, my parents (Luis Fernando Herbozo Alvarado and Gloria Laura Contreras Tupac-Yupanqui), and my family for their unwavering love, patience, and support throughout my Ph.D. journey.

Authorship Attribution Statement

This thesis contains material that has been published or submitted for publication. Relevant chapters include individual authorship attribution statements identifying the publication details and my contribution to the work.

In addition to the authorship attribution statements above, in cases where I am not the corresponding author of a published item, permission to include the published material has been granted by the corresponding author.

Luis Fernando Herbozo Contreras Date: 28 May 2026

As supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

Prof. Omid Kavehei Date: 28 May 2026

Generative AI Statement

During the preparation of this thesis, I declare the use of OpenAI's ChatGPT-5.2 for the purposes of text enhancement and language refinement of the author's original work. The use of this generative AI tool included paraphrasing, improving sentence structure, refining academic expression, and correcting grammar, spelling, and clarity.

The author confirms that where text was modified by generative AI, the content was reviewed for possible errors, inaccuracies, and bias. The author takes full responsibility for the submitted thesis, confirms that the intellectual content and research are their own, and has used generative AI within the parameters of use set by the University of Sydney's guidance for research students

Contents

Statement of Originality i

Data Availability ii

Abstract iii

Acknowledgements iv

Authorship Attribution Statement v

Generative AI Statement vi

Contents vii

List of Figures xiii

List of Tables xvi

Nomenclature xviii

1 Introduction 1

1.1 Research Motivation 2

1.2 Problem Statements 3

1.3 Research questions 3

1.4 Thesis Organisation 4

1.5 Terminology 7

1.6 List of publications 7
1.6.1 Publications and Source Codes 7
1.6.2 Other publications and Source Codes 8
1.6.3 Conferences 9

2 Neuromorphic Neuromodulation: A Perspective 11
2.1 Introduction 13
2.1.1 Driving the next generation of on-device AI-revolution in electroceuticals through neuromorphic neuromodulation conceptualisation 20
2.1.2 Application for neuromorphic neuromodulation 24
2.1.3 Evaluating Standards in a Self-Sufficient Responsive Neuromorphic System 25
2.1.4 Case Studies for AURA: Foundations of Efficient, Low-Power, and Biologically Inspired Models for Seizure Detection 29
2.1.5 Challenges and Opportunities of Neuromorphic AI 34
2.2 Conclusions 38

3 Biologically Plausible Liquid Spiking Neurons for EEG-Based Seizure Detection 40
3.1 Introduction 42
3.1.1 Background 44
3.1.2 Novelty and significance 46
3.2 Methods 46
3.2.1 Datasets 46
3.2.2 Pre-Processing 49
3.2.3 SNN structure 50
3.2.4 Implementation details 53
3.2.5 Performance metrics 53
3.2.6 Power Consumption 53
3.3 Results 54

3.3.1 Training and Validation 54
3.3.2 Spiking vs Non-Spiking Generalisation 55
3.3.3 Less is Effective: Memory Reduction Leads To Efficiency Even In Generalisation. 56
3.3.4 Power Consumption of LTC-SNN-FPTT model 58
3.3.5 AI-SNN Feature Extraction 59
3.4 Discussion 60
3.5 Conclusion 62

4 Time Domain Multiscale Processing with Dendritic Spiking Networks 63

4.1 Introduction 65
4.1.1 Bio-Inspired Neural Networks 67
4.1.2 The proposed system 67
4.1.3 Novelty and significance 68
4.2 Methods 68
4.2.1 Datasets 68
4.2.2 Pre-Processing 70
4.2.3 SNN structure 71
4.2.4 Implementation details 75
4.2.5 Performance metrics 76
4.3 Experimental Results 76
4.3.1 Scalp-EEG. 76
4.3.2 Intracranial-EEG 79
4.4 Timescale analysis of the dendrite branches 80
4.4.1 Learned Timing Factors of Dendrite Branches and Membrane Potential 83
4.5 Estimation of Power Consumption 83
4.6 Discussion 85
4.7 Conclusion 86

5 Unified Dendro-Liquid Framework for Spatio-Temporal Seizure Detection. 88

5.1 Introduction . . . 90

5.1.1 Advancing Towards Bio-Realistic Models in Seizure Detection 91

5.1.2 Novelty and Significance . . . 93

5.2 Datasets . . . 94

5.2.1 Scalp-EEG . . . 95

5.2.2 Intracranial-EEG . . . 95

5.3 Methods . . . 96

5.3.1 Pre-processing . . . 96

5.3.2 Spiking Neural Network . . . 96

5.3.3 Backpropagation through time . . . 98

5.3.4 Performance Metrics . . . 101

5.3.5 Implementation Details . . . 101

5.4 Results . . . 102

5.4.1 Neural Network Analysis . . . 102

5.4.2 In-Sample Generalisation . . . 103

5.4.3 Edge-Inference Raspberry Pi . . . 106

5.4.4 Robustness test . . . 107

5.4.5 Ablation experiments . . . 108

5.4.6 Evaluating SNN Performance Across Different Training Steps 109

5.4.7 Energy Consumption Estimation . . . 110

5.5 Limitations . . . 113

5.5.1 Learning Rule . . . 113

5.5.2 Extending to Out-of-sample generalisation . . . 113

5.5.3 Overhead of learnable parameters . . . 113

5.6 Conclusion . . . 113

6 Adaptive Neuromorphic Continual Edge Training for Seizure Detection and Prediction 115

6.1 Introduction . . . 117

6.2 Methods . . . 122

6.2.1 Datasets . . . 122

6.2.2 Pre-Processing . . . 123

6.2.3 Detection vs Prediction . . . 124

6.2.4 ANN Structure . . . 124

6.2.5 Quantisation Aware Training . . . 125

6.2.6 Spiking Neural Network Structure . . . 125

6.2.7 Edge Learning . . . 125

6.2.8 Few Shot Learning . . . 126

6.2.9 Performance Metrics . . . 127

6.2.10 Implementation Details . . . 127

6.3 Results . . . 128

6.3.1 Feature Extraction . . . 128

6.3.2 Pre-Post Quantisation . . . 130

6.3.3 Quantisation Learning Rate effects on Neuron Activations . . . 131

6.3.4 Detection Out-Of-Sample Framework . . . 131

6.3.5 Prediction Framework . . . 132

6.3.6 Power, Memory and Latency Analysis . . . 135

6.3.7 Robustness . . . 137

6.3.8 Complexity: Training and Inference . . . 137

6.4 Discussion . . . 140

7 Learnable functions led to Interpretable Systems. 144

7.1 Introduction . . . 146

7.1.1 Background . . . 148

7.1.2 Limitations of the Previous Studies . . . 150

7.1.3 Novelty and Significance . . . 151
7.2 Methods . . . 152
7.2.1 Datasets . . . 152
7.2.2 The KAN-EEG Structure . . . 153
7.2.3 Pre-Processing . . . 155
7.3 Experiments and Results . . . 156
7.3.1 Training and Validation in Sample . . . 156
7.3.2 Robustness . . . 157
7.3.3 Compact Networks Leads to Out-of-sample Continental generalisation . . . 158
7.3.4 Computation Time in Training . . . 160
7.4 Discussion . . . 160
7.4.1 Scope and Model Limitations . . . 162
7.5 Conclusion . . . 163

8 Concluding Remarks **164**
8.1 Thesis Contributions . . . 164
8.2 Future Research Directions . . . 165

List of References **168**

List of Figures

1.1 Thesis Structure 6

2.1 Brief history of neurostimulation devices and possible future directions. 16
2.2 Conventional (Von Neumann) vs Neuromorphic Architectures 22
2.3 Neuromodulation approaches. 26
2.4 Bio-inspired learning rules for neuromorphic neuromodulation 32
2.5 Energy–memory trade-offs in Von Neumann and neuromorphic architectures 34
2.6 Power breakdown of a Neuromorphic Device 35

3.1 Contrast between BNN, DNN, SNN. 45
3.2 Summary of the RPAH Dataset 47
3.3 Seizure detection system towards a low-power and accurate neuromorphic neuromodulation system 48
3.4 Summary of the TUH dataset for both training and validation 54
3.5 State-of-the-art comparison on TUH dataset 55
3.6 Generalisation performance of large vs small LTC-SNN-FPTT 57
3.7 Performance on the RPAH dataset by memory reduction 58
3.8 EPILEPSIAE Scalp-EEG results with the small LTC-SNN 59
3.9 EPILEPSIAE Scalp-EEG results across brain regions in patients with varying seizure foci 60
3.10 AI-Interpretability of our model 61

4.1 Biological inspired neuron model 69

4.2 Summary of datasets . . . 71
4.3 Neuromorphic detecting system using dLIF . . . 74
4.4 Timescale analysis in the dLIF neural network . . . 82
4.5 Histogram of the learned timing factors of dendrites branches and soma of our neural network . . . 84
4.6 Pareto plot of energy consumption versus memory usage and AUROC 85

5.1 Overview of four spiking neuron models . . . 91
5.2 Datasets used in this work . . . 92
5.3 Our current framework and vision of our model . . . 94
5.4 Analysis of learned timing factors and membrane potentials . . . 104
5.5 Training and validation results across the TUH epilepsy dataset. . . . 106
5.6 Black-out channels results on the TUH and CHB-MIT dataset. . . . 109
5.7 Ablation experiment results in the three datasets used . . . 110
5.8 Timesteps analysis in the TUH, CHB-MIT and EPILEPSIAE used for this work . . . 111

6.1 Edge-based personalised seizure detection and prediction frameworks 118
6.2 Results for the initial training and prediction stage on GPU . . . 129
6.3 Figure of merit and latency vs number of shots for the three datasets 136
6.4 Pareto plot of energy consumption versus memory usage and latency for three datasets . . . 136
6.5 Robustness test . . . 138
6.6 Log–log scatter plot of training vs inference computational complexity 139
6.7 Average AUROC, FPR, and Sensitivity across all patients and datasets (EPILEPSIAE, CHB-MIT, Freiburg). . . . 141
6.8 Performance metrics as a function of shots for patient in EPILEPSIAE dataset. . . . 142
6.9 Performance metrics in the prediction framework as a function of shots for each patient. . . . 143

7.1 Representation of the past and future in neuro-AI . . . 147

7.2 Summary of the datasets being used 154
7.3 KAN-EEG seizure system 156
7.4 In-Sample results on the TUH dataset 157

List of Tables

2.1 A benchmark of neuromorphic chips. This list is not extensive but considers the breakthroughs. 23

3.1 Different types of neural networks applied to different bio-signals with applications such as detection and prediction. 52

3.2 Details of RPAH dataset for 192 patients 56

3.3 Generalisation of our LTC-SNN-FPTT against non-SNN (ConvLSTM) in the RPAH dataset 56

3.4 Power consumption of the LTC-SNN-FPTT model 59

4.1 Summary of hyperparameters settings for each dataset 75

4.2 AUROC across input domains on the TUH dataset 77

4.3 TUH Results for down-sampling test 78

4.4 CHB-MIT results 79

4.5 FB results 80

4.6 iEEG-EPILEPSIAE results 81

5.1 Summary of hyperparameter settings for each dataset. 102

5.2 State-of-the-art performance comparison on the TUH dataset. 105

5.3 CHB-MIT results: AUROC comparison of models using time-frequency and time-domain inputs. 105

5.4 EPILEPSIAE-iEEG Results 107

5.5 Theoretical equations of energy computation per layer. 111

5.6 Spiking rate of the layers in the liquid-dendrite model 112

6.1 Summary of EEG datasets used in this work. 130
6.2 Comparison of CNN and our work for seizure detection on the EPILEPSIAE dataset. 133
6.3 Comparison of CNN, FGL, and Our Work. 134
6.4 Comparison of CNN and our work for seizure prediction on the FB Dataset. 135
6.5 Training complexity per model 139

7.1 Different types of neural networks applied to EEG signals for epileptic seizure detection. 150
7.2 Model assessment robustness 158
7.3 Out-of-sample generalisation results Area Under the Receiver Operating Characteristic curve (AUROC) on the RPAH dataset. 161
7.4 Out-of-sample generalisation results AUROC on the EPILEPSIAE dataset. 161

Nomenclature

List of Acronyms

Clinical Neurotechnology, Signals, Datasets, and Governance

AES	American Epilepsy Society Seizure Prediction Challenge
AFE	Analog Front End
BAC	Body Area Communication
CHB-MIT	Children's Hospital Boston–MIT EEG dataset
DBS	Deep Brain Stimulation
ECAP	Evoked Compound Action Potential
ECG	Electrocardiogram
ECoG	Electrocorticography
EEG	Electroencephalogram
EPAD	Epilepsy Personal Assistant Device
FB	Freiburg dataset
FDA	Food and Drug Administration
FoM	Figure of Merit
fps	Frames per second
HFO	High-Frequency Oscillation
iEEG	Intracranial EEG
IMD	Implantable Medical Device
LFP	Local Field Potential
LPCH	Lucile Packard Children's Hospital
MNIST	Modified National Institute of Standards and Technology dataset
RNS	Responsive Neurostimulation System
RPAH	Royal Prince Alfred Hospital
SCS	Spinal Cord Stimulation
SOP	Seizure Occurence Period
SPH	Seizure Prediction Horizon
TNMG	Telehealth Network of Minas Gerais
TUH	Temple University Hospital EEG Corpus
VNS	Vagus Nerve Stimulation

Neuromorphic Computing, Models, Training, and Evaluation

AC	Accumulate
ACC	Accuracy
AE	Autoencoder
AI	Artificial Intelligence
ANN	Artificial Neural Network
ASIC	Application-Specific Integrated Circuit
ASRNN	Adaptive Spiking Recurrent Neural Network
AUPRC	Area Under the Precision–Recall Curve
AUROC	Area Under the Receiver Operating Characteristic curve
AURA	Adaptive, Unlabeled, and Real-time Approximate-Learning
BCE	Binary Cross-Entropy
BNN	Binary Neural Network
BP	Backpropagation
BPTT	Backpropagation Through Time
BSS	Blind Source Separation
CE	Cross-Entropy
CMOS	Complementary Metal–Oxide–Semiconductor
CNN	Convolutional Neural Network
ConvLSTM	Convolutional Long Short-Term Memory
CPU	Central Processing Unit
CSNN	Convolutional Spiking Neural Network
DC	Direct current
DBN	Deep Belief Network
dLIF	Dendritic Leaky Integrate-and-Fire
DNN	Deep Neural Network
DWT	Discrete Wavelet Transform
FFT	Fast Fourier Transform
FPR	False Positive Rate
FPTT	Forward Propagation Through Time
FPGA	Field-Programmable Gate Array
FGL	Future Guided Learning
GPU	Graphics Processing Unit
GRU	Gated Recurrent Unit
ICA	Independent Component Analysis
IF	Integrate-and-Fire
KAN	Kolmogorov-Arnold Network
LIF	Leaky Integrate-and-Fire
LR	Learning Rate
LSTM	Long Short-Term Memory
LTC	Liquid Time-Constant
LTC-SNN	Liquid Time-Constant Spiking Neural Network

MAC	Multiply-Accumulate
MNE	MNE-Python library
MLP	Multilayer Perceptron
MP	Max Pooling
MSE	Mean Squared Error
NCPS	Neural Circuit Policies
NWN	Nanowires Networks
QAT	Quantization-Aware Training
RAM	Random Access Memory
ReLU	Rectified Linear Unit
RNN	Recurrent Neural Network
R-STDP	Reward-Modulated Spike-Timing-Dependent Plasticity
SConvLSTM	Spiking Convolutional Long Short-Term Memory
SNN	Spiking Neural Network
SoC	System-on-Chip
STDP	Spike-Timing-Dependent Plasticity
STFT	Short-Time Fourier Transform
STLmax	Short-term maximum Lyapunov exponents
SVM	Support Vector Machine
TE	Teager Energy
TSSP	Time-Series Signal Processing
VLSI	Very Large Scale Integration

Chapter 1

Introduction

Closed-loop neurostimulation promises transformative therapies for neurological disorders by detecting pathological brain activity and delivering targeted stimulation in real time. However, despite major advances in sensing and actuation technologies, the intelligence that drives these systems remains a critical bottleneck. Most contemporary approaches rely on large, cloud-dependent artificial intelligence models and offline updates, imposing fundamental limitations on latency, power consumption, data privacy, and long-term adaptability-constraints incompatible with the realities of implantable and wearable neurotechnology.

This thesis advances the concept of neuromorphic neuromodulation, in which neuromorphic artificial intelligence operates as the control and decision-making layer for responsive neurotherapies. Rather than developing neuromodulation techniques themselves, this work focuses on the intelligence required to enable fully embedded closed-loop systems capable of sensing, inferring, adapting, and making therapeutic decisions directly on-device under strict constraints in power, memory, latency, and autonomy.

Epilepsy is used throughout this thesis as a clinically relevant and technically demanding case study. Seizure dynamics are highly non-stationary, vary substantially across and within patients, and occur under extreme class imbalance. At the same time, responsive neurostimulation for epilepsy represents one of the most mature closed-loop

clinical paradigms. This combination makes epilepsy an ideal testbed for investigating how neuromorphic intelligence can support real world responsive therapies.

1.1 Research Motivation

The motivation for this thesis arises from a fundamental mismatch between how modern artificial intelligence is designed and how clinical neurotechnology must operate.

On one hand, state-of-the-art AI systems achieve impressive performance by scaling model size, data volume, and computational resources. On the other hand, implantable and wearable neurostimulation devices demand the opposite: ultra-low power operation, strict memory limits, minimal latency, strong privacy, and long-term autonomous function. Bridging this gap requires a shift away from accuracy-at-all-costs paradigms toward efficiency-first intelligence to intelligence that is explicitly designed for deployment, not just for benchmark performance.

At the same time, the clinical need for adaptive, personalised neurotechnology is growing. Neurological disorders such as epilepsy exhibit long-term variability driven by disease progression, medication changes, and lifestyle factors. Static models trained offline cannot accommodate this reality. What is needed are systems that can learn continuously, adapt safely, and operate independently of cloud infrastructure, particularly in sensitive medical contexts where latency, reliability, and privacy are paramount.

Neuromorphic computing offers a compelling foundation for this shift. By exploiting event-driven processing, temporal dynamics, and biologically inspired learning mechanisms, neuromorphic systems promise orders-of-magnitude reductions in power and memory consumption compared with conventional digital AI. Yet, despite this promise, most neuromorphic research has focused on perception tasks such as vision, leaving closed-loop clinical intelligence largely unexplored, with performance that would translate into an open-loop system with a high false-positive rate and high variability.

1.2 Problem Statements

Despite recent progress in AI-assisted seizure detection, several key problems remain unresolved:

- Existing seizure detection systems often lack generalisation across patients and datasets, relying heavily on patient-specific optimisation rather than automation.
- Many approaches rely on handcrafted or computationally intensive feature extraction, which increases system complexity and limits robustness.
- Conventional deep learning architectures are poorly aligned with the constraints of edge and neuromorphic hardware, particularly in terms of memory usage and energy efficiency.
- Interpretability of AI-assisted seizure detection models remains limited, which hinders clinical trust and adoption.
- Continuous and adaptive learning mechanisms suitable for real-time, closed-loop seizure detection and prediction are still underexplored.

1.3 Research questions

1. How can neural architectures with intrinsic liquid time-constant dynamics be designed and extended into neuromorphic spiking neural networks to enable accurate, generalisable, and low-power biomedical signal analysis—particularly EEG-based seizure detection under strict constraints on memory, computation, and energy consumption?
2. Can under-explored compartmental biological mechanisms, such as dendritic processing and heterogeneous temporal dynamics, explicitly model the temporal complexity of EEG and iEEG signals for seizure evolution without relying on complex signal preprocessing or feature extraction?

3. Does the integration of liquid time-constant dynamics and dendritic computation enable ultra-compact spiking neural architectures that achieve superior energy efficiency, under kilobyte-scale memory budgets, and robust seizure detection from raw signals?

4. Can neuromorphic edge-learning systems be trained directly on low-power hardware to continuously adapt to patient-specific seizure signatures through online learning, enabling personalised, closed-loop seizure detection and prediction in an autonomous manner?

5. Can learnable activation functions enable compact, interpretable, and high-performance neural models for EEG-based seizure detection, while reducing memory usage and training complexity as a step toward future neuromorphic translation?

1.4 Thesis Organisation

This thesis is structured as a thesis by publication and is organised to reflect a progressive research trajectory, moving from conceptual foundations to algorithmic development, neuromorphic implementation, and closed-loop clinical translation. Each chapter addresses one or more of the research questions outlined in Section 1.3. Fig. 1.1 summarises the relationships among the chapters and the overall structure of the thesis.

- Chapter 2 presents a conceptual and technical perspective on the next generation of closed-loop neuromodulation systems enabled by neuromorphic computing. While the proposed framework is broadly applicable to neurological disorders, the chapter focuses specifically on epilepsy as the primary clinical use case. By leveraging neuromorphic principles at both the algorithmic and hardware levels, this chapter explores the feasibility of fully embedded, self-responsive therapeutic systems that operate without continuous data transmission. This

vision provides the foundational motivation for the subsequent chapters and frames the key challenges addressed throughout the thesis.

- Chapter 3 addresses the Research Question 1 by introducing liquid time-constant spiking neural networks for EEG-based seizure detection. This chapter investigates how intrinsic learnable time-membrane constants can be embedded directly into neural architectures to improve robustness, generalisation, and efficiency under strict constraints on memory, computation, and energy consumption.

- Chapter 4 extends this neuromorphic perspective by introducing dendrite-inspired computation in spiking neural networks. This chapter explores how compartmental biological mechanisms and heterogeneous temporal processing can explicitly model the temporal complexity of EEG and iEEG signals without reliance on complex signal preprocessing or feature extraction. These contributions primarily address Research Question 2.

- Chapter 5 integrates the contributions of the previous chapters into a unified framework. This chapter presents an ultra-compact neuromorphic architecture that combines liquid time-constant dynamics with dendritic computation to achieve robust seizure detection under kilobyte-scale memory budgets and severe power constraints. By demonstrating superior energy efficiency and compactness, this chapter directly addresses Research Question 3.

- Chapter 6 focuses on continual online learning in resource-constrained environments by addressing Research Question 4. This chapter proposes neuromorphic edge-learning mechanisms that enable seizure detection and forecasting systems to adapt to patient-specific dynamics using few-shot transfer learning.

- Chapter 7 investigates learnable activation functions as a pathway toward compact, interpretable, and high-performance neural models for EEG-based seizure detection. Rather than claiming neuromorphic deployment, this chapter evaluates the feasibility of function-based representations to improve performance

while reducing memory usage and training complexity. These findings address Research Question 5 and establish efficient activation mechanisms as enabling components for future neuromorphic translation.

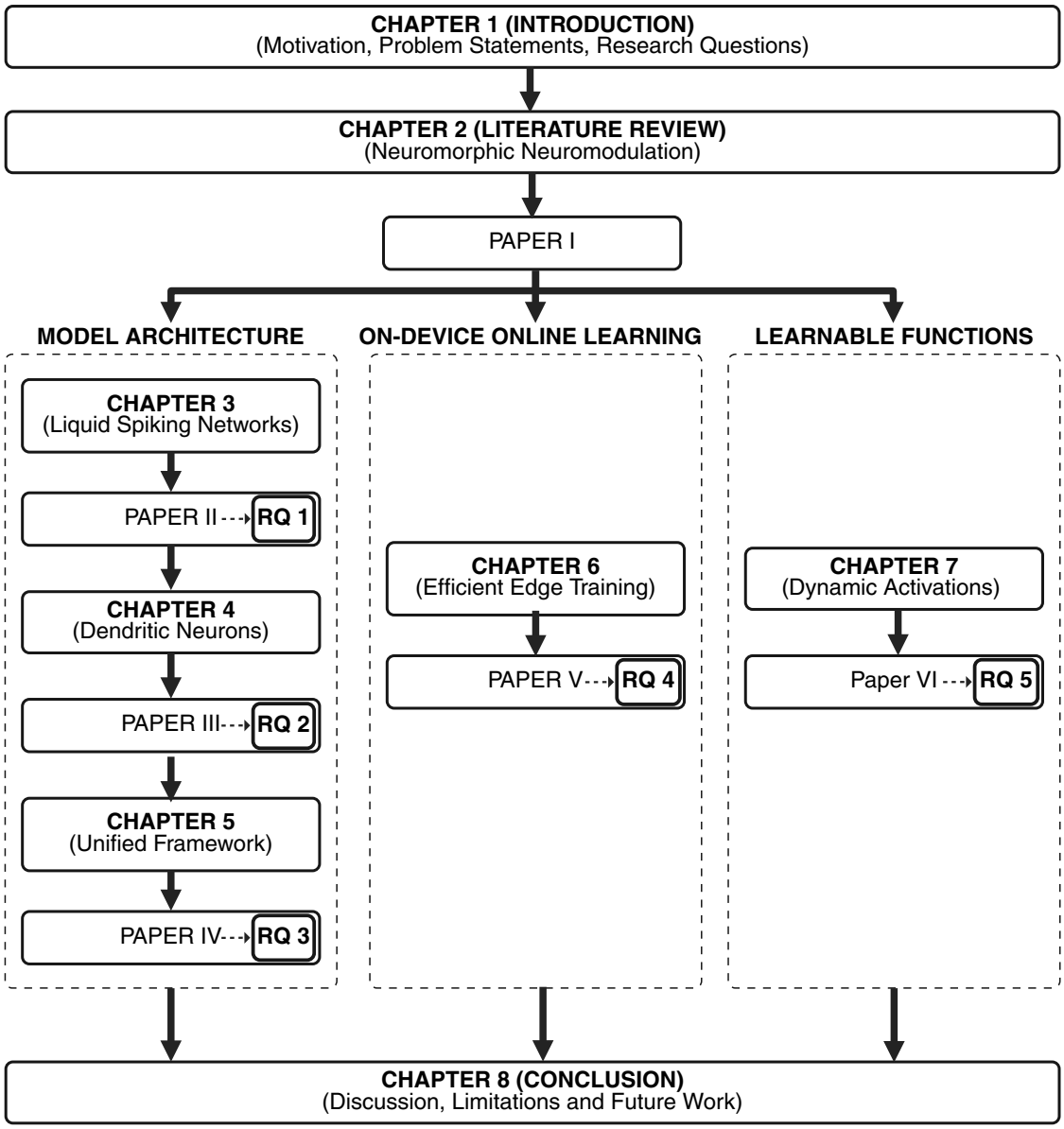

Figure 1.1 – Schematic representation of the thesis structure. Each Chapter, with it's own corresponding paper answer a Research Question (RQ).

1.5 Terminology

Although acronyms are defined throughout the thesis, this subsection clarifies how several related concepts are used in this work. Throughout this thesis, Artificial Neural Network (ANN) refers broadly to conventional architectures implemented using dense numerical computation, while Convolutional Neural Network (CNN) denotes convolutional architectures within this family. Spiking Neural Network (SNN) refers to models that encode and process information using spike events and neuronal dynamics. Neuromorphic-compatible refers to models designed to align with neuromorphic constraints, such as sparse/event-driven computation, compact memory use, low-precision operations, or local learning, but does not necessarily imply deployment on neuromorphic hardware. Neuromorphic deployment is used only when a model is mapped to a neuromorphic hardware platform. On-chip learning is reserved for cases where model adaptation or training is performed directly on the target hardware platform.

1.6 List of publications

1.6.1 Publications and Source Codes

The following publications are included as the main chapters for this thesis.

- **Herbozo Contreras, L. F.**, Duy Truong, N., Eshraghian, J. K., Xu, Z., Huang, Z., Bersani-Vincenzo, T., Aguilar, I., Leung W. H., Nikpour, A. and Kavehei, O. (2024).“Neuromorphic neuromodulation: Towards the next generation of closed-loop neurostimulation.” *PNAS Nexus*, page 488. DOI:10.1093/pnasnexus/pgae488

- **Herbozo Contreras, L. F.**, Huang, Z., Yu, L., Nikpour, A. and Kavehei, O. (2024). “Biological plausible algorithm for seizure detection: Toward AI-enabled electroceuticals at the edge.” *APL Machine Learning*, 2(2). DOI:10.1063/5.0192875. https://github.com/NeuroSyd/Seizure-LTC-FPTT.

- **Herbozo Contreras, L. F.**, Yu, L., Huang, Z., Zhang, Z., Nikpour, A. and Kavehei, O. (2025). "Tiny dLIF: A Dendritic Spiking Neural Network Enabling a Time-Domain Energy-Efficient Seizure Detection System." *IOP Neuromorphic Computing and Engineering*, 5(1), 014015.DOI:10.1088/2634-4386/adc0b9. https://github.com/NeuroSyd/TinydLIF

- **Herbozo Contreras, L. F.***, Cui, J.*, Yu, L., Huang, Z., Nikpour, A. and Kavehei, O. (2025). "KAN-EEG: Towards Replacing Backbone-MLP for an Effective Seizure Detection System." *Royal Society Open Science*, 12(3), 240999. DOI:10.1098/rsos.240999 https://github.com/NeuroSyd/KAN-EEG.

- **Herbozo Contreras, L. F.**, Yu, L., Huang, Z., Nikpour, A. and Kavehei, O. (2025). "Spiking neural networks with liquid time-constant dynamics and dendritic branches for efficient time-domain edge-based seizure detection." *IOP Machine Learning Health*, 1(1), 015001.DOI:10.1088/3049-477X/adeb8e. https://github.com/NeuroSyd/Liquid-Dendrite

- **Herbozo Contreras, L. F.**, Yu, L., Huang, Z., Aguilar, I., Nikpour, A., and Kavehei, O. (2025). "Neuromorphic Neuromodulation: A Low-Power Edge-Training Framework for the Future of Personalized and Closed-Loop Neurostimulation." *Under Review; Minor* https://github.com/NeuroSyd/Akida-Seizure

1.6.2 Other publications and Source Codes

- Huang, Z.*, **Herbozo Contreras, L. F.***, Yu, L., Truong, N. D, Nikpour, A. and Kavehei, O. (2024). "S4D-ECG: A shallow state-of-the-art model for cardiac abnormality classification." *Cardiovascular Engineering and Technology*, 1-12. DOI: 10.1007/s13239-024-00716-3. https://github.com/NeuroSyd/S4D-ECG

- Huang, Z.*, **Herbozo Contreras, L. F.***, Leung, W. H., Yu, L., Truong, N. D, Nikpour, A. and Kavehei, O. (2024). "Efficient Edge-AI Models for Robust ECG

Abnormality Detection on Resource-Constrained Hardware." *Journal of Cardiovascular Translational Research*, 1-14. DOI: 10.1007/s12265-024-10504-y. https://github.com/NeuroSyd/NCP

- Huang, Z., Yu, L., **Herbozo Contreras, L. F.**, Eshraghian, K., Truong, N. D., Nikpour, A., and Kavehei, O. (2025). "Advancing privacy-aware machine learning on sensitive data via edge-based continual µ-training for personalized large models." Machine Learning: Science and Technology, 6(1), 015025. https://github.com/NeuroSyd/microTrainer

- Huang, Z., Yu, L., **Herbozo Contreras, L. F.**, and Kavehei, O. (2025). "Efficient and secure µ-training and µ-fine-tuning for edge-based TinyML with future-guided self-distillation." Machine Learning: Health, 1(1), 015002. https://github.com/NeuroSyd/uTraining

- Aguilar, I., Bersani-Veroni, T., **Herbozo Contreras, L. F.**, Nikpour, A., Querlioz, D., and Kavehei, O (2025). "Continuous metaplastic training on brain signals " *npj Unconventional Computing* 2(1), 9.
DOI: 10.1038/s44335-025-00025-5
https://github.com/NeuroSyd/MetaplasticEEG

- Yu, L., **Herbozo Contreras, L. F.**, Huang, Z., Yang, Y., Chen, B., and Kavehei, O. (2025). Hearables: Bioelectronics technological challenges and opportunities. Wearable Technologies, 6, e49.

* Joint first authorship

1.6.3 Conferences

- Huang, Z.*, **Herbozo Contreras, L. F.***, Leung, W. H., Yu, L., Truong, N. D., Nikpour, A. and Kavehei, O. (2023). "Efficient Edge-AI Models for Robust ECG Abnormality Detection on Resource-Constrained Hardware." *tinyML Research Symposium*, Poster Presentation, Seoul, South Korea, 16 November 2023.

- **Herbozo Contreras, L. F.**, Huang, Z., Yu, L., Nikpour, A. and Kavehei, O. (2024). "Biological plausible algorithm for seizure detection: Toward AI-enabled electroceuticals at the edge. " *NeuroEng Workshop at Western Sydney University*, Poster Presentation, Sydney, Australia, 11-13 December 2023..

- **Herbozo Contreras, L. F.**, Yu, L., Huang, Z., Nikpour, A. and Kavehei. O (2024). "Liquid-Dendrite Model: A 153 KB 8-bit Quantized Bio-Inspired Spiking Neural Network Model for Time-Domain Efficient Seizure Detection." *tinyML Research Symposium*, Poster Presentation, Taipei, Taiwan, 6 November 2024.

- **Herbozo Contreras, L. F.**, Yu, L., Huang, Z., Aguilar, I., Nikpour, A., and Kavehei, O. (2025). "Neuromorphic Neuromodulation: A Low-Power Edge-Training Framework for the Future of Personalized and Closed-Loop Neurostimulation. *10th IEEE International Conference on Rebooting Computing (ICRC)*, Oral Presentation, San Diego, United States, 15-16 December 2025.

Chapter 2

Neuromorphic Neuromodulation: A Perspective

The content presented in this chapter is published as:

- **Herbozo Contreras, L. F.**, Duy Truong, N., Eshraghian, J. K., Xu, Z., Huang, Z., Bersani-Vincenzo, T., Aguilar, I., Leung W. H., Nikpour, A. and Kavehei, O. (2024). "Neuromorphic neuromodulation: Towards the next generation of closed-loop neurostimulation." *PNAS Nexus*, page 488. DOI:10.1093/pnasnexus/pgae488.

Statement of Contributions of Joint Authorship

- Luis Fernando Herbozo Contreras (Candidate): First author, formal analysis, investigation, methodology, resources, software, visualisation, validation, writing original draft, writing review & editing of the manuscript.
- Nhan Duy Truong: Contribute to editing of the manuscript.
- Jason K Eshraghian: Contribute to editing of the manuscript.
- Zhangyu Xu: Contribute to editing of the manuscript.
- Zhaojing Huang: Contribute to editing of the manuscript.
- Thomas Bersani-Vincenzo: Contribute to editing of the manuscript.
- Isabelle Aguilar: Contribute to editing of the manuscript.
- Wing Hang Leung: Contribute to editing of the manuscript.
- Armin Nikpour (Co-Supervisor): Result investigation, project mentorship, resources, supervision, validation, writing review & editing.
- Omid Kavehei (Principal Supervisor): Conceptualisation, funding acquisition, investigation, methodology, project administration, resources, supervision, visualisation, validation, writing review & editing.

In addition to the statements above, in cases where I am not the corresponding author of a published item, permission to include the published material has been granted by the corresponding author.

Luis Fernando Herbozo Contreras Date: 28 May 2026

As supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

Professor Omid Kavehei Date: 28 May 2026

Chapter Summary

Neuromodulation techniques have emerged as promising approaches for treating a wide range of neurological disorders, precisely delivering electrical stimulation to modulate abnormal neuronal activity. Although AI-driven methods may support more adaptive and personalised responsive neurostimulation, it presents a challenging proposition, as real-time (low-latency) processing, low power consumption, and heat constraints are limiting factors. Another limitation is the dependence on back-telemetry to external systems (e.g., cloud-based medical mesosystems and ecosystems). While this can be a solution, integrating continuous learning within implantable neuromodulation devices for several applications, such as seizure prediction in epilepsy, is an open question. Neuromorphic architectures may provide a promising direction for efficient on-chip analysis of neural signals and AI-assisted personalised treatments. With more than three orders of magnitude reduction in the total data required for data processing and feature extraction, the high power- and memory-efficiency of neuromorphic computing, combined with hardware-firmware co-design, can be considered a solution for resource-constrained implantable neuromodulation systems. This chapter discusses *neuromorphic neuromodulation* as a potential direction for future closed-loop responsive systems. It highlights its potential to enable implantable brain-machine microsystems for patient-specific treatment.

2.1 Introduction

Electrical brain stimulation has evolved significantly over the past half-century. It started in the 1950s when it was found that emotional responses can be triggered by electrical brain stimulation Delgado et al. (1954). Penfield and Jasper's work was pivotal in mapping cortical functions, which they used to enhance the understanding of seizure semiology Penfield and Jasper (1954). Since then, there has been an increasing number of studies on the safety of brain stimulation and its applications as therapy of intractable epilepsy, spinal cord injury, psychiatric illness, Parkinson's dis-

ease, dystonia, refractory depression and Alzheimer's disease Goddard et al. (1969); Gordon et al. (1990); Kim et al. (2021); Kupsch et al. (2006); Laxton et al. (2010); Lo and Widge (2017); Lozano et al. (2016); Mayberg et al. (2005); Piper et al. (2022). However, there is yet to be an effective, scalable, personalised, and truly responsive stimulation solution for refractory epilepsy or neurological diseases in general. The market share of neurostimulation devices was more than US$6B in 2020 and is projected to pass US$11B by 2026 [1]. Key manufacturers of neurostimulation devices include Medtronic, Boston Scientific, Abbott, LivaNova, Nevro, NeuroPace, Beijing Pins, and Synapse Biomedical. Fig. 2.1 depicts the history of implantable neurostimulation devices and the trend in advanced neurostimulation. Although it does not perform neurostimulation, we consider the first pacemaker Kerzenmacher (2013) an important milestone on the roadmap, as it shares the same core idea: electrical stimulation. A decade after the first pacemaker, in 1967, the first implantable stimulation device was introduced for chronic pain relief. Since then, neurostimulation has shown consistent effectiveness in reducing chronic pain Hofmeister et al. (2020). This is followed by the first implantable defibrillator reported in 1980 Mirowski et al. (1980). Neurostimulation has been explored for its potential as a treatment or therapy for other diseases such as epilepsy, Parkinson's disease, Alzheimer's disease, and spinal cord injury. The year 1997 marks the first Food and Drug Administration (FDA)-approved Vagus Nerve Stimulation (VNS) device in treating intractable epilepsy Lulic et al. (2009); Schachter and Saper (1998). The device, NeuroCybernetic Prosthesis, is based on the finding that stimulating the vagus nerve modulates cortical activity via thalamocortical pathways, though the precise mechanism is not yet fully understood Rao (2021). Deep Brain Stimulation (DBS) was first used in 1980 for the reduction of tremors Brice and Mclellan (1980) and has since become an effective treatment of Parkinson's disease with impressive clinical outcomes in terms of motor and non-motor effects and quality of life improvements Deuschl et al. (2013); Kim et al. (2021). Neuromodulation has been used for more acute conditions, such as spinal cord injuries, where epidural electrical stimulation is applied to activate specific sensorimotor functions Wenger et al. (2014). Closed-loop VNS has shown

[1] Neurostimulation Devices Market (2021). Information can be found here: link

promising evidence of the prolonged effects in restoring neural circuitry with a study on rats Ganzer et al. (2018). Less commonly, neurostimulation has also been explored with other conditions such as psychiatric illness Lo and Widge (2017) and loss of control eating Wu et al. (2020a). Recent advancements in neuromodulation underscore its growing flexibility in targeting specific brain regions or networks, facilitated by various administration methods, including one-time treatments, continuous delivery, and adaptive responses to physiological changes. These advancements are pivotal for enhancing the efficacy and versatility of neuromodulation devices, particularly through the integration of advanced algorithms and responsive feedback mechanisms, as projected by 2035 Denison and Morrell (2022).

Future brain stimulation devices may increasingly incorporate advanced algorithms that combine predictive models and responsive feedback mechanisms: In responsive neurostimulation, a neurostimulator device is surgically implanted within the patient's brain or near the affected area. This device has electrodes that continuously monitor the brain's electrical activity in real-time. It is programmed to detect abnormal electrical patterns or seizure onset based on predefined algorithms. One example is to treat epilepsy by continuously monitoring iEEG and providing stimulation only when epileptiform activity is detected. NeuroPace developed the first Responsive Neurostimulation System (RNS) systems for epilepsy that detect abnormal brain activity and respond in real-time Krucoff et al. (2021). This reduces the amount of stimulation required and improves the accuracy of the treatment. This closed-loop system utilises databases, modelling, and machine learning to enhance performance, while data collection requires additional telemetry and storage. Another example is the Mayo Epilepsy Personal Assistant Device (EPAD) that combines an implanted device with iEEG telemetry, electrical stimulation, behavioural state classifiers, remote parameter control, a handheld computational device, and a cloud training for managing neurological diseases Balzekas et al. (2021); Nejedly et al. (2019). Consequently, data-driven methods are likely to play an increasingly important role in the development of future neuromodulation systems.

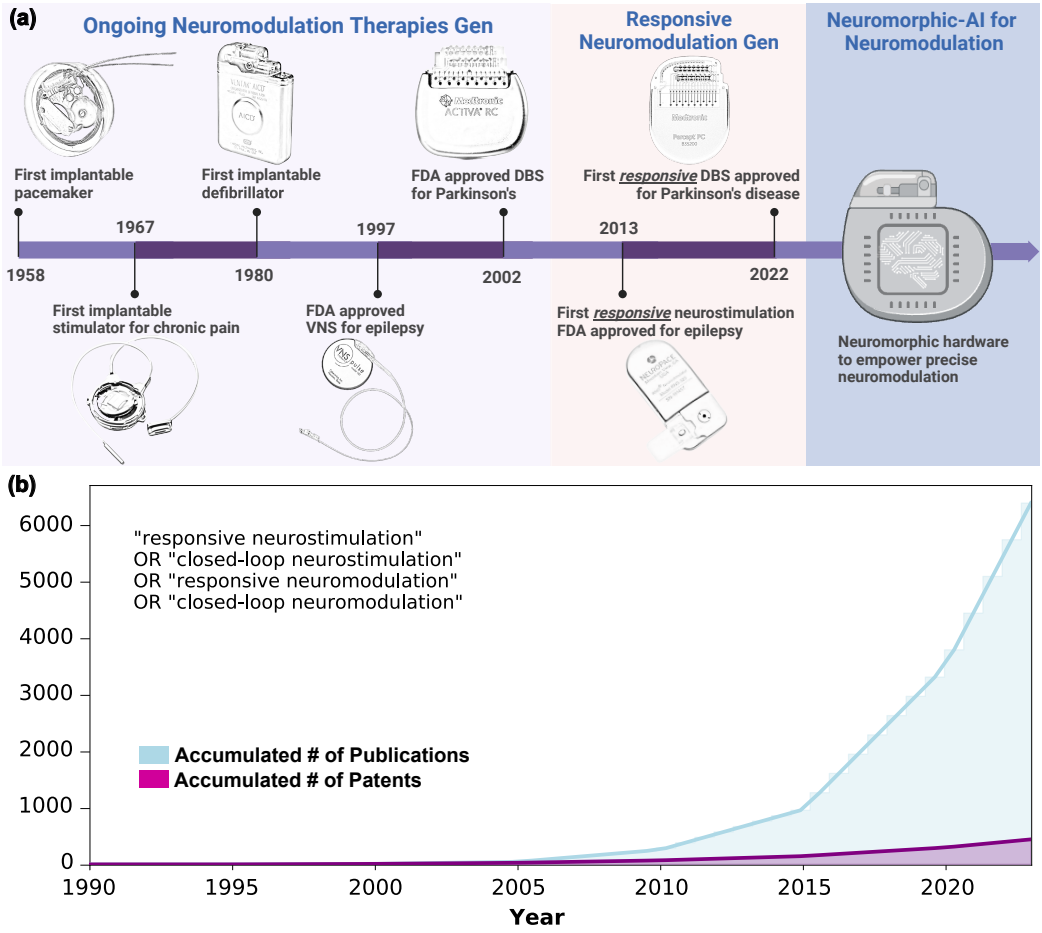

Figure 2.1 – Brief history of neurostimulation devices and possible future directions. The first generation of neuromodulation devices primarily delivered constant electrical stimulation to targeted brain regions. Responsive stimulation generation represents a significant leap forward, incorporating closed-loop systems that dynamically adjust stimulation parameters based on real-time feedback from neural activity or physiological markers. Inspired by brain computing, a possible future direction is neuromorphic neuromodulation, which may support more precise and energy-efficient therapeutic interventions if challenges in reliability, validation, and clinical deployment are addressed (a). Accumulated number of publications in responsive or closed-loop neurostimulation (b). An evident gap exists in closed-loop systems about the requirement for on-chip devices capable of continuous learning.

Challenges with current responsive neurostimulation devices: Almost 55% of neurostimulation devices are intended for pain management. For more acute diseases like epilepsy, despite promising reports showing a reduction of around 50% in seizure frequency with responsive neurostimulation Bergey et al. (2015); Heck et al. (2014), people with refractory epilepsy still develop seizures that prevent them from joining the workforce or performing certain daily activities (e.g., driving). There are several reasons for the low efficacy in preventing/suppressing seizures. Firstly, the stimulation is activated based on the detection of anomalies, presumably epileptic seizures, in brain signals which are manually pre-defined by physicians Geller et al. (2017); Heck et al. (2014); Jobst et al. (2017); Kokoszka et al. (2018); Ma et al. (2020); Singhal et al. (2018); Skarpaas et al. (2019). We argue that stimulating the brain after seizure onset is detected is sub-optimal. Activation of stimulation before the onset is likely more effective in preventing seizures. This idea was proposed back in 2003 Litt et al. (2003) and has been taken up in several patents Harrer et al. (2020); Wingeier and Tcheng (2016), but has not yet been tested in clinical trials. What prevents such a system from being effective is the lack of a high-performance model for predicting neurological events, such as seizures. Seizure prediction is more challenging than the detection counterpart, and its performance relies on long-term and patient-specific EEG recording Cook et al. (2013); Duun-Henriksen et al. (2020); Stirling et al. (2021); Weisdorf et al. (2019). Nevertheless, having access to ultra-long-term EEG recordings is just one part of the overall solution for robustly developing seizure prediction models. The EEG recording needs to be labelled such that it can be used to train a machine learning model, e.g., a deep neural network. This process is not only time-consuming but also requires neurologists to manually read and label, and must be regularly repeated as underlying physiological patterns change over time Sun and Morrell (2014). This corresponds to data set drift. Furthermore, interpreting brain data obtained chronically or in real time requires advanced analytics that rely on deep learning algorithms and intensive computational capabilities, which are unsuitable for current on-chip hardware and software approaches for on-chip learning Kremen et al. (2018).

Can these devices be smarter, extraordinarily energy-efficient and perform truly real-time closed-loop therapy? Neurotechnology research has seen a surge in startups and companies over the past decade, but on-chip computation is currently limited to simple signal processing and feature extraction. Existing systems such as the RNS, Percept family™: (PC/RC) and Summit RC+S Carrette et al. (2015); Jimenez-Shahed (2021); Stanslaski et al. (2018) rely on external systems with advanced machine learning algorithms for accurate symptom tracking. [2] For instance, the investigational Medtronic Summit RC+S utilises an embedded dual Linear Discriminant Classifier that consumes 5 μW/channel, and its parameters can be upgraded through telemetry. Percept family™ includes the PC and RC. The PC device incorporates BrainSense™ technology, specifically designed to acquire brain signals, known as Local Field Potential (LFP)s, utilising the implanted DBS lead. Regarding the utilisation of BrainSense™ technology for a patient with Parkinson's Disease, the system typically consumes a moderate amount of energy over two months when BrainSense™ is incorporated. This energy usage is expected to sustain the device for 5 years. The FDA approved the latest innovation of the Percept™(RC) in early 2024, which includes rechargeable neurostimulation. The RC is the smallest and thinnest dual-channel neurostimulator available for DBS, offering at least 15 years of service life with consistent, fast recharge performance [3]. Brain Interchange ONE is the first version of the CorTec Brain Interchange technology, which is currently approved for the first study by the FDA. It consists of a bidirectional, closed-loop neuromodulation device that supplies energy via induction and records and stimulates brain activity across 32 channels. This device relies on external artificial intelligence running on designed software [4].

The RNS system continuously monitors Electrocorticography (ECoG) at the seizure focus and delivers closed-loop electrical stimulation when abnormal (epileptiform) patterns are detected. Two versions have been available in the market: The RNS-300 and RNS-320, with the latter incorporating the most recent advancements. Both

[2] Some of these techniques are approved only for investigational use.
[3] FDA approves Medtronic Percept™ RC neurostimulator with BrainSense™, 2024 technology
[4] Cortec: Brain Interchange One, 2015. Link: Here

versions depend on the telemetry component used for communication, storage, and access to historical neurostimulator data. However, one significant design feature, or lack thereof, in these devices is the demand for external transmission of information, as continuous data telemetry drains their battery quickly. This system, which is said to be a continuous monitoring system, is only partially true. For instance, one of the RNS models can only record up to 4 minutes of ECoG and can be scheduled to repeat this up to 4 times within a 24-hour period. That means information about the whole day is not present; only instances of times are. This device also passively records multiple seizures, which helps develop detection algorithms tailored to the patient. By analysing these recordings, algorithms can identify seizure patterns and apply responsive stimulation using techniques such as line length and half-wave detection. Existing deep learning models outperform such algorithms. These alternatives face challenges, such as increased power consumption from wireless data telemetry and significant feedback-loop latency (several hundred milliseconds) compared to an on-device equivalent. Integrating this alternative approach could reduce the effectiveness of closed-loop stimulation and increase the need for more frequent battery replacements or recharges of implanted batteries. The goal is to design the next generation of intelligent neuromodulation systems with greater on-chip computing, improved energy efficiency, and overall miniaturisation Yoo and Shoaran (2021).

In this perspective, we aim to explain why neuromorphic computing may represent a potential solution for making embedded smart electroceutical devices. These groundbreaking devices use electrical impulses to precisely modulate the body's neural circuits Famm et al. (2013); Long et al. (2021); Maeng et al. (2022). We discussed the practical advantages of our approach, feasible applications, case studies with potential for improvement, and the challenges and opportunities in this emerging field.

2.1.1 Driving the next generation of on-device AI-revolution in electroceuticals through neuromorphic neuromodulation conceptualisation

Data telemetry is power hungry: The rapid advancement of AI and neural networks has enabled computers to exhibit impressive cognitive abilities. However, reducing computational costs and achieving brain-like efficiency remain challenging. Deep neural networks form the basis of state-of-the-art AI as it stands, and these networks rely on computing systems, from the transistors to hugely memory-intensive graphics processing units Graphics Processing Unit (GPU), which consume substantial energy in their general-purpose and conventional computing architectures. Training these networks on energy-intensive servers yields high accuracy but also high energy consumption. For example, running a model on an intelligent glass-embedded processor would exhaust its battery (2.1 Wh) in 25 minutes Venkataramani et al. (2016). This high power consumption makes such systems unsuitable for bio-electronic medicine applications, which prioritise low energy usage. External data processing in implantable devices requires wireless data telemetry, which is limited by bandwidth, communication range, interference, and, crucially, energy requirements. As real-time processing is necessary, such external interaction would hinder timely responses to signal features and raise potential efficacy issues. Conversely, on-device (edge) computing solutions enable the immediate processing of recorded signals and facilitate closed-loop interventions Kiourti and Nikita (2012); Movassaghi et al. (2014). Expanding edge computing capabilities beyond inference-only to on-device learning would significantly enhance the personalisation and efficacy of these devices.

Can we take inspiration from the brain through neuromorphic? The human brain possesses a remarkable computational power ranging from 10^{13} to 10^{16} operations per second, with a power consumption of approximately 20 W [5]. In contrast, a

[5]This is an indicative figure based on whole body metabolic studies.

computer performing a classification task requires about 250 W. The brain consists of billions of neurons ($\sim 9\times10^9$) connected by trillions of synapses ($\sim 3\times10^{14}$), allowing for information processing at a rate of approximately 6×10^{16} bits per second Fares et al. (2022); Martins et al. (2019). Recent investigations explore the prospects of neuromodulation over the past decade Afsaneh and Zarei Ghobadi (2023), discussing the potential of neuromorphic chips for implanted body-machine systems that mimic the co-location of logic and memory, hyper-connectivity, and parallel processing of the human brain, as shown in Fig. 2.2. The field of neuromorphic computing has seen significant advancements in industry and academia Fares et al. (2022). Several notable neuromorphic chips, like IBM's TrueNorth and Intel's Loihi, cater to specific applications with dedicated software ecosystems. In the European Union Human Brain Project, chips such as BrainScales, SpiNNaker, NeuroGrid, IFAT, and DYNAPs excel at tasks such as object detection and medical image analysis. There's a growing focus on versatile neuromorphic platforms that integrate hardware and software, such as the Tianjic chip, which supports both SNN and traditional ANN. Spinnaker serves as a general-purpose accelerator for diverse workloads. These chips employ digital, analog, or mixed-signal configurations based on their functional needs Benjamin and et al. (2014); Blouw et al. (2019); Davies and other authors (2018); Furber et al. (2014); Getty et al. (2021); Merolla and et al. (2014); Moradi et al. (2017); Mostafa et al. (2015); Schemmel et al. (2021, 2010); Thakur and et al. (2018). Table 2.1 overviews some of the most prominent current neuromorphic chips. Notably, there is a wide range of neuromorphic chips, but we consider the most commercially available for demonstration. Research notes limitations in current devices for peripheral nervous system stimulation and suggests neuromorphic circuits as an ideal solution for enhancing bioelectric medicine. Adaptive closed-loop systems based on neuromorphic engineering can improve symptom control by continuously monitoring physiological signals and adapting in real time Donati and Indiveri (2023). These systems use mixed-mode analog and digital transistors and consume ultra-low power. Neuromorphic engineering can overcome bandwidth and power-consumption limitations, thereby improving neural data acquisition and processing Chicca et al. (2014). Analog neuromorphic front-ends offer a low-power solution for high-bandwidth neural

recording and multi-channel processing needs. They process analog signals directly, converting them into spikes for SNN use. Recent advances in neuromorphic computing enable on-chip training with minimal power consumption and a small device footprint Frenkel et al. (2020), demonstrating on-chip training in a 32-mm^2 silicon area, achieving 95.3% accuracy on the Modified National Institute of Standards and Technology dataset (MNIST) dataset, slightly lower than off-chip training's 97.5% accuracy. Existing neuromorphic chips, such as the well-known IBM TrueNorth and Intel Loihi, are general-purpose chips that support various network architectures and configurable parameters (e.g., number of layers, kernel sizes, etc.). However, their versatility comes at a cost: higher power consumption and greater heat dissipation. For the implementation of neural networks with learning capability, a neuromorphic chip should be fully optimised for a specific application if continuous active learning is to be coupled with a medical device, especially implants with strict temperature constraints.

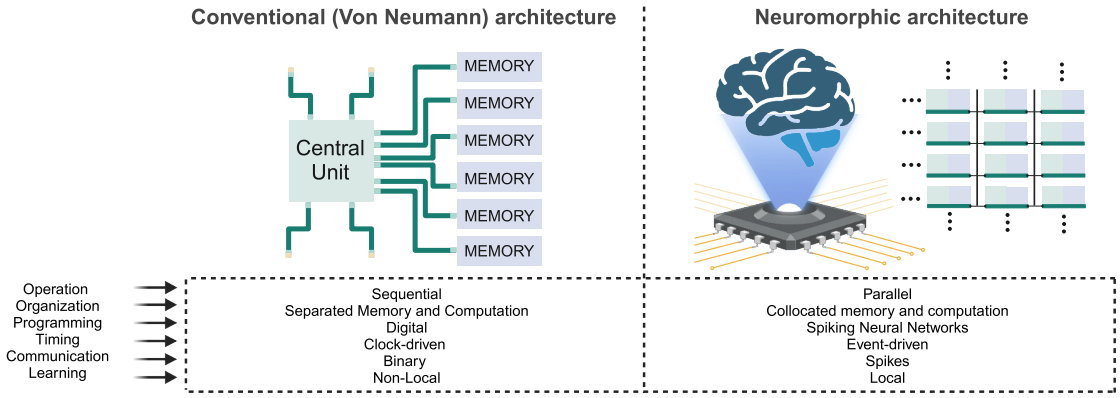

Figure 2.2 – Conventional (Von Neumann) vs Neuromorphic Architectures Conventional computers rely on sequential, clock-driven (synchronous) binary operations, separating memory and computation units. In contrast, the human brain employs event-driven (asynchronous), neural action potentials (spikes), with a great network capacity for parallel processing and the capability for local learning mechanisms. These basic, seemingly shallow, yet fundamental distinctions contribute to the brain's inherent superiority in energy efficiency, positioning it as a promising avenue for the development of custom or general-purpose integrated circuits and computing architectures.

Table 2.1 – A benchmark of neuromorphic chips. This list is not extensive but considers the breakthroughs.

Features / Chips	CPU	TrueNorth	Loihi2	Tianjic	SpiNNaker	Brain ScaleS-2	DYNAP SEL	Akida	Human Brain
Event-driven	No	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Analog or Digital	Digital	Digital	Digital	Digital	Analog	Analog	Analog	Analog	Analog
Node (nm)	5	28	7	28	130	65	28	28	-
In-memory computing	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Neurons / Synapses count	-	1M/256M	1M/120M	40K/10M	16k/8M per chip	-	1K/78K	1.2M/10B	20B/200T
Cores	Vary	4096	128	156	18 per chip	8	5	80	-
Neurons / Synapses per Core	-	256/64K	8K/900K	256/64K	800/1M	512/130K	256/	15K/125M	-
On-device learning	Backprop / Spike-Timing-Dependent Plasticity (STDP)	No	STDP /Surrogate backprop	No	STDP	STDP, Reward-Modulated Spike-Timing-Dependent Plasticity (R-STDP), homeostatic plasticity	STDP	Few-shot training	Diverse learning
Network compatibility	ANN	SNN	SNN	Hybrid	Hybrid	SNN	SNN	ANN, SNN	SNN
Key properties	Limited to on-chip training	First commercial neuromorphic chip	Fully programmable chip	First hybrid chip	Scalable SNN simulation	Hybrid Plasticity	One plastic core	Continuous and few-shot learning	-
ASIC / FPGAs	-	ASIC	ASIC	ASIC	ASIC	BOTH	Field-Programmable Gate Array (FPGA)s	ASIC	-
GSOPS / W	-	400	$\sim 10 \times$ Loihi1[†]	649-1278	0.033	> 10	33	Akida	-
Memory type	-	SRAM/ DRAM	SRAM /DRAM	SRAM	SRAM	SRAM	SRAM	SRAM	Different brain regions
Routing schemes	-	Grid/2D Mesh	Grid/2D Mesh	2D Mesh	Grid/2D Mesh	Hierarchical Tree	Hierarchical Tree	Grid/2D Mesh	Stochastic
Refs.		Akopyan et al. (2015); Young et al. (2019)	[◇]	Deng et al. (2020); Pei et al. (2019)	Furber et al. (2014); Human-BrainProject (2023)	Grübl et al. (2020); HumanBrain-Project (2023); Pehle et al. (2022)	Moradi et al. (2017); Rueckert (2020)	Demler (2019)	Martins et al. (2019, 2016)

[†] There is no direct data available but it is stated to be 10 times more than Loihi1.

[◇] Intel Neuromorphic Computing - Loihi 2 Technology Brief, 2023. Link: Here.

2.1.2 Application for neuromorphic neuromodulation

Fig. 2.3 demonstrates three main categories of neuromodulation devices, including those which are commercially available (a-b) and under investigation (c). In an open-loop system (Fig. 2.3(a)), stimulation parameters such as amplitude, frequency, and duty cycle are pre-determined by a clinician. Stimulation persists unless the patient or the clinician manually turns it off. The device can be reprogrammed during a subsequent patient visit if the stimulation is ineffective. In contrast, a closed-loop system triggers stimulation by responding to physiological changes. The system in Fig. 2.3(b) continuously senses the patient's state (e.g., EEG signals) and streams it wirelessly to a portable device (or a bedside computer) that is in charge of analysing the signals by using threshold-based rules or machine learning models, and turning it into a control signal to the stimulation Chen et al. (2021); Denison and Santa (2017); Giftakis et al. (2014); Kremen et al. (2018); Mivalt et al. (2022); Sladky et al. (2022); Snyder and Leyde (2008); Stanslaski et al. (2018); Stirling et al. (2021). In epilepsy management, a portable device activates stimulation upon detecting or predicting a seizure onset. This closed-loop system, aided by external computing, minimises unnecessary stimulation compared to traditional approaches. However, continuous data streaming to external resources consumes substantial power, limiting battery life or increasing device size. Moreover, wireless communication between the device and external computing systems poses challenges, including connection loss, interference, and security risks. Fig. 2.3(c) illustrates how an alternative closed-loop system can address the complications from the continuous data stream of standard closed-loop systems by incorporating an on-device computing unit to the neuro-modulation device without reliance on external computational power to host the control algorithms Harrer et al. (2020); Pepin and Kotzev (2021); Rhew et al. (2014). In these closed-loop system models (Fig. 2.3(b) and (c)), the optimisation of models (on-device or off-device) must still be performed regularly and involves a human expert and/or cloud computation to adapt with the changes in the patient's conditions and/or in the underlying disease Sun and Morrell (2014). This implies that the patient's data needs to be stored in the external device (Fig. 2.3(b)) or in the implantable device

(Fig. 2.3(c)) and regularly be uploaded to the cloud. The data will also need to be analysed or labelled by a human expert so it can be used to update threshold-based rules or to retrain the machine learning models. It should be noted that none of the aforementioned methods offer on-device training and re-training, and require expert involvement for regular retraining Ashourvan et al. (2020), which limits the scalability of the system to a large number of patients Karuppiah Ramachandran et al. (2018). The idea of on-device active learning proposed in Xiao et al. (2017) relied on an ideal detection and deterministic feature extraction technique to actively train a prediction model without expert intervention or external computational resources. However, we argue that deterministic feature extraction may lose efficacy over time as the underlying disease evolves. Our alternative is neuromorphic neuromodulation, a computationally self-sufficient closed-loop system, shown in Fig 2.3(d). Our proposed system eliminates the need for continuous data telemetry and reliance on external computational resources. We believe our self-contained system can provide an ultimate personalised closed-loop neuromodulation system. The vision is ambitious but not impractical.

2.1.3 Evaluating Standards in a Self-Sufficient Responsive Neuromorphic System

Here, we discuss some fundamental standards and criteria that our system adheres for future neuromodulation devices.

Physiological event detection becoming more trustworthy: The field of automatic annotation of physiological data has seen significant advancements, with recent developments approaching the accuracy and reliability of human experts. Notable examples of these advancements include the detection and classification of arrhythmias Hannun et al. (2019), the identification of epileptiform discharges Jing et al. (2020), and the marking of seizures Scheuer et al. (2021). These technological improvements not only enhance automated health monitoring, thereby reducing clin-

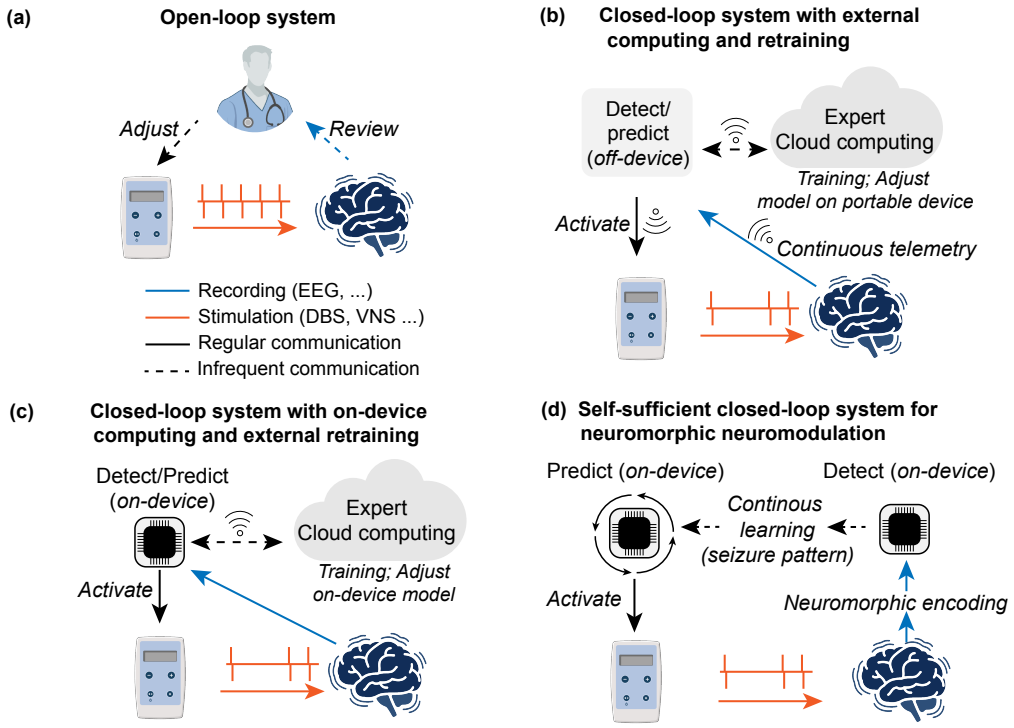

Figure 2.3 – Neuromodulation approaches. (a) An open-loop system with an expert who occasionally reviews the effectiveness of the system and adjusts the stimulation parameters accordingly. Such a system employs cyclic stimulation regardless of the target's current state (e.g., brain state). (b) A closed-loop system with external computing for accessing the state of the target to condition the stimulation. The external computing component can be a portable device, e.g., a tablet, or a local computer. The recording device continuously streams data (e.g., EEG signals) to an external computing system, where trained algorithms are executed to determine the target's state. The deployed algorithms on the external computing component can be updated occasionally by involving a review from an expert(s) and big data/cloud computing (retraining). (c) The computing component is embedded within the device, which eliminates the need for *continuous* streaming of data to the outside world Harrer et al. (2020); Pepin and Kotzev (2021). However, the on-device algorithms need to be updated periodically to reflect changes in physiological signals (e.g., alterations in seizure patterns in patients with epilepsy). The device must also have sufficient memory to store the signals for the expert(s) to review and for retraining in the cloud. (d) A neuromorphic neuromodulation system where the medical device can run and retrain its algorithm by itself without relying on external computing resources. The system utilises labels generated by a detection model with performance on par with a human expert Hannun et al. (2019); Jing et al. (2020); Scheuer et al. (2021) and a continuous learning strategy that serves as a recorder to train a prediction model. The rapid advances in neuromorphic computing Bohnstingl et al. (2019); Park et al. (2019) have enabled on-device active learning. It's noteworthy that such a system will intermittently transmit relevant snapshots or markers of the recorded biosignal to comply with safety and efficacy standards.

icians' burden, but also create opportunities to leverage unlabeled data. For instance, algorithm-generated labels can be utilised to train other predictive models without the need for human expert intervention. While detecting seizures during or immediately after the onset has proven more successful than predictive methods, these advancements have led to the development of Adaptive, Unlabeled, and Real-time Approximate-Learning (AURA) platform Yang et al. (2022a).

On-device learning: AI systems utilizing Application-Specific Integrated Circuit (ASIC) with parallel multiply-and-accumulate Multiply-Accumulate (MAC) units exhibit better inference and energy efficiency than GPU- based systems. However, performing MAC operations requires intensive data transfer between these units and data buffers, limiting energy efficiency and restricting them to functioning solely in inference mode, whereas the human brain has the remarkable ability to learn continuously. Consequently, on-device learning emerges as a significant characteristic of neuromorphic systems. On-chip learning is indispensable for tailoring and personalising smart devices to cater to individual user requirements. Moreover, it bolsters privacy by eliminating the need to transmit user data to the cloud Frenkel et al. (2023); Ivanov et al. (2022); White et al. (2022); Zhang et al. (2020). At the core of AURA, on-device learning is achieved through a high-performance physiological event (e.g., epileptic seizure) detection model that serves as an algorithmic "human expert," generating labels on the fly as the signal arrives. The generated labels are paired with recorded signals from a loop recorder to be used as a training dataset for a predictive model (e.g., seizure forecasting). It is worth noting that while the detection model or label generator must perform well, it does not necessarily need to be perfect. In fact, imperfect labelling from a mix of clinicians and medical students with varying levels of experience has been shown to remain effective in training a deep learning model to achieve high-level seizure detection accuracy and generalisation. AI models face limitations due to misaligned metrics with clinical needs, requiring validation through prospective data and real-world data testing Surianarayanan et al. (2023).

Embrace multi-modal signals: As part of physiological monitoring, it is usual that there are multiple signals being recorded. For instance, in the VNS device, multimodal approaches are also used in some cases, such as incorporating a heart rate sensor to activate stimulation when the heart rate exceeds a predetermined threshold, since some seizures are associated with an acceleration in heart rate. Combining signals from multiple sources can improve the performance of a detection/prediction model Greene et al. (2007); Valderrama et al. (2010). It is important to note that, depending on performance requirements, power consumption, and/or heat dissipation, the detection model of AURA may use a different set of sensory modalities from the prediction counterpart.

Integrating Bio-Inspired Algorithms for Energy-Efficient Electroceutical Systems SNNs provide an alternative approach by mimicking the behaviour of biological neurons and offering potential energy-efficiency advantages, making them suitable for resource-constrained environments such as edge devices Surianarayanan et al. (2023). Training approaches can be categorised into different methods. One common approach is to directly train the SNN itself using surrogate gradient descent. Another approach involves training a traditional ANN and then mapping it into an SNN. The ANN is trained using conventional techniques, and the resulting trained weights and connections are then transferred to the SNN. Reservoir computing is another technique used in SNNs, in which the network is structured into an input layer, a reservoir layer, and a readout layer. The reservoir acts as a dynamic memory, and the readout layer is trained to interpret the reservoir's activity. Lastly, spike-timing-dependent plasticity is an approach based on synaptic plasticity, in which the weights of connections are adjusted according to the relative spike timings between pre- and post-synaptic neurons. These training approaches provide different strategies for training SNNs, each with its own advantages and applications. Kulkarni and Rajendran (2018); Lee et al. (2016); Mead (1990, 2020); Schliebs and Kasabov (2013); Schuman et al. (2022); Stöckl and Maass (2021); Tanaka et al. (2019). The AURA system is built on conventional SNN architectures and training. Studies have investigated seizure detection using closed-loop direct neurostimulation devices in

epilepsy with neuromorphic chips by successfully transferring a CNN to TrueNorth, demonstrating accurate detection with low memory usage and efficient runtime, with power consumption below 40μW. However, it is noted that CNNs' dependency on back-propagation can result in issues such as catastrophic forgetting and heightened computational costs Dümpelmann (2019); Esser et al. (2016); Kiral-Kornek et al. (2018).

Back-propagation: implausible biological way and issues with neuromorphic hardware: ANNs use back-propagation and gradient descent to adjust synaptic weights, but this leads to several issues, including catastrophic forgetting, weight-symmetry problems Liao et al. (2016), freezing of neural activity Whittington and Bogacz (2019), and non-local weight updates Lillicrap et al. (2020). Back-propagation is also vulnerable to adversarial attacks Akrout (2019) and requires excessive computational hardware in Very Large Scale Integration (VLSI) Jabri and Flower (1992). Novel solutions, such as the forward-forward algorithm, aim to address these problems Dellaferrera and Kreiman (2022); Hinton (2022). However, its current scope is limited to static datasets like CIFAR-10, with slightly worse test errors than the current back-propagation framework. Nonetheless, in the following case studies, we used biologically plausible solutions to address those problems and potentially served as a framework for the proposed perspective solution.

2.1.4 Case Studies for AURA: Foundations of Efficient, Low-Power, and Biologically Inspired Models for Seizure Detection

In the realm of seizure detection/prediction, the quest for efficient, low-power models that employ lifelong learning has drawn attention to bio-inspired algorithms, which can enhance the performance of the AURA system. Lifelong learning refers to a system's ability to continuously learn and adapt to new information, much as biological systems do. Within this context, our exploration unveils three distinct yet

impact cases of studies. The following models serve as exemplars of effectiveness and power efficiency, showcasing a novel approach to seizure detection with the potential for accurate neuromodulation. Leveraging principles derived from biological systems, these algorithms demonstrate prowess in lifelong learning, exhibiting adaptability and responsiveness.

Case 1: Continual Learning with Artificial Metaplastic Models: In computational neuroscience, the stability-plasticity dilemma revolves around an AI model's ability to acquire new memories while retaining existing ones. Synaptic plasticity, the basis of learning, involves neuronal connections adjusting their strength over time. This paradox is addressed through synaptic artificial meta-plasticity, a bio-inspired approach to continuous learning. Using a binarised neural network Binary Neural Network (BNN), researchers implemented synaptic meta-plasticity to mitigate catastrophic forgetting in multi-task learning. This method modulates hidden weights via a function *fmeta(Wh)*, potentially applicable to neuromorphic platforms Laborieux et al. (2021); Mermillod et al. (2013). We applied these principles to our model, MetaEEG, a low-power neuromorphic proof-of-concept for lifelong learning of EEG seizures. We propose a BNN with artificial metaplasticity for stream learning, targeting real time wearable physiological data streams.. We trained our model on the TUH dataset, dividing it into 300 subsets and sequentially presenting each to the model for 20 epochs. Every 5 subsets, we tested the model on unseen test data. We evaluated its performance on different EEG signatures to assess its adaptability to significant changes in seizure patterns. By generating 5 synthetic EEG datasets with distinct seizure signatures, we demonstrated the model's ability to adapt without forgetting prior patterns, achieving an AUROC of nearly 0.80 Aguilar et al. (2025). This feasibility proof paves the way for future studies that integrate artificial meta-plastic behaviour with SNN compatibility for seizure prediction, addressing buffer limitations in the AURA system.

Case 2: Effective, Sparse, Interpretable and Low-Power Liquid Time Constant Based Models: Liquid Time Neural Networks are a class of time-continuous

recurrent neural network models that exhibit stable, bounded behaviour, thereby improving performance on time-series prediction tasks. Their low complexity allow for a better representation of the hidden states, and adapting to changing conditions such as autonomous driving and medical time-series data Banerjee et al. (2009); Beveridge and Pereira (2022); Chahine et al. (2023); Hasani et al. (2022, 2021, 2023); Huang et al. (2024a); Lechner et al. (2020). We used Liquid Time Constant in different scenarios such as models on shallow bio-inspired models and spiking neural version with a forward-propagation through time algorithm Yin et al. (2023). We developed a dynamic spiking model for seizure detection across continental datasets utilising spiking liquid-neuron networks with Forward Propagation Through Time (FPTT) Herbozo Contreras et al. (2024a). By training and validating in the TUH dataset, we evaluated generalisation in the RPAH dataset, achieving an AUROC of 0.83 in 192 patients, which is slightly higher than a conventional algorithm based on Convolutional Long Short-Term Memory (ConvLSTM) Yang et al. (2022b). By reducing the model's memory requirement by 10 times, we examined its robustness and found it to perform at the level of a large dynamic SNN, with an AUROC of 0.82. Subsequently, we applied a scaled-down model, achieving an AUROC of 0.83 on the EPILEPSIAE Dataset. We provided an estimate of the model's power consumption of 3.1 μJ/Inf (per inference) on Loihi. Further exploration should include these dynamic models with ECoG and LFP data.

Case 3: On-Device Fine-Tuning with Spiking Liquid-Base Models: Feasible application on hardware: We initiated by exploring a Neural Circuit Policies (NCPS) model variant that employs a liquid time constant for Arrhythmia detection, incorporating on-device fine-tuning Huang et al. (2024b). Our setup utilised the Radxa Zero SBC with an Amlogic 905Y2 processor and 4GB LPDDR4 memory. Refinement of the model on the Radxa Zero involved leveraging a pretrained model, initially trained for 2 epochs on the Telehealth Network of Minas Gerais (TNMG) using an GPU. Adjustments were implemented, including reducing the batch size to 8 during fine-tuning to utilise 72% of memory. The fine-tuning phase consisted of training on a dataset of 640 data points across five epochs, followed by validation on a

subset of 320 data points. The evaluation of the fine-tuned model involved comparing its performance with that of the base model trained on GPU. Throughout the training iterations, the model displayed notable improvements in performance metrics. The average F1 score increased from 0.46 to 0.56, and the AUROC enhanced from 0.65 to 0.73. A fine-tuned model was tested on a larger dataset of 1280 samples, where significant improvements in F1 score and AUROC were observed, increasing from 0.31 and 0.63 to 0.45 and 0.72, respectively. These findings have motivated further investigation into its suitability for deployment on neuromorphic chips. Integrating these study cases could be seen in a bio-inspired model as detailed in Fig 2.4.

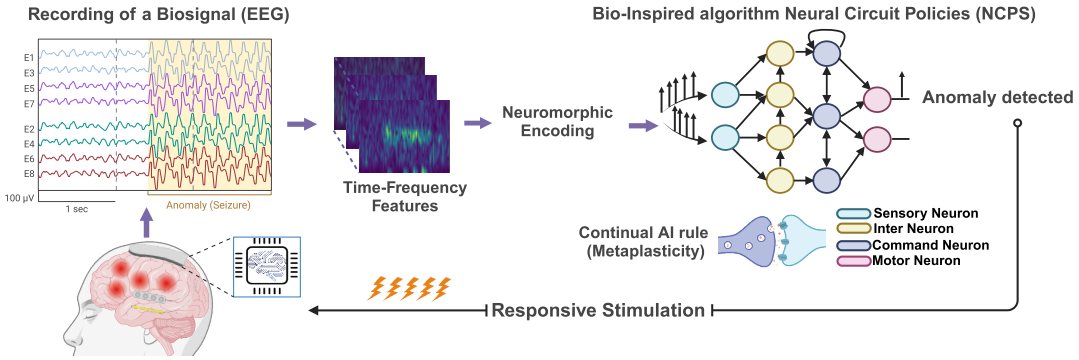

Figure 2.4 – Bio-Inspired Learning Rules for Neuromorphic Neuromodulation represents a cutting-edge paradigm in the field of neural systems. This innovative approach enables the development of on-device learning capabilities, thereby facilitating the seamless integration and real-time processing of continuous bio-signals. By leveraging this neuromorphic system, it becomes possible to dynamically adapt to extracted features and convert them into spikes. These spikes are then fed into a shallow, sparse, and bio-inspired algorithm that utilises a continuous learning rule for precise adjustment, ultimately yielding responsive stimulation tailored to each individual patient. This approach eliminates the reliance on cloud computing.

Estimation of power consumption of a fully integrated AURA system: A study conducted with SNN on EEG datasets (Freiburg, CHB-MIT, Epilepsiae) for seizure detection was proposed by Yang et al. (2023) where they demonstrate the capabilities of neuromorphic approaches to reduce the memory usage and energy consumption from ten to thousands of magnitude in comparison like running in conventional GPU devices with conventional AI algorithms. The results of this work are

demonstrated in Fig 2.5. Drawing on the aforementioned findings and case studies, these instances serve as a basis for projecting the power usage of a fully integrated system. For a fully integrated detection/prediction system, we assume the input signal EEG comprises 10 channels and is sampled at 128 Hz. For simplicity, both detection and prediction networks use the same ConvLSTM architecture proposed in Yang et al. (2022b), consisting of three ConvLSTM layers followed by two fully connected layers. The detection and prediction algorithms use input windows of 10 and 30 seconds, respectively. The inputs are divided into 50% overlapping 1-second sub-windows, which are fed to a ConvLSTM network. Using the Loihi neuromorphic chip as a reference and a real-time batch size of 1 (one input is processed at a time), the cost of inference for a single data sample is 25 mJ and 77 mJ for the detection and prediction algorithms, respectively. Since detection occurs every 10 seconds and prediction every 30 seconds, the energy can be amortised over time, with the total power consumption for inference of both networks being 5.1 mW (2.6 mW for detection and 2.5 mW for prediction). Note that this power consumption can be reduced by parallelising inference into batches and distributing them over a longer time interval. For example, with a batch size 32, the total inference power consumption becomes 160 μW. Regarding the training of the prediction network, since the architecture is fixed, the backward pass can be performed simultaneously with the forward pass in deterministic forward-mode auto-differentiation; therefore, the gradient calculation cost is similar to the inference cost. Given the network has 31.5M parameters and the energy to update each additional weight is 120 pJ (Loihi), the total cost of weight updates is 3.78 mJ. The total energy required to train the prediction algorithm is (77 mJ + 3.78 mJ) or 81.78 mJ. This training step occurs every 30 seconds, so its power consumption is 2.73 mW (batch size=1) or 85 μW (batch size=32). With a custom design, the EEG Analog Front End (AFE)'s power consumption can be optimised to less than 3 μW per channel Do Valle et al. (2016). Considering the AFE solely, using a commercially available rechargeable Li-Po battery with a size volume of 2.7×30×34 mm^3 and a capacity of 240 mAh at 3.7 V Co. (2021), the system can be powered for at least 24 hours before a recharge. The battery life can be considerably extended by performing calculations in batches, with the only trade-off

being a slight delay in obtaining results. Considering the size of the detection and prediction networks and the synaptic density of the TrueNorth chip, we estimate a required area of ∼62 mm^2 to implement the entire system. A power breakdown of an inspired neuromorphic device is provided in Fig 2.6. To this end, there may be opportunities to improve our algorithm by adopting a more biologically plausible approach. Studies have achieved efficient, energy-saving training of time-domain signals by incorporating dendrites into spiking neurons Zheng et al. (2024), potentially eliminating the dependency on Time-Series Signal Processing (TSSP) blocks. Leveraging the model's capacity to interpret biosignal data conserves power, prolonging the lifespan of implantable devices.

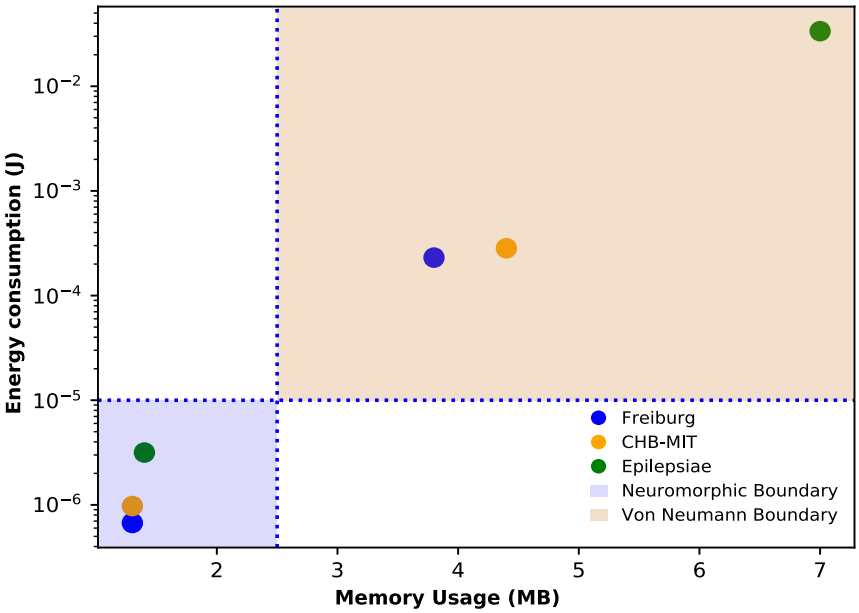

Figure 2.5 – Energy–memory trade-offs in Von Neumann and neuromorphic architectures. Memory usage and energy consumption are measured during an inference-only task, comparing a conventional AI model, ConvLSTM, running on a von Neumann computing architecture with a spiking ConvLSTM model deployed on a neuromorphic architecture.

2.1.5 Challenges and Opportunities of Neuromorphic AI

We will need a more robust mapping of the neural circuits underlying the pathophysiology associated with the treatment. At the signal level, we will need better decoders

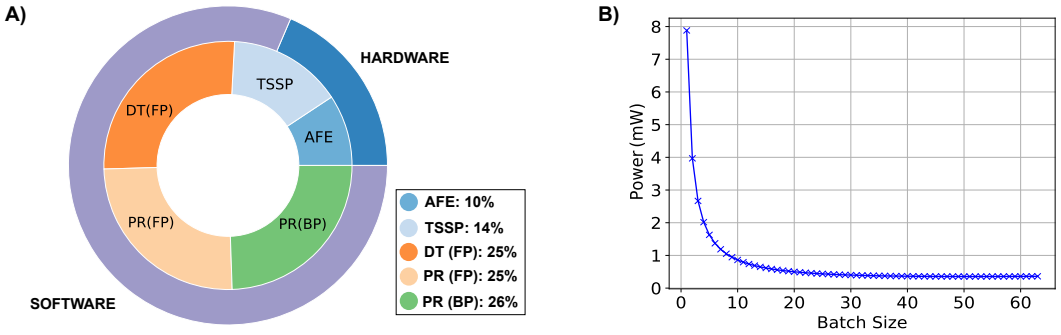

Figure 2.6 – Power breakdown of a Neuromorphic Device. Traditional algorithms used in neuromorphic often utilise TSSP methods like Fast Fourier Transform (FFT) or Short-Time Fourier Transform (STFT) for feature extraction to enhance performance. All parts discussed are considered as one system (A). Power reduction can be addressed through algorithm (software) development, which can be further improved by using larger batch sizes in processing operations and incorporating more biologically plausible algorithms for future implantable neuromorphic devices (B). AFE: Analog Front End, DT: Detection, PR: Prediction, FP: Forward Propagation, BP: Backwards Propagation.

of the neural language associated with the pathophysiological states and more precise therapeutic patterns of electrical impulses targeting the rate, even the timing of spikes Donati and Indiveri (2023); Donati and Valle (2024). Generating such adaptive and precise neuromodulators will require a multidisciplinary effort: the development of neuromorphic circuits for real-time spike processing will translate the biological understanding of what is happening at the neural level in health and disease Michmizos et al. (2017). The absence of standardised benchmarks in neuromorphic algorithm development makes it difficult to compare and assess hardware systems for specific applications. Neuromorphic hardware development involves extensive research into new materials and devices. Choosing appropriate materials is crucial for developing neuromorphic chips for neuromodulation, with carbon-based nanostructures suggested for bio-compatible probes and FDA-approved parylene utilised for neuromorphic building blocks Minnekhanov et al. (2019); Sangwan and Hersam (2020). Memristive systems provide prospects for the hardware realisation of ANNs for wearable and biomedical applications. Opportunities are present for software-hardware co-design, tailoring hardware to specific applications. It explores analog and mixed-signal computing, mimicking the stochastic nature of biological neural computation. Nanowires

Networks (NWN)s offer a promising hardware approach to emulate the brain's physical structure, including neurons and synapses. NWNs mimic synaptic metaplasticity, strengthening synaptic pathways for memory consolidation. This highlights their potential for neuromorphic systems, which are crucial for practical applications such as robotics and sensor edge devices Loeffler et al. (2023); Sandamirskaya (2022). Neuromorphic processors, with their low power consumption, are set to play a crucial role in various edge-computing and edge-learning applications in autonomous systems, robotics, remote sensing, implantable, wearables, and the Internet of X Things, where the X can be medical, industrial, etc Christensen et al. (2022); Schuman et al. (2022).

Promises of Neuromorphic AI: A Complementary Metal–Oxide–Semiconductor (CMOS)-based neuromorphic device detects epileptic seizures by analysing LFP signals Ronchini et al. (2021). It enables closed-loop intervention to reduce seizure duration using an SNN with a 64.98±30.92 ms delay and consumes <50 pW for ictal event detection. NET-TEN, a subsequent technology, enhances neuromorphic processors by reducing area and power consumption, making it suitable for implantable devices. Another study presents a first feasible real-time neuromorphic detection system for high-frequency oscillations, which utilises a mixed-signal neuromorphic computing system with high sensitivity Ronchini et al. (2023); Sharifshazileh et al. (2021). Integrating neuromorphic technologies into neuroprosthetic devices could offer a promising strategy for enhancing the development of more intuitive human-machine interfaces, by improving performance and embedding capability Buccelli et al. (2019); Chiappalone et al. (2022); Corradi and Indiveri (2015); Donati and Valle (2024). An study has unveil epilepsy seizure prediction system using deep learning and big data, compatible with low-power neuromorphic chip and wearable integration via closed-loop therapies Broccard et al. (2017); Donati et al. (2019); Kiral-Kornek et al. (2018).

Risk of false alarms or unnecessary stimulation: Ongoing Neuromodulation has shifted towards adaptive, closed-loop stimulation from traditional open-loop methods. The key challenge now is ensuring stimulation is activated precisely when

needed. Based on research findings showing that chronic brain stimulation can be performed safely with appropriate control of charge density Gordon et al. (1990); Morrell (2006), we can allow the stimulation activation system to have as high a sensitivity as possible with an increased number of false positives as a trade-off. A responsive stimulation with many false positives can be considered equivalent to an open-loop system performing cyclic stimulation, provided that charge density and stimulation frequency/duration are controlled. Saluda Medical conducted a pioneering study on responsive neurostimulation therapies, the first FDA-approved double-blind trial for Spinal Cord Stimulation. Patients were randomly assigned to either Evoked Compound Action Potential (ECAP)-controlled closed-loop stimulation or fixed-output, open-loop stimulation. Results showed a 21% higher success rate in the closed-loop group at both 3 and 12 months, without adverse effects Mekhail et al. (2020, 2023).

Elimination of continuous data communication: Wireless data communication can consume half or more of the total power consumption of the whole EEG recording implant Sawan et al. (2013); Yin et al. (2013). Neuralink reduces the frequency of sending data outside to every 25 ms and places a rechargeable battery and an inductive charger in the implant [6]. Continuous data communication or a brain-computer interface is critical for disease diagnosis. It is also unavoidable in responsive, closed-loop neurostimulation systems, where some computation (e.g., training the event detection/prediction model) must be performed on an external device or in the cloud Chen et al. (2021). The system's event detection/prediction model requires periodic re-training with recent data to adapt to patients' physiological changes. We propose that if the implant autonomously learns from its own data to adapt to changes and becomes patient-specific, external data communication and model training on external hardware can be eliminated, except for debugging. This approach aligns with interventional medicine, in which the device autonomously treats the condition based on diagnostic information.

[6] Neuralink: Monkey MindPong, 2021. https://www.bbc.com/news/technology-56688812

Data Security, Privacy, Dangers of these techniques, Hazards and Pitfalls: Neuromodulation devices must be developed to prevent this data from being abused or hacked. Issues to be addressed include how long these data should be stored, where they should be stored, and who is in charge. If data can be "written to" the brain, we need systems to guard against undesirable intrusions. Access to data provided by a medical device can be empowering for patients. This allows them to access their health data and receive alerts for concerning events, such as seizures. As examples, studies aimed to improve the security of insulin pump devices for diabetic patients. One employed an on-chip neural network, while the other proposed an efficient deep learning method to counter fake glucose dosages Rathore et al. (2017, 2018). Unauthorised access to the device, often called *brainjacking*, could allow an attacker to manipulate the stimulation parameters or even cause harm to the patient. Interference with the wireless communication between the device and external equipment could disrupt therapy. Ensuring the security of communication in Implantable Medical Device (IMD) is a critical issue for patient safety, with several research groups focusing on addressing challenges for a reliable solution due to factors such as the device's battery life, adaptability, and the required level of security to avoid malicious software Pugh et al. (2018); Pycroft et al. (2016); Surianarayanan et al. (2023). Integrating Body Area Communication (BAC) into AI systems reduces reliance on external telemetry while enabling the capture of daily activity snapshots for safety-efficiency standards. This enhances self-sufficiency, privacy, and security Tang et al. (2020).

2.2 Conclusions

Current challenges for designing implantable stimulation devices or electroceuticals, in general, include implant volume, safety, energy consumption, limited capacity in signal processing and the need for data telemetry Clément (2019); Yoo and Shoaran (2021). We envision that effective, responsive neuromodulation should be computationally self-sufficient for active on-chip learning, eliminating the need for regular

telemetry. Recent advancements in neuromorphic computing are critical to realising our vision. We argue that neuromorphic computing in combination with highly low-power microelectronics for sensing Denison et al. (2007); Qian et al. (2011) and stimulation Ker et al. (2011) will enable the emergence of a neuromorphic neuromodulation device as a long-term solution for intractable neurological diseases.

Chapter 3

Biologically Plausible Liquid Spiking Neurons for EEG-Based Seizure Detection

This chapter introduces the first implementation of liquid time-constant spiking neural networks for epilepsy, evaluating their feasibility for out-of-sample EEG seizure detection in resource-constrained scenarios.

The content presented in this chapter is published as:

- Herbozo Contreras, L. F., Huang, Z., Yu, L., Nikpour, A. and Kavehei, O. (2024). "Biological plausible algorithm for seizure detection: Toward AI-enabled electroceuticals at the edge." *APL Machine Learning*, 2(2). DOI:10.1063/5.0192875.

Statement of Contributions of Joint Authorship

- Luis Fernando Herbozo Contreras (Candidate): First author, formal analysis, investigation, methodology, resources, software, visualisation, validation, writing original draft, writing review & editing of the manuscript
- Zhaojing Huang: contributed to writing review & editing of the manuscript
- Leping Yu: contributed to writing review & editing of the manuscript
- Armin Nikpour (Co-Supervisor): conceptualisation, data curation, funding acquisition, investigation, project administration, resources, supervision, validation, writing review & editing
- Omid Kavehei (Principal Supervisor): conceptualisation, data curation, funding acquisition, investigation, methodology, project administration, resources, supervision, visualisation, validation, writing review & editing.

In addition to the statements above, in cases where I am not the corresponding author of a published item, permission to include the published material has been granted by the corresponding author.

Luis Fernando Herbozo Contreras Date: 28 May 2026

As supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

Professor Omid Kavehei Date: 28 May 2026

Chapter Summary

EEG based diagnostics and monitoring tools, such as scalp-EEG, subscalp-EEG, stereo-EEG, or sub/epi-dural EEG recordings (aka. electrocorticography, ECoG), are used widely in different settings as the gold standard techniques to perform seizure identification, localisation and more primarily in epilepsy, or suspected epilepsy in patients. Techniques like subscalp EEG and ECoG offer long-term brain interaction, potentially replacing traditional electroceuticals with smart closed-loop therapies. However, these systems require continuous on-device training due to real-time demands and high power consumption. Inspired by brain architecture, biologically plausible algorithms, such as some neuromorphic computing, show promise in addressing these challenges. In this chapter, we utilised Liquid Time-Constant Spiking Neural Network (LTC-SNN) with FPTT to detect seizures in scalp-EEG. We trained and validated our model on the TUH dataset and tested its generalisation on out-of-sample data from the RPAH and EPILEPSIAE datasets. Our model achieved high AUROC scores of 0.83 in both datasets. We assessed robustness by reducing the memory size by 90% and obtained overall AUROC values of 0.82 in the RPAH dataset and 0.83 in the EPILEPSIAE dataset. Our model achieved outstanding results of 3.1 μJ per inference and a 20% firing rate during training. This enables the incorporation of bio-inspired, efficient algorithms for future on-device training, addressing challenges such as memory, power consumption, and efficiency.

3.1 Introduction

Closed-loop intervention has shown promise in promptly identifying anomalies and providing real-time responses, such as employing interventions like Spinal Cord Stimulation (SCS) Mekhail et al. (2023) and epilepsy management Mortazavi et al. (2021); Scheid et al. (2022). Recently, Saluda Medical conducted the first and only double-blind, FDA-approved study for SCS comparing closed-loop vs open-loop modalities. They showed that closed-loop intervention outperforms open-loop intervention Levy

et al. (2019); Mekhail et al. (2020). RNS continuously monitors the ECoG at the seizure focus and delivers closed-loop electrical stimulation when abnormal (epileptiform) patterns are detected. It can record up to 4 minutes of data, repeated up to 4 times daily, resulting in fragmented snapshots rather than a complete day's information. This leads to the loss of crucial data. The system relies on passive recordings of seizures to develop patient-specific detection algorithms for responsive stimulation, utilising criteria such as line length, half-wave area, and peak-to-peak amplitude. However, recent advancements in deep learning models have shown superior performance in detecting seizure onset compared to these algorithms. This raises an important question: *Is it possible to adapt algorithms directly on the chip?*

Artificial intelligence has made significant strides in healthcare, ushering in a new era of innovation and patient care. Its impact on the healthcare industry has been overwhelmingly positive, with numerous benefits spanning various aspects of medical practice, research, and patient outcomes Amann et al. (2020); Davenport and Kalakota (2019); Yu et al. (2018). Deep neural networks are the foundation of cutting-edge artificial intelligence, relying on computing systems ranging from transistors to memory-intensive GPU. These systems consume significant amounts of energy in their conventional computing architectures. Training deep neural networks on energy-intensive servers yields impressive accuracy but also high energy consumption. However, this elevated power usage makes them unsuitable for bio-electronic medical applications where low energy consumption is a priority. To achieve energy-efficient machine intelligence, there is potential in developing specialised hardware that emulates the structure of biological brains, enabling more localised cognitive capabilities for processing, identifying, and responding to personalised features. An alternative approach is offered by SNN, which mimics the behaviour of biological neurons and has the potential to offer energy-efficiency advantages. This makes them well-suited for environments with limited resources, such as edge devices Surianarayanan et al. (2023). However, SNN comes with a trade-off. It can suffer from reduced accuracy due to sparse information around the network, to reduce power consumption. This phenomenon is very prevalent in neuromorphic domains, especially in models using

Leaky Integrate-and-Fire (LIF) Neurons, which lack a dynamic time constant. Therefore, an exploration of more smart SNN is needed. As SNN are a type of recurrent neural network, they can also suffer from vanishing gradients due to backpropagation and increased linear memory over timesteps.

3.1.1 Background

Table 3.1 offers a comprehensive survey of pertinent research works within the field, each intimately associated with our present investigation. In the study by Zhao et al. (2020) on BNN, a salient accomplishment is a commendable balance between computational efficiency and performance, as evidenced by an impressive 95.5% accuracy Zhao et al. (2020). However, the dataset used is defined by its narrow focus on a few patients. It raises concerns about how well the model's performance generalises to a broader range of patients and out-of-sample datasets. The intrinsic constraints associated with dataset diversity may contribute to overfitting in the outcomes. Conversely, despite their superior performance compared to SNN and BNN, deep neural networks often exhibit higher power consumption and significantly greater memory requirements. Consequently, these characteristics render them unsuitable for on-device deployment. A notable contrast is observed in SNN models tailored for seizure detection. These models manage to strike a balance between power consumption, memory size, and accuracy. For instance, Yang et al. (2023) achieved an impressive AUROC score of 92.7% ,89.0% and 81.1% for different datasets while simultaneously maintaining power consumption of 974.2 nJ Yang et al. (2023). The diagram depicted in Fig. 3.1 illustrates three distinct neural network architectures. Across all three network types, SNN offer an equilibrium between accuracy and energy efficiency. For the SNN, three distinct types of neurons were explored: Integrate-and-Fire (IF), LIF, and the LTC-SNN. Among these neuron types, LTC-SNN are time-continuous recurrent neural network models with stable, bounded behaviour, thereby improving performance on time-series tasks. Their low complexity allows for a better representation of the hidden states and adapting to changing conditions such as autonomous driving and medical time-series data Banerjee et al. (2009); Beveridge and Pereira (2022);

Chahine et al. (2023); Hasani et al. (2022, 2021, 2023); Lechner et al. (2020). The Integration of Liquid Time-Constant (LTC) with SNN outperformed others model performance concerning accuracy and model size trade-offs in different tasks Yin et al. (2023).

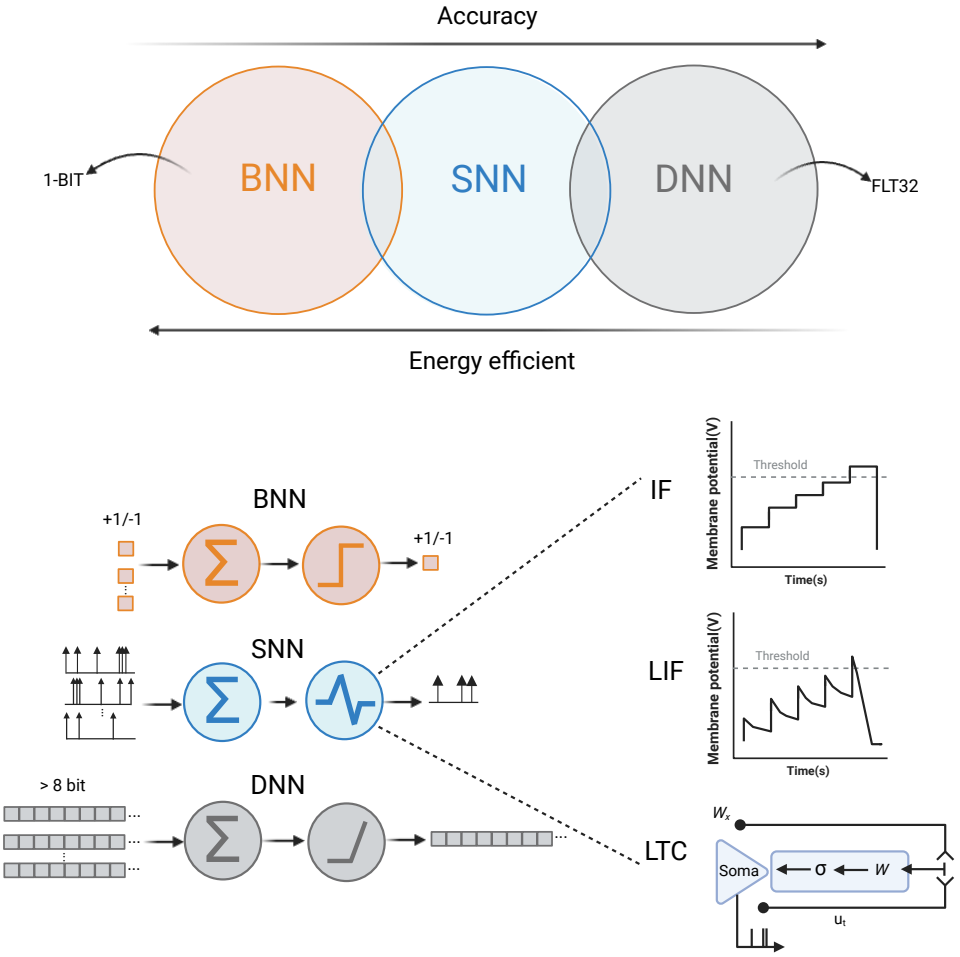

Figure 3.1 – Contrast between BNN, DNN, SNN. In the realm of neural networks, a spectrum of architectures exists, each with its unique characteristics. Thanks to their complex, layered structures, Deep Neural Network (DNN)s have long been recognised for their high accuracy across a wide range of tasks. On the other end of the spectrum, BNNs have attracted attention for their energy-efficient binary weights and activations, albeit often at the expense of some accuracy. Nestled intriguingly in between these extremes is SNNs. These networks mimic the communication patterns of neurons in the brain, enabling them to process information in discrete spikes. SNNs stand out for their unique ability to achieve a seamless balance between accuracy and energy efficiency.

3.1.2 Novelty and significance

Training Dynamic LTC-SNN with FPTT has been proposed to leverage on-device learning capabilities in a neuromodulation closed-loop system, where sending data outside is power-hungry and accuracy can degrade Herbozo Contreras et al. (2024b). Generalisation with SNNs is challenging, as even in-sample tests often show reduced accuracy, and many studies lack a large enough cohort to provide a meaningful assessment of the model's performance. Therefore, we proposed a pseudo-prospective study with 192 patients from the RPAH neurology dataset and 30 patients from the EPILEPSIAE dataset. We tested the lower complexity of our model on a reduced memory size of our first model, used by 90% in both datasets. We expect this contribution to be the first to incorporate biologically inspired models with an accurate neuromorphic encoding that enables on-device learning, thanks to their efficiency, low power, and memory efficiency for seizure detection at an implantable device level.

3.2 Methods

3.2.1 Datasets

Three datasets are used in this work: the TUH, the EPILEPSIAE, and the RPAH datasets. Fig. 3.4 summarises the TUH dataset used, providing key statistics about the data it contains. This dataset is divided into two categories: the first column represents the training set, and the second column represents the validation set, providing a comprehensive overview of the dataset's contents. It includes information such as the number of files, sessions, and patients in the dataset, as well as specific details about seizures, including the number of files, sessions, and patients with seizure data. The TUH dataset serves as the primary training dataset, providing a diverse range of seizure signals for various seizure types. With its large volume of files, the TUH dataset from America offers vast data for robust model training. The scalp-EPILEPSIAE and RPAH datasets of adult EEG are used for our inference tests. Furthermore, both datasets share common characteristics, such as identical

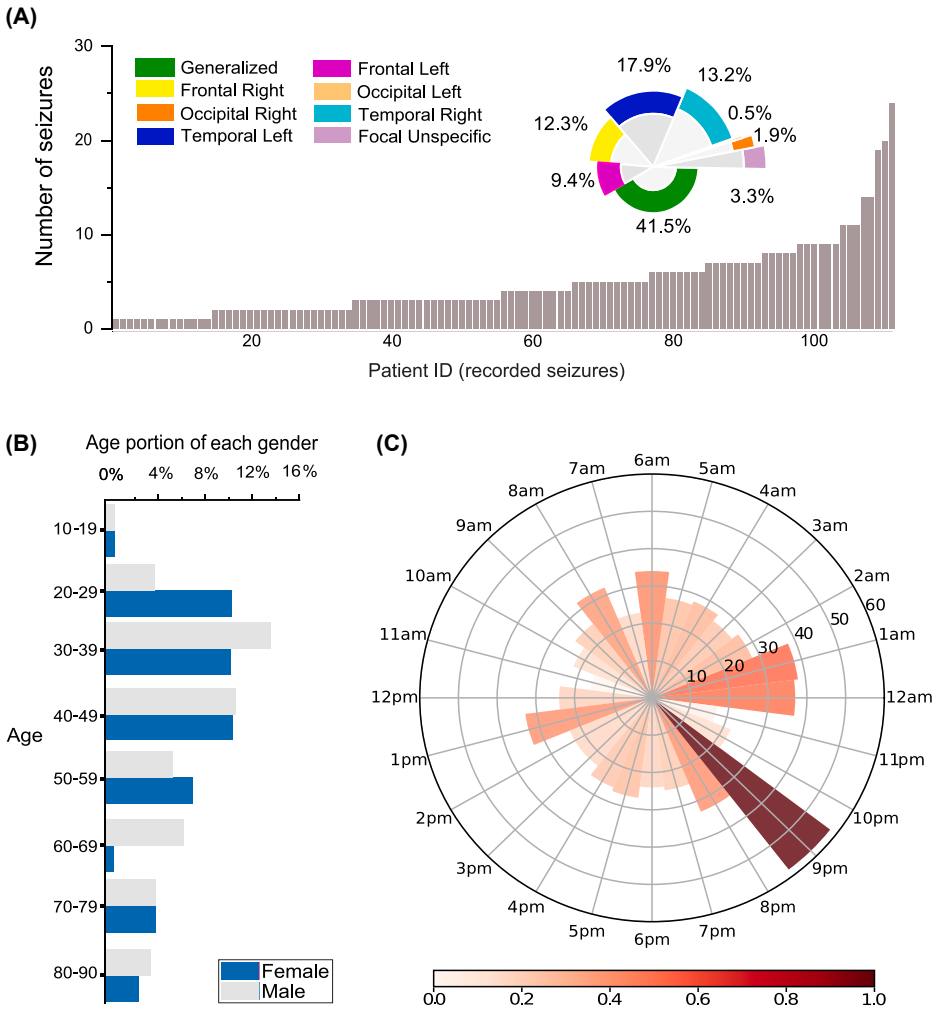

Figure 3.2 – Summary of the RPAH Dataset. (A) Distribution of seizure frequency and types in patients with confirmed seizures. (B) Age and gender distribution. (C) Seizure occurrence during a 24-hour cycle between 2011-2019 Yang et al. (2022b).

montages and the inclusion of adult patient data, ensuring consistency and comparability across analyses. Overall, integrating these complementary datasets enhances the comprehensiveness and reliability of the AI model for seizure detection. Fig. 3.2 visually represents the distribution of the RPAH datasets across three distinct domains: seizure type and frequency, age and gender, and seizure occurrence within a 24-hour cycle. It is worth noting that seizure detection is not continuous monitoring

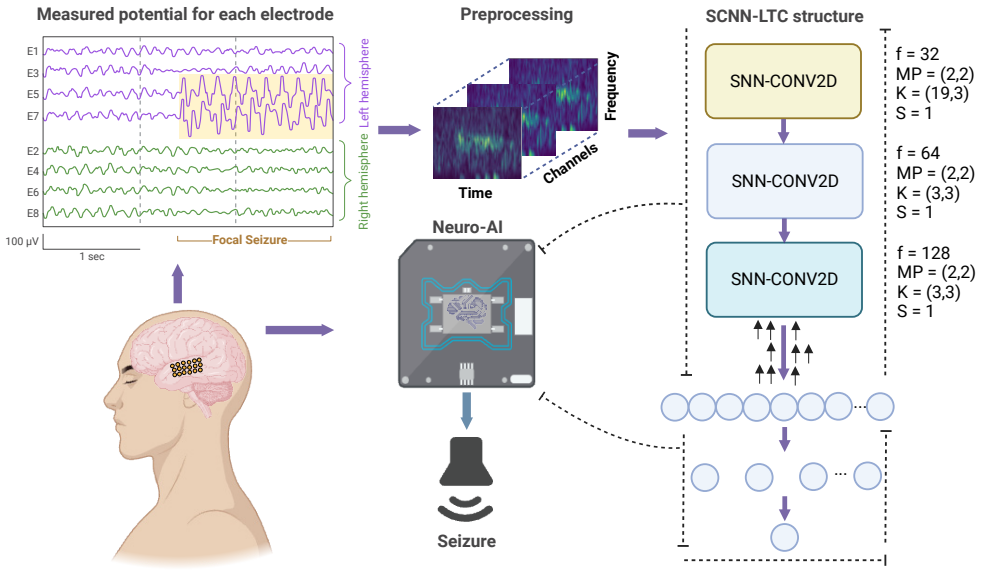

Figure 3.3 – Seizure detection system towards a low-power and accurate neuromorphic neuromodulation system. The dynamics of LTC demonstrate potential applications in these domains. While we analysed scalp-EEG datasets, consideration should be given to utilising ECoG for future studies. We decided to employ a convolutional-based model and selected parameters like Filters (f), Max Pooling (MP), Kernel (K), and Strides (S) based on their efficacy in detecting seizures.

but rather provides a clue as to where the seizure is most likely to occur. During the nocturnal hours, particularly between 8:00 PM and 9:00 PM, a heightened seizure frequency event is anticipated, contrasting with a more dispersed distribution during daylight hours. The RPAH structurally and reliably maintained data of its adult epilepsy patients who came from across Australia. In this work, we select 9 years (2011–2019) of data to test, comprising nearly 14,590 hours of EEG data from 192 patients, with a total of 1006 sessions, each with an average recording length of approximately 15 hours. The number of patients in the data was 212, of whom we excluded 20 for different reasons, including too many seizures (measured in more than 11 seizures/24 hours) (5 patients), missing electrode data (14 patients), or seizures that could only be confirmed by video information (1 patient). The Australian dataset is about 16 times larger than the US training dataset, with much longer interictal periods (between seizures), which is much more realistic and makes our false negatives much more reliable. The RPAH is one of Australia's major hospitals, with one of the

longest, if not the longest, EEG recordings in the country through its Comprehensive Epilepsy Services. We obtained ethical approval to facilitate our access to this clinical data.

3.2.2 Pre-Processing

To address challenges with raw EEG data, we utilised two signal processing techniques: Independent Component Analysis (ICA) and the STFT. Initially, the EEG signals were divided into 12-second segments, and the ICA algorithm was applied to decompose the signals into 19 independent components using Blind Source Separation (BSS). ICA separates EEG signals into statistically independent components, represented in Equ. (3.1),

$$T \approx MA^T \ , \tag{3.1}$$

where T contains the EEG data, M contains the time information, and A contains the weights for topographic maps. We used Pearson's correlation to identify independent sources strongly associated with eye movement, detected from two EEG channels, 'FP1' and 'FP2'. These identified sources related to eye movement were removed from the independent components, yielding EEG signals free of eye movement artefacts. Subsequently, we applied the STFT to the cleaned EEG signals. This involved using a window length of 250 (equivalent to 1 second) with a 50% overlap and eliminating the Direct current (DC) component of the transform. As a result, the data had dimensions of $(N \times 23 \times 125)$, where N represents the number of electrodes, 23 represents the time index, and 125 represents the frequencies. The pre-processing pipeline is performed offline and is therefore not included in the energy calculations. This ensures that the data is appropriately prepared before being fed into the LTC-SNN model for further training/inference, as depicted in Fig. 3.3.

3.2.3 SNN structure

We implement a Dynamic LTC-SNN with FPTT to efficiently process temporal data. The SNN architecture consisted of three CNN layers, enabling us to capture complex spatial features in the input data. The liquid time constant is a crucial parameter that introduces dynamic behaviour to our neural network's processing, influencing how information flows through the network over time and mimicking the temporal dynamics observed in biological neural systems, especially pyramidal neurons, across different brain areas. By integrating CNN with this liquid time constant within the FPTT framework, our SNN could respond to temporal patterns in data, making it particularly well-suited for tasks involving dynamic sequences or time-varying inputs. The system structure is illustrated in Fig. 3.3.

Liquid Time Constant Spiking Neuron Structure

The LTC-SNN was modeled as an adaptive spiking neuron. Its time constants, governing factors like the membrane time constant (τ_m) and adaptation time constant (τ_a), aren't static but dynamically adapt based on internal variables such as membrane potential (u) and deviation (b). These time constants are determined differently based on the network type. Equ. (3.2) and Equ. (3.3) are used to integrate convolution and dense neural networks with LTC-SNN, respectively. The parameter $\sigma()$ is a necessary scaling function that ensures smooth changes during the learning process. Equ. (3.4) describes the parameter θ_t as the adaptive threshold, mimicking the more realistic nature of adapting spiking neural networks in the brain. Equ. (3.5) and Equ. (3.6) describe the membrane potential and spiking rate behavior, governed by (τ_m). Equ. (3.7) is defined as resetting the membrane potential. Overall, the LTC-SNN neuron is defined accordingly based on the following set of equations:

$$\tau_a \text{ update} : \rho^i = \exp\left(-\frac{dt}{\tau_a^i}\right) = \sigma\left(\text{Dense}_a[x_t, b_{t-1}^i]\right) \ , \tag{3.2}$$

$$\tau_m \text{ update} : \alpha^i = \exp\left(-\frac{dt}{\tau_m^i}\right) = \sigma\left(\text{Dense}_m[x_t, u_{t-1}^i]\right) \ , \tag{3.3}$$

$$\theta_t \text{ update} : b_t^i = \rho^i b_{t-1}^i + (1-\rho^i)s_{t-1}^i;\ \theta_t^i = 0.1 + 1.8b_t^i\ , \tag{3.4}$$

$$u_t \text{ update} : du^i = -u_{t-1}^i + x_t;\ u_t^i = \alpha^i u_{t-1}^i + (1-\alpha^i)du^i\ , \tag{3.5}$$

$$\text{spike } s_t : s_t^i = f_s(u_t^i, \theta^i)\ ,\ \text{and} \tag{3.6}$$

$$\text{resetting} : u_t^i = u_t^i(1-s_t^i) + u_{\text{rest}}s_t^i\ . \tag{3.7}$$

FPTT Training

Algorithm 3.1: Training LTC-SNN with FPTT

Input: Training Data $B = \{x_t, y_t\}$, Timesteps T, Epochs E
Input: Learning rate η and optimiser
W, and $\overline{W} = W$; `// Initialized Weights`
for *each* $e \in E$ **do**
  $u_t, s_t = u_{t_0}, s_{t_0}$; `// Neuron States Initialisation`
  $B = \{x_t, y_t\}$; `// Shuffling Data`
  for *each* $t \in T$ **do**
    $s_{h,t}, u_{h,t} = \hat{f}_s(x_{t-1}, [u_{h,t-1}, s_{h,t-1}] \| W)$; `// Update States`
    $\hat{y}_t = \hat{f}_s(s_{h,t}, [u_{o,t}, s_{o,t}] \| v)$; `// Target Predictions`
    $\ell_t(W) = \ell_{\text{BCE}}(y_t, \hat{y}_t)$; `// BCE Loss`
    $\ell_{\text{dync}}(W) = \ell_t(W) + \frac{\alpha}{2}\|W - \overline{W}_t - \frac{1}{2\alpha}\nabla\ell_{t-1}(W_t)\|^2$; `// Dynamic Loss`
    $W : W_{t+1} = W_t - \eta\nabla_W \ell(W)|_{W=W_t}$; `// Update Weights`
    $\overline{W} : \overline{W}_{t+1} = \frac{1}{2}(\overline{W}_t + W_{t+1}) - \frac{1}{2}\alpha\nabla\ell_t(W_{t+1})$; `// Update` $\overline{W}$

FPTT has been used to mitigate a notable problem encountered when training deep or shallow spiking neural networks with Backpropagation Through Time (BPTT). In particular, BPTT has a linearly increasing memory cost as a function of sequence length $\Omega(T)$ and suffers from vanishing or exploding back-propagating gradients, which limits its applicability on long-time sequences. Another advantage of using the FPTT algorithm is the ability to use a higher batch size. The batch size we used was around 512, meaning that the more data we put in a batch, the less power it takes to train in a given amount of time. FPTT allows efficient gradient update and

memory storage. A dynamic regularizer loss is critical for the algorithm to effectively train with LTC-SNN models, as described in the Algorithm 3.1.

Table 3.1 – Different types of neural networks applied to different bio-signals with applications such as detection and prediction.

Neural Network	Application‡	Pre-Processing+	Dataset◇	Metrics†	Power	Memory Usage
BNN	SP Zhao et al. (2020)	Raw data	CHB-MIT AES	AUC: 97.0% AUC: 91.5%	-	672 KB*
	SD Ghazali et al. (2022)	DWT	U-Bonn	ACC: 98.5%	-	-
	ARR-CL Wu et al. (2020b)	Fixed-Length Full-Length Bucket Padding	PhysioNet CinC	F1: 86.8%	-	260 KB
SNN	SD Yang et al. (2023)	STFT-ICA	iEEG-FB CHB-MIT EPILEPSIAE	AUC: 92.7% AUC: 89.0% AUC: 81.1%	671.8 nJ 974.2 nJ 1142 nJ	1.3 MB 1.3 MB 1.4 MB
	SD Burelo et al. (2022a)	Filtering baseline ADM	HFO-Zurich	ACC: 80.0%	555.6 μW	-
	ARR-CL Yan et al. (2021)	Raw data	MIT-BIH	ACC: 79.0%	0.077 W	-
DNN	SD×, Yang et al. (2022b)	STFT-ICA	TUH EPILEPSIAE RPAH	AUC: 84.0% AUC: 81.0% AUC: 81.7%	283.3 μJ*	6 MB*
	SD Saab et al. (2020)	12s-window	Stanford-H LPCH	AUC: 93.0% AUC: 94.0%	-	-

‡ Applications: Seizure Prediction (SP), Seizure Detection (SD), Electrocardiogram Classification (ARR-CL).

+ Discrete Wavelet Transform (DWT), Asynchronous Delta Modulation (ADM), Short-Time Fourier Transform STFT, Independent Component Analysis ICA.

◇ Dataset: American Epilepsy Society Seizure Prediction Challenge (American Epilepsy Society Seizure Prediction Challenge (AES)), Children's Hospital Boston Children's Hospital Boston–MIT EEG dataset (CHB-MIT), University of Bonn (U-Bonn), Intracranial electroencephalography iEEG: Freiburg Hospital FB, European Epilepsy Dataset (EPILEPSIAE), Temple University Hospital TUH, Royal Prince Alfred Hospital RPAH, Lucile Packard Children's Hospital Lucile Packard Children's Hospital (LPCH), Stanford Hospital (Stanford-H).

† Metrics: F1 Score, Area Under the Curve (AUC or AUROC), and Accuracy (ACC).

* Represents an estimation as this information was not provided directly.

× In this Table, the work by Yang et al. (2022b) is the only study that provides an out-of-sample generalisation.

3.2.4 Implementation details

We employed Binary Cross-Entropy (BCE) as our loss function to evaluate the accuracy of our binary classification predictions. This loss function, calculated as BCE with logits, quantifies the dissimilarity between our model's predicted probability distribution and the actual binary labels (0 or 1) in our dataset, and it is dynamic and regularised.

3.2.5 Performance metrics

AUROC is used to evaluate model performance. It captures both the sensitivity and specificity in a threshold-free score. We perform out-of-sample testing with the RPAH and EPILEPSIAE datasets to test the model's generalisation robustness. As seizure detection in a neuromorphic-based algorithm aims to reduce energy usage for always-on monitoring, generalisation, and accurate detection, we also provide power-consumption benchmarks and memory reduction for a deep LTC-SNN.

3.2.6 Power Consumption

We utilised the Keras-spiking Python package v0.3.0 from Nengo to estimate the power consumption of our model in both GPU and Loihi [1]. The backend power profiling tool accounts for the cost of spike emission, spike routing on the many-core architecture, memory access, state updates, etc. The following assumptions are made when estimating the energy used by the proposed model on a particular device:

- The total energy is calculated by the sum of the energy used on the synaptic operation and neuron update.
- The energy consumption on a specific device is calculated based on the published data for GPU Degnan et al. (2015) and for Loihi Davies and other authors (2018).

[1] KerasSpiking: A Keras-based library for spiking neural networks

- Though additional overhead (e.g., on-device data transfer) can impact the energy usage of a model in practice, only internal model computations are included in the calculation.

3.3 Results

3.3.1 Training and Validation

Following the methodology outlined earlier, our model was trained and validated using the TUH dataset. We successfully attained an impressive AUROC score of 0.89. Our findings, as illustrated in Fig. 3.5, highlight the superior performance of our spiking model compared to non-spiking models. Notably, our model's performance rivals that of contemporary approaches.

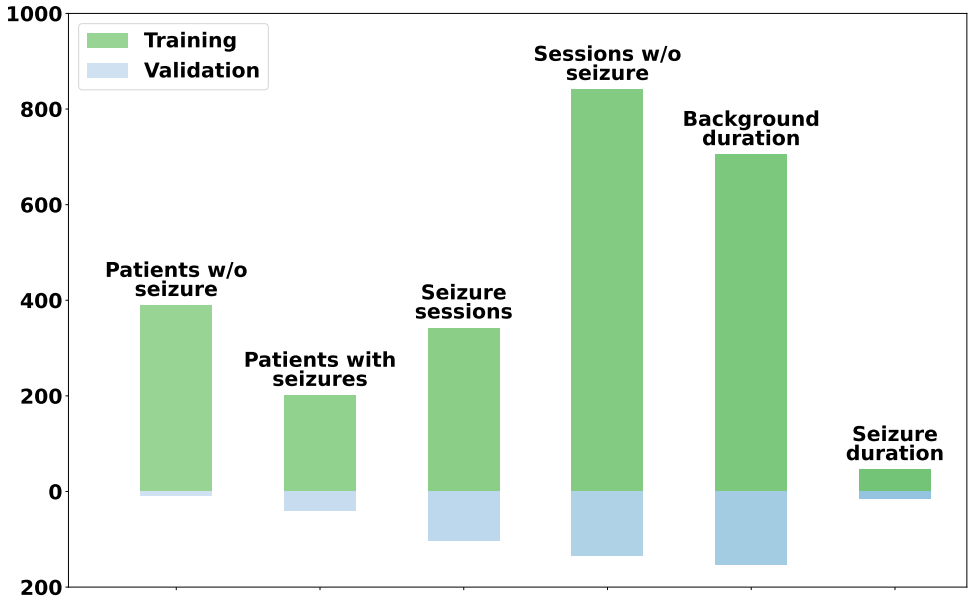

Figure 3.4 – Summary of the TUH dataset for both training and validation.

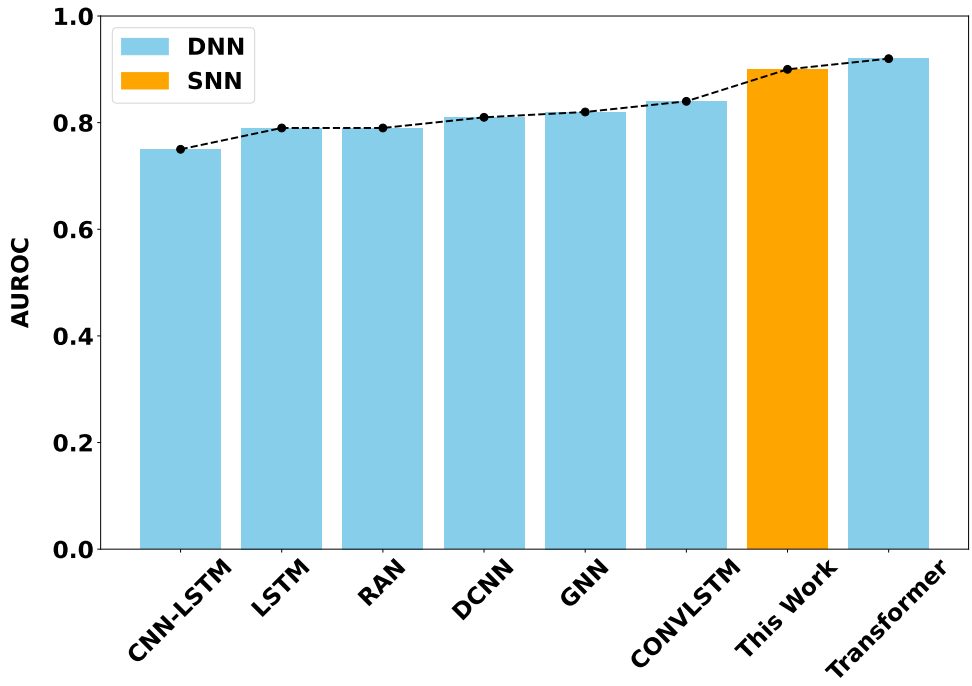

Figure 3.5 – State-of-the-art comparison on TUH dataset against our Liquid Spiking Neural Model. Every baseline was assessed with a window size of 12s and used either time, frequency, or time-frequency analysis Ahmedt-Aristizabal et al. (2020); Ma et al. (2023); Tang et al. (2022); Wang et al. (2017); Yang et al. (2022b).

3.3.2 Spiking vs Non-Spiking Generalisation

A fundamental question in evaluating our model is: *Can we effectively generalise its performance to datasets beyond those used for training?*. Our initial experiment compares our spiking model with a non-spiking model to address this. To assess generalisation, we initially used the RPAH dataset comprising 192 patients from 2011 to 2019 (see Table 3.2 for details). Our comprehensive LTC-FPTT model (40 MB) achieved an overall AUROC of 0.83, whereas the non-spiking version ConvLSTM model attained an AUROC of 0.82. Detailed year-wise results are available in Table 3.3.

Table 3.2 – Details of RPAH dataset for 192 patients.

Year	**Total duration** (hours)	**Number of sessions**	**Number of patients**	**Number of patients with seizure**
2011	1114.69	75	11	9
2012	1752.62	117	23	14
2013	2090.99	118	22	9
2014	1792.13	111	24	12
2015	2075.35	139	28	16
2016	1506.03	101	19	14
2017	2171.21	174	24	16
2018	1181.05	100	20	15
2019	907.60	71	21	6
Overall	**14591.6**	**1006**	**192**	**111**

Table 3.3 – Generalisation of our LTC-SNN-FPTT against non-SNN, ConvLSTM, in the RPAH dataset. Despite their differing architectures, LTC-SNN-FPTT is at the level of a non-SNN.

Year	**Spiking Model**	**non-Spiking Model** Yang et al. (2022b)
2011	0.88	0.90
2012	0.92	0.91
2013	0.92	0.89
2014	0.78	0.73
2015	0.79	0.82
2016	0.81	0.85
2017	0.80	0.68
2018	0.80	0.72
2019	0.80	0.78
Overall	**0.83**	**0.82**

3.3.3 Less is Effective: Memory Reduction Leads To Efficiency Even In Generalisation.

Another notable question is: *Can we reduce our model's memory usage without sacrificing generalisation performance?*. Initially, we employed a 40 MB model for generalisation in the RPAH dataset. However, we made a noteworthy finding regarding the robustness of the LTC-SNN that *is less effective*. We reduce the number of hid-

den neurons by 80% and, consequently, shrink the entire model's memory size by 90% (to 4 MB). Results depicted in Fig. 3.6 show that the performance of the small model stands at 0.82. This demonstrates only a 1% drop in performance compared to a larger model. These findings underscore the robustness of our model, even at a reduced scale.

Furthermore, we reduced the model's size by 95% and 98% relative to its original size on the RPAH dataset, as depicted in Fig. 3.7. This enabled us to understand the effects of even minor reductions in model capacity. The graph shows that even with a reduced model size of 1 MB, performance decreases by approximately 10-15%.

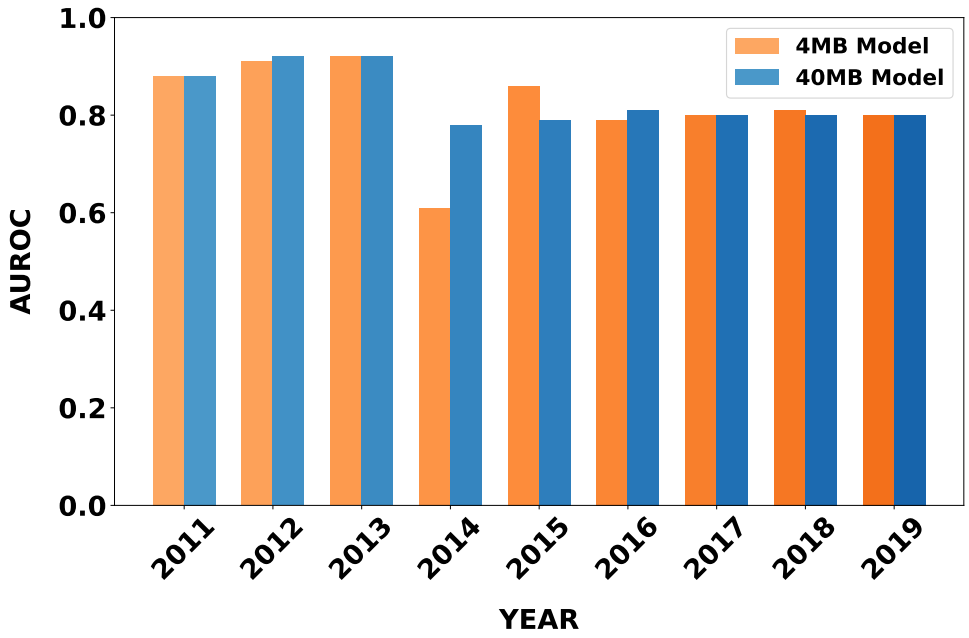

Figure 3.6 – Generalisation performance of large vs small LTC-SNN-FPTT models. This demonstration highlights the effectiveness present on a 4 MB model, despite being reduced by 90% of its current size for 192 patients.

To further support this finding, we conducted additional testing on the EPILEPSIAE dataset with the small LTC-SNN-FPTT model. The overall AUROC was 0.83, indicating a consistently positive trend across patients. Detailed results can be observed in Fig. 3.8

In Fig. 3.9, the detection of seizures on the EPILEPSIAE Scalp-EEG from different seizure foci based on patient ID is presented. The figure illustrates distinct brain

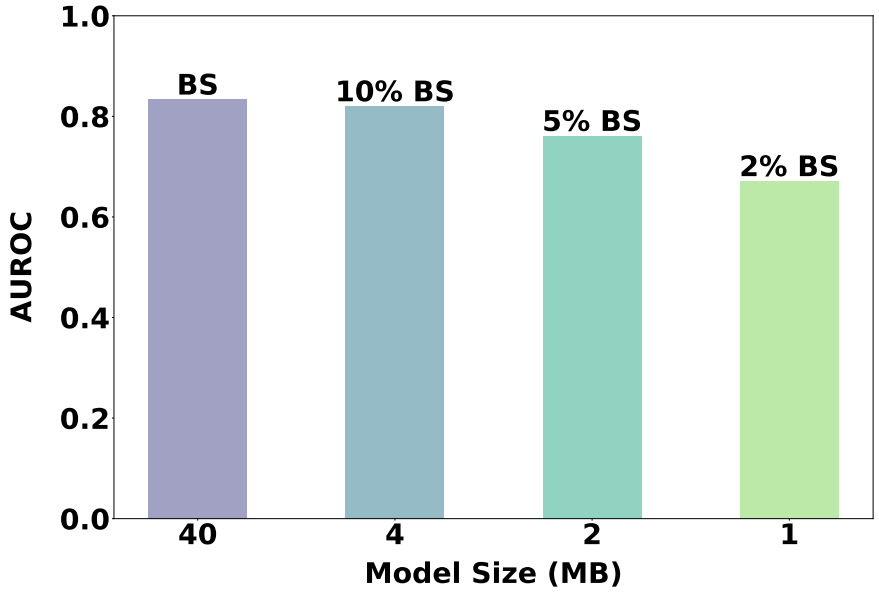

Figure 3.7 – **Performance on the RPAH dataset by memory reduction** of the model size achieved through reduction in the number of hidden neurons. Percentage values represent the total memory compared to the Baseline (BS) model.

regions, with a notable concentration of activity in the temporal lobe, where seizures primarily originate within the dataset. These findings underscore the model's ability to discriminate a wide range of seizures across the brain associated with seizure onset.

3.3.4 Power Consumption of LTC-SNN-FPTT model

We estimate 3.1 μJ per inference for power consumption. Detailed consumption by the overall model is established in Table 3.4. We also obtained the firing rate of our model during training, with only 20% of the total information spiking for accurate detection. This highlights energy-efficient behaviour in the seizure model. The model consists of 400,000 parameters for 100 hidden neuron models and 3.5 million for 500 hidden neurons, with memory sizes of 4 MB and 40 MB, respectively.

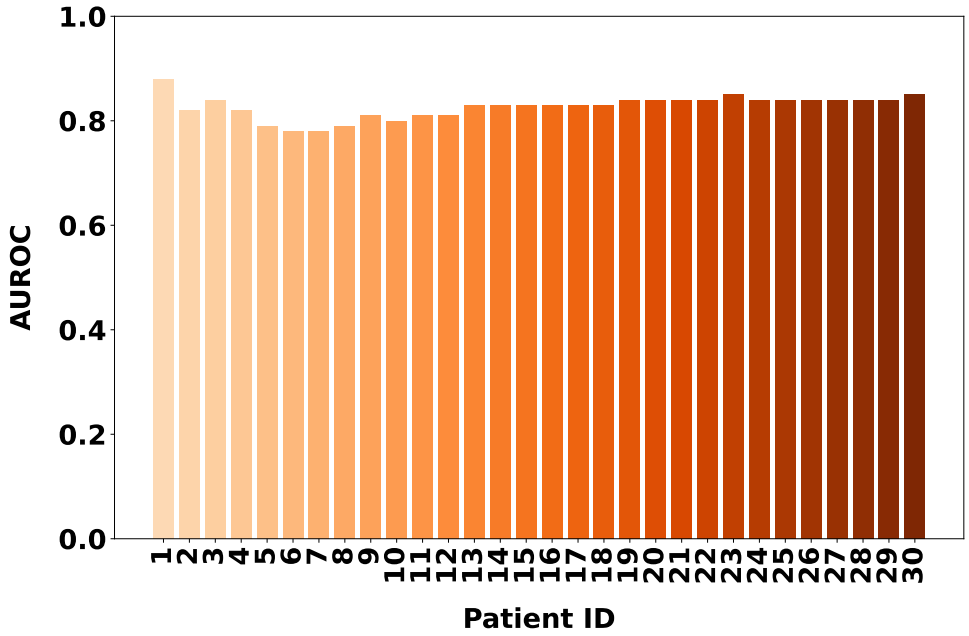

Figure 3.8 – EPILEPSIAE Scalp-EEG results with the small LTC-SNN showing robustness and across generalisation.

Table 3.4 – Power consumption of the LTC-SNN-FPTT model detailed by each layer. GPU is shown in μJ/Inf. and Loihi in nJ/Inf.

Layer (type)	μ**J/Inf.** (GPU)	n**J/Inf.** (Loihi)
CNN2D-LTC	3400	3100
CNN2D-LTC	28	14
CNN2D-LTC	3.8	1.3
Recurrent Neural Network (RNN)-LTC-DENSE	1.5	0.5
RNN-LTC-DENSE	1.5	0.5
Total energy/Inf.	**3420**	**3100**

3.3.5 AI-SNN Feature Extraction

We provided a feature extractor for our detection model to differentiate our work from most studies that are based solely on numerical representations and lack AI interpretability. This helped us improve the accuracy. We wanted to assess whether the model accurately learns the intended features we input into the STFT arrays. We expected to see activation and spiking at lower frequencies in a seizure file. We extract the encodings of one of the convolution layers for feature extraction. In this

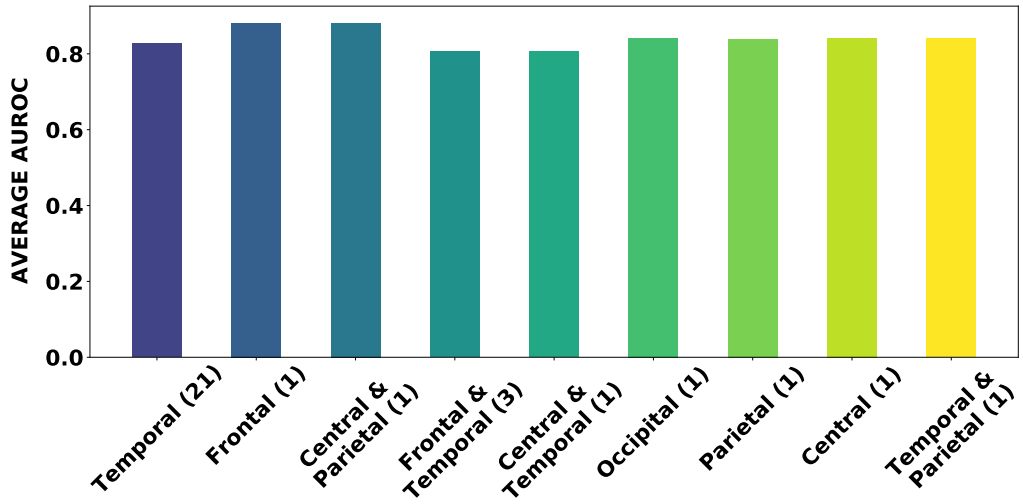

Figure 3.9 – EPILEPSIAE Scalp-EEG results across brain regions in patients with varying seizure foci. Numbers in the brain regions axis denote patient count for each seizure focus, independent of seizure frequency. While the dataset may be limited in certain seizure foci, it is essential to emphasise the efficacy of the model.

case, we use the first convolution layer represented in Fig. 3.10. The parameter fc1 represents the dimensions of the filters of the first convolutional layer. h stands for the height we select as a representation for the time bins, and w stands for the width as a part of the frequency domain. (A) represents the membrane potential activity in a random channel of the EEG data, and (B) represents the spiking activation in the same channel. It is clear how accurate this model can be, as its activity falls within the lower boundaries of the frequency filters. With this representation, we can confirm that LTC-SNN models are robust and can avoid shortcut learning Geirhos et al. (2020). It is also worth noting that the spiking rate is effective, yielding an accurate, low-power algorithmic solution. Less is better.

3.4 Discussion

Our next step is to incorporate this biologically plausible algorithm for seizure prediction. We foresee that this algorithm will be robust against catastrophic forgetting, as it does not rely on back-propagation, which is known to be prone to this

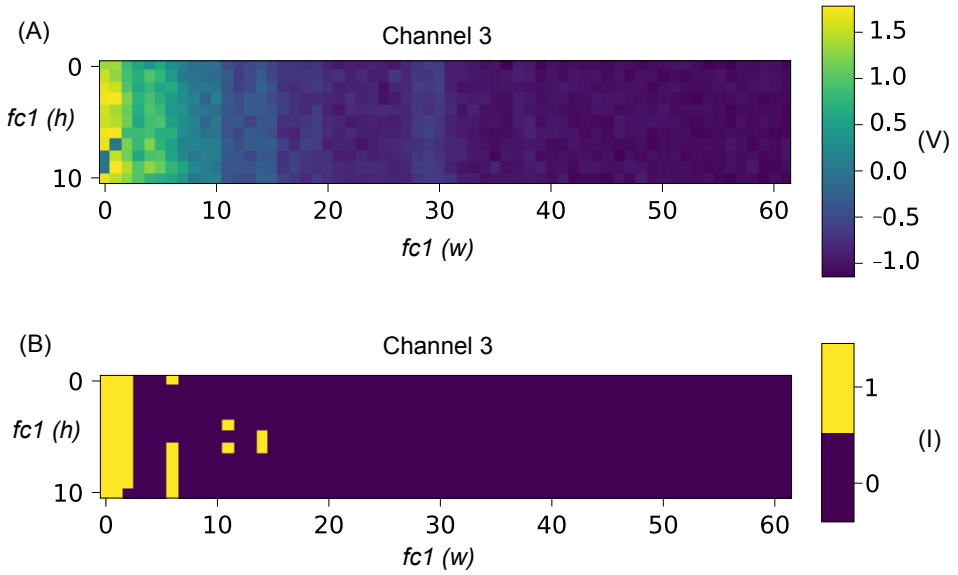

Figure 3.10 – AI-Interpretability of our model. The neuromorphic encoding within the primary convolutional layer is extracted, followed by a random channel selection from the TUH dataset. Membrane potentials are extracted from this channel (A). The membrane potential follows firing activation in the same layer (B).

phenomenon Schak and Gepperth (2019). BPTT updates RNN parameters on an instance by back-propagating the error in time over the entire sequence length, leading to poor trainability due to the well-known gradient explosion/decay phenomena. Empirically, FPTT outperforms BPTT on several well-known benchmark tasks like MNIST, DVS-GESTURE Yin et al. (2023), and now generalisation on EEG data.

Although the proposed LTC-SNN is motivated by neuromorphic temporal processing, the model is trained and evaluated in this chapter on conventional digital hardware. Therefore, the contribution of this chapter is primarily algorithmic, demonstrating the benefit of learnable temporal dynamics for seizure detection rather than direct deployment on a neuromorphic chip. Mapping such liquid time-constant dynamics to current neuromorphic platforms would require either hardware support for heterogeneous/adaptive neuronal time constants which can be costly or an approximation using simpler supported neuron models.

3.5 Conclusion

In this chapter, we combined the advantages of a biologically plausible algorithm with feature extraction techniques that exhibit exceptional performance within the confines of the training dataset and generalise to new, unseen data. Its robustness is evident in its successful application to tasks such as detection within EEG data. Moreover, the algorithm's strong performance extends to pseudo-prospective data, as demonstrated by its effectiveness in RPAH and EPILEPSIAE. One of the most notable features of this algorithm is its ultra-low power consumption, estimated to be on the order of μJ per inference. This attribute is critical in contemporary computational applications, emphasising the algorithm's potential to operate efficiently in resource-constrained environments. Furthermore, its notable memory efficiency enables additional compression without compromising accuracy, providing a significant advantage for practical implementation and deployment. The promising results and the algorithm's adaptability position it as a viable and effective solution for real-world applications, particularly in EEG data analysis and detection tasks. Further exploration and implementation of this algorithm hold promise for addressing complex computational challenges while prioritising energy efficiency and maintaining high-performance standards.

Chapter 4

Time Domain Multiscale Processing with Dendritic Spiking Networks

This chapter introduces compartmental dendrite spiking models for seizure detection, evaluating whether biologically motivated temporal processing can improve model compactness and energy efficiency across time-domain EEG and iEEG datasets.

The content presented in this chapter is published as:

- Herbozo Contreras, L. F., Yu, L., Huang, Z, Zhang, Z., Nikpour, A. and Kavehei, O. (2024).“Tiny dLIF: a dendritic spiking neural network enabling a time-domain energy-efficient seizure detection system.” *Neuromorphic Computing and Engineering*, 5(1). DOI:10.1088/2634-4386/adc0b9.

Statement of Contributions of Joint Authorship

- Luis Fernando Herbozo Contreras (Candidate): First author, formal analysis, investigation, methodology, resources, software, visualization, validation, writing original draft, writing review & editing of the manuscript
- Leping Yu: contributed to writing review & editing of the manuscript
- Zhaojing Huang: contributed to writing review & editing of the manuscript
- Ziyao Zhang: contributed to writing review & editing of the manuscript
- Armin Nikpour (Co-Supervisor): conceptualisation, data curation, funding acquisition, investigation, project administration, resources, supervision, validation, writing review & editing
- Omid Kavehei (Principal Supervisor): conceptualisation, data curation, funding acquisition, investigation, methodology, project administration, resources, supervision, visualisation, validation, writing review & editing.

In addition to the statements above, in cases where I am not the corresponding author of a published item, permission to include the published material has been granted by the corresponding author.

Luis Fernando Herbozo Contreras Date: 28 May 2026

As supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

Prof. Omid Kavehei Date: 28 May 2026

Chapter Summary

Reliable seizure detection, localisation and modulation from (EEG) recordings require models capable of capturing multiscale temporal dynamics across heterogeneous recording modalities, including scalp-EEG and iEEG. However, these techniques often rely on feature extraction methods, such as STFT, to improve efficiency in seizure detection. Drawing inspiration from brain architecture, we investigate biologically plausible algorithms, specifically emphasising time-domain inputs with low computational overhead. Our novel approach features two hidden-layer dendrites with Dendritic Leaky Integrate-and-Fire (dLIF) spiking neurons, containing fewer than 300K parameters and occupying a mere 1.5 MB of memory. Our proposed network is tested and successfully generalised on four datasets from the USA and Europe, recorded with different front-end electronics. USA datasets are scalp-EEG in adults and children, and European datasets are iEEG in adults. All datasets are from patients living with epilepsy. Our model exhibits robust performance across different datasets through rigorous training and validation. We achieved AUROC scores of 81.0% and 91.0% in two datasets. Additionally, we obtained Area Under the Precision–Recall Curve (AUPRC) and F1 Score metrics of 91.9% and 88.9% for one dataset, respectively. We also conducted out-of-sample generalisation by training on adult patient data, and testing on children's data, achieving an AUROC of 75.1% for epilepsy detection. This highlights its effectiveness across continental datasets with diverse brain modalities, regardless of montage or age specificity. It underscores the importance of embracing system heterogeneity to enhance efficiency, thus eliminating the need for computationally expensive feature engineering techniques like Fast Fourier Transform FFT and STFT.

4.1 Introduction

Researchers are actively developing AI models for seizure detection. For instance, studies have combined ICA and STFT for pre-processing, followed by ConvLSTM

blocks, achieving an AUROC of 0.84 in the largest epilepsy dataset in United States Yang et al. (2022b). Similarly, by utilising domain adaptation and STFT pre-processing, authors have reported an AUROC of 0.75 ± 0.03 Peng et al. (2022). Moreover, a study has introduced anchored-STFT to enhance temporal-spectral resolution trade-off in EEG analysis Ali et al. (2022). Despite their strong performance, these models often demand significant memory and power resources, making them impractical for deployment on edge devices. Neuromorphic AI, inspired by the brain, differs significantly from the traditional von Neumann architecture, which utilises separated memory and processing units, consuming power during data transmission. Instead, a neuromorphic chip employs physical neurons interconnected by physical synapses, implementing co-located memory processing in a non-volatile manner. This design drastically reduces the necessity to move data across the circuit, leading to substantial improvements in speed and energy efficiency Marković et al. (2020). Seizure detection has advanced with the integration of SNN and neuromorphic hardware, which mimic the brain's computational methods for real-time, low-power processing. It has demonstrated effectiveness in handling both scalp Yang et al. (2023) and intracranial EEG data Costa et al. (2024). For instance, a Spiking Convolutional Long Short-Term Memory (SConvLSTM) demonstrated high AUROC scores in seizure detection, experiencing only a 10% drop in performance while offering nearly 100–1000x power savings compared to traditional methods Yang et al. (2023). Seizure Prediction with neuromorphic computing has been achieved using an Convolutional Spiking Neural Network (CSNN), achieving an AUROC of 0.95 and reducing complexity by 98.58% compared to a traditional neural network Tian et al. (2021). Remarkably, energy-efficient seizure classification has been achieved with CSNNs, with power consumption as low as 1.28 µJ per classification, making them suitable for wearable devices Muneeb and Kassiri (2023). A CMOS-based neuromorphic system have successfully captured different brain modalities like LFP and biomarkers such as High-Frequency Oscillation (HFO), for detecting ictal events under 50 pW Burelo et al. (2022b); Ronchini et al. (2021). Mixed-signal neuromorphic processors have further demonstrated real-time seizure detection and provided new insights by encoding these events Bartels et al. (2025); Gallou et al. (2024). However, most of these studies are limited to

point-neuron models, and more biologically realistic models have yet to be thoroughly explored for seizure detection.

4.1.1 Bio-Inspired Neural Networks

Studies have been moving toward models that extend beyond merely mimicking somas. Neuron models incorporating dendritic computations are under active research, and neuromorphic hardware also integrates dendritic processing capabilities. For example, stochastic dendrites have been implemented in a spiking neural network, facilitating online learning through dendritic computations Cartiglia et al. (2022). Dendritic mechanisms operate across multiple timescales, enabling complex computations such as coincidence detection, filtering, input segregation, nonlinear processing, and logical operations Ariav et al. (2003); Mel and Schiller (2004). Therefore, dendrites are essential for accurately modeling neuronal integration and output at the single-cell level, contributing significantly to the computational power of neural networks. Simulators have been proposed that enable dendrites to operate semi-independently of the soma and perform complex functions, enhancing the computational capabilities of the model Pagkalos et al. (2023), but they are not scaled to SNNs.

4.1.2 The proposed system

We introduce a tiny dLIF SNN that embraces dendritic computations by representing them as small RC circuits to provide heterogeneity to the network Zheng et al. (2024). This integration improves computational efficiency and balances performance. The model is based on two hidden layers, with dendritic inputs that act as rich inputs, considering how biological neurons receive, process, and transmit information. The simplicity of this model allows for approximately 1.5 MB of memory utilisation with less than 300K parameters. The biological-inspired neuron model is described in Fig. 4.1.

4.1.3 Novelty and significance

This work leverages the potentials of dLIF models for seizure detection, which enables:

1. Direct analysis in the time domain without relying on traditional time-frequency transformation techniques like STFT or FFT. This approach simplifies the pre-processing analysis, potentially reducing power and hardware requirements.
2. Tiny model by its minimal requirement of only two hidden layers, enhancing its efficiency below 300K parameters.
3. Down-sampling scalp-EEG signal (i.e 250 Hz to 125 Hz) yield marginal drop in performance, paving the way for rapid, energy-efficient training systems.
4. Efficacy across different brain activity recordings is demonstrated through rigorous testing using both scalp EEG and iEEG datasets.
5. Out-of-sample seizure detection by training with adult patients montages and testing in children montages.

4.2 Methods

4.2.1 Datasets

There are 4 datasets used in this work: TUH, CHB-MIT, FB, and the iEEG EPILEPSIAE dataset. The TUH dataset is divided into two categories: one set of training and one set of validation, offering a comprehensive overview of the dataset's contents. It includes information such as the number of patients with and without seizure sessions. It also shows the inter-ictal and ictal periods used (Background and Seizure duration). With its large volume of files, the TUH dataset from America offers vast data and served as a meaningful scalp-EEG dataset as it is the largest dataset worldwide. Most of these patients are also adults.

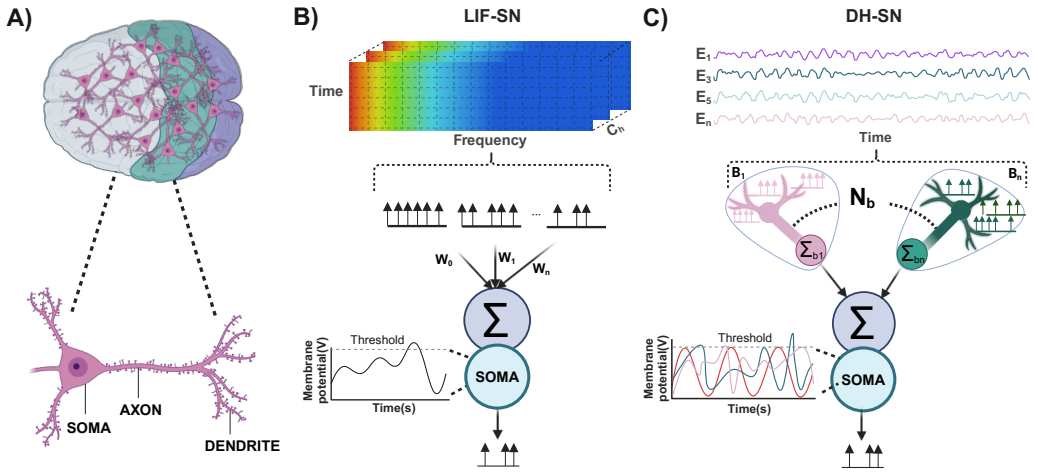

Figure 4.1 – Biological Inspired Neuron Model. The main components of the brain are neurons, which have different parts as the dendrites, soma, and axon (A). Different types of biological-inspired neural networks, like LIF Neurons, are built from the Soma and do not consider the dendrites as part of their system. Therefore, it minimises the heterogeneity of an actual brain. Those models often use Time-Frequency Analysis to enhance efficiency (B). However, this analysis can be computational and power costly; therefore, adding heterogeneity by incorporating dendrites in the LIF-Spiking Neurons improves time-domain series processing. $B_{1\text{-}n}$: Branches, N_b: Number of Branches, W_n = Synaptic Weights.

We performed out-of-sample generalisation in the CHB-MIT dataset for a non-montage and age-dependent systems. We randomly selected 17 patients. The TUH dataset contains 19 channels, while the CHB-MIT dataset consists of 22 channels. To ensure compatibility during inference, given that our model was trained on the TUH dataset, we aligned the input channels between both datasets. The TUH dataset includes the following montages: Fp1-F7, Fp2-F8, F7-T3, F8-T4, T3-T5, T4-T6, T5-O1, T6-O2, T3-C3, C4-T4, C3-Cz, Cz-C4, Fp1-F3, Fp2-F4, F3-C3, F4-C4, C3-P3, C4-P4, P3-O1, P4-O2. The CHB-MIT dataset comprises montages such as: 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, P7-T7, T7-Ft9, Ft9-Ft10, Ft10-T8. We selected the overlapping channels/montages between the two datasets and reordered the CHB-MIT dataset to align with the channel order of the TUH dataset. For the remaining unmatched montages, we masked them and randomly removed 3 channels to reduce the amount of channels in the CHB-MIT dataset. This was the last step

to ensure the same number of montages as the TUH dataset. We compared with previous studies that utilise different model structures, with time-frequency domain and in-sample testing.

We explored further how our model can recognise and train with iEEG signals. Thereby, we utilise the FB and EPILEPSIAE datasets. We trained and tested in-sample in 14 patients as these datasets do not possess the same montages and different amounts of channels across the EPILEPSIAE, varying from 30 to 114 electrodes, utilising a sampling rate of 256 Hz. The EPILEPSIAE dataset contains high-quality, long-term EEG, iEEG, and concurrently recorded ECG data. Intracranially implanted strips, grids, and/or stereotactically (stereo-EEG) implanted depth electrodes are used in patients with invasive recordings. The EPILEPSIAE dataset possessed very long interictal periods (background data between ictal events), which is a much more realistic iEEG dataset and makes our false negatives highly reliable. The FB dataset contains 21 intractable epilepsy patients with iEEG recordings. In this dataset, data was recorded from 6 selected electrodes, at a sampling rate of 256 Hz, where three are epileptogenic regions, and the rest are from other remote areas. 12 patients with 311.4 hours recording length are tested because of the availability of the dataset. A summary detailed of these datasets can be seen in Fig. 4.2.

4.2.2 Pre-Processing

We utilised ICA for scalp EEG artifacts removal. Initially, the EEG signals were split into 12-second segments, and the ICA algorithm was applied to decompose the signals into 19 independent components using BSS. ICA separates EEG signals into statistically independent components, represented in Equ. (4.1),

$$T \approx MA^T \tag{4.1}$$

where T contains the EEG data, M contains the time information, and A contains the weights for topographic maps. We used Pearson's correlation to identify independent

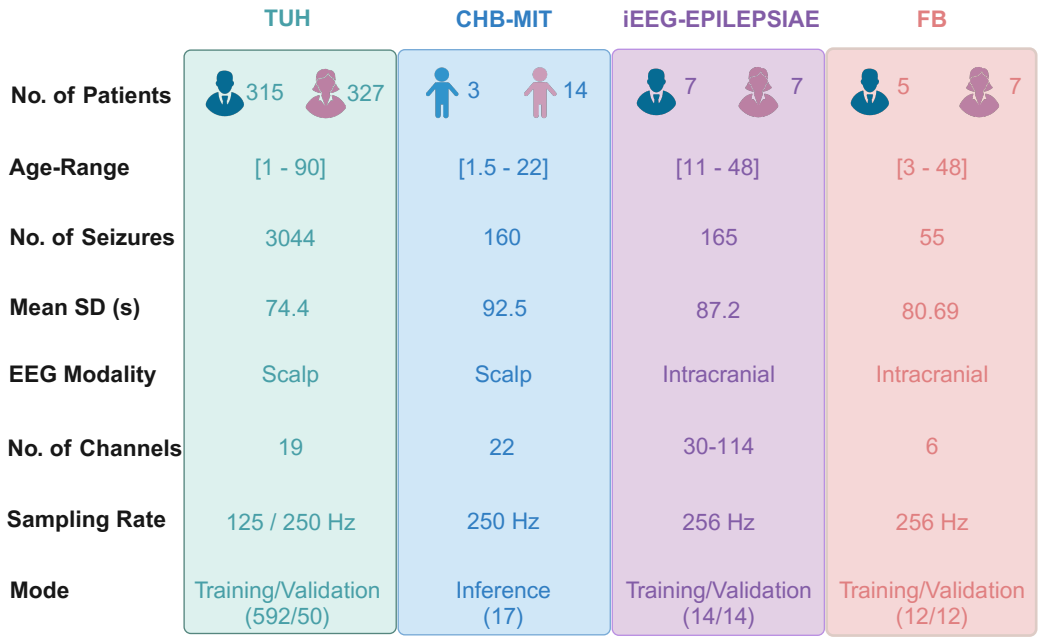

Figure 4.2 – **Summary of datasets** being employed in this work. Mean SD: Mean Seizure Duration.

sources strongly associated with eye movement, detected from two EEG channels, 'FP1' and 'FP2'. These identified sources related to eye movement were removed from the independent components, yielding EEG signals free of eye movement artefacts. Subsequently, we implemented power-line noise removal at 50 Hz for both the FB and iEEG-EPILEPSIAE datasets, employing a notch filter from the MNE Package. For the TUH and CHB-MIT datasets, we applied a notch filter at 60 Hz. This measure was imperative for detecting noise within the specified frequency range via power spectrum analysis. The pre-processing pipeline is performed offline and is therefore not included in the energy calculations.

4.2.3 SNN structure

dLIF Spiking Neuron Structure:

The dLIF neuron model enhances the classic LIF-based spiking neuron by introducing multi-timescale memory on dendrites, governed by Equ. (4.2), (4.3), (4.4) where u

represents the soma's membrane potential, β is the timing factor, R is membrane resistance, d is the dendritic branch index, and u_{th} is the firing threshold Zheng et al. (2024).

$$u^{t+1} = \beta u^t + (1-\beta)\sum_d R i_d^{t+1} - o^t u_{\text{th}} \tag{4.2}$$

$$i_d^{t+1} = \alpha_d i_d^t + (1-\alpha_d) I_d^{t+1} \tag{4.3}$$

$$o^{t+1} = H(u^{t+1} - u_{\text{th}}) \tag{4.4}$$

$$\alpha_d = sigmoid(\hat{\alpha}_d), \beta = sigmoid(\hat{\beta}) \tag{4.5}$$

The Heaviside function $H(\cdot)$ regulates spiking. Synaptic input on the d_{th} dendritic branch is the sum of feedforward and recurrent inputs as shown in Equ. (4.6), defined by sparse vectors W_d and U_d Zheng et al. (2024). To reduce the impact of negative timing factors in Equ. (4.5), the parameters α and β should be constrained between 0 and 1, which can be accomplished using a sigmoid function.

$$I_d^{t+1} = < W_d, X^{t+1} > + < U_d, o^t > \tag{4.6}$$

Extending this to dLIF SNN involves incorporating layer information as shown in Equs. (4.7), (4.8), (4.9), with synaptic currents on the d_{th} branch as shown in Equ. (4.10), determined by matrix forms of feed-forward and recurrent synaptic weights (W_d and U_d), which are again sparse, considering only valid synapses connected to dendritic branches in the next layer Zheng et al. (2024).

$$u^{t+1,l} = \beta \odot u^{t,l} + (1-\beta^l) \odot R \sum_d i_d^{t+1,l} - o^{t,l} u_{\text{th}} \tag{4.7}$$

$$i_d^{t+1,l} = \alpha_d^l \odot i_d^{t,l} + (1 - \alpha_d^l) \odot I_d^{t+1,l} \tag{4.8}$$

$$o^{t+1,l} = H(u^{t+1,l} - u_{\text{th}}) \tag{4.9}$$

$$I_d^{t+1,l} = W_d^l o^{t+1,l-1} + U_d^l o^{t,l} \tag{4.10}$$

Learning rule:

The BPTT algorithm has been adopted to better handle the unique characteristics of SNN. During training, model parameters such as synaptic weights (W) and timing factors ($\hat{\alpha}, \hat{\beta}$) are automatically learned to optimise network performance based on the loss function Zheng et al. (2024). The Equ. (4.11), (4.12), (4.13) shows the details of the BPTT where δ denotes the gradient of the loss function L concerning specific variables Zheng et al. (2024). The BPTT for dLIF SNN utilises gradient descent with the chain rule to update parameters while dealing with non-differentiable spiking activities using a soft multi-Gaussian curve shown in Equ. (4.14) as a surrogate gradient function, with parameters like γ, h which will influence the magnitude, σ, s which will influence the width Zheng et al. (2024).

$$\delta u^{t,l} = \beta^l \odot \delta u^{t+1,l} + H' \odot \delta o^{t,l} \tag{4.11}$$

$$\delta i_d^{t,l} = (1 - \beta^l) R \odot \delta u^{t,l} + \alpha_d^l \odot \delta i_d^{t+1,1} \tag{4.12}$$

$$\delta o^{t,l} = -u_{th} \delta u^{t+1,l} + \sum_d W_d^{l+1^T} (1 - \alpha_d^{l+1}) \odot \delta i_d^{t,l+1} + \sum_d U_d^{l^T} (1 - \alpha_d^l) \odot \delta i_d^{t+1,l} \tag{4.13}$$

$$H' = \frac{\delta o^t}{\delta u^t} = \gamma(1+h)N(u^t|U_{th},\sigma^2) - \gamma h N(u^t|\sigma,(s\sigma)^2) - \gamma h N(u^t|-\sigma,(s\sigma)^2) \quad (4.14)$$

Gradients of parameters are shown in Equ. (4.15), (4.16), (4.17) are computed using this surrogate gradient function, particularly around the firing threshold (u_{th}) where neurons emit spikes, enabling efficient training and improved performance of dLIF-SNNs Zheng et al. (2024).

$$\delta W_d^l = \sum_t (1-\alpha_d^l) \odot \delta i_d^{t,l} o^{t,l-1^T}, \delta U_d^l = \sum_t (1-\alpha_d^l) \odot \delta i_d^{t+1,l} o^{t,l^T} \quad (4.15)$$

$$\delta\hat{\beta}^l = \sum_t \delta\beta^{t,l} \odot (1-\delta\beta^{t,l}), \delta\beta^{t,l} = u^{t-1,l} \odot \delta u^{t,l} - R\sum_d i_d^{t,l} \odot \delta u^{t,l} \quad (4.16)$$

$$\delta\hat{\alpha}_d^l = \sum_t \delta\hat{\alpha}_d^{t,l} = \sum_t \alpha_d^{t,l} \odot (1-\delta\alpha_d^{t,l}), \alpha_d^{t,l} = i_d^{t-1,l} \odot \delta i_d^{t,l} - I_d^{t,l} \odot \delta i_d^{t,l} \quad (4.17)$$

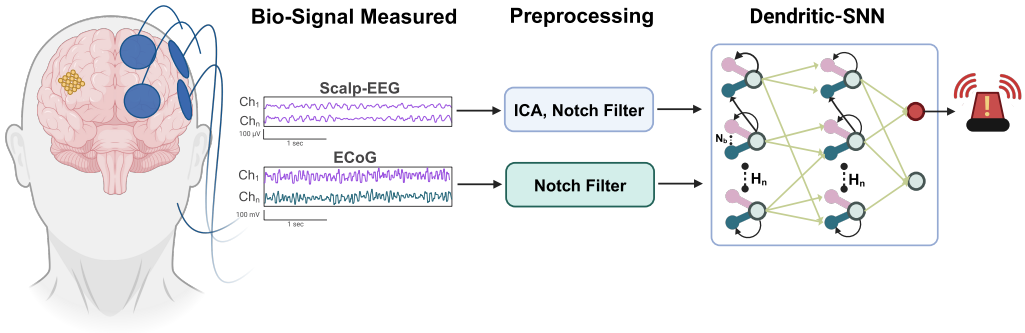

Figure 4.3 – Neuromorphic Detecting System using dLIF. Bio-signals such as scalp-EEG and iEEG are preprocessed using a notch filter for power-line noise removal or ICA to separate blind-source components and then fed into a single two-layer dLIF for detecting seizures. N_b: Number of branches (4), H_n: Hidden Neurons (100).

4.2.4 Implementation details

For our experiments, as they are based on non-spiking dataset, the real valued inputs are directly provided to the dLIF. The first layer serves as an encoding layer, transforming non-spiking inputs into spiking outputs before passing them to subsequent layers. The parameters of this initial layer are trained along with the rest of the network to achieve optimal performance Wu et al. (2019). For each compartmental neuron (comprising soma and dendrites), pre-synaptic inputs are distributed randomly across the dendritic branches, with non-overlapping input indexes and an approximately equal number of inputs per branch. We conducted a series of experiments and ultimately decided to choose 4 branches as performance was enhanced by utilising this. Extending beyond or below do not add further improvements. We employed cross-entropy as our loss function and a softmax function of the model's predictions to assess different performance metrics. The detection layer of our model is responsible for determining whether a seizure is present based on the input data. A detailed diagram of the system is represented in Fig. 4.3. The model was trained and tested on a V100 GPU. A summary of hyper-parameters been used and information around the model we used is specified in Table. 4.1.

Table 4.1 – Summary of hyperparameters settings for each dataset

Dataset	TUH	CHB-MIT	EPILEPSIAE	FB
Input Channel	19	22 (19)	30-114	6
Output Channel	2	2	2	2
Sampling rate(Hz)	125/250	250	256	256
Time Window(s)	12	12	12	12
Dendritic Branches	4	4	4	4
Type of Connections	Recurrent	Recurrent	Recurrent	Recurrent
Network structure	I-100R-100R-2	I-100R-100R-2	I-100R-100R-2	I-100R-100R-2
Loss	CE	CE	CE	CE
Batch size	200	200	256	256
Optimiser	Adam	Adam	Adam	Adam
Learning rate (LR)	1e-2	1e-2	1e-2	1e-2
LR decay (per 100)	0.1	0.1	0.1	0.1
LR decay type	Step	Step	Step	Step
Tau initialiser	Uniform	Uniform	Uniform	Uniform
Tau initialiser Limit Range	[0,4]	[0,4]	[0,4]	[0,4]
Reset Mechanism	Soft	Soft	Soft	Soft

4.2.5 Performance metrics

Since our training is supervised, the dataset we use to train and evaluate the model includes labels for each time window (12s). A seizure is detected when the model correctly predicts an ictal event (class 1) based on the ground truth, while non-seizure events are labeled as class 0. Seizure datasets are highly imbalanced, with significantly more non-seizure than seizure data, as seizures are infrequent and brief. In a 12-second time window over a day, seizure periods are a small fraction of the data. Using conventional metrics (i.e accuracy) in this context is misleading, as it favors the majority class, failing to capture the model's ability to detect seizures accurately, which is the main objective. Therefore, we utilised AUROC. It captures both the sensitivity and specificity in a threshold-free score. Another metric used is AUPRC which quantifies the trade-off between precision and recall at various threshold settings. We also provide precision, recall, and F1-score metrics to evaluate the model performance across different datasets. Precision measures the proportion of correctly identified positive cases among all cases predicted as positive by the model. Recall measures the proportion of correctly identified positive cases among all actual positive cases in the dataset. The F1-score is the harmonic mean of precision and recall. It provides a single metric that balances both precision and recall. In our experiments, it is worth noting that we performed in-sample testing with the TUH, FB, iEEG EPILEPSIAE and CHB-MIT datasets. However, we also performed out-of-sample testing in the CHB-MIT dataset.

4.3 Experimental Results

4.3.1 Scalp-EEG.

Training and Validation:

Following the methodology outlined earlier, our model was trained and validated using the TUH dataset. We successfully attained an AUROC score of 81.4%. Our findings,

as illustrated in Table. 4.2 highlight the superior performance of our spiking model compared to non-spiking models. Our model's performance rivals existing approaches within the time and time-frequency domains, with fewer layers and memory needs, and in a spiking fashion. Table 4.3 provides a comprehensive overview of the model performance across a range of metrics.

Preliminary findings in down-sampling the signal leads to marginal drop in performance:

In our investigation, we examined the effects of reducing the sampling rate in the TUH dataset on our model's performance. Our results showcased a significant level of robustness, with minimal drops in performance observed across key metrics such as AUROC, AUPRC, and Precision. Notably, while there was a slight decline of approximately 7% in Recall and 4% in F1-Score, the overall performance remained beneficial. These findings support the hypothesis that our model prioritises detection biomarkers located in lower frequencies over higher frequencies. This assertion is further substantiated by previous research Lundstrom et al. (2021). Our results are represented in Table 4.3.

Table 4.2 – AUROC across input domains on the TUH dataset. A 12-s window is used for baseline.

Model	Sampling Rate (Hz)	Network Type	Input Domain	AUROC
CNN-LSTM Tang et al. (2022)	250	ANN	Time	0.71
GNN Tang et al. (2022)	250	ANN	Time	0.73
CNN-LSTM Tang et al. (2022)	250	ANN	Frequency	0.75
DCNN Tang et al. (2022)	250	ANN	Time	0.78
LSTM Tang et al. (2022)	250	ANN	Frequency	0.79
ResNN Tang et al. (2022)	250	ANN	Frequency	0.79
DCNN Tang et al. (2022)	250	ANN	Frequency	0.81
GNN Tang et al. (2022)	250	ANN	Frequency	0.81
dLIF (this work)	**250**	**SNN**	**Time**	**0.81**
dLIF (this work)	**125**	**SNN**	**Time**	**0.81**
Conv-LSTM Yang et al. (2022b)	250	ANN	Frequency	0.84
LTC-FPTT Herbozo Contreras et al. (2024a)	250	SNN	Frequency	0.89
Transformer Ma et al. (2023)	250	ANN	Frequency	0.91

Table 4.3 – TUH Results for down-sampling test. It is worth noting that a decrease of 50% in the original sampling frequency leads to similar performance.

EEG Sampling Rate	AUROC	AUPRC	Precision	Recall	F1-Score
250 Hz	81.4	90.1	84.4	81.3	83.1
125 Hz	81.0	89.2	84.2	74.9	79.2

In-sample and Out-of-Sample generalisation in children: Non-Montage and Non-Age Specific:

This section investigated in-sample and out-of-sample generalisation in children, using a non-montage, non-age-specific approach. In-sample validation involves splitting a single dataset into training and validation subsets. In contrast, Out-of-sample validation involves training a model on one dataset and then evaluating it on a different dataset, offering a stronger measure of generalisation. Our research methodology involved rigorous testing utilising the CHB-MIT dataset, a widely recognised repository for Paediatric EEG data. Unlike previous models predominantly operating in the time-frequency domain, our novel approach focuses solely on the time domain, offering a unique perspective on EEG analysis in Paediatric populations. Notably, our model adopts a spiking paradigm, further distinguishing it from existing methodologies. In our first experiment, we conducted an in-sample validation. In Table 4.4, we demonstrate the effectiveness of our model by achieving an AUROC of 94.2%, which surpasses a previous spiking neural network architecture that utilised a time-frequency input signal. While our results did not outperform traditional models such as CNN and ConvLSTM, our model achieved competitive performance, with a marginal drop of 4-5% in accuracy.

For our second experiment, we emphasise the importance of out-of-sample validation, in contrast to the conventional in-sample evaluations commonly used. Through meticulous analysis, we reveal how our framework, emphasising time-domain evaluation and in-sample testing, differs from existing methodologies. Importantly, while our model does not demonstrate superiority in all aspects, it offers valuable insights by considering spiking models and temporal input data often overlooked in prior research. Results can be observed in Table 4.4. By considering more biologically

Table 4.4 – CHB-MIT results. Comparison of AUROC across different models with time-frequency *vs* time-domain input (ours) and in-sample *vs* out-of-sample (ours).

Patient ID	Gender	In-Sample				Out-of-Sample
		Time-Frequency Domain			Time Domain	Time Domain
		CNN Yang et al. (2023)	**ConvLSTM** Yang et al. (2023)	**SConvLSTM** Yang et al. (2023)	**dLIF-SNN**	**dLIF-SNN**
1	F	100.0	100.0	96.4	98.6	71.9
3	F	100.0	100.0	92.3	100.0	51.1
5	F	100.0	100.0	95.3	86.0	66.4
6	F	94.9	93.4	88.8	93.2	51.1
7	F	97.0	96.5	92.0	96.7	87.3
8	M	99.4	98.2	72.5	98.3	79.8
9	F	99.7	100.0	96.1	98.1	98.3
10	M	100.0	100.0	97.5	97.8	86.6
11	F	98.8	98.9	81.2	97.7	98.6
15	M	99.9	99.7	73.0	95.1	50.7
17	F	88.2	91.0	87.7	94.3	88.1
18	F	98.5	96.9	86.0	97.3	63.0
19	F	95.4	99.8	95.9	92.3	72.0
20	F	97.1	97.7	87.9	96.0	68.8
21	F	99.3	100.0	92.5	81.4	65.7
22	F	100.0	100.0	96.6	90.3	95.6
23	F	100.0	100.0	95.9	88.7	95.0
Average	3M/14F	98.13	98.36	89.9	94.2	75.9

M: Male, F: Female.

plausible networks, the dLIF SNN can address highly challenging tasks, although its performance is lower than that of models outside the spiking domain, which utilise time-frequency analysis and are validated only in-sample.

4.3.2 Intracranial-EEG

Training and Validation:

We conducted extensive analyses using data from both the FB and EPILEPSIAE datasets. Our examination involved multiple patients from the FB dataset, in which we assessed our model's performance in the time domain and compared it with that of alternative models. Despite utilising only temporal input, our model exhibited efficacy on par with non-SNN and conventional SNN methods, achieving an average AUROC score of 93.4. For the EPILEPSIAE dataset, we performed multiple training as each patient exhibited different channel numbers. This evaluation yielded meaningful outcomes across various evaluation metrics, including AUPRC, Sensitivity (Recall), Precision, F1 Score, and AUROC, with impressive values of 92.0,

Table 4.5 – FB results. Comparison of AUROC across different models with time-frequency vs time-domain(ours).

Patient ID	Gender	Time-Frequency Domain			Time Domain
		CNN Yang et al. (2023)	**ConvLSTM** Yang et al. (2023)	**SConvLSTM** Yang et al. (2023)	**dLIF-SNN**
1	F	100.0	100.0	100.0	100.0
3	M	100.0	100.0	93.7	80.2
4	F	100.0	100.0	99.9	96.1
5	F	85.0	84.8	84.4	100.0
6	F	86.8	90.6	92.4	53.3
14	F	80.5	92.6	85.6	100.0
15	M	93.3	99.1	89.5	100.0
16	F	93.4	96.1	85.7	99.0
17	M	100.0	100.0	97.3	94.9
18	F	100.0	100.0	98.5	99.1
20	M	94.8	99.1	94.5	99.2
21	M	98.8	97.5	91.4	99.7
Average	5M/7F	94.4	96.7	92.7	93.4

M: Male, F: Female.

85.0, 95.0, 89.0, and 89.9, respectively. Previous studies Yang et al. (2023) that utilised seizure detection models with the iEEG EPILEPSIAE dataset have reported various AUROC scores, including a SNN that achieved 81.1, and other approaches such as ConvLSTM, Teager Energy (TE), and Short-term maximum Lyapunov exponents (STLmax), which reached AUROCs of 91.8, 67.6, and 69.7, respectively, by utilising frequency-domain features. Notably, our time-domain-based approach outperformed three of these models (SNN, TE, and STLmax), highlighting the effectiveness of our method without reliance on frequency-domain analysis. These results, along with detailed comparative findings presented in Table 4.5, and in Table 4.6, underscore the promising capabilities of our proposed approach in the analysis of iEEG signals.

4.4 Timescale analysis of the dendrite branches

We provide a detailed analysis of the timescale dynamics of the soma and dendritic branches across the layers of our dLIF neural network. Focusing particularly on the second spiking layer, we examined the behaviour of two randomly selected neurons using a seizure recording from the TUH dataset. Our findings, illustrated in Fig. 4.4,

Table 4.6 – iEEG-EPILEPSIAE results. Training and testing occur independently for each patient.

Patient ID	Gender	NoE	AUPRC	Sensitivity	Precision	F1 Score	AUROC
1	F	66	78.1	71.2	70.0	71.1	77.2
2	F	30	100.0	100.0	100.0	100.0	100
3	M	114	96.1	92.3	100.0	96.1	96.8
4	F	75	97.2	93.1	100.0	96.2	92.7
5	M	82	80.3	50.3	100.0	67.1	77.6
7	M	60	98.1	92.1	100.0	96.2	98.9
8	M	38	60.2	50.1	68.0	58.1	56.3
9	M	92	99.2	97.2	100.0	98.1	99.8
10	M	56	94.1	75.3	100.0	86.6	85.1
11	F	46	100.0	100.0	100.0	100.0	100
12	F	124	99.1	95.1	100.0	97.1	95.2
13	F	62	98.0	88.3	100.0	94.2	94.0
14	F	121	95.4	89.2	97.3	93.1	93.2
15	M	58	91.1	93.1	89.1	91.2	91.7
Average	-	-	91.9	84.8	94.6	88.9	89.8

M: Male, F: Female, NoE: Number of invasive electrodes.

reveal distinct firing patterns among different neurons. Specifically:

1. Neuron 11: Dendritic branch 4 exhibited a strong correlation with changes in the overall firing rate (Fig. 4.4B). This suggests that this particular branch plays a significant role in modulating the neuron's spiking activity.

2. Neuron 90: Similarly, an extended activity in dendritic branch 4 aligned with a sustained increase in firing rate (Fig. 4.4C), indicating a consistent influence on the neuron's output.

These observations highlight the importance of specific dendritic branches in shaping the network's temporal dynamics during seizure detection. By analysing the learned timescales, we demonstrate how different timescales in the soma and dendritic branches contribute to improved time series processing.

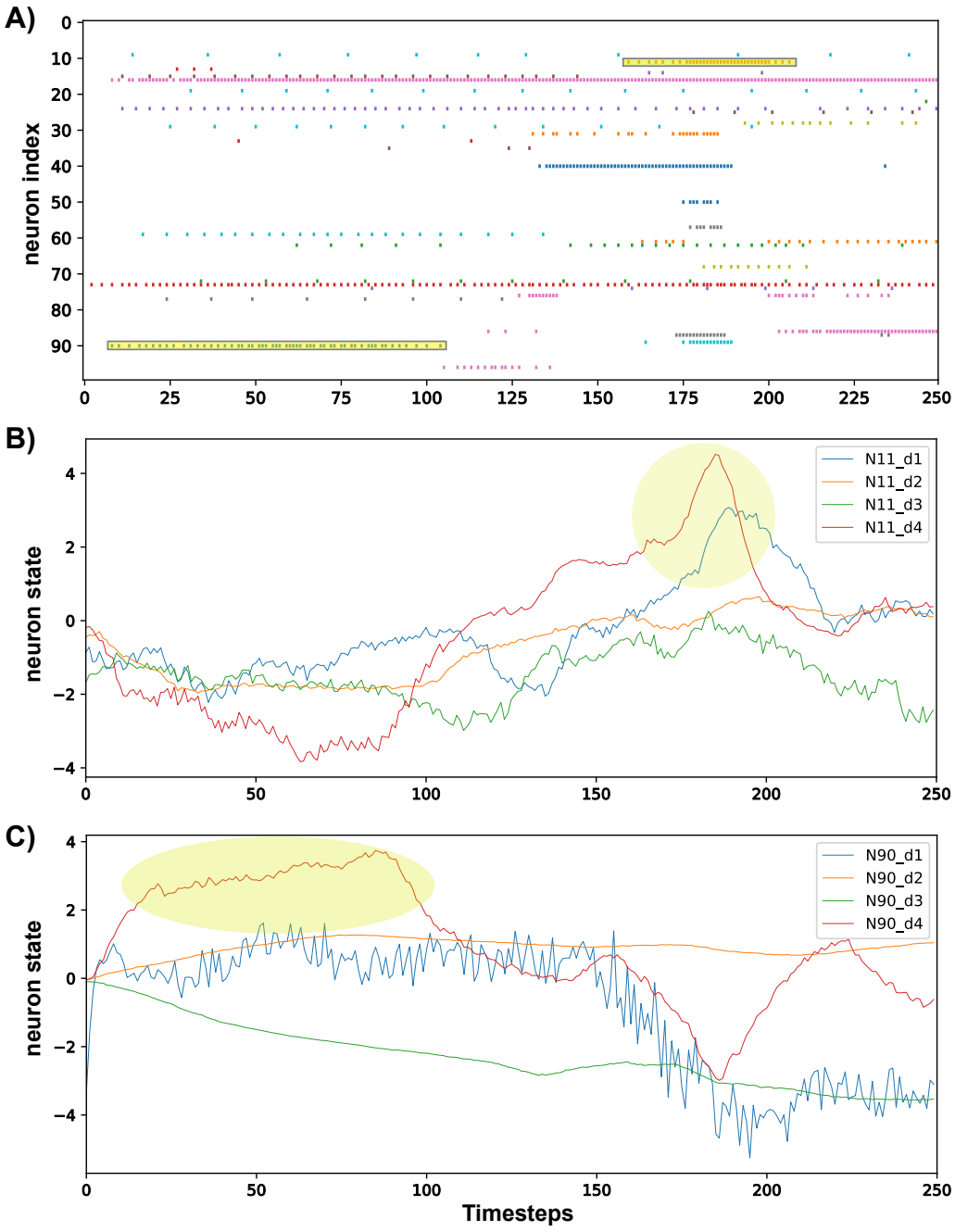

Figure 4.4 – Timescale analysis in the dLIF neural network. Analysis of the second spiking layer's output using a seizure file from the TUH dataset reveals distinct firing patterns (A). The dendrite branch 4 output of neuron 11 shows a strong correlation with changes in the overall firing rate, suggesting a key role in the network's spiking behavior (B). A similar relationship is observed in neuron 90, where branch 4's output extended activity corresponds (C).

4.4.1 Learned Timing Factors of Dendrite Branches and Membrane Potential

We analysed the learned timing factors and examined how branches with heterogeneous behaviour improve performance for different brain signal modalities. Fig. 4.5 shows the timing factors of the branches for both scalp (A) and intracranial (B) recordings. After training, we observe that a higher density of the timing factor (α) has been preserved in both datasets, which may contribute to long-term memory consolidation. For the timing factor (β) at the soma level, we can observe a similar trend. Our hypothesis is supported by previous studies Perez-Nieves et al. (2021), which demonstrated that heterogeneous time constants (such as those observed in dendritic branches) within a layer enable the network to integrate incoming spikes across different time scales. Similarly, Zheng et al. (2024) demonstrated that long-term memory with high timing factors has been addressed depending on the nature of slow-fast input signals.

4.5 Estimation of Power Consumption

We employed two strategies to estimate the power consumption of our neural network. For our first power consumption estimation, we multiplied the total spikes per timestep by 120 pJ, the typical energy cost per spike for the Loihi chip Davies and other authors (2018). Timesteps were based on the time window for processing EEG data and the frequency sampling rate. During inference mode, we recorded a total of 87 spikes/step across layers 1 and 2 (which contain dendritic branches). Through this method, we found an energy cost of merely 10.4 nJ per timestep. Recognising some limitations of this approach, we adopted a second methodology based on prior studies Chen et al. (2023), which involved calculating the number of MAC and Accumulate (AC) operations. The total number of computational operations between dendrites and synapses in our model is twice the number of branches times the number of neurons. Each AC operation can be regarded as the 19.6% of a MAC

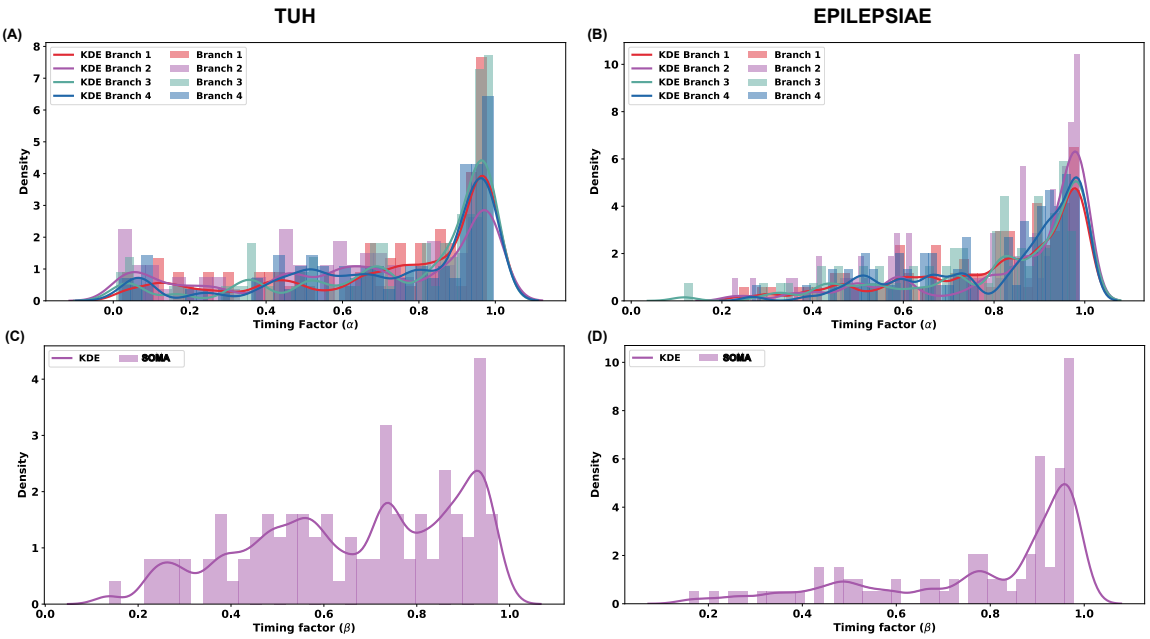

Figure 4.5 – Histogram of the learned timing factors of dendrites branches and soma of our neural network for scalp and intracranial recordings. The timing factor (α) for the branches in both datasets shows a consolidation of high timing factors in the scalp (A) and intracranial (B) datasets. A similar trend in the timing factors of the soma membrane potential (β) can be observed in both datasets (C-D). KDE: Kernel Density Estimation.

operation Horowitz (2014). Considering a technology of 45 nm, and at energy consumption of 0.9 pJ and 4.6 pJ for ACs and MACs respectively, we estimated a total energy consumption of 6.38 nJ per time step. Furthermore, we compared our results with the previous model reported Yang et al. (2023). We set the timesteps as 23 for comparison. The results in Fig. 4.6 indicate favorable energy and memory efficiency. Despite not achieving the lowest memory usage, our approach demonstrates significant improvements in memory and energy efficiency, achieving enhancements of 3 to 5 times in memory efficiency and demonstrating an improvement in energy efficiency ranging from three to five orders of magnitude compared to the ConvLSTM model across both datasets.

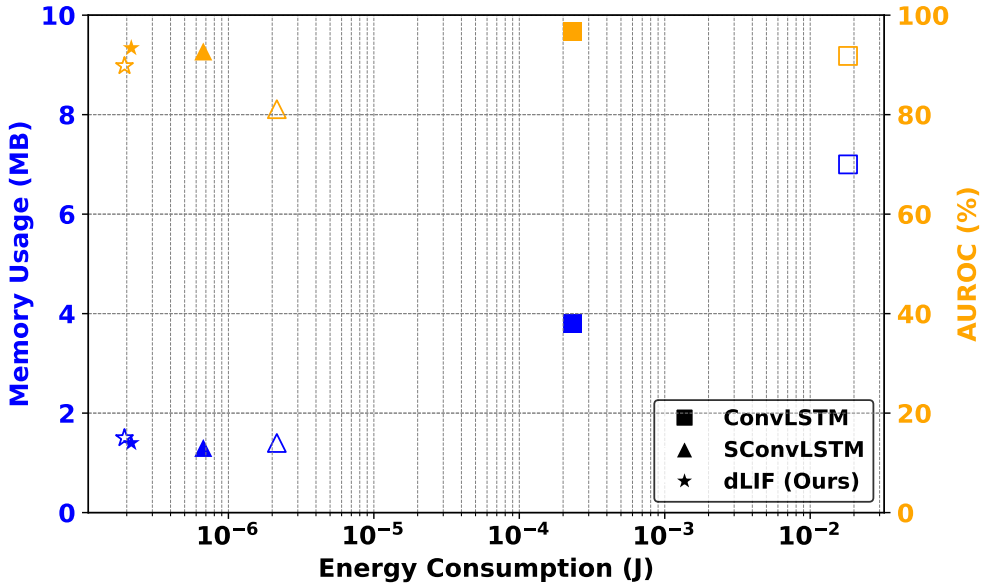

Figure 4.6 – Pareto plot of energy consumption versus memory usage and AUROC for three models (ConvLSTM Yang et al. (2022b), SConvLSTM Yang et al. (2023), and dLIF) tested for 23 timesteps. Filled markers represent the results for the EPILEPSIAE dataset, while non-filled markers represent the FB dataset.

4.6 Discussion

Our results demonstrate robustness in detecting seizures using a compact dendritic spiking neural network for both iEEG and scalp-EEG signals. The model, requiring only 1.5MB of memory and fewer than 300K parameters, is well-suited for edge deployment, which is significantly lower compared to typical effective seizure detection systems. Our model achieved an AUROC of 81% for the TUH dataset, comparable to other baseline models that use both time-domain and time-frequency domain features. Importantly, it operates without feature extraction techniques like STFT, which are computationally expensive and power-intensive. This makes our approach particularly viable for real-time edge deployments. Moreover, reducing the sampling rate showed minimal performance decline across metrics, further demonstrating the model's robustness and enabling faster training. In the iEEG domain, our model excelled by achieving a higher AUROC and outperforming previous SNN models by almost 1%, highlighting its ability to handle high-sampling rate inputs effectively.

Memory and Energy efficiency against other baseline models has been demonstrated. Challenges remain in improving efficiency, particularly with BPTT used during training. BPTT is computationally intensive and susceptible to gradient vanishing/explosion. To overcome these challenges, we aim to incorporate FPTT, which has shown promise in previous EEG seizure detection studies Herbozo Contreras et al. (2024a). This approach aims to enhance robustness for continuous learning with lower computational overhead, advancing our goal towards smart implantable system Yang et al. (2022a).

Although the dendritic spiking models in this chapter are evaluated on conventional digital hardware, related dendrite-inspired SNN models have recently been mapped onto hybrid-paradigm neuromorphic hardware such as TianjicX, where spiking operations support soma dynamics and non-spiking operations support dendritic dynamics Zheng et al. (2024). This indicates that dendritic mechanisms can be made hardware-compatible under suitable neuromorphic architectures. However, their immediate deployment remains platform-dependent, as current commercial neuromorphic chips vary in their support for compartmental dynamics, synaptic operations, and network topology.

4.7 Conclusion

Here, we assessed the advantages of compartmental models that enabled the analysis of seizure detection within the time-domain. By only utilising basic pre-processing and non-power consuming techniques such as power-line noise in real-world scenarios of iEEG datasets such as FB and EPILEPSIAE, and ICA on scalp-EEG datasets as the meaningful TUH dataset, we have demonstrated the capabilities to exhibit efficient performance while maintaining a low-power consumption spectrum. By only utilising two hidden dendritic layers, these results indicate the feasibility of more embedded AI applications for long-term data, demonstrating that more dynamic and biological neural networks are necessary. Future studies will merge liquid-time constant spiking neurons Herbozo Contreras et al. (2024a), which incorporate dynamic

membrane time constants that are learnable parameters, with dendritic neurons for more dynamic features.

Chapter 5

Unified Dendro-Liquid Framework for Spatio-Temporal Seizure Detection.

This chapter combines liquid temporal dynamics from Chapter 3 and dendritic computation from Chapter 4 into an ultra-compact neuromorphic architecture for energy-efficient on-device seizure detection inference.

The content presented in this chapter is published as:

- Herbozo Contreras, L. F., Yu, L., Huang, Z., Nikpour, A. and Kavehei, O. (2025) "Spiking neural networks with liquid time-constant dynamics and dendritic branches for efficient time-domain edge-based seizure detection." *IOP Machine Learning Health*, 1(1), 015001.DOI:10.1088/3049-477X/adeb8e.

Statement of Contributions of Joint Authorship

- Luis Fernando Herbozo Contreras (Candidate): First author, formal analysis, investigation, methodology, resources, software, visualization, validation, writing original draft, writing review & editing of the manuscript
- Leping Yu: contributed to writing review & editing of the manuscript
- Zhaojing Huang: contributed to writing review & editing of the manuscript
- Isabelle Aguilar: contributed to writing review & editing of the manuscript
- Armin Nikpour (Co-Supervisor): conceptualisation, data curation, funding acquisition, investigation, project administration, resources, supervision, validation, writing review & editing
- Omid Kavehei (Principal Supervisor): conceptualisation, data curation, funding acquisition, investigation, methodology, project administration, resources, supervision, visualisation, validation, writing review & editing.

In addition to the statements above, in cases where I am not the corresponding author of a published item, permission to include the published material has been granted by the corresponding author.

Luis Fernando Herbozo Contreras Date: 28 May 2026

As supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

Prof. Omid Kavehei Date: 28 May 2026

Chapter Summary

Accurate seizure detection from heterogeneous EEG recordings requires models capable of capturing spatial interactions across channels and temporal dependencies across multiple timescales. While dendritic architectures provide efficient multiscale processing in the time-domain, and liquid time-constant dynamics offer adaptive temporal learning, their integration remains underexplored in compact edge-deployable systems. This chapter introduces Liquid-Dendrite, a novel bio-inspired model for seizure detection, leveraging LTC-SNN and dLIF with heterogeneous time-constants. The model comprises two hidden layers with dendritic neurons and one layer of liquid-time constant networks. Our model achieves a memory efficacy of 535 KB with 130 K trainable parameters. The model was tested across the most noteworthy epilepsy datasets for scalp EEG (TUH and CHB-MIT) and iEEG (EPILEPSIAE). Our model demonstrated commendable performance, achieving AUROC scores of 83%, 96%, and 93%, respectively, outperforming some existing models in an energy and memory-efficient way. Moreover, we conducted a robustness test by blacking out EEG channels at the inference stage, where we showed the ability of our network to work with fewer channels. We deployed our tiny model and perform inference at the edge of the Raspberry Pi 5 without the need for additional quantisation. This highlights the potential of Neuro-Inspired AI for efficient, small-scale, and energy-embedded AI systems across different brain modalities.

5.1 Introduction

Inspired by the brain's structure, Neuromorphic AI contrasts sharply with the traditional von Neumann architecture, where memory and processing are separated, leading to power inefficiencies during data transmission. Neuromorphic chips, however, integrate memory and processing within neurons and synapses, improving speed and energy efficiency and on-device training Marković et al. (2020). While some models utilise LIF neurons, their efficacy often falls short, partly due to the simplified nature

of these models. A promising direction is incorporating biophysical brain models, particularly those that account for dendrites, which are critical for neuronal computation Yamazaki et al. (2022). Researchers have also utilised Adaptive Spiking Recurrent Neural Network (ASRNN), which have demonstrated significant potential in this area, offering enhanced computational capabilities with learnable membrane time constants Yin et al. (2020, 2023).

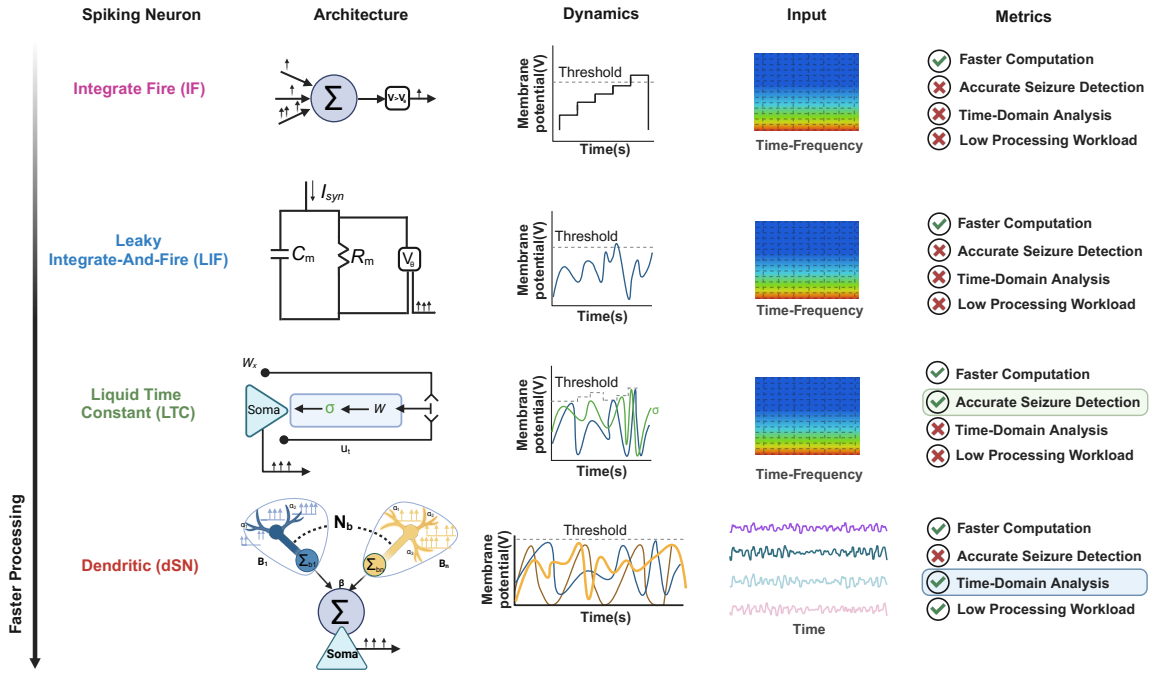

Figure 5.1 – Overview of four spiking neuron models: Integrate & Fire IF, Leaky Integrate & Fire LIF, Liquid Time Constant LTC, and dLIF showing their respective architectures (middle column), typical membrane potential dynamics (third column), input representations (fourth column), and performance metrics (right column). As spiking neurons becomes more bio-realistic, they are able to perform faster computation, accurate seizure detection and within the time-analysis.

5.1.1 Advancing Towards Bio-Realistic Models in Seizure Detection

Recent advancements in seizure detection have moved beyond traditional LIF models to embrace more bio-realistic approaches that better mimic the dynamics of the

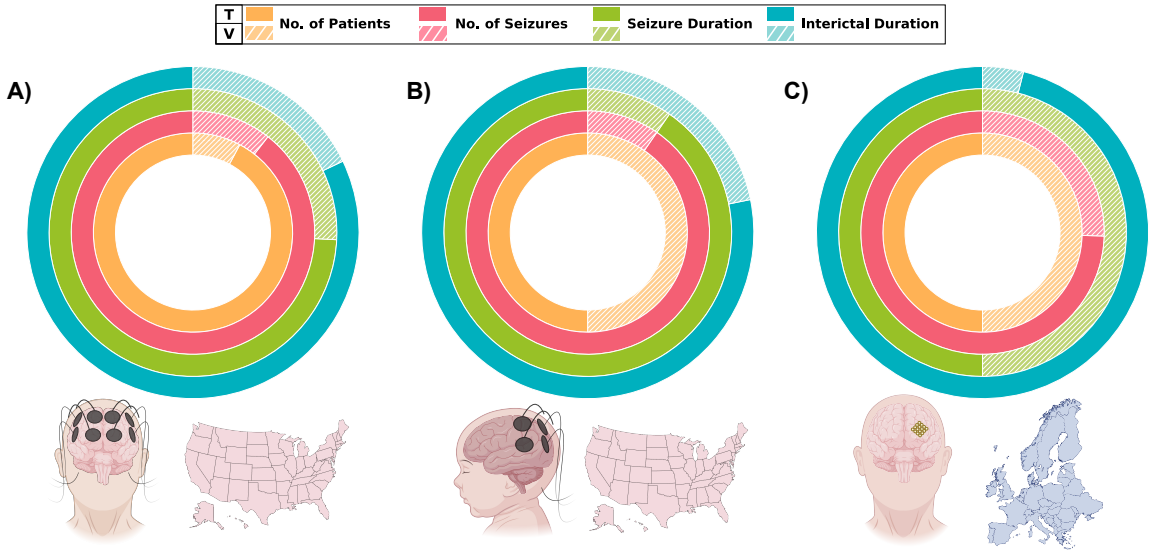

Figure 5.2 – Datasets used in this work. TUH, a large scalp-EEG dataset from adult patients in the USA (A). CHB-MIT, a pediatric scalp-EEG dataset from the USA (B). iEEG-EPILEPSIAE, the largest long-term intracranial-EEG dataset in Europe (C). T: Training, V: Validation.

brain. For instance, Herbozo Contreras et al. (2024a) introduces a biologically plausible algorithm that employs Liquid-Time Constant Spiking Neurons with FPTT training. This approach enhances interpretability and energy efficiency and demonstrates strong generalisation across various datasets, achieving an AUROC of approximately 0.90 on validation samples and above 0.82 for out-of-sample generalisation, all while reducing the computational cost. Complementing this, Herbozo Contreras et al. (2024c) leverages dendritic branches to introduce heterogeneity within the system. By moving away from conventional time-frequency analysis methods such as the STFT and FFT, this approach facilitates more practical time-domain seizure analysis. It also minimises energy consumption without compromising detection performance, demonstrating high resilience across different brain signal modalities, including scalp and intracranial EEG. Fig. 5.1 offers a comparative overview of four spiking neuron models: Integrate & Fire IF, LIF, LTC, and dLIF. The IF model, characterised by its simple stepwise membrane potential dynamics, falls short of biological plausibility. The LIF model improves on this by incorporating membrane potential decay, yet it remains primarily focused on time-frequency analysis. In contrast, the LTC

model further enhances computational efficiency by adapting its time constants while maintaining high seizure-detection accuracy. Finally, the dLIF model employs dendritic processing to enable direct time-domain analysis, offering faster computation, more precise seizure detection, and reduced processing workload. The trend towards increased bio-realism, as illustrated in Fig. 5.1, significantly boosts the efficiency of real-time seizure detection.

5.1.2 Novelty and Significance

This work integrates LTC neurons with dendritic neurons, leveraging distinct time constants within a unified spiking neural network to enhance computational efficiency and improve seizure detection accuracy. This system avoids the computational intensity of traditional techniques, proving robust across various brain modalities and enabling low-latency inference in real-world applications. Our work is the first to integrate distinct time constants within a spiking neural network. The following points outline our key research highlights:

1. This is the first contribution in incorporating heterogeneous and liquid-time constants within spiking neural networks.
2. The model significantly reduces computational requirements, with a memory footprint of 535 KB and 130 K trainable parameters.
3. The model demonstrates reliable performance across different brain-signal modalities.
4. The model remains robust even with missing data in channels.
5. The approach ensures inference on edge devices with a latency of 0.81 seconds per batch without additional quantisation.

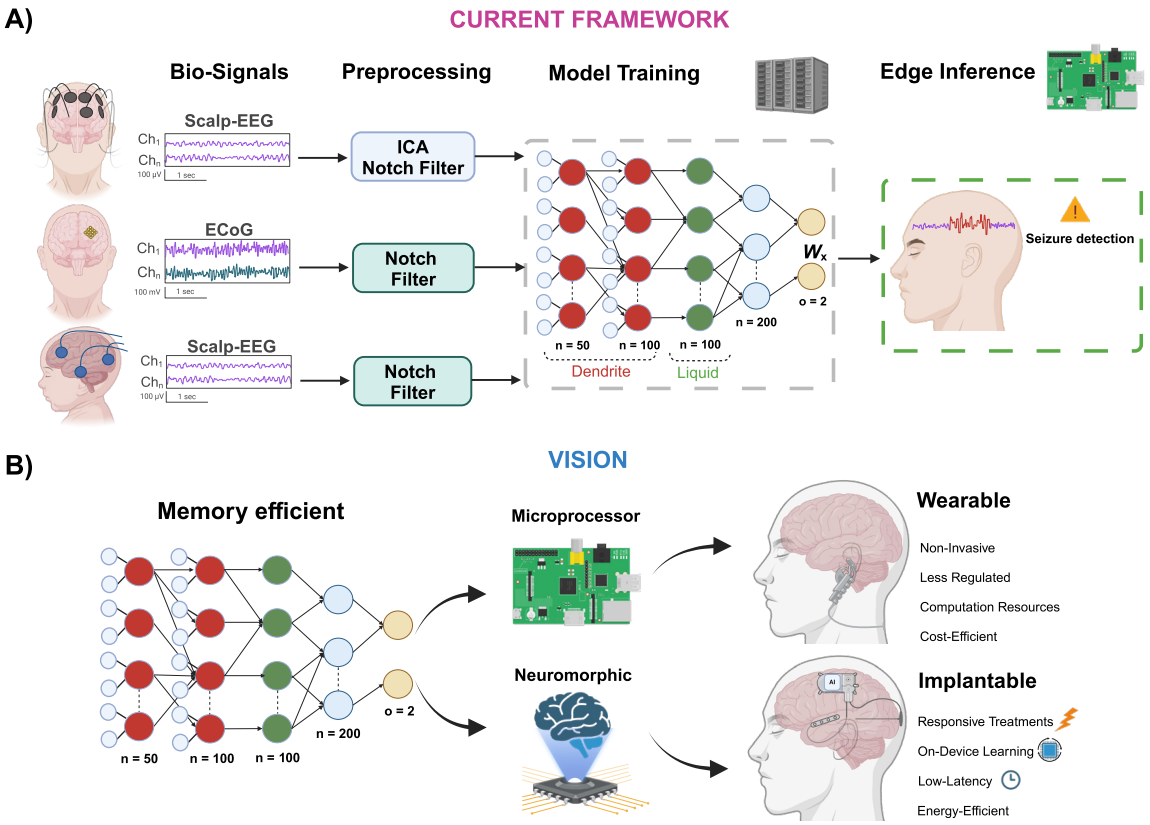

Figure 5.3 – Our current framework and vision of our model. We independently trained our model on three datasets, ensuring a consistent time-domain preprocessing pipeline across all. The model consists of two layers of dendritic spiking neurons followed by a single spiking neuron layer. Training was conducted on a GPU, while inference was performed efficiently on a Raspberry Pi (A). Our approach demonstrated both memory efficiency and high accuracy, highlighting its potential for deployment and on-device training in wearable and implantable neurostimulation devices to address seizure detection challenges (B).

5.2 Datasets

Here, we utilised 3 datasets: TUH, CHB-MIT, and the iEEG EPILEPSIAE dataset, as shown in Fig. 5.2.

5.2.1 Scalp-EEG

TUH:

TUH is the largest scalp-EEG dataset available worldwide, comprising recordings from a diverse cohort of 642 patients (315 male and 327 female), aged 1 to 90 years. This dataset contains 3,044 seizures with a mean seizure duration of 74.4 seconds, recorded across 19 channels using a scalp EEG modality. The dataset is split into training and validation sets, with 592 sessions allocated to training and 50 to validation, making it an extensive and balanced resource for developing and validating machine learning models.

CHB-MIT:

The CHB-MIT dataset is a widely used scalp-EEG dataset primarily focused on pediatric patients. We selected 17 patients, which aligns with the baseline results for comparison. It contains recordings of 3 males and 14 females, ranging from 1.5 to 22 years. The dataset includes 160 seizures with an average duration of 92.5 seconds. Recordings were made using a scalp EEG modality across 22 channels, all sampled at 256 Hz.

5.2.2 Intracranial-EEG

EPILEPSIAE:

We trained and tested the model using data from 14 patients, as the datasets differ in montage and channel count across the EPILEPSIAE dataset, which ranges from 30 to 114 electrodes and has a sampling rate of 256 Hz. The EPILEPSIAE dataset includes high-quality, long-term EEG, iEEG, and simultaneously recorded Electrocardiogram (ECG) data. Invasive recordings are obtained from patients using implanted strips, grids, or depth electrodes, stereo-EEG. The dataset contains extensive interictal

periods (background data between seizures), making it a more realistic representation of iEEG data and ensuring that our false negatives are highly reliable.

5.3 Methods

5.3.1 Pre-processing

We applied ICA to eliminate artifacts from scalp EEG signals. Initially, the EEG data were divided into 12-second segments, and the ICA algorithm was used to decompose them into 19 independent components through BSS. ICA facilitates the decomposition of EEG signals into statistically independent components, as described by Equ. (5.1):

$$T \approx MA^{\mathrm{T}} \ . \tag{5.1}$$

In this equation, T represents the EEG data, M denotes the time-related information, and A corresponds to the weight matrix for the topographic maps. Pearson's correlation was used to identify independent components highly correlated with eye movements, detected using the 'FP1' and 'FP2' EEG channels Dammers et al. (2008). The identified elements associated with eye movement were removed, yielding EEG signals free of such artifacts. Following this, we removed powerline noise from the MNE-Python library (MNE) package from the TUH and CHB-MIT datasets using a notch filter at 60 Hz. The pre-processing pipeline is performed offline and is therefore not included in the energy calculations.

5.3.2 Spiking Neural Network

The proposed method is modelled as an adaptive spiking recurrent neural network incorporating dendrites with heterogeneous time constants and liquid-time-constant membrane potentials, which are inherently learnable.

Dendrite Spiking Neuron Structure:

The dLIF neuron model improves upon the traditional LIF-based spiking neuron by incorporating multi-timescale memory within the dendrites. This enhancement is described by Equ. (5.2), Equ. (5.3), and Equ. (5.4), where u represents the membrane potential of the soma, β represents the timing factor, R denotes the membrane resistance, d represents the dendritic branch index, and u_{th} represents the firing threshold Zheng et al. (2024).

$$u^{t+1} = \beta u^t + (1-\beta)\sum_d R i_d^{t+1} - o^t u_{\text{th}} \ , \tag{5.2}$$

$$i_d^{t+1} = \alpha_d i_d^t + (1-\alpha_d) I_d^{t+1} \ , \tag{5.3}$$

$$o^{t+1} = H(u^{t+1} - u_{\text{th}}) \ . \tag{5.4}$$

The Heaviside function $H(\cdot)$ governs the spiking activity. The synaptic input on the dendritic branch combines feedforward and recurrent inputs Zheng et al. (2024).

$$\alpha_d = \text{sigmoid}(\hat{\alpha}_d), \beta = \text{sigmoid}(\hat{\beta}) \ . \tag{5.5}$$

To avoid negative timing factors in Equ. (5.5), α and β should be restricted within values of 0 and 1, achieved by a sigmoid function.

Liquid Time Constant Spiking Neuron Structure:

These time constants are determined differently depending on the type of network being used Yin et al. (2023). Equ. (5.6) and (5.7) integrate the LTC-SNN behavior. The scaling function $\sigma()$ is essential to ensure smooth transitions during learning. Equ. (5.8) defines θ_t as an adaptive threshold, which helps simulate the realistic

adaptive behaviour observed in biological spiking neural networks. Equ. (3.5) and Equ. (3.6) describe how the membrane potential and spiking rate are regulated by τ_m, while Equ. (3.7) defines the resetting of the membrane potential after a spike. Fig. 5.1 shows the liquid-dendrite neural architecture.

$$\tau_a \text{ update} : \rho^i = \exp\left(-\frac{dt}{\tau_a^i}\right) = \sigma\left(\text{Dense}_a[x_t, b_{t-1}^i]\right) \ , \tag{5.6}$$

$$\tau_m \text{ update} : \alpha^i = \exp\left(-\frac{dt}{\tau_m^i}\right) = \sigma\left(\text{Dense}_m[x_t, u_{t-1}^i]\right) \ , \tag{5.7}$$

$$\theta_t \text{ update} : b_t^i = \rho^i b_{t-1}^i + (1-\rho^i)s_{t-1}^i;\ \theta_t^i = 0.1 + 1.8b_t^i \ , \tag{5.8}$$

$$u_t \text{ update} : du^i = -u_{t-1}^i + x_t;\ u_t^i = \alpha^i u_{t-1}^i + (1-\alpha^i)du^i \ , \tag{5.9}$$

$$\text{spike } s_t : s_t^i = f_s(u_t^i, \theta^i) \ , \text{ and} \tag{5.10}$$

$$\text{resetting} : u_t^i = u_t^i(1-s_t^i) + u_{\text{rest}} s_t^i \ . \tag{5.11}$$

5.3.3 Backpropagation through time

The BPTT algorithm has been adopted to better handle the unique characteristics of both spiking neural networks.

LTC Neuron:

Consider a recurrent neural network defined by the differential Equ. (5.12):

$$\hat{y}_t,\ h_t = \text{NN}(x_t, h_{t-1}) \tag{5.12}$$

where x_t denotes the input at time t, $\hat{y}_t$ represents the prediction, and h_t is the hidden state. Then, the gradient with respect to the weights w is computed by considering the effect of the input x_t on all future losses $l_t, l_{t+1}, \ldots, l_T$. The gradient of the total loss L is given by Equ. (5.13):

$$\frac{\partial L}{\partial w} = \sum_{t=1}^{T} \frac{\partial l_t}{\partial w} = \sum_{t=1}^{T} \sum_{i=t}^{T} \frac{\partial l_i}{\partial h_i} \frac{\partial h_i}{\partial w} \tag{5.13}$$

Recognising that each hidden state h_t depends on the weights and contributes to all subsequent losses, the expression can be reformulated as Equ. (5.14):

$$\frac{\partial L}{\partial w} = \sum_{t=1}^{T} \left(\sum_{i=t}^{T} \frac{\partial l_i}{\partial h_t} \right) \frac{\partial h_t}{\partial w} \tag{5.14}$$

Further, by unrolling the recurrence and incorporating the chain of gradients from future time steps, we obtain Equ. (5.15):

$$\frac{\partial L}{\partial w} = \sum_{t=1}^{T} \left\{ \sum_{i=t}^{T} \left(\prod_{j=i}^{T-1} \frac{\partial l_{j+1}}{\partial l_j} \right) \frac{\partial l_i}{\partial h_t} \right\} \frac{\partial h_t}{\partial w} \tag{5.15}$$

The weight update is then performed via the rule defined at Equ. (5.16):

$$w_{\text{new}} \leftarrow w_{\text{old}} - \frac{\partial L}{\partial w} \tag{5.16}$$

At convergence, the training process minimises the loss L such that the optimal weights w^* satisfy Equ. (5.17).

$$\frac{\partial L(w^*)}{\partial w} \approx 0. \tag{5.17}$$

dLIF Neuron:

During training, model parameters such as synaptic weights (W) and timing factors ($\hat{\alpha}, \hat{\beta}$) are automatically learned to optimise network performance based on the loss

function Zheng et al. (2024). The Equ. (5.18), (5.19), (5.20) shows the details of the BPTT where δ denotes the gradient of the loss function L concerning specific variables Zheng et al. (2024). The BPTT for dLIF utilises gradient descent with the chain rule to update parameters while dealing with non-differentiable spiking activities using a soft multi-Gaussian curve shown in Equ. (5.21) as a surrogate gradient function, with parameters like γ, h which will influence the magnitude, σ, s which will influence the width Zheng et al. (2024).

$$\delta u^{t,l} = \beta^l \odot \delta u^{t+1,l} + H' \odot \delta o^{t,l} \tag{5.18}$$

$$\delta i_d^{t,l} = (1-\beta^l)R \odot \delta u^{t,l} + \alpha_d^l \odot \delta i_d^{t+1,1} \tag{5.19}$$

$$\delta o^{t,l} = -u_{th}\delta u^{t+1,l} + \sum_d W_d^{l+1^T}(1-\alpha_d^{l+1}) \odot \delta i_d^{t,l+1} + \sum_d U_d^{l^T}(1-\alpha_d^l) \odot \delta i_d^{t+1,l} \tag{5.20}$$

$$H' = \frac{\delta o^t}{\delta u^t} = \gamma(1+h)N(u^t|U_{th},\sigma^2) - \gamma h N(u^t|\sigma,(s\sigma)^2) - \gamma h N(u^t|-\sigma,(s\sigma)^2) \tag{5.21}$$

The gradients of the parameters, as presented in Equ. (5.22), (5.23), and (5.24), are computed using this surrogate gradient function. This approach is particularly effective around the firing threshold (u_{th}), where neurons generate spikes, facilitating efficient training and enhancing the performance of the dLIF Zheng et al. (2024).

$$\delta W_d^l = \sum_t (1-\alpha_d^l) \odot \delta i_d^{t,l} o^{t,l-1^T}, \delta U_d^l = \sum_t (1-\alpha_d^l) \odot \delta i_d^{t+1,l} o^{t,l^T} \tag{5.22}$$

$$\delta\hat{\beta}^l = \sum_t \delta\beta^{t,l} \odot (1-\delta\beta^{t,l}), \delta\beta^{t,l} = u^{t-1,l} \odot \delta u^{t,l} - R\sum_d i_d^{t,l} \odot \delta u^{t,l} \tag{5.23}$$

$$\delta\hat{\alpha}_d^l = \sum_t \delta\hat{\alpha}_d^{t,l} = \sum_t \alpha_d^{t,l} \odot (1 - \delta\alpha_d^{t,l}), \alpha_d^{t,l} = i_d^{t-1,l} \odot \delta i_d^{t,l} - I_d^{t,l} \odot \delta i_d^{t,l} \quad (5.24)$$

5.3.4 Performance Metrics

The dataset used for model training and evaluation includes labeled data for each 12-second time window as previously papers used as a baseline reference Herbozo Contreras et al. (2024a,c); Saab et al. (2020); Tang et al. (2022); Wang et al. (2017); Yang et al. (2023, 2022b). A seizure is identified when the model accurately predicts an ictal event (class 1) based on the ground truth, while non-seizure events are labeled as class 0. Seizure datasets are often imbalanced, with far more non-seizure data points, since seizures are rare and brief. Over a day, seizure occurrences represent only a small fraction of the data within a 12-second window. Using traditional metrics like accuracy is misleading in this scenario, as they tend to favor the majority class and do not effectively measure the model's ability to detect seizures, the primary goal. For this reason, we rely on the AUROC, which evaluates both sensitivity and specificity in a threshold-independent manner.

5.3.5 Implementation Details

We independently trained the model on both scalp EEG and iEEG data for 200 epochs, with a learning rate of 0.001 for the base parameters of the dendrite model. Adam was selected as the optimiser. The time membrane constants for the dendrite layers were initialised uniformly. We employed Xavier uniform distribution for the liquid model to initialise the learnable weights. To distinguish between the seizure (1) and non-seizure (0) classes, we used Mean Squared Error Mean Squared Error (MSE) as the loss function. Additionally, in our experiments with the scalp-EEG dataset, we reduced the sampling rate to challenge our model with lower-resolution data. This modification significantly reduced memory computation and enabled training to be approximately twice as fast. Lastly, we conducted edge inference on the Raspberry

Pi 5. Fig. 5.3(A) illustrates the overall workflow, detailing both the training and inference stages. In addition, Fig. 5.3(B) presents a conceptual framework in which our model is applicable to wearable and implantable systems, targeting both monitoring and responsive stimulation using edge-based AI. This approach minimises the need for communication with cloud-based AI systems. Training parameters for each dataset can be found in Table 5.1, and the training algorithm based on Back-Propagation through time is detailed in Algorithm 5.1.

Table 5.1 – Summary of hyperparameter settings for each dataset.

Dataset	TUH	CHB-MIT	EPILEPSIAE
Input Channel	19	22	30-114
Output Channel	2	2	2
Sampling rate(Hz)	125/250	256	256
Time Window(s)	12	12	12
Dendritic Branches	4	4	4
Connection Type	Recurrent	Recurrent	Recurrent
Network structure	I-50R-100R-100R-2	I-50R-100R-100R-2	I-50R-100R-100R-2
Loss	Cross-Entropy (CE)	CE	CE
Batch size	200	200	256
Optimiser	Adam	Adam	Adam
Learning rate Learning Rate (LR)	1e-2	1e-2	1e-2
LR decay (per 100)	0.1	0.1	0.1
LR decay type	Step	Step	Step
Tau initialiser	Uniform	Uniform	Uniform
Tau Range	[0,4]	[0,4]	[0,4]
Reset Mechanism	Soft	Soft	Soft

5.4 Results

5.4.1 Neural Network Analysis

Timing Factors:

We decided to support our findings by extracting crucial information through the analysis of both the heterogeneous timing factors of the dendrite branches and the liquid timing factors. In Fig. 5.4, the first 2 layers, which are based on the dendrite spiking neurons, show a higher density at higher timing factors, which are related to the heterogeneous system Perez-Nieves et al. (2021); Zheng et al. (2024). Interestingly,

Algorithm 5.1: Training Liquid-Dendrite with BPTT

Input: Training data $B = \{x_t, y_t\}$, Timestep T, Epoch E, Number of branches N_b.
Input: Learning rate η and optimiser
Initialise weights: W, v ; `// Initial weights`
for *each* $e \in E$ **do**
  $B = \{x_t, y_t\}$; `// Shuffle training data`
  $u_0, s_{0_d}, s_0, u_{L_0}, s_{L_0}$; `// Initial neuron states`
  for *each* $t \in T$ **do**
    Update Dendrite Hidden States:;
    $s_{h,t}, u_{h,t}, s_{h,t_d} = \hat{f}_s(x_{t-1}, [u_{h,t-1}, s_{h,t-1}, s_{h,t-1_d}])$;
    Update Liquid Hidden States:;
    $s_{L_{h,t}}, u_{L_{h,t}} = \hat{f}_s(x_{t-1}, [u_{L_{h,t-1}}, s_{L_{h,t-1}}])$;
    $\hat{y}_t = \hat{f}_s(s_{h,t}, [u_{o,t}, s_{o,t}])$; `// Predictions`
    $\ell_W = \sum_{t=1}^{T} \ell(y_t, \hat{y}_t)$; `// Loss`
    $W = W - \eta \nabla_W \ell(W)$; `// Update Weight`

the liquid time factors show a normal distribution around 0.5. This could be due to the nature of the dense functions that govern their values during training.

Membrane Potential:

We extracted membrane potentials for two randomly selected classes. Class 0, which is non-seizure, shows lower membrane potential, which, in turn, makes the system spike less. For class 0 (seizure data), there are more membrane potential activations. One explanation for these findings is that the model is more sensitive to identifying seizure data due to the higher voltage/amplitude of the input signal. At the same time, the state of class 0 has a lower amplitude than a seizure file. The simulating steps correspond to 12 seconds of data sampled at 125 Hz.

5.4.2 In-Sample Generalisation

We evaluated in-sample generalisation using three distinct datasets: two from scalp EEG (one adult and one pediatric) and one from iEEG. In-sample validation was

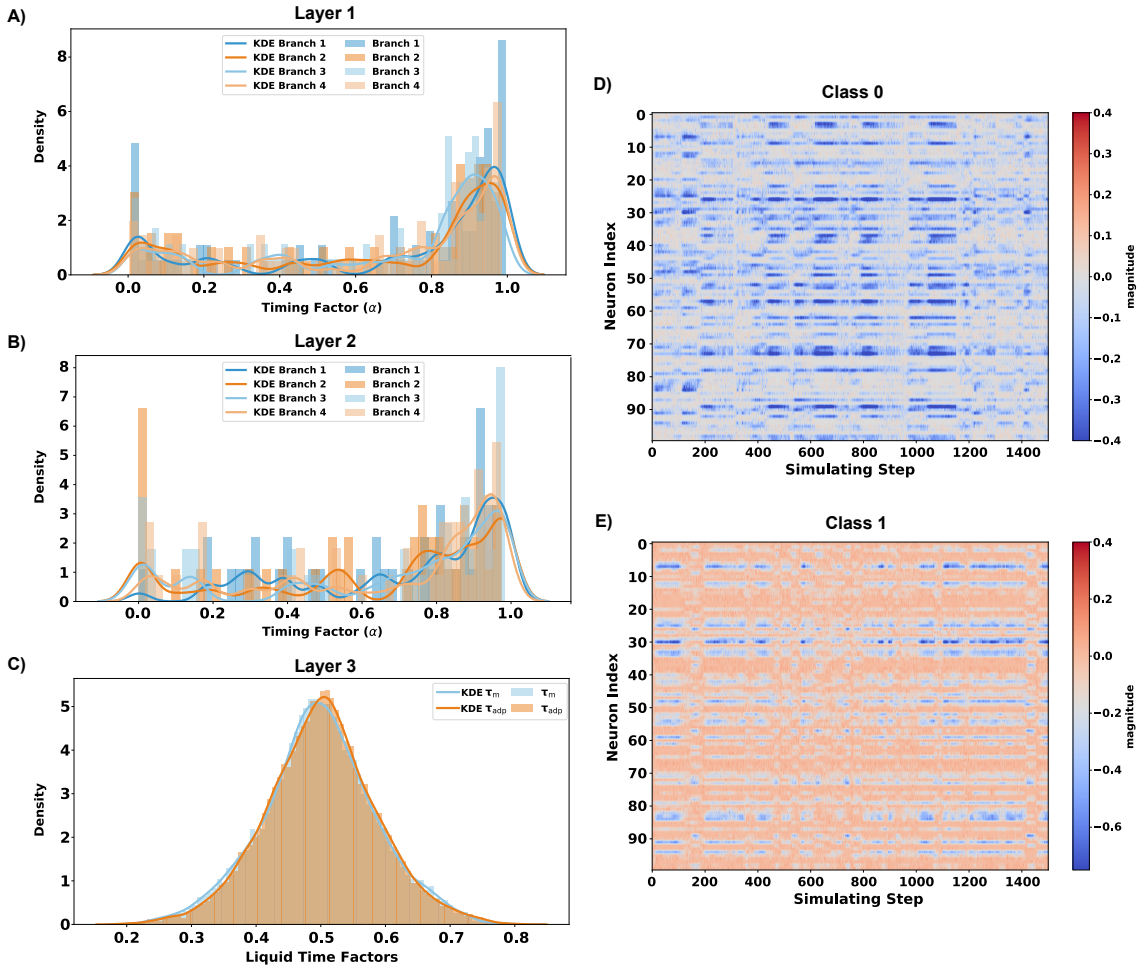

Figure 5.4 – Analysis of learned timing factors and membrane potentials. Our neural network is composed of two layers, each based on dendrites, with the timing factor of each branch showing higher consolidation for higher timing factors in both layer 1 (A) and layer 2 (B). Layer 3 has 2 liquid time factors, and overall density follows a normal distribution as both factors are sigmoid (C). Membrane potentials before the last 2 neurons for a random class 0 (D) and a random class 1 different (E) show distinct accumulations, which are inherently related to the dynamics of the input data.

performed by partitioning each dataset into training and validation subsets, and our findings were compared against existing benchmarks using the same baseline conditions.

Scalp-EEG:

Our novel architecture achieved an AUROC of **83%** in the time domain on the TUH dataset. As depicted in Fig. 5.5, our model has stable training loss, and metrics for AUPRC and AUROC are also provided. Table 5.2 shows that our model excels in performance despite using a time domain, spiking network, and a very small memory size with fewer parameters and a lower sampling rate. For the CHB-MIT pediatric dataset, we achieved an AUROC of **96%**, which we compared against another state-of-the-art model under the same basis as depicted in Table 5.3. The number of patients and patient ID reported were selected to provide a direct comparison with previous studies utilising the same cohort.

Table 5.2 – State-of-the-art performance comparison on the TUH dataset.

Model	GraphNN Tang et al. (2022)	ConvLSTM Yang et al. (2022b)	Transformer Ma et al. (2023)	LTC-FPTT Herbozo Contreras et al. (2024a)	TinydLIF Herbozo Contreras et al. (2024c)	This work
Signal Domain	Time-Frequency	Time-Frequency	Time-Frequency	Time-Frequency	Time	**Time**
AUROC	0.82	0.84	0.92	0.89	0.81	**0.83**
Network Type	ANN	ANN	ANN	SNN	SNN	**SNN**
Memory Size	-	≈ 4.5MB	-	4MB	1.5MB	**535 KB**
Parameters	-	≈ 190K	-	≈ 500K	300K	**130K**
Data Type	Float32	Float32	Float32	Float32	Float32	**Float32**
Pre-Processing	FFT	ICA, STFT	STFT	ICA, STFT	ICA	**ICA**
EEG Sampling Rate	200	250 Hz	250 Hz	250 Hz	125 Hz	**125 Hz**

Table 5.3 – CHB-MIT results: AUROC comparison of models using time-frequency and time-domain inputs.

Patient ID	Gender	Time-Frequency Domain			Time Domain	
		CNN Yang et al. (2023)	**ConvLSTM** Yang et al. (2023)	**SNN ConvLSTM** Yang et al. (2023)	**Tiny-dLIF** Herbozo Contreras et al. (2024c)	**dLIF-LTC (This work)**
1	F	100.0	100.0	96.4	98.6	100.0
3	F	100.0	100.0	92.3	100.0	100.0
5	F	100.0	100.0	95.3	86.0	100.0
6	F	94.9	93.4	88.8	93.2	88.2
7	F	97.0	96.5	92.0	96.7	100.0
8	M	99.4	98.2	72.5	98.3	96.2
9	F	99.7	100.0	96.1	98.1	91.2
10	M	100.0	100.0	97.5	97.8	100.0
11	F	98.8	98.9	81.2	97.7	98.3
15	M	99.9	99.7	73.0	95.1	81.9
17	F	88.2	91.0	87.7	94.3	97.5
18	F	98.5	96.9	86.0	97.3	93.7
19	F	95.4	99.8	95.9	92.3	99.6
20	F	97.1	97.7	87.9	96.0	98.4
21	F	99.3	100.0	92.5	81.4	90.1
22	F	100.0	100.0	96.6	90.3	97.6
23	F	100.0	100.0	95.9	88.7	96.0
Average	3M/14F	98.13	98.36	89.9	94.2	95.8

M: Male, F: Female.

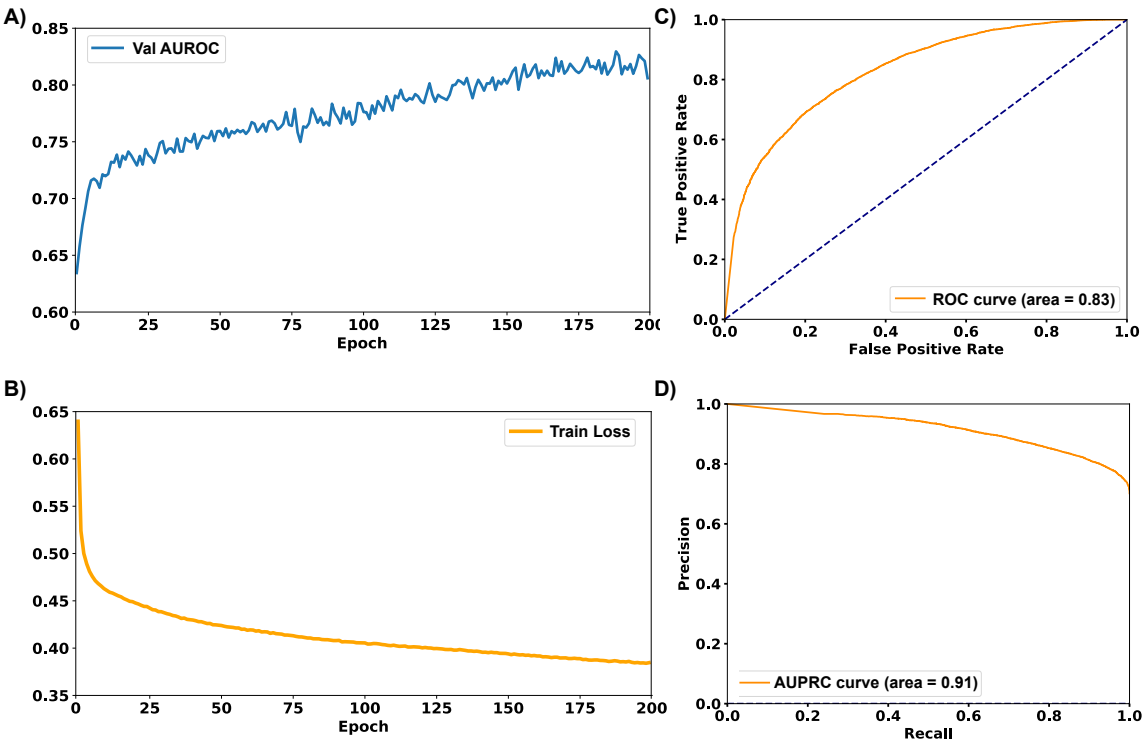

Figure 5.5 – Training and validation results across the TUH epilepsy dataset. The AUROC (A) validation metric improves over time. Train Loss converges (B). AUROC (C) and AUPRC (D) curves are shown and indicate optimal performance of our network.

Intracranial-EEG Results:

When applied to iEEG data, the model achieved an AUROC of **93%**, comparable to larger models that rely on more complex feature extraction techniques and traditional artificial neural networks. Overall, we outperform state-of-the-art methods tested in the Time-Frequency Domain and trained in the ANN domain by almost 2%, with a lower memory footprint. Results per patient are shown in Table 5.4.

5.4.3 Edge-Inference Raspberry Pi

We performed inference on the edge by utilising a Raspberry Pi 5. Although this version has more memory, our Random Access Memory (RAM) consumption remained limited to 2 GB. We chose the Raspberry Pi 5 because it supports the latest PyTorch

Table 5.4 – EPILEPSIAE-iEEG Results: AUROC comparison of models using time-frequency and time-domain inputs.

Patient ID	Gender	Time-Frequency Domain				Time Domain	
		ConvLSTM Yang et al. (2022b)	**TE** Iasemidis et al. (2003)	**STLmax** Hutson et al. (2018)	**SNN ConvLSTM** Yang et al. (2023)	**Tiny-dLIF** Herbozo Contreras et al. (2024c)	**dLIF-LTC (Ours)**
1	F	90.3	65.6	66.3	81.7	77.2	87.6
2	F	98.1	83.3	82.4	89.0	100.0	79.0
3	M	96.6	78.5	87.4	93.5	96.8	100.0
4	F	98.7	85.4	86.6	91.0	92.7	100.0
5	M	90.8	61.6	58.5	76.4	77.6	92.9
7	M	80.0	64.7	62.5	70.1	98.9	98.7
8	M	98.5	83.4	80.6	91.5	56.3	93.6
9	M	80.4	61.6	65.8	68.2	99.8	91.2
10	M	83.4	49.2	56.4	75.0	85.1	95.5
11	F	98.4	64.2	63.8	90.6	100	84.9
12	F	100.0	78.2	77.6	83.1	95.2	96.4
13	F	86.4	82.8	87.0	83.0	94.0	93.0
14	F	94.5	47.6	51.5	76.1	93.2	95.8
15	M	91.3	62.6	61.5	73.5	91.7	98.9
Average	7M/7F	91.9	69.1	70.5	81.2	89.8	93.4

M: Male, F: Female.

version and Python framework. To be able to facilitate the process at the edge, we decoupled the graph computations from the previously trained model, which was originally executed on a GPU. This approach enabled the Raspberry Pi Central Processing Unit (CPU) to process key parameters, such as voltage, membrane potential, and currents at both the soma and dendritic compartments, thereby maintaining computational compatibility and efficiency.

Latency:

We measured the inference time for 1 second of EEG data comprising 256 readings sampled at 125 Hz. By averaging the time taken to process all batches of 10,000 EEG signals, we observed an inference time of 0.81 seconds per batch. We recorded an average on-the-fly processing time of 3.1 ms per EEG reading, without a drop in performance.

5.4.4 Robustness test

We conducted a robustness experiment by randomly removing channels. The primary motivation was to evaluate the model's resilience in real-world scenarios, particularly for wearable devices, where certain channels may become disrupted or lose information

over time. A critical question we aimed to address was: **Can our model maintain commendable performance despite missing channel information?**

TUH dataset:

We experimented to evaluate the robustness of our neural network on the TUH dataset by randomly blacking out between 10% and 90% of the EEG channels during each batch. This approach introduces additional variability and builds confidence in the model's resilience to channel reduction. The results in Fig. 5.6 demonstrate that the neural network maintains strong performance even as the number of available channels decreases. Notably, with a 50% reduction in channels, the model achieved approximately 76% accuracy. This performance is competitive compared to other neural networks Ahmedt-Aristizabal et al. (2020); Tang et al. (2022); Wang et al. (2017), even when utilising the maximum number of channels.

CHB-MIT dataset:

Similarly, we conducted an experiment. Fig. 5.6 shows the average of all patients in 5 trials. We randomly blackout channels from 10% and 90% to evaluate the performance on the children dataset. Although results are slightly worse than TUH dataset, our model tends to have an acceptable resilience up to a 50% of reduction channels, which highlights the importance of including liquid-time constants and heterogeneous behaviour in the neural network. The findings on both datasets highlight the advantages of energy-efficient systems where channel dependency can be addressed by utilising this model.

5.4.5 Ablation experiments

We conducted a series of ablation experiments on our neural network to support the findings that adding both dendrites and liquid neurons in a neural network is essential for efficiency. Fig. 5.7 illustrates the ablation study conducted on the three datasets

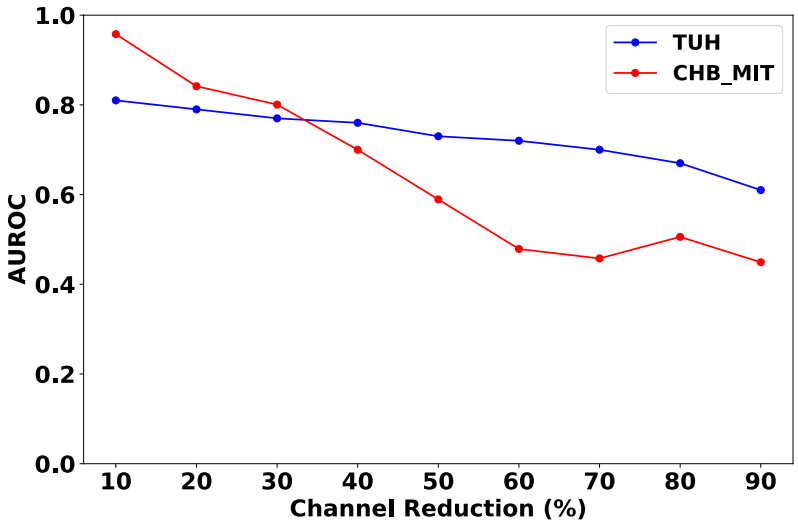

Figure 5.6 – Black-out channels results on the TUH and CHB-MIT dataset.

used in this work. The bars represent the AUROC for each ablated configuration (left y-axis), while the dashed orange line highlights the associated memory footprint in megabytes (right y-axis). As the ablations progress from right to left, simpler architectures (those omitting specific neuron types or layers) exhibit reduced memory consumption but lower AUROC values. Conversely, configurations that retain both liquid and dendritic neurons demonstrate progressively higher AUROC, signifying more accurate seizure detection, at the expense of a small increase in memory. Notably, our base model achieves the highest performance across all three datasets and requires only about 550 KB of memory. These results underscore the importance of both Liquid and Dendritic Neurons in optimising seizure detection accuracy.

5.4.6 Evaluating SNN Performance Across Different Training Steps

We conducted experiments to evaluate the impact of varying training timesteps T on our spiking neural network. Given that our analysis is performed in the time domain, these experiments assess the model's ability to capture long-term dependencies within both scalp and intracranial signals. As described in Fig. 5.8, the network begins to exhibit commendable performance at higher timesteps, with further increases in

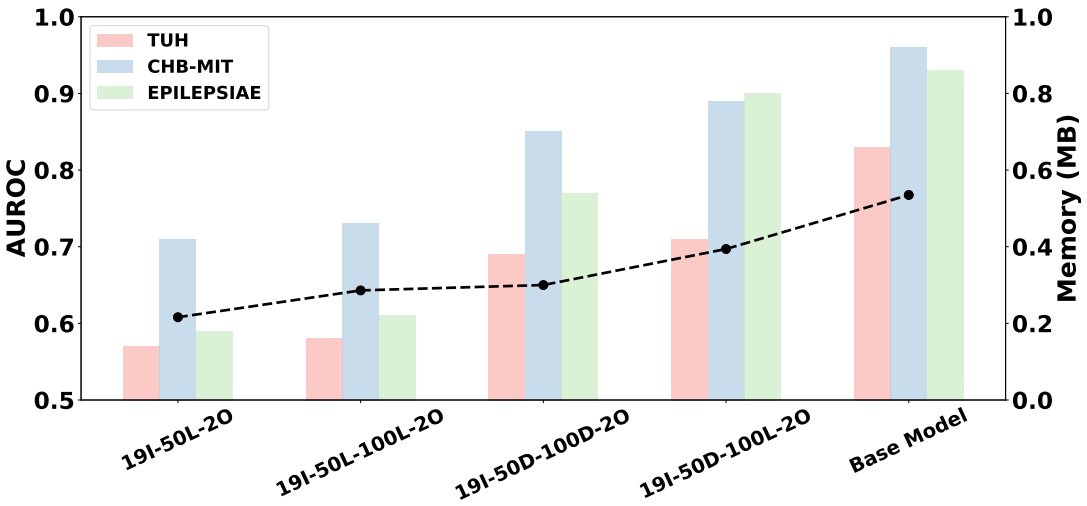

Figure 5.7 – Ablation experiment results in the three datasets used: TUH, CHB-MIT, and EPILEPSIAE. An effective seizure detection occurs when Liquid and Dendrite Neurons work together. Results reported at epoch 150.

timesteps correlating with improved accuracy, thereby confirming its capacity to fully encapsulate the temporal information present across all datasets. We also studied the impact of longer timesteps, and found that, compared with $T = 100$, the highest AUROC we reported was achieved with a latency factor within an order of magnitude of 2x, with almost 14 times greater. These findings highlight a trade-off that does not grow exponentially and are therefore essential for the model to capture long-term dependencies in tasks such as brain signals.

5.4.7 Energy Consumption Estimation

Theoretical Estimations:

We provided an extensive list of theoretical calculations for each of the neurons utilised. As the neurons are used in a recurrent manner, the LTC-SNN overhead is calculated and is approximately similar to the Adaptive Recurrent Spiking Neural Network Yin et al. (2020). The dLIF is modelled as a LIF neuron, but requires additional calculations due to the multiplexing of dendritic branches. Table 5.5 breaks down the equation based on the computation overhead for them, which also depends

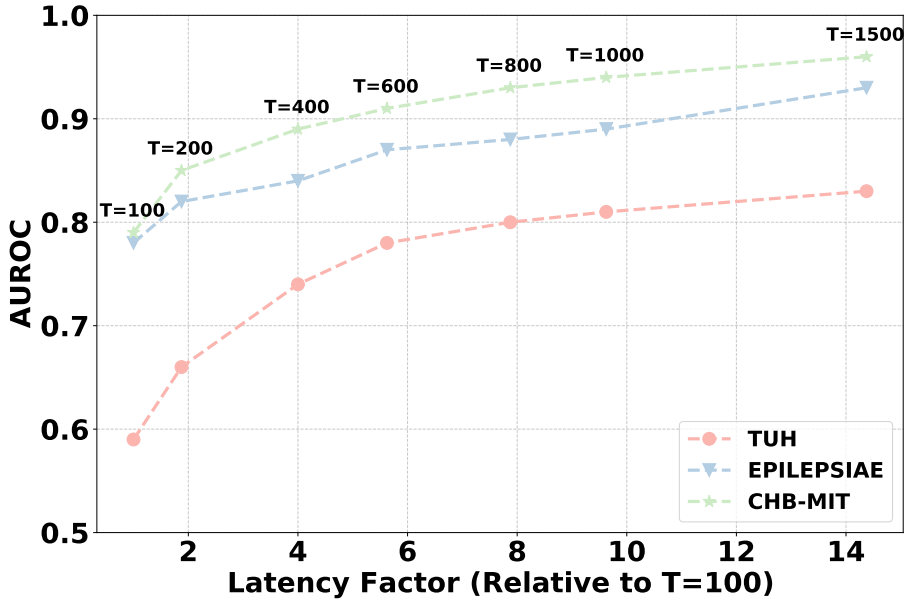

Figure 5.8 – Timesteps analysis in the TUH, CHB-MIT and EPILEPSIAE used for this work. The neural network can handle long timestream dependencies in both scalp and intracranial datasets. T: Timesteps.

on the input and output firing rates for each layer.

Table 5.5 – Theoretical equations of energy computation per layer.

Model	Parameter	Energy	Total$^{\diamond}$
RNN-LIF†	Current(Equ. (5.10))	$(mnF_{rin} + n^2F_{rout})E_{AC}$	$(mnF_{rin} + n^2F_{rout})E_{AC} + nE_{MAC}$
	Voltage(Equ. (5.9))	nE_{MAC}	
LTC-SNN	Current(Equ. (5.10))	$(mnF_{rin} + n^2F_{rout})E_{AC}$	$(mnF_{rin} + (n^2 + 2n)F_{rout})E_{AC} + 3nE_{MAC}$
	Voltage(Equ. (5.9))	nE_{MAC}	
	Adaptive(Equ. (5.8))	$n(E_{MAC} + E_{AC}F_{rout})$	
	Threshold(Equ. (5.8))	$n(E_{MAC} + E_{AC}F_{rout})$	
dLIF	Current(Equ. (5.3))$^{\otimes}$	$(mnF_{rin} + (2D+1)n^2F_{rout})E_{AC}$	$(mnF_{rin} + (2D+1)(n^2 + 2n)F_{rout})E_{AC} + nE_{MAC}$
	Voltage(Equ. (5.2))	nE_{MAC}	

† Recurrent Leaky Integrate & Fire Neuron RNN-LIF.
$^{\diamond}$ Multiplication and Accumulation Operations (E_{MAC}), Accumulation Operations (E_{AC}), Dendrite Branches (D), Feedforward input (n), Recurrent input (m), Firing output rate (F_{rout}), Firing input rate (F_{rin}).
$^{\otimes}$ The current is the total amount for both soma and dendrite compartments.

Spiking rate shows sparsity:

We calculated the spiking rates for each of the neurons by averaging the spikes that occur for the layers of the dendrites and liquid neurons. For the TUH dataset, we

found that the average firing rate per timestep is around 19% and 21% for classes 0 and 1, respectively. For the CHB-MIT dataset, the spiking rate was 10% higher for both classes. A high firing rate is observed in the EPILEPSIAE dataset, with values around 40%. One explanation for this phenomenon is the brain signal's inherent characteristics. As the EPILEPSIAE dataset has a higher sampling rate and a higher number of electrodes than both scalp EEGs, it is reasonable to expect more firing rate patterns to be discriminated. As well, the signal amplitude of iEEG is higher than the scalp EEGs, which further supports the generality of our neural network to adapt to different firing rates. These results, which are depicted in Table 5.6, are indicative of an effective sparse system while achieving high AUROC values for the three datasets.

Table 5.6 – Spiking rate of the layers in the liquid-dendrite model across the three datasets used in this work.

Dataset	Layer	Spikes/T		Firing Rate †	
		Class 0	Class 1	Class 0	Class 1
TUH	**dLIF-1**	10	7	19%	14%
	dLIF-2	14	19	14%	19%
	LTC-SNN-1	23	29	23%	29%
	Average	16	18	19%	21%
EPILEPSIAE	**dLIF-1**	29	21	58%	42%
	dLIF-2	34	39	34%	39%
	LTC-SNN-1	40	31	40%	31%
	Average	34	30	41%	37%
CHB-MIT	**dLIF-1**	22	25	44%	50%
	dLIF-2	18	25	18%	25%
	LTC-SNN-1	31	33	31%	33%
	Average	24	28	28%	33%

† Average Firing Rate is calculated based on no. of spikes per timestep (T) divided by total #neurons.

5.5 Limitations

5.5.1 Learning Rule

Currently, our method employs BPTT for training. In future work, we plan to explore FPTT, which will enable the analysis of more time steps with reduced computational complexity, thereby enhancing efficiency. Additionally, the training speed of our method prompts us to consider alternative training methodologies that could accelerate the process, which can be enhanced with encoding schemes Hu et al. (2023). Moreover, quantisation-aware training can lead to faster inference times Eissa et al. (2023).

5.5.2 Extending to Out-of-sample generalisation

Our current training approach does not yet address out-of-sample generalisation tasks. However, the model's performance across a diverse dataset of different brain signals indicates its potential effectiveness in other tasks.

5.5.3 Overhead of learnable parameters

One of our next aims is to deploy our model for training on FPGA devices, such as the PYNQ-Z2, to support efficient, scalable edge computing solutions Pes et al. (2024). However, the liquid part of the model involves different learnable parameters, which are modelled in an exponential factor. To achieve this, we will need to optimise the learnable parameters to minimise additional computational load.

5.6 Conclusion

The proposed method demonstrates efficiency even with a smaller model size, making it well-suited for demanding tasks such as seizure detection. We further challenged

our model by analysing data purely in the time domain rather than relying on conventional time-frequency-domain feature extraction. This approach simplifies the data processing pipeline and offers significant advantages towards an envisioned on-the-edge training. Further, a reduction in the sampling rate didn't affect performance. This methodology can be particularly beneficial for neuromodulation devices, where privacy concerns may arise from the transfer of sensitive information to cloud servers. By enabling real-time, on-device learning and analysis, our approach can mitigate these concerns while ensuring efficient, effective performance in critical applications.

Chapter 6

Adaptive Neuromorphic Continual Edge Training for Seizure Detection and Prediction

This chapter shifts from efficient neuromorphic inference to patient-specific on-device training, demonstrating few-shot transfer learning on a neuromorphic chip for seizure detection and prediction.

The content presented in this chapter is under revision as:

- Herbozo Contreras, L. F., Yu, L., Huang, Z., Aguilar, I., Nikpour, A., and Kavehei, O. (2025). "Neuromorphic Neuromodulation: A Low-Power Edge-Training Framework for the Future of Personalised and Closed-Loop Neurostimulation". *Under Review; Minor*

Statement of Contributions of Joint Authorship

- Luis Fernando Herbozo Contreras (Candidate): First author, formal analysis, investigation, methodology, resources, software, visualisation, validation, writing original draft, writing review & editing of the manuscript
- Leping Yu: contributed to writing review & editing of the manuscript
- Zhaojing Huang: contributed to writing review & editing of the manuscript
- Isabelle Aguilar: contributed to writing review & editing of the manuscript
- Armin Nikpour (Co-Supervisor): conceptualisation, data curation, funding acquisition, investigation, project administration, resources, supervision, validation, writing review & editing
- Omid Kavehei (Principal Supervisor): conceptualisation, data curation, funding acquisition, investigation, methodology, project administration, resources, supervision, visualisation, validation, writing review & editing.

In addition to the statements above, in cases where I am not the corresponding author of a published item, permission to include the published material has been granted by the corresponding author.

Luis Fernando Herbozo Contreras Date: 28 May 2026

As supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

Prof. Omid Kavehei Date: 28 May 2026

Chapter Summary

Closed-loop neurostimulation requires intelligent systems capable of operating autonomously under constraints on latency, energy consumption, and privacy for effective seizure treatments. However, these devices rely on cloud-based AI, which introduces significant challenges, including high false-positive rates, latency issues, and privacy concerns. In this chapter, we present a neuromorphic framework for real-time seizure detection and prediction, implemented directly on a neuromorphic System-on-Chip (SoC) with on-chip learning. By leveraging SNN and few-shot edge learning, our system enables continual, patient-specific adaptation without requiring data transmission or cloud connectivity. For a detection framework, a model pre-trained on the TUH dataset is deployed on the BrainChip Akida processor and personalised using long-term EPILEPSIAE recordings. For prediction, we demonstrate robust performance on both scalp EEG and iEEG from the CHB-MIT and Freiburg datasets, using a leave-one-seizure-out strategy for edge training. Across datasets, our approach achieved superior performance to state-of-the-art neural networks in metrics such as AUROC, Sensitivity, and False Positive Rate (FPR), while drastically reducing the memory footprint through quantisation and lowering energy consumption by orders of magnitude. This work establishes a foundation for the next-generation of closed-loop neurostimulation systems.

6.1 Introduction

A substantial proportion of individuals with drug-resistant seizures require therapeutic strategies beyond pharmacological intervention. Although advances have improved outcomes, approximately 30–40% of individuals remain refractory to medication, highlighting the need for alternative interventions Banerjee et al. (2009); Fisher et al. (2014). Neurostimulation modalities such as the RNS have demonstrated efficacy in reducing seizure frequency in drug-resistant cohorts by sensing and sending electrical stimulation to the brain. Nevertheless, these systems exhibit several inher-

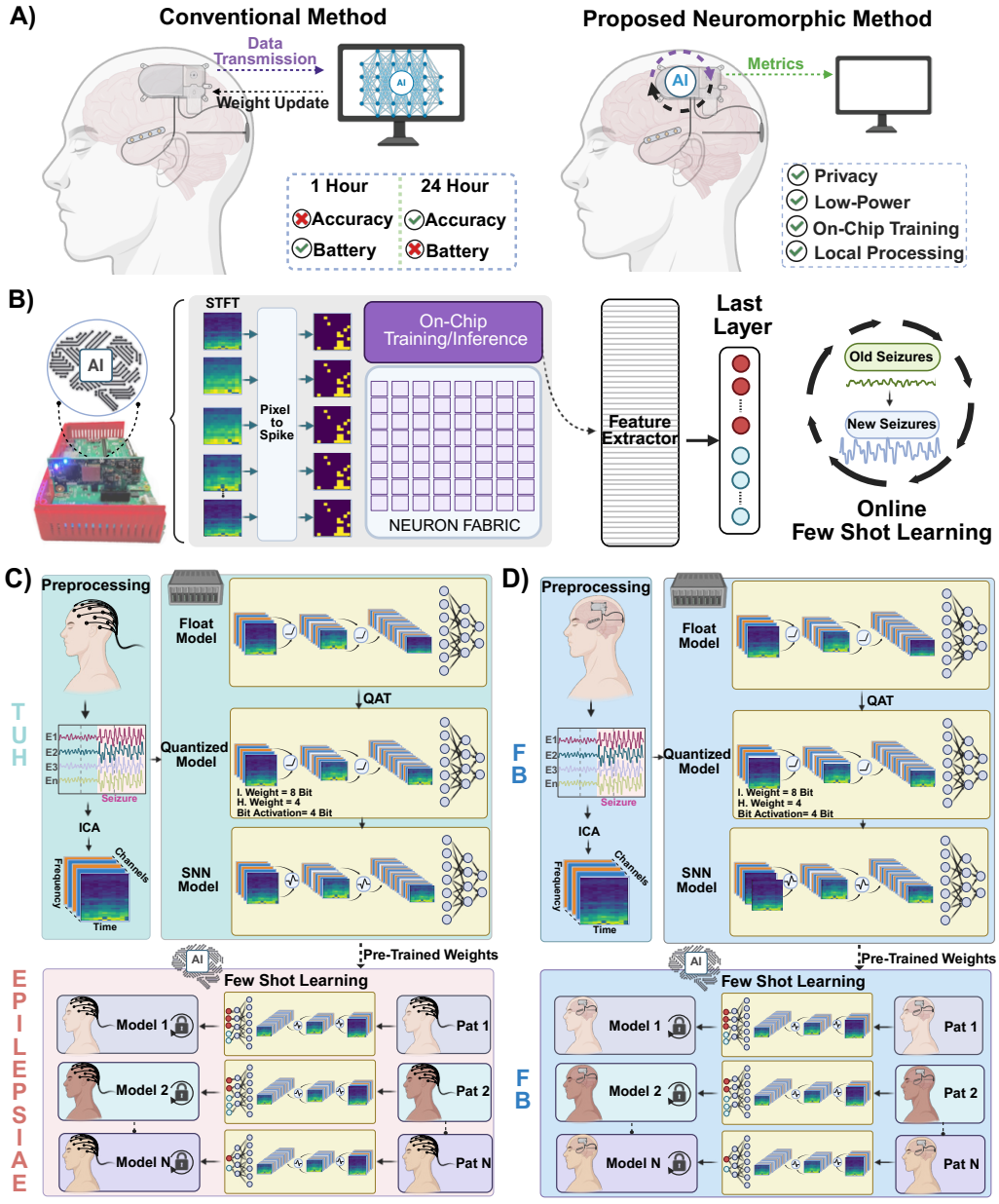

Figure 6.1 – Edge-based personalised seizure detection and prediction frameworks. Conventional Methods for treating epilepsy rely on AI outside the device, while our proposed method envisions an efficient neuromorphic-AI system-on-chip closed-loop neurostimulation device without compromising power consumption (A). Continuous STFT of EEG images is spike-encoded and passed through a frozen, pre-trained convolutional feature extractor at the edge. Only the final classifier layer is updated on-chip via online few-shot learning, reducing compute and memory overheads (B). For the detection framework, a CNN is trained on a large-scale dataset TUH, quantised to mixed precision and converted to a spiking neural network SNN, and edge-adapted with patient-specific data from a long-term recording dataset (EPILEPSIAE) for patient-specific (C). For the prediction pipeline, the same methodology is applied to the Freiburg and CHB-MIT datasets. Because of limited patient numbers, individualised models are pre-trained rather than a global model and subsequently refined on previously unseen patient data at the edge, yielding an adaptable cloud-independent approach (D).

ent limitations.

First, existing neurostimulation platforms rely on pre-trained detection models that often inadequately capture the considerable inter-patient heterogeneity in seizure morphology and dynamics, leading to elevated rates of both false positives and false negatives. Second, reliance on cloud-based inference pipelines results in communication latency during acute seizure events and introduces potential privacy vulnerabilities Jarosiewicz and Morrell (2021). Third, power and bandwidth constraints on implantable devices restrict iEEG telemetry to approximately one hour per day. While extending this window of data could enhance model training, it would accelerate battery depletion and compromise device longevity.

Although model generalisation measured via performance on distinct out-of-sample datasets is the prevailing metric for assessing adaptability to novel data, it does not necessarily predict performance on the continuously evolving biosignals of a single patient Ghorbani et al. (2024). Accordingly, we advocate shifting the focus from population-level generalisation to individual-level personalisation, in which a detection and prediction algorithm continuously refines itself based on an individual's unique neural signature at the edge. Personalised medicine, which aligns therapeutic strategies to an individual's physiological profile, has demonstrated superior clinical utility across numerous domains. In epilepsy management, real-time, patient-specific adaptation of seizure detection and forecasting algorithms promises to revolutionise neurostimulation efficacy. However, realising such personalisation necessitates overcoming challenges in real-time signal processing, ultra-low-power operation, and resource-constrained computation Yang et al. (2022b). To realise this shift from cloud-based AI to on-device intelligence, it is essential to review existing system-on-chip SoC solutions for seizure monitoring and their current limitations. Studies have demonstrated impressive advances in energy-efficient online training, achieving sensitivities (true positive rate) above 90%, specificities (true negative rate) exceeding 95–99%, and false-alarm rates below one event per hour while operating within the μJ-per-classification range Chua et al. (2022); Hsieh et al. (2022); Liu et al. (2021); Shin et al. (2022); Wang et al. (2021); Zhang et al. (2022). These architec-

tures are based on CNN Zhang et al. (2022), Support Vector Machine (SVM) and tree-based processors Hsieh et al. (2022); Shin et al. (2022), and unsupervised online classifiers Chua et al. (2022) illustrate the maturity of low-power seizure detection hardware. Nevertheless, several common limitations remain:

- **Slow adaptation:** Training and retraining cycles typically occur over seconds or longer.
- **Limited adaptability:** Most designs are detection-only systems without continual personalisation, and evaluations are often confined to short-term datasets (i.e CHB-MIT) and small patient cohorts, limiting insights into long-term variability.
- **Frame-based processing:** Computation remains clock-driven, operating on fixed feature windows instead of leveraging the sparse, event-driven nature of neural activity.
- **Incomplete reporting:** Performance metrics are primarily limited to sensitivity, specificity, and accuracy, with little emphasis on threshold-independent measures such as AUROC, reducing cross-study comparability.

These factors highlight the need for architectures capable of millisecond-scale continual learning, operation over long-term iEEG and scalp-EEG recordings, and efficient event-based computation capabilities. Neuromorphic computing offers a compelling pathway to meet these demands by replicating the brain's event-driven, energy-efficient processing paradigm. Compared to conventional digital architectures, neuromorphic systems can reduce computational energy consumption by several orders of magnitude Donati and Indiveri (2023), while supporting on-chip learning for continuous model adaptation without cloud reliance Herbozo Contreras et al. (2024a,b). Recent work has demonstrated online epileptic seizure detection on neuromorphic systems Burelo et al. (2022a); Gallou et al. (2024), though existing implementations are restricted to inference and lack support for continual learning. Such on-chip adaptability is essential to assimilate new data incrementally without catastrophic

forgetting, ensuring sustained performance during physiological drift, whether due to electrode repositioning, evolution of seizures, or pharmacological effects Aguilar et al. (2025); De Lange et al. (2021).

The key question motivating this work is whether it is possible to design a unified framework that bridges the gap between the state-of-the-art SoC's and the adaptive, event-driven principles of neuromorphic computing for the next generation of neurostimulation devices. In this work, we present a personalised neuromorphic framework for on-device seizure detection and prediction that performs few-shot learning on low-power hardware. This approach provides a scalable, privacy-preserving, and low-latency solution for adaptive neurostimulation. Our contributions include:

- Continual on-chip learning enables sustained performance improvement as new data arrive, reducing false-positive rates while enhancing sensitivity in both pipelines for 3 epilepsy datasets (2 long-term recording and 1 short-term recording).
- Rapid personalisation is achieved with as little as three minutes of patient-specific seizure data (15-shot learning).
- On-device processing enables real-time adaptation and data privacy.
- Hardware-efficient training confined to the final layer reduces computational cost.
- Energy and latency improve by 100 times compared with conventional architectures.
- Robustness in handling missing or corrupted EEG channels, simulating real-world scenarios of partial signal loss.

6.2 Methods

6.2.1 Datasets

Four datasets (1 iEEG and 3 Scalp-EEG) were employed in this chapter. For the detection framework, we used the two largest EEG datasets: the TUH Corpus and the EPILEPSIAE datasets. The TUH dataset, comprising over 642 patient recordings, serves as the principal training resource due to its wide variety of seizure types and recording conditions. From this dataset, we partitioned 592 patients for model training and reserved 50 independent sessions for inference validation. Due to its large scale and heterogeneity, the TUH Corpus provides a robust foundation for initial model development. Subsequently, we utilised the EPILEPSIAE dataset, comprising continuous, long-term scalp EEG recordings from 30 patients, to train and develop personalised models via few-shot transfer learning. Given the rarity of ictal events (seizure events) in real-world monitoring, the EPILEPSIAE dataset exhibits a highly imbalanced task, with inter-ictal (non-seizure) segments vastly outnumbering ictal segments. This imbalance presents both a challenge and an opportunity: it reflects the conditions of clinical monitoring while forcing our adaptive algorithms to contend with realistic data distributions during on-device personalisation. Both datasets have the same number of channels, which makes them suitable for this analysis.

For the Prediction Framework, we used two datasets: The CHB-MIT and the FB dataset iEEG. The CHB-MIT dataset comprises scalp EEG recordings from 23 pediatric patients, totaling 844 h of continuous data and 163 seizures Shoeb (2009). Recordings were acquired using 22 electrodes at a sampling rate of 256 Hz. Interictal periods are defined as intervals that begin at least 4 h before seizure onset and end at least 4 h after seizure termination. In cases where multiple seizures occur in close succession, specifically within 30 minutes of each other, they are treated as a single event, with the onset time of the earliest seizure taken as the combined seizure onset. For the seizure prediction task, we focus only on patients experiencing fewer than 10 seizures per day, as predicting seizures for individuals averaging one every 2 h is less clinically relevant. Applying these criteria yields 13 patients with adequate data

defined as having at least three leading seizures and 3 h of interictal recordings.

The FB dataset comprises iEEG recordings from 21 patients diagnosed with refractory epilepsy. Due to restricted availability, recordings from 13 patients were utilised. Signals were acquired at a sampling rate of 256 Hz. For each patient, six channels were selected from intracranial contacts, including three located within epileptogenic regions and three from non-epileptogenic (remote) areas. The dataset provides at least 50 minutes of pre-ictal activity and 24 hours of interictal recordings per patient. All the datasets used in this work are described in Table 6.1.

6.2.2 Pre-Processing

To address challenges inherent in raw EEG signals, we employed two preprocessing techniques: ICA and the STFT. EEG recordings were first segmented into 12-second intervals, after which ICA was applied to perform BSS, decomposing the signal into 19 statistically independent components. This process is represented in Equ. (6.1),

$$T \approx MA^T \tag{6.1}$$

where T denotes the EEG signal matrix, M captures the temporal dynamics, and A contains the spatial weights corresponding to the topographic maps. To identify components related to ocular artifacts, we computed the Pearson correlation between each independent source and the 'FP1' and 'FP2' EEG channels. Components exhibiting strong correlations with eye movements were removed, yielding artifact-reduced EEG signals. Following this cleaning step, we applied the STFT with a 1-second window (250 samples) and 50% overlap, while also discarding the DC component. This transformation resulted in data shaped as $(N \times 23 \times 125)$, where N is the number of electrodes, 23 corresponds to the time steps, and 125 represents the frequency bins. Prior to edge training, dataset preprocessing is performed offline and is therefore excluded from the energy calculations.

6.2.3 Detection vs Prediction

We addressed two distinct tasks: seizure detection and seizure prediction (forecasting). As illustrated in Fig. 6.2(A), the EEG data are categorised into three distinct states for analysis: interictal, pre-ictal, and ictal. In detection, the objective is to discriminate between background activity (interictal) and seizure activity (ictal) at the time the seizure occurs. Here, the model essentially classifies whether the signal currently belongs to a seizure event. In contrast, prediction focuses on anticipating seizures before onset. The task involves distinguishing background activity from two additional states within the pre-ictal State: Seizure Occurence Period (SOP) and the Seizure Prediction Horizon (SPH). The SOP refers to the time window in which a seizure is expected to begin. The interval between an alarm and the start of the SOP is defined as the SPH. A prediction is considered correct if the seizure onset occurs after the SPH and within the SOP. Conversely, a false alarm occurs when the system issues a positive prediction but no seizure occurs during the SOP. Once triggered, an alarm remains active until the SOP concludes.

6.2.4 ANN Structure

We designed a compact CNN tailored for EEG spectrograms Truong et al. (2018). The model comprises three convolutional blocks: each includes a 2D convolution (filters: 16, 32, 64), batch normalisation, ReLU activation, and max pooling. The first layer uses a kernel across all EEG channels ($N \times 3$) to capture spatial-frequency patterns where N is the number of channels. The output is flattened and passed through a dense layer (128 units) with ReLU and dropout (0.5), followed by a softmax classifier. An important distinction of Akida is that it only allows odd kernels to be computed. In the EPILEPSIAE Dataset, the channels are odd numbers, whereas in the CHB-MIT and FB datasets, the channels are even. To meet the neuromorphic requirements, we select one channel (0) from the CHB-MIT and FB datasets whose STFT magnitudes were very different from those of other channels, to ensure on-device fitting by processing the channels as odd kernels.

6.2.5 Quantisation Aware Training

The BrainChip Akida architecture was used to run our work for personalisation at the edge. However, Akida supports only integer arithmetic for both inference and on-chip learning, ruling out any floating-point operations. To address these hardware constraints, we employed Quantization-Aware Training (QAT) using an 8–4–4–1 precision scheme, where inputs are quantised to 8 bits, hidden layer weights to 4 bits, and all intermediate activations to 4 bits. In the output layer, precision is further reduced to 1 bit by setting the last activation function to binary, which aligns with Akida's built-in on-chip learning mechanism and enables efficient integer-only weight updates during continual training. This tailored quantisation approach, together with kernel modifications to enforce odd dimensions within the EEG channels, preserves the model's representational capacity while fully complying with the architectural restrictions of the Akida processor.

6.2.6 Spiking Neural Network Structure

The SNN is mapped from the quantised ANN, where the conventional activation functions Rectified Linear Unit (ReLU) are replaced with spiking event representations. The architecture works in a manner analogous to an event-based vision sensor, converting pixel activity into discrete events through AUROC for input encoding. Each Akida core can support tens of thousands of neurons. The exact number of instances that can be instantiated depends on the model architecture and mapping strategy.

6.2.7 Edge Learning

The Edge Learning mechanism in Akida enables efficient on-chip adaptation through a two-phase process. In the first phase, a pre-trained and quantised backbone is deployed to the hardware, with all weights frozen to preserve the extracted feature representations. In the second phase, only the final classification layer remains trainable

directly on the device. By confining parameter updates to this layer, the computational and memory demands of the training process are substantially reduced. This approach enables rapid, energy-efficient personalisation at the edge, eliminating the need for full network retraining and making the system particularly well-suited for resource-constrained environments.

An important hyperparameter in the edge learning process was the selection of neurons in the final training layer. Rather than fixing the number of neurons to match the number of target classes, this layer acts as a clustering mechanism, treating the neuron count itself as a tunable parameter. To optimise this, we employed a function that reserved approximately 10-20% of the training dataset as an estimation set for at least one class (i.e., the seizure class). This subset is used to analyse neuron activation behavior and to calibrate neuron allocation, helping determine whether increasing the neuron count improves representational learning or simply leads to redundancy and excessive memory utilisation on the Akida processor. Based on this analysis, we evaluated neuron counts of 100, 200, 300, 400, 500, 750, 1000, 1250, 1500, 2000, and 3000 to identify the optimal trade-off between accuracy, learning stability, and on-chip memory efficiency.

6.2.8 Few Shot Learning

Few-shot learning is a machine learning paradigm where models are trained to recognise new classes or patterns using only a very small number of labelled examples per class. Unlike conventional supervised learning, which relies on large datasets and extensive retraining, few-shot methods typically leverage prior knowledge encoded in a pre-trained backbone. In the case of BrainChip Akida, few-shot learning is implemented directly at the edge through its last-layer training mechanism. In our framework, 1 shot is equivalent to approximately 12 seconds of seizure activity. By providing just one or a few such examples, the final trainable layer of Akida was able to cluster and adapt to the new seizure representations. This approach is particularly advantageous for seizure detection and prediction, where obtaining large amounts

of labeled data is often difficult and seizure events may be rare or highly variable across patients. Few-shot learning enables the system to rapidly personalise to an individual patient's seizure characteristics directly on-device, without the need for cloud connectivity or full retraining of the backbone network.

6.2.9 Performance Metrics

The dataset used for model training and evaluation includes labelled data for each 12-second time window, as previously used in papers as a baseline reference Herbozo Contreras et al. (2024a,c); Saab et al. (2020); Tang et al. (2022); Yang et al. (2023). A seizure is identified when the model accurately predicts an ictal event (class 1) based on the ground truth, while non-seizure events are labelled as class 0. Seizure datasets are often imbalanced, with far more non-seizure data points, since seizures are rare and brief. Over a day, seizure occurrences represent only a small fraction of the data within a 12-second window. Using traditional metrics like accuracy is misleading in this scenario, as they tend to favour the majority class and do not effectively measure the model's ability to detect seizures, the primary goal. For this reason, we rely on the AUROC, which evaluates both sensitivity and specificity in a threshold-independent manner. We also provided additional metrics such as Sensitivity and FPR.

6.2.10 Implementation Details

For detection, we first trained the model on the TUH dataset for 200 epochs, with a learning rate of 0.005. Adam was selected as the optimiser. To distinguish between the seizure (1) and non-seizure (0) classes, we used MSE as the loss function. We then quantised the model as previously described, using a much lower learning rate than the initial one. We transformed the ANN model into an SNN by replacing the ReLU activations with SNN activations and transferring the pre-trained weights. We trained the quantised SNN on the EPILEPSIAE dataset in a few-shot learning manner. This is also an out-of-sample task as the dataset used does not belong to the

same cohort as the Pre-Trained dataset. 30 independent models were created by only training the last layer of the SNN in the BrainChip Akida. We evaluated the impact of few-shot learning on it, with shots ranging from 1 to 15, to examine the effect of long-term data. Each shot is composed of a seizure and a non-seizure class. Similarly, for the prediction task, we followed the same overall approach. The key difference lies in the initialisation: for both the FB and CHB-MIT datasets, models were trained individually per patient rather than as a global model, reflecting the limited number of patients and seizures available. For edge learning, we adopted a leave-one-seizure-out strategy, ensuring robust evaluation while preserving personalisation. An overview of the entire process is shown in Fig. 6.1(C-D).

6.3 Results

Fig. 6.1 illustrates our proposed edge-based personalised seizure detection and prediction framework. Unlike conventional epilepsy management methods that rely on cloud- or server-based AI, our approach integrates a neuromorphic system-on-chip with existing closed-loop neurostimulation devices to enable on-device intelligence without exceeding stringent power budgets. The framework employs spike-encoded continuous STFT representations of EEG signals, processed through an offline-trained convolutional feature extractor on the BrainChip Akida, with only the final classifier adapted online to patient-specific data. For seizure detection, a CNN trained on the TUH corpus is quantised, converted to an SNN, and edge-refined using the EPILEPSIAE dataset. For seizure prediction, a similar pipeline is applied to the FB and CHB-MIT datasets, with models pre-trained on a GPU and then updated at the edge, ensuring adaptability and independence from cloud connectivity.

6.3.1 Feature Extraction

An essential preliminary step for enabling edge learning is ensuring that backbone models capture physiologically meaningful neural dynamics. We first demonstrate the

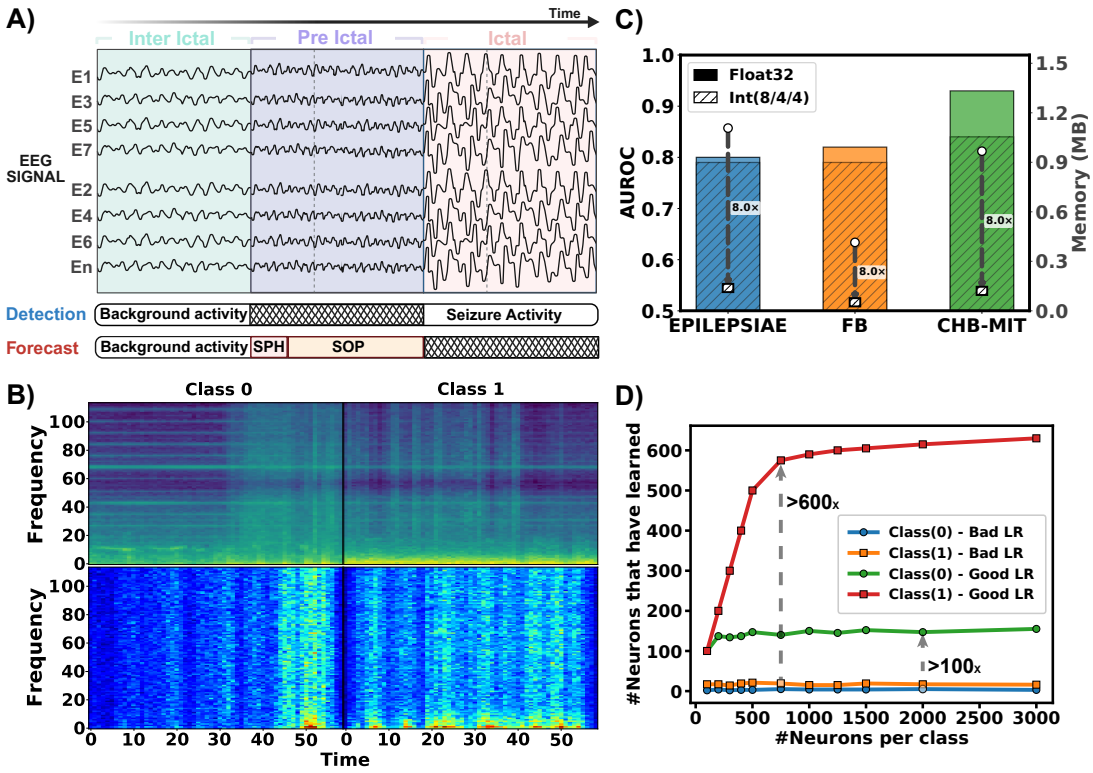

Figure 6.2 – Results for the initial training and prediction stage on GPU. When analysing seizure events, three phases are particularly critical: interictal, preictal, and ictal. These phases form the basis of two distinct tasks: Detection (discriminating between interictal and ictal states) and forecasting (discriminating between interictal and preictal states) (A). We applied STFT to the EEG data to extract two-class representations (0 and 1), and Grad-CAM visualisations revealed distinct activation patterns associated with non-seizure and seizure dynamics (B). Across all datasets, our models achieved commendable AUROC scores with floating-point precision, with only slight reductions after quantisation, while still yielding up to 8 times the memory reduction (C). During quantisation, lowering the learning rate relative to the initial training phase proved essential, enabling more effective consolidation of previously learned neuronal activations before transitioning to edge learning. Notably, proper initialisation with an appropriately tuned learning rate boosted per-class neuron activations from 100 to over 600 times(D).

Table 6.1 – Summary of EEG datasets used in this work.

Dataset◇	Type	Rate (Hz)	No. of Patients	No. of Channels	No. of Seizures	Interictal time(h)	Task	Edge Train	Temporal Duration
TUH	Scalp	250	642	19	3050	859.7	Detection	No	Short-Term
EPILEPSIAE	Scalp	250	30	19	276	4598.2	Detection	Yes	Long-Term
CHB-MIT	Scalp	256	13	22	64	209	Prediction	Yes	Short-Term
FB	Intracranial	256	13	6	59	311.4	Prediction	Yes	Long-Term

◇ Temple University EEG Corpus (TUH), Boston Children's Hospital–MIT CHB-MIT and Freiburg Hospital (FB).

STFT representations of a dataset for the non-seizure and seizure classes. We then extracted saliency maps from the trained model using Grad-CAM for both classes, revealing consistent patterns of model attention, with a pronounced emphasis on lower-frequency regions. Results are shown in Fig. 6.2(B). This observation aligns with prior findings that low-frequency activity and slow-wave components are key biomarkers associated with seizures in scalp EEG recordings Lagarde et al. (2016).

6.3.2 Pre-Post Quantisation

In the initial stage, we evaluated model performance on the TUH, FB, and CHB-MIT datasets using AUROC, recall, and precision. Model weights with validation scores exceeding or achieving 0.80 (0.80, 0.82, and 0.93, respectively) were selected for weight transfer. Following quantisation, performance declined only marginally for the TUH and FB datasets (1% and 3%, respectively), whereas the CHB-MIT dataset exhibited a more pronounced reduction of 9%. Despite this, memory requirements were reduced by approximately 8 times across all datasets as an outcome of the adopted quantisation scheme. Comparative AUROC results for both non-quantised and quantised models, alongside memory consumption (MB), are presented in Fig. 6.2(C).

6.3.3 Quantisation Learning Rate effects on Neuron Activations

We set the learning rate to 10 times lower than the original learning rate used to compile during the training GPU phase. Based on our observations, this has the greatest impact on edge deployability. A bad learning rate (to be similar to the stage 1 of training) leads to fewer neurons being learned per class, as observed in Fig. 6.2(D) for non-seizure and seizure classes, respectively. We observed as the number we input, with a good initialisation of LR there was neurons learned for both classes. This behaviour also suggests that, beyond a certain point, performance improvements became marginal, and the number of neurons was capped. In this way, resources were allocated efficiently, reducing memory usage on the BrainChip Akida hardware without compromising performance. For instance, as illustrated in Fig. 6.2(D), neuron activation becomes marginal beyond 1000 neurons, at which point the neuron for the last layer is capped to 1000.

6.3.4 Detection Out-Of-Sample Framework

Using the transferred weights, we further trained the model on the EPILEPSIAE dataset directly at the edge. To provide a comprehensive performance assessment, we introduce an Figure of Merit (FoM) that combines multiple evaluation metrics into a single score. The FoM is defined as:

$$\text{FoM} = \text{AUROC} \times \text{Sensitivity} \times (1 - \text{FPR}), \tag{6.2}$$

where (1-FPR) is chosen to penalize false alarms. Sensitivity reflects a True Positive Rate. This formulation enables a balanced evaluation of competing objectives: accuracy, reliability, and robustness. Fig. 6.3 presents the FoM Equ. (6.2) as a function of the number of shots, with corresponding latency measurements plotted in parallel in all datasets. As the number of shots increases, all performance metrics exhibit con-

sistent improvement. By the 15th shot, the AUROC, sensitivity, and FPR attained values of 0.85, 0.87, and 0.20, respectively. These personalised outcomes were subsequently compared against a generalisation baseline, which reflects performance under inference-only conditions without further adaptation and with floating point precision (float32). The personalised models demonstrated superior performance, yielding relative gains of 2% in AUROC, 4% in sensitivity, and a 6% reduction in FPR. These findings demonstrate that continuous, on-device learning with patient-specific data is both beneficial and critical for achieving effective, personalised seizure detection. The dataset holds particular significance as it comprises long-term continuous recordings in which seizure events constitute less than 1% of the total duration, with interictal segments comprising nearly the entire dataset. The results for all patients against a baseline are shown in Table 6.2. Results of Few-Shot Learning across all Patients are detailed in 6.8.

6.3.5 Prediction Framework

For the prediction task, we employed two datasets to compare different baselines: the scalp EEG dataset CHB-MIT and the iEEG FB dataset. The inclusion of the FB dataset also aligns with the vision of self-sustaining AI at the edge for future implantable devices. The experimental setup differed from the TUH-EPILEPSIAE framework, where models were first trained on TUH using a high-performance GPU and subsequently adapted at the edge on the EPILEPSIAE dataset. In contrast, for both CHB-MIT and FB, models were trained initially on a GPU, and then a leave-one-seizure-out approach was adopted for edge learning. This strategy reserved one seizure per patient for on-chip training while leaving sufficient samples for post-training evaluation, ensuring robust assessment of edge adaptation across both datasets. For preprocessing, we followed the same pipelines in two studies Gunasekaran et al. (2025); Truong et al. (2018).

For the CHB-MIT dataset, we observed notable improvements and smaller degradation during the GPU training phase (Fig. 6.2(C)). Using the pre-trained individual

Table 6.2 – Comparison of CNN and our work for seizure detection on the EPILEPSIAE dataset.

Patient	No. of Seizures	Interictal Hours	CNN Truong et al. (2018)			OUR WORK		
			AUROC	Sens	FPR	AUROC	Sens	FPR
Pat1	11	164.5	**91.4**	87.9	**0.16**	85.0	**92.0**	0.22
Pat2	8	177.3	67.4	87.1	0.58	**81.4**	**100.0**	**0.41**
Pat3	8	143.2	**89.6**	**94.4**	0.28	82.8	85.8	**0.23**
Pat4	5	167.6	**86.5**	76.6	0.15	83.4	**79.1**	**0.13**
Pat5	8	266.2	81.4	63.8	**0.15**	**82.1**	**97.7**	0.42
Pat6	8	135.3	**93.4**	86.7	**0.10**	89.9	**86.7**	0.13
Pat7	5	118.0	83.3	88.2	0.32	**98.1**	**100.0**	**0.03**
Pat8	22	115.5	83.5	96.2	0.31	**93.3**	**100.0**	**0.13**
Pat9	6	93.9	96.7	90.3	**0.08**	**96.7**	**93.8**	0.11
Pat10	11	137.9	58.7	48.4	**0.27**	**60.9**	**95.8**	0.55
Pat11	14	137.9	**85.8**	85.3	**0.26**	81.9	**86.8**	0.29
Pat12	9	159.6	**92.1**	**96.3**	0.26	87.9	91.7	**0.17**
Pat13	8	157.9	66.8	**91.2**	0.61	**83.7**	79.4	**0.17**
Pat14	6	162.0	**84.6**	**81.4**	0.22	84.3	77.3	**0.16**
Pat15	5	118.6	**90.5**	80.8	**0.13**	88.2	**100.0**	0.29
Pat16	6	92.8	90.6	92.3	0.15	**95.3**	**100.0**	**0.11**
Pat17	9	159.0	**85.2**	**94.7**	0.37	83.2	84.1	**0.13**
Pat18	7	178.1	**71.9**	**76.2**	**0.33**	64.1	66.7	0.34
Pat19	22	160.7	93.7	86.8	0.10	**94.5**	**96.7**	**0.10**
Pat20	7	164.5	**91.2**	**100**	0.29	79.4	70.8	**0.19**
Pat21	8	159.3	74.5	**83.3**	0.41	**88.0**	80.0	**0.08**
Pat22	7	137.7	75.0	75.0	0.30	**76.4**	**79.3**	**0.30**
Pat23	9	237.4	**93.3**	89.5	**0.16**	85.8	**91.3**	0.30
Pat24	10	94.2	82.3	64.3	0.09	**93.0**	**82.9**	**0.07**
Pat25	8	159.0	55.5	**81.5**	0.65	**71.7**	77.3	**0.20**
Pat26	9	158.8	**84.8**	**84.1**	**0.21**	84.5	82.8	0.22
Pat27	9	159.3	73.5	**74.2**	0.38	**79.7**	73.9	**0.19**
Pat28	9	162.1	95.2	83.6	**0.07**	**97.0**	**91.3**	0.08
Pat29	10	160.6	80.5	69.2	0.23	**81.1**	**75.1**	**0.19**
Pat30	12	159.5	**96.1**	**94.3**	0.17	93.6	81.8	**0.09**
Total	276	4598.2	83.2	83.5	0.26	**84.9**	**86.7**	**0.20**

weights, we subsequently trained patient-specific models at the edge. As the number of shots increased up to 15, performance stabilised with AUROC, sensitivity, and FPR reaching 0.87, 0.89, and 0.16, respectively (see Fig. 6.7). To benchmark our approach, we compared performance against two state-of-the-art methods. The first baseline was a widely cited CNN architecture for seizure prediction Truong et al. (2018), trained in full-precision and evaluated only at inference. This model achieved AUROC and sensitivity scores of 0.84 and 0.80, respectively, whereas our method im-

proved upon these results by 3% in AUROC and 9% in sensitivity. The second baseline was the recently proposed Future Guided Learning (FGL) framework Gunasekaran et al. (2025), which integrates detection models to support prediction training and has demonstrated competitive performance against established benchmarks. Notably, our approach outperformed FGL across all metrics, achieving relative improvements of 3% in AUROC, 1% in sensitivity, and 7% in FPR. Results are shown in Table 6.3. Results of Few-Shot Learning across all Patients in Fig. 6.9(B).

Table 6.3 – Comparison of CNN, FGL, and Our Work on CHB-MIT dataset.

Patient	No. of Seizures	Interictal Hours	CNN Truong et al. (2018)			FGL Gunasekaran et al. (2025)			OUR WORK		
			AUROC	Sens	FPR	AUROC	Sens	FPR	AUROC	Sens	FPR
Pat1	7	17	92	85.7	0.24	98	89	0.04	**100**	**100**	**0.00**
Pat2	3	22.9	34	33.3	**0.00**	88	83	0.10	**93.8**	**100**	0.09
Pat3	6	21.9	95	**100**	0.18	**97**	90	**0.01**	72.6	64.5	0.46
Pat5	5	13	88	80	0.19	**96**	**97**	**0.07**	85.5	81.4	0.16
Pat9	4	12.3	**74**	50	**0.12**	57	**94**	0.60	70.8	90.4	0.48
Pat10	6	11.1	56	33.3	**0.00**	**81**	**65**	0.26	67.1	46.4	0.22
Pat13	5	14	95	80	0.14	100	99	0.00	**100**	**100**	**0.00**
Pat18	6	23	**93**	**100**	0.28	69	88	0.43	85.4	90.1	**0.27**
Pat19	3	24.9	**98**	**100**	**0.00**	76	89	0.48	83.4	95.6	0.09
Pat20	5	20	97	100	0.25	94	86	0.13	**98.5**	**100**	**0.01**
Pat21	4	20.9	89	**100**	0.23	90	73	**0.12**	**92.2**	98.5	0.15
Pat23	5	3	99	100	0.33	67	100	0.50	**100**	**100**	**0.00**
Total	64	209	84.2	80.19	0.16	84.4	87.8	0.23	**87.4**	**88.9**	**0.16**

For the FB dataset, results show a similar trend to the previous model, with smaller performance gains but greater stability in metrics across shots (see Fig. 6.7). This limitation can be attributed in part to the BrainChip Akida architecture's requirement to use odd kernel sizes, which led us to mask one channel to conform to the hardware constraints. As discussed earlier, the first convolutional kernel spans the available channels, and given that the FB dataset includes only six intracranial channels, this restriction likely contributed to the diminished improvement. Despite these constraints, our edge-trained model achieved AUROC, sensitivity, and FPR values of 0.84, 0.90, and 0.23, respectively, at 15 shots. Our sensitivity increased by nearly 10% compared to a state-of-the-art CNN for seizure prediction. However, our false positive rate remained higher (0.23 vs. 0.06), and AUROC was approximately 4–5% lower than the CNN baseline. It is important to note, though, that the baseline CNN was trained in full precision (float32) and utilised all available intracranial channels,

whereas our approach was implemented under the quantisation and one-channel fewer constraints of the Akida neuromorphic hardware. Results are demonstrated in Table 6.4. Results of Few-Shot Learning across all Patients are detailed in Fig. 6.9(A).

Table 6.4 – Comparison of CNN and our work for seizure prediction on the FB Dataset.

Patient	No. of Seizures	Interictal Hours	CNN Truong et al. (2018)			OUR WORK		
			AUROC	Sens	FPR	AUROC	Sens	FPR
Pat1	4	23.9	100	100	0.00	**100**	**100**	**0.00**
Pat3	5	23.9	100	100	0.00	**100**	**100**	**0.00**
Pat4	5	23.9	100	100	0.00	**100**	**100**	**0.00**
Pat5	5	23.9	65	40	0.13	**90.4**	**89.7**	**0.10**
Pat6	3	23.8	**100**	**100**	**0.00**	71.6	62.6	0.18
Pat14	4	22.6	**82**	50	**0.27**	65.2	**72.1**	0.36
Pat15	4	23.7	99	**100**	0.02	**99.8**	98.1	**0.00**
Pat16	5	23.9	**86**	80	**0.17**	68.1	**100**	0.67
Pat17	5	24.0	94	80	**0.00**	**96.4**	**91.8**	0.07
Pat18	5	24.8	**100**	**100**	**0.00**	78.6	83.7	0.37
Pat19	4	24.3	51	50	**0.16**	**58.3**	**87.1**	0.71
Pat20	5	24.8	**76**	60	**0.04**	67.9	**85.7**	0.52
Pat21	5	23.9	100	100	0.00	**100**	**100**	**0.00**
Total	59	311.4	**88.7**	81.54	**0.06**	84.3	**90.1**	0.23

6.3.6 Power, Memory and Latency Analysis

The power consumption and inference efficiency of the Akida system were systematically evaluated using the MetaTF power-profiling API during inference. The floor power, representing the device's dynamic baseline power consumption while idle, was measured at 900.5 mW. During active inference, the system achieved an average throughput of 70.41 Frames per second (fps). Each frame corresponds to the STFT representation of a 12-second EEG window. The inference latency per frame was 14.2 ms, which equates to approximately 1.18 ms per second of raw EEG data. The energy consumption per inference was measured at 12.86 mJ per frame of a 12-second EEG window, which is equivalent to 1.07 mJ per 1 second of EEG data, reflecting the low-power profile of the neuromorphic hardware. These results demonstrate the suitability of the Akida platform for real-time, energy-efficient processing of EEG data in edge-based seizure detection and prediction applications. Results are detailed

in Fig. 6.4, where we compare running the CNN backbone on a conventional CPU architecture and our approach on Akida, showing latency and energy per inference. We demonstrate that our neuromorphic approach enables power consumption that is two to three orders of magnitude lower.

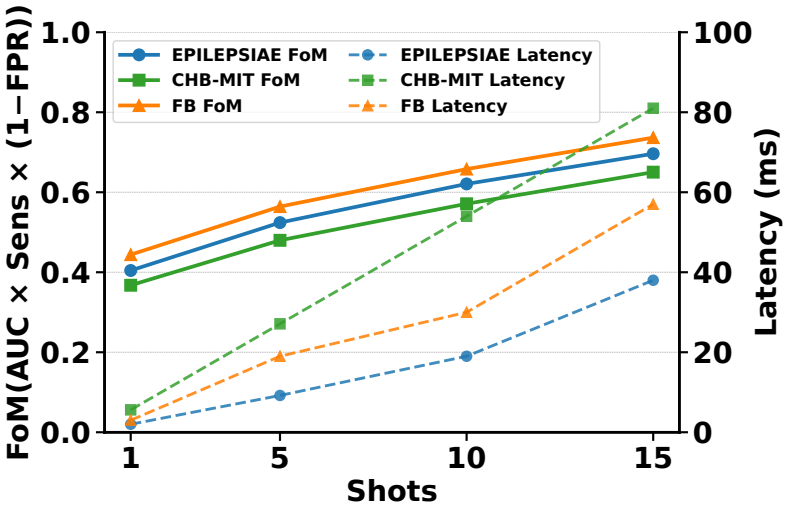

Figure 6.3 – Figure of Merit FoM and latency vs number of shots for the three datasets. FoM is defined as $AUROC \times Sensitivity \times (1 - FPR)$. Higher shot counts yield consistent FoM gains, while latency rises marginally, indicating improved performance without meaningful computational cost.

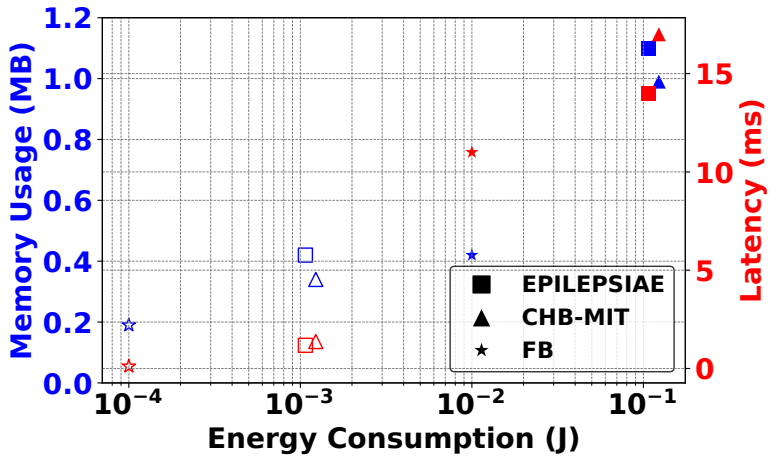

Figure 6.4 – Pareto plot of energy consumption versus memory usage and latency for three datasets (EPILEPSIAE, CHB-MIT, and FB). Filled markers represent the results for metrics in CPU, while non-filled markers represent those at the Edge with Akida.

6.3.7 Robustness

A critical validation in this work was to assess the robustness of our model in the spiking domain under reduced input information. We evaluated this on the EPILEPSIAE and CHB-MIT datasets, each comprising 19 EEG channels, by systematically blacking out 20%, 40%, and 60% of the channels. As illustrated in Fig. 6.5, the CHB-MIT dataset demonstrated stable performance despite increasing channel loss and further benefited from a higher number of shots. In contrast, the EPILEPSIAE dataset showed a more pronounced degradation as blackout levels increased. Nevertheless, resilience improved considerably with additional shots. We attribute this effect to the long-term nature of EPILEPSIAE recordings, which serve as a more realistic and challenging benchmark than CHB-MIT. Overall, our model achieved strong performance across both datasets, demonstrating robustness to channel dropout.

6.3.8 Complexity: Training and Inference

Our analysis of computational complexity shows that the proposed Akida-based edge learning framework is more efficient than conventional deep learning approaches. Traditional CNNs and Long Short-Term Memory (LSTM)s exhibit high training costs due to their dependence on large parameter updates and, in the case of recurrent models, sequential operations that limit parallelisation Tang et al. (2022). Similarly, SNNs trained with surrogate gradients require costly BPTT, while even local learning rules still scale linearly with sequence length Yin et al. (2021). SVMs, though non-sequential at inference, demand prohibitively expensive training with cubic dependence on the number of samples Hsieh et al. (2022); Zhang et al. (2022). By contrast, our proposed edge learning approach reduces training time to a cluster with linear complexity in the feature dimension, while inference scales with the number of neurons and the feature extractor. Crucially, this approach enables efficient, low-latency operation. These findings demonstrate that EDGE-AI on neuromorphic hardware offers a favourable trade-off between performance and computational cost, positioning neuromorphic as a viable solution for real-time, resource-constrained biomedical

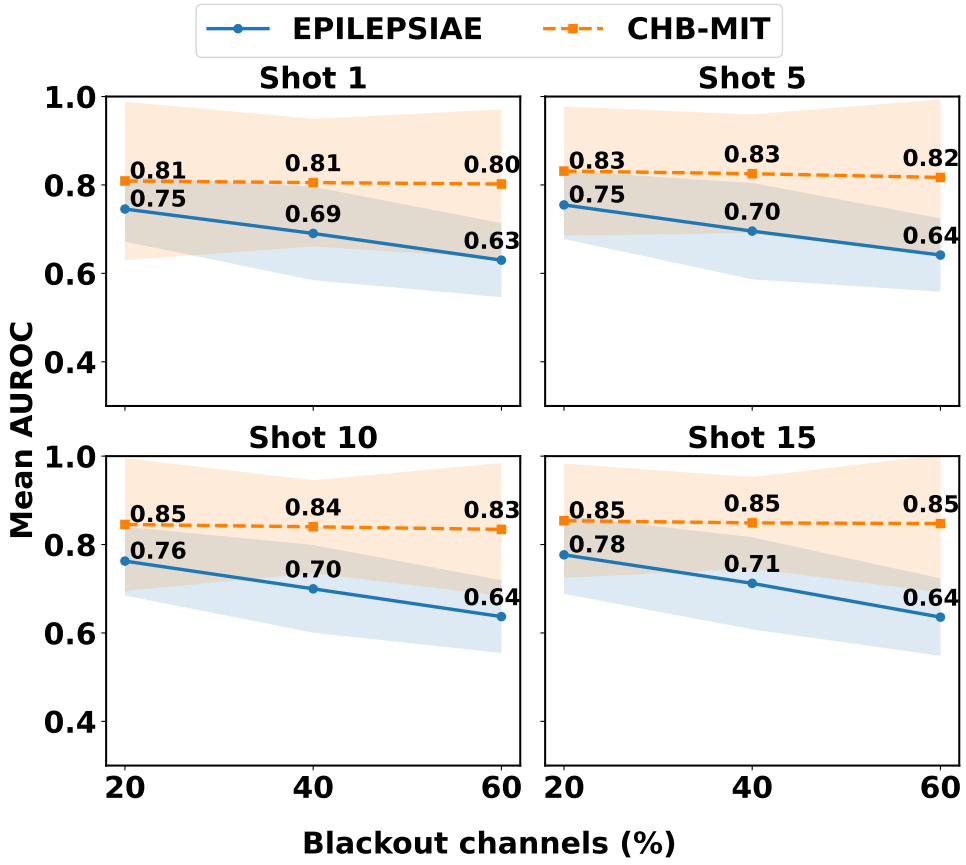

Figure 6.5 – **Robustness test** of our approach for seizure detection and prediction across datasets under progressive channel blackout with three repetitions. Performance is shown for the EPILEPSIAE dataset (blue, long-term clinical data) and the CHB-MIT dataset (orange) across different training shots (1, 5, 10, 15) as increasing proportions of EEG channels were set to zero, highlighting resilience to missing-channel conditions.

applications. The complexity formulas are listed in Table 6.5. We benchmark our approach against representative state-of-the-art seizure detection/prediction models whose architectures and hyperparameters are documented in the literature and evaluated on at least three EEG datasets. In Fig. 6.6, we demonstrate that our approach is computationally efficient for both training and inference as it occupies the lower-left region of the plot.

Table 6.5 – Training complexity per model

Algorithm	Training Complexity$^{\otimes}$
CNN$^{\diamond}$	$\mathcal{O}\left(\sum_{\ell=1}^{L} C_{\text{in}}^{(\ell)} C_{\text{out}}^{(\ell)} K_\ell^2 H_\ell W_\ell\right)$
LSTM†	$\mathcal{O}\left(\sum_{\ell=1}^{L} T \cdot 4\left((M^{(\ell)})^2 + M^{(\ell)} D^{(\ell)}\right)\right)$
ConvLSTM†	$\mathcal{O}\left(\sum_{\ell=1}^{L} T \cdot 4\left[C_{\text{in}}^{(\ell)} C_{\text{out}}^{(\ell)} K_\ell^2 H_s W_s + \left(C_{\text{out}}^{(\ell)}\right)^2 K_\ell^2 H_s W_s\right]\right)$
SNN†	$\mathcal{O}\left(\sum_{\ell=1}^{L} T \cdot \text{cost}_{\text{layer}}^{(\ell)}\right)$
SConvLSTM†	$\mathcal{O}\left(\sum_{\ell=1}^{L} T \cdot 4\left[Fr_{x,\ell}\, C_{\text{in}}^{(\ell)} C_{\text{out}}^{(\ell)} K_\ell^2 H_s W_s \;+\; Fr_{h,\ell}\left(C_{\text{out}}^{(\ell)}\right)^2 K_\ell^2 H_s W_s\right]\right)$
SVM	$\mathcal{O}\left(n^2 m + n^3\right)$
Proposed	$\mathcal{O}\left(n \cdot U\right)$

$^{\diamond}$ Trained with Backpropagation (BP).
† Trained with Backpropagation Through Time BPTT.
$^{\otimes}$ L:Layer; H:Height; W:Width; K:Kernel; C:Filter; T:Sequence length; M:Hidden size; D:Input Size; n:Neurons; m:Feature dim.; Fr_{in},Fr_{out}:Spike Rate; U:Input Spikes.

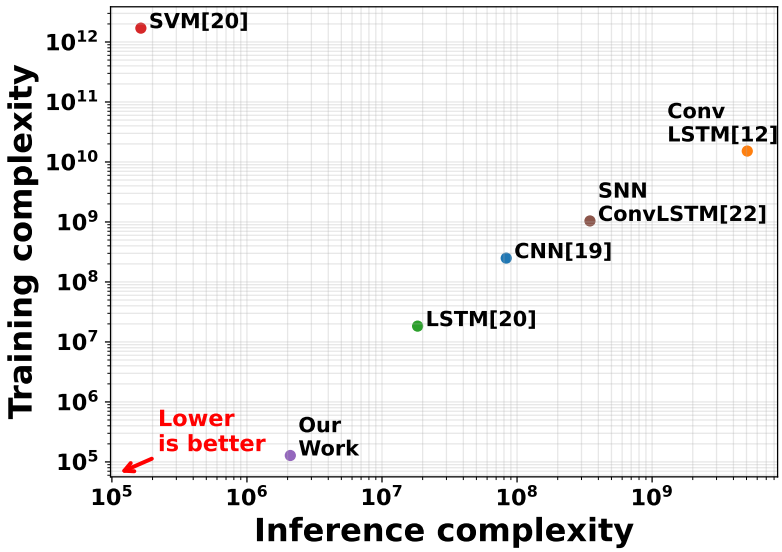

Figure 6.6 – **Log–log scatter plot of training vs inference computational complexity** for representative EEG seizure detection/prediction models. Each point reflects hyperparameters reported in prior work. Conventional method of training these networks are BP (CNN) and BPTT (LSTM, SConvLSTM, ConvLSTM). For inference complexity, the feature extractor is taken into account as a CNN-SNN.

6.4 Discussion

A novel framework for seizure detection and prediction directly on the edge was presented, enabling continual learning with lower computational complexity, reduced energy consumption, and fully unsupervised adaptation while processing streaming EEG data. These characteristics make the framework particularly suitable for long-term neuromorphic neuromodulation, where embedded intelligence must operate reliably and efficiently, with minimal dependence on cloud resources. While the results demonstrate strong potential, several opportunities remain for further improvement.

- In this work, we do not consider different structures of different neural networks, which can lead to better results. Nonetheless, by using a simple architecture, we observe improved performance compared to more complex models Truong et al. (2018); Yang et al. (2022b).
- BrainChip Akida only allows certain models and layers to run at the Edge for training. It can be more beneficial to incorporate more advanced algorithms that can run on the edge Contreras et al. (2025); Shen et al. (2025); Yang et al. (2023); Yin et al. (2021, 2023).
- Different encoding mechanisms were not explored in this work. Studies have shown that different techniques in processing the EEG data into spikes are fundamental to capture temporal dynamics on seizure data Sharifshazileh et al. (2021) and also a reduction in power consumption and continual learning Pes et al. (2024).
- Although training only the last layer offers a computationally efficient adaptation strategy, it may be insufficient to fully address distribution shifts or the broader challenges of continual learning. The current methodology approach mainly addresses plasticity, while stability is less explicitly considered. Future work could therefore explore local learning mechanisms, such as metaplasticity, to better regulate the stability-plasticity trade-off Aguilar et al. (2025).

Future work will focus on designing architectures capable of better capturing temporal and spatial dependencies, as shown in recent studies Herbozo Contreras et al. (2024c), thereby achieving higher performance without increasing memory or computational requirements. Overall, this work provides a promising foundation for next-generation AI-driven electroceuticals.

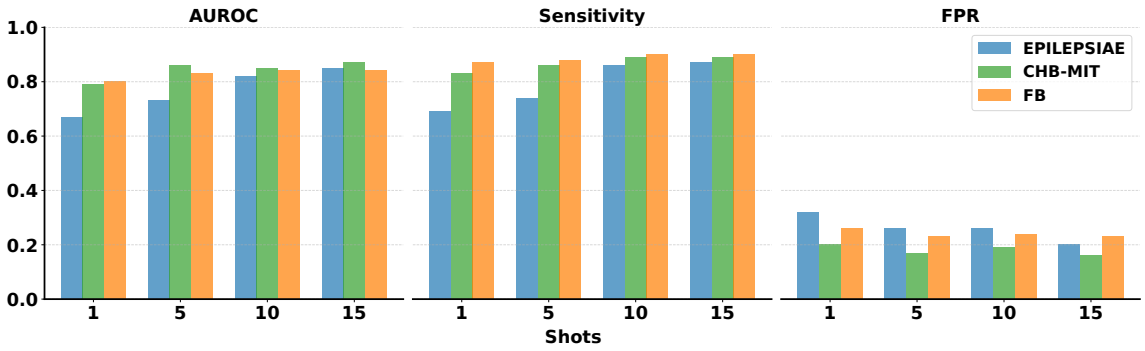

Figure 6.7 – Average AUROC, FPR, and Sensitivity across all patients and datasets (EPILEPSIAE, CHB-MIT, Freiburg).

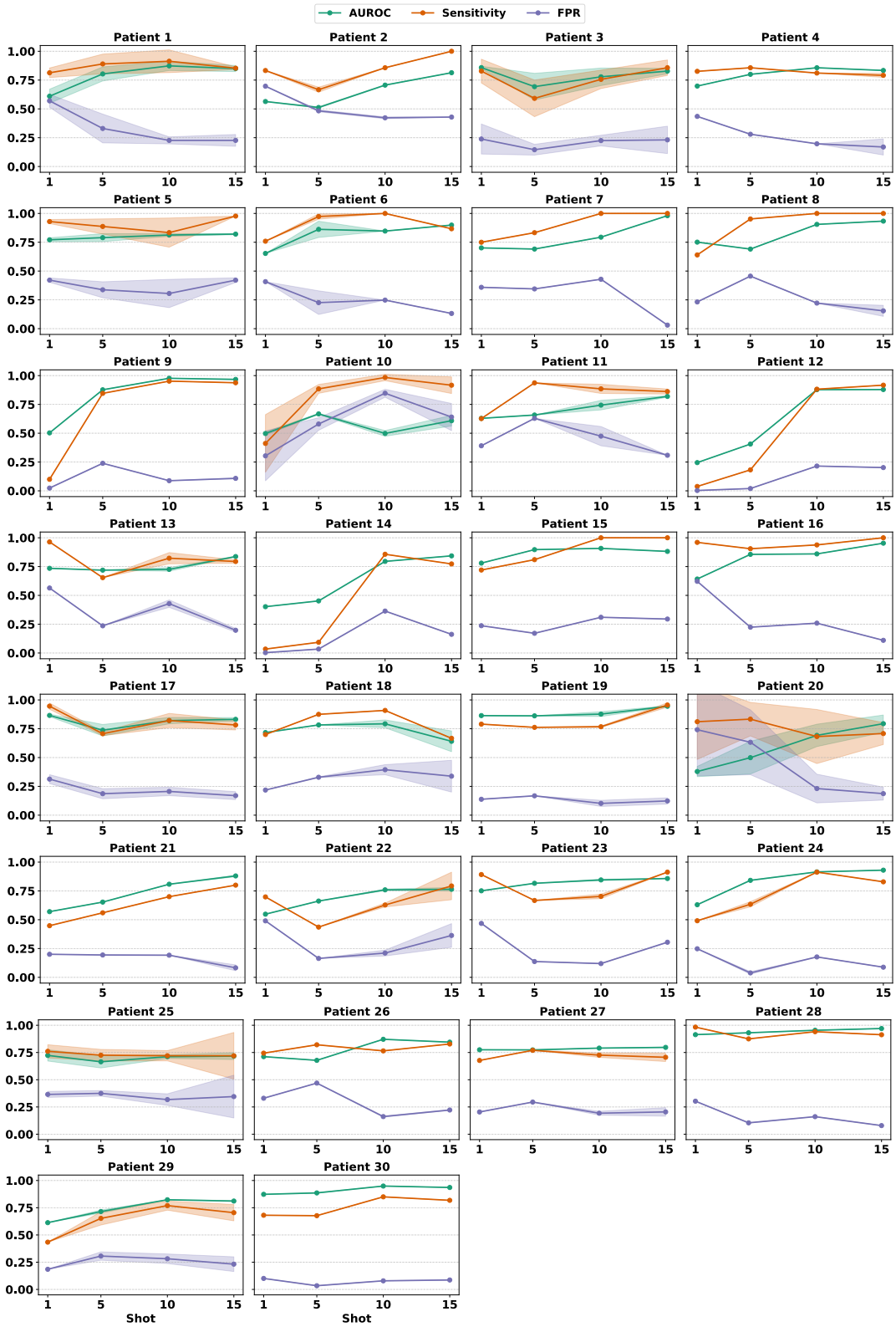

Figure 6.8 – Performance metrics as a function of shots for patient in EPILEPSIAE dataset. Results averaged over 3 independent runs.

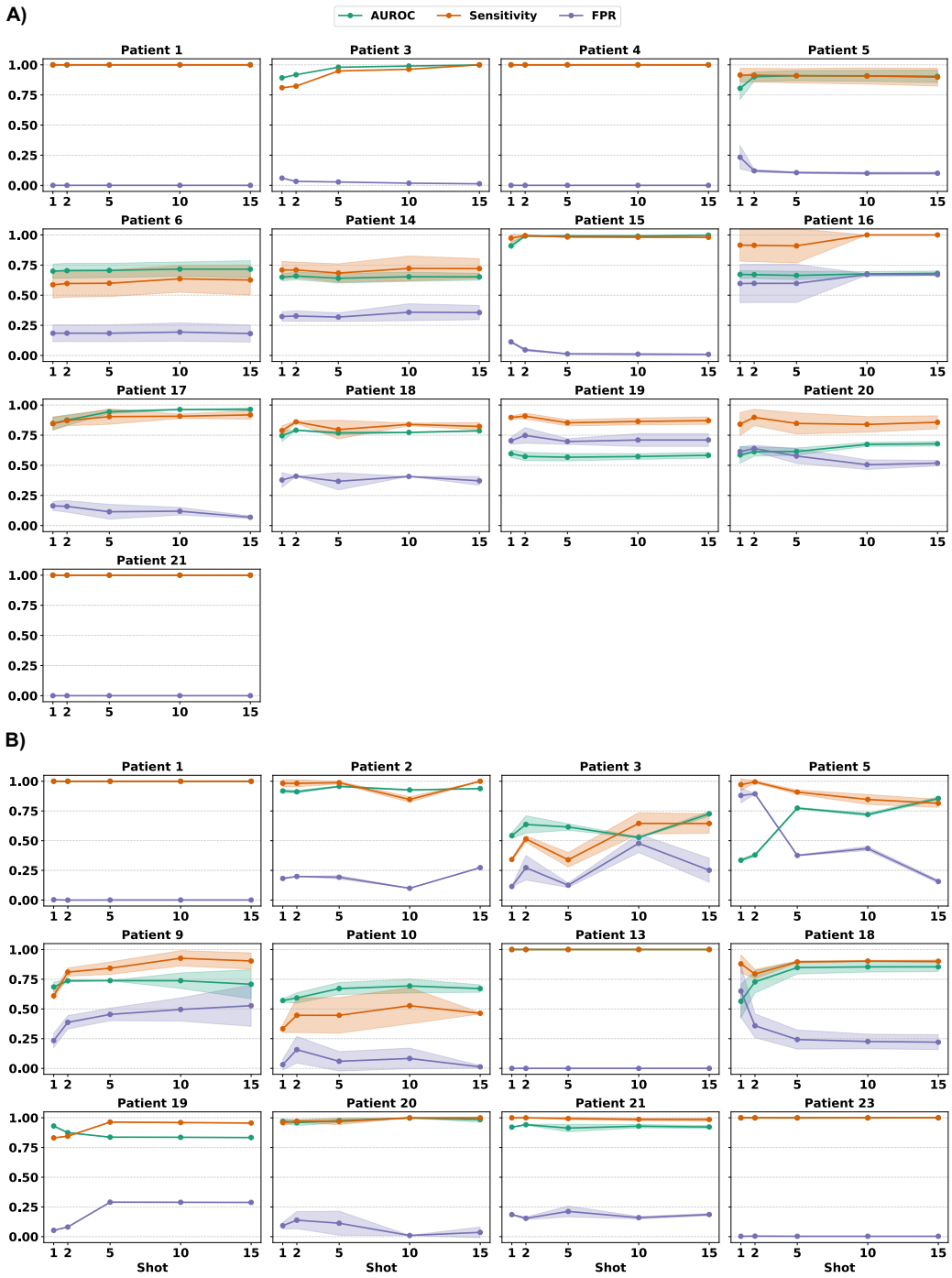

Figure 6.9 – Performance metrics in the prediction framework as a function of shots for each patient in the FB (A) and CHB-MIT (B) datasets. All results are averaged over 3 independent runs.

Chapter 7

Learnable functions led to Interpretable Systems.

This chapter extends the thesis beyond spiking neuromorphic models by introducing the first implementation of Kolmogorov-Arnold Network (KAN) for seizure detection, evaluating their potential as compact and interpretable function-based representations that may enable future neuromorphic architectures.

The content presented in this chapter is published as:

- Herbozo Contreras, L. F.*, Cui, J*, Yu, L., Huang, Z., Nikpour, A. and Kavehei, O. (2025). "KAN–EEG: towards replacing backbone–MLP for an effective seizure detection system." *Royal Society Open Science*, 12(3). DOI:10.1098/rsos.240999.

Statement of Contributions of Joint Authorship

- Luis Fernando Herbozo Contreras (Candidate): Co-First author, formal analysis, investigation, methodology, resources, software, visualisation, validation, writing original draft, writing review & editing of the manuscript
- Jiashuo Cui: Co-First author, investigation, visualisation, validation, writing original draft, writing review & editing of the manuscript
- Leping Yu: contributed to writing review & editing of the manuscript
- Zhaojing Huang: contributed to writing review & editing of the manuscript
- Armin Nikpour (Co-Supervisor): conceptualisation, data curation, funding acquisition, investigation, project administration, resources, supervision, validation, writing review & editing
- Omid Kavehei (Principal Supervisor): conceptualisation, data curation, funding acquisition, investigation, methodology, project administration, resources, supervision, visualisation, validation, writing review & editing.

In addition to the statements above, in cases where I am not the corresponding author of a published item, permission to include the published material has been granted by the corresponding author.

Luis Fernando Herbozo Contreras Date: 28 May 2026

As supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

Prof. Omid Kavehei Date: 28 May 2026

Chapter Summary

The landscape of artificial intelligence AI research is witnessing a transformative shift with the emergence of the KAN, a novel architectural paradigm aimed at redefining the structural foundations of AI models based on Multilayer Perceptron (MLP). Through rigorous experimentation and evaluation, we introduce the KAN-EEG model, a tailored design for efficient seizure detection. Our proposed network is tested and successfully generalised on three different datasets, one from the USA, one from Europe, and one from Oceania, recorded with different front-end hardware. All datasets are scalp-EEG recordings from adults with epilepsy. Our empirical findings reveal that while both architectures demonstrate commendable performance in seizure detection, the KAN model exhibits high-level out-of-sample generalisation across datasets from diverse geographical regions, underscoring its inherent efficacy and adaptability at the backbone level. Furthermore, we demonstrate the resilience of the KAN architecture to model size reduction and shallow network configurations, highlighting its versatility and efficiency by preventing over-fitting in-sample datasets. This work advances our understanding of innovative neural network architectures and underscores the pioneering potential of KANs in critical domains such as medical diagnostics.

7.1 Introduction

Artificial intelligence AI has undeniably made significant strides in healthcare, becoming a new era of innovation and patient care. Its impact on the healthcare industry has been overwhelmingly positive, with numerous benefits spanning various aspects of medical practice, research, and patient outcomes Amann et al. (2020); Davenport and Kalakota (2019); Yu et al. (2018). Among the various AI techniques, machine learning models, particularly neural networks, have shown great promise in analysing and interpreting complex medical data. MLP networks are the backbones of today's AI architectures Cybenko (1989); Haykin (1998) and have been extensively used for

their effectiveness in detecting and classifying abnormalities in biosignals. The MLP, a feed-forward artificial neural network, consists of multiple layers of neurons that process input data through weighted connections and activation functions. It excels at capturing relationships within the data, making it well-suited for analysing the intricate patterns in medical signals. Despite its advantages, the MLP has limitations, particularly in model interpretability and efficiency. To address these challenges, a new model, the KAN, has been proposed as a promising alternative to MLPs Liu et al. (2025). Unlike MLPs, which use fixed activation functions on the hidden layers, KANs employ learnable activation functions, replacing linear weights with univariate functions parametrised as splines. This architectural difference allows them to achieve greater accuracy and interoperability.

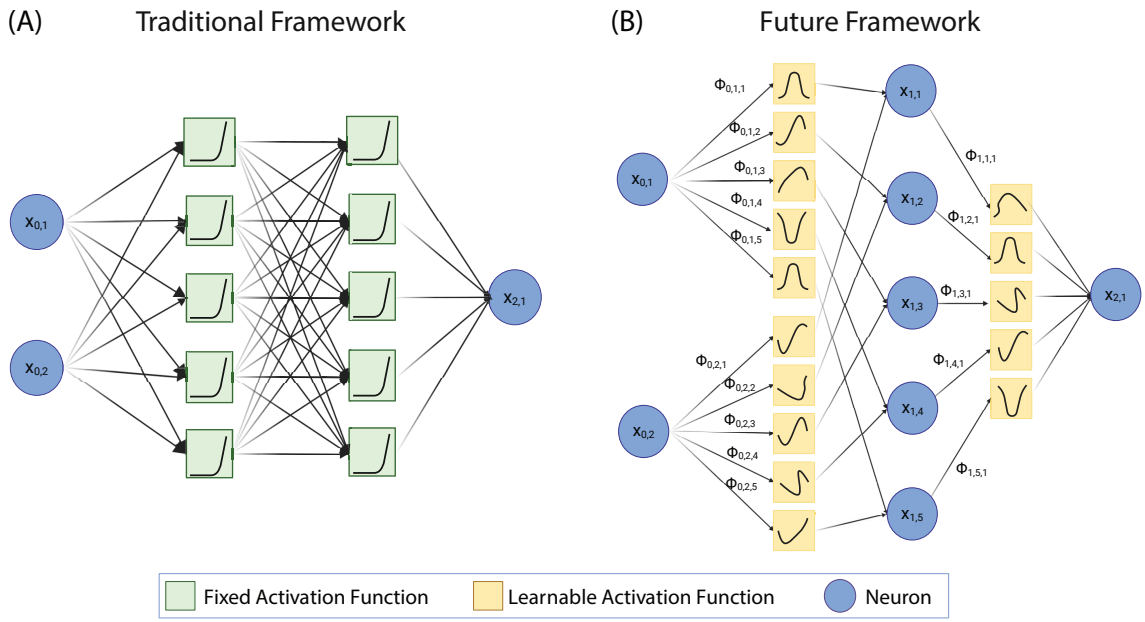

Figure 7.1 – Representation of the past and future in neuro-AI. The multilayer perceptron was built based on non-learnable activation functions and relies on deep, fully connected networks for accurate performance (A). On the contrary, KAN emerges as a potential solution for more explainable, shallow, and efficient architectures at the core level (B).

KANs have demonstrated the ability to outperform MLPs with smaller network sizes, making them more computationally efficient Liu et al. (2025). Despite its advantages, the MLP has limitations, particularly in model interpretability and efficiency. Fur-

thermore, it has demonstrated the ability to outperform MLPs with smaller network sizes, making them more computationally efficient Liu et al. (2025). In the context of EEG signal analysis, KANs offer potential advantages over MLPs. Their learnable activation functions and efficient representation of data structures enable better handling of the complex, high-dimensional data characteristic of EEG signals. This can improve detection accuracy and more insightful interpretations of the underlying neural activity. The subsequent sections will delve deeper into the comparative performance of our proposed architectures and MLPs for detecting epileptic seizures, highlighting their strengths and potential in this critical application. The diagram in Fig. 7.1 delineates the structural differences between the KAN and the MLP architectures. The KAN structure demonstrates learnable activation functions that can constitute interpretable systems and address MLP architecture challenges through enhanced capabilities for continual learning and efficiency at shallow, sparse connections.

7.1.1 Background

Epileptic seizure detection has seen substantial progress with the rise of machine learning models, particularly those utilising EEG signals. EEG tests record the electrical activity produced by neurons, typically noninvasively, by placing multiple electrodes on the scalp Niedermeyer and da Silva (2005). EEGs play a crucial role in detecting epileptic foci and categorising epilepsy types, such as focal, generalised, and unknown seizures Birjandtalab et al. (2017); Fisher et al. (2014); Truong (2020). These recordings are essential for diagnosing epilepsy, monitoring ongoing conditions, making predictions, and effectively responding to neurostimulation. Recent advancements in deep learning have introduced a range of architectures to improve EEG-based seizure detection. CNNs are prominently used, transforming EEG signals into different dimensional forms for detailed analysis Abdelhameed et al. (2018); Avcu et al. (2019); Bouaziz et al. (2019); Covert et al. (2019); Hossain et al. (2019); Zuo et al. (2019). RNN, including their advanced versions like LSTM and Gated Recurrent Unit (GRU), are adept at capturing temporal dependencies in EEG data, making them suitable

for sequential data analysis Chen et al. (2018); Fukumori et al. (2019); Geng et al. (2020); Vidyaratne et al. (2016). Unsupervised learning methods, such as Deep Belief Network (DBN) Le et al. (2018); Turner et al. (2014) and Autoencoder (AE) Emami et al. (2019); Golmohammadi et al. (2019); Shah et al. (2017), are employed to extract and reconstruct features from raw signals.

Table 7.1 summarises various traditional models applied to different bio-signal applications, preprocessing methods, datasets, and indicators. In addition, hybrid models that combine CNNs with RNNs or AEs leverage the spatial feature extraction capabilities of CNNs and the temporal modelling strengths of RNNs or the reconstruction abilities of AEs Choi et al. (2019); Fang et al. (2018); Roy et al. (2018); Wen and Zhang (2018); Yang et al. (2022b); Yuan et al. (2018). For example, ConvLSTM networks integrate the spatial feature extraction power of CNNs with the temporal sequence modelling capabilities of LSTMs Yang et al. (2022b). This integration has improved the model's seizure-detection performance.

Transformer-based networks, a more recent innovation, incorporate attention mechanisms to better capture complex patterns in EEG data, further improving detection accuracy Ma et al. (2023); Pedoeem et al. (2022); Peh et al. (2023); Sun et al. (2022). Despite these significant advancements, traditional seizure detection models still encounter challenges in generalisation and real-time implementation. Models trained on large datasets often struggle with low AUROC and high false-positive rates, limiting their clinical applicability. Thus, there is a pressing need for models that can balance sensitivity and specificity and be broadly applicable across diverse patient populations and real-world scenarios Yang et al. (2022a,b).

Although the traditional MLP-based model has significantly contributed to the development of EEG-based seizure detection systems, the emergence of KAN and other advanced models provides a promising direction for improving the accuracy, efficiency, and applicability of epilepsy detection technology. The following section provides a detailed comparison between MLP and KAN, emphasising the advantages and potential of KAN for EEG- based epilepsy detection.

Table 7.1 – Different types of neural networks applied to EEG signals for epileptic seizure detection.

Model	Application	Pre-Processing[+]	Dataset [◇]	Metrics [†]	Limitations
3D-CNN + RNN Choi et al. (2019)	Seizure detection	STFT 2D mapping	CHB-MIT SNUH	ACC: 99% ACC: 99%	Small cohort
ConvLSTM Yang et al. (2022b)	Seizure detection	ICA, STFT	TUH EPILEPSIAE RPAH	AUC: 84% AUC: 81% AUC: 82%	Post processing dependency
SConvLSTM Yang et al. (2023)	Seizure detection	STFT	FB CHB-MIT EPILEPSIAE	AUC: 93% AUC: 89% AUC: 81%	Low accuracy
TSD Ma et al. (2023)	Seizure detection	STFT	TUH	AUC: 92%	No out-of-sample test
ResNet-LSTM Qiu et al. (2023)	Seizure detection	ICA, STFT	U-Bonn	ACC: 97%	Memory usage
GraphS4mer Tang et al. (2023)	Seizure detection Sleep staging ECG classification	STFT	TUH DOD-H ICBEB	AUC: 91% F1: 82% AUC: 98%	Memory increases quadratically
SVM Alotaibi et al. (2021)	Seizure detection	DWT	CHB-MIT	ACC: 98%	Small cohort

[+] Pre-Processing methods: Short-Time Fourier Transform STFT, Independent Component Analysis ICA, Discrete Wavelet Transform DWT.
[◇] Datasets: Temple University Hospital TUH, Children's Hospital Boston CHB-MIT, Royal Prince Alfred Hospital RPAH, Seoul National University Hospital (SNUH), University of Bonn (U-Bonn), Freiburg FB, Dreem Open Dataset-Healthy (DOD-H), International Conference on Biomedical Engineering and Biotechnology (ICBEB).
[†] Metrics: Accuracy ACC, Area Under the Curve (AUC or AUROC), Macro-Averaged F1-Score (Macro-F1).

7.1.2 Limitations of the Previous Studies

Previous models used in machine learning-based systems, particularly those relying on multi-layer perceptron MLP architectures, have been constrained by their reliance on fixed activation functions. These functions, such as ReLU or sigmoid, are predefined and static, limiting the model's flexibility and adaptability. Furthermore, traditional MLP architectures lack the capability for continuous learning, meaning that once a model is trained, it cannot adapt based on new incoming data without retraining. This limitation poses a significant challenge for real-time applications or scenarios where the data distribution changes over time. Without the ability to continuously learn and adapt, the models can become less accurate, leading to performance degradation. Therefore, there is growing interest in developing architectures

that dynamically adjust their activation functions and support continuous learning, enabling more flexible, robust, and efficient models for complex, real-world applications.

7.1.3 Novelty and Significance

This work introduces a novel backbone architecture for advancing the field of seizure detection, focusing on addressing challenges related to generalisation and real-time implementation. To the best of our knowledge, this is the first study to utilise a novel non-backbone MLP for healthcare applications and the first to be utilised in Seizure Detection. We propose the utilisation of KANs as a key component for identifying epileptic seizures from pre-recorded EEG signals. We use the STFT to effectively process EEG data. Rather than solely emphasising the challenges in generalisation and implementation, we highlight the innovative potential of efficient-shallow architectures adaptable as a foundational framework for future AI models in seizure detection. Through our experimentation with the 3 different continental datasets, we aim to showcase the inherent efficacy of KANs and their capacity to serve as an efficient architecture for developing advanced seizure detection systems. We found the following advantages for our proposed method: (**Higher Accuracy**) Preliminary findings suggest that our model achieves AUROC values comparable to those of traditional MLP models, indicating similar performance in seizure detection. This comparative analysis suggests that while KANs do not necessarily outperform MLP models, they offer a comparable level of reliability, thereby contributing to the advancement of seizure detection methodologies; (**Efficiency**) Our proposed architecture requires a smaller network size. Thereby making them computationally efficient and suitable for real-time applications; (**Less Training, High performance**) Our proposed architecture achieves high accuracy while requiring only a small portion of the training dataset, outperforming models that typically require larger datasets for similar results; (**Out-of-sample seizure detection**) generalisation beyond the training dataset was assessed by training our model on the USA dataset and evaluating its performance on independent datasets from Europe and Oceania, using the same trained weights.

The outcomes indicate encouraging support for this test.

7.2 Methods

7.2.1 Datasets

Three datasets were used in this work: the TUH, the scalp-EPILEPSIAE, and the RPAH datasets. Fig. 7.2 summarises the TUH dataset, detailing key statistics. This dataset is divided into training and validation sets, offering comprehensive insights into the number of hours utilised for patients with seizures and non-seizures, and the number of patients who present with seizures and non-seizures. The TUH dataset is the primary training dataset, providing a diverse range of seizure signals for various seizure types. With its large volume of files, the TUH dataset from the USA offers extensive data for robust model training. We utilised 400 hours of EEG data for our training process, comprising 120,000 samples with a 12-second window. This dataset included approximately 75% background activity and 25% seizure activity. Notably, this represents a significantly smaller data volume than other models. Despite the reduced dataset size, we successfully validated the KAN-EEG model's efficiency in seizure detection with reduced training data. The validation was conducted using 192 hours of EEG data, corresponding to 57,306 samples with a 12-second window, and maintained the same 75%-25% ratio of background to seizure information.

The EPILEPSIAE and RPAH datasets, consisting of adult EEG, are used for inference tests. Both datasets share common characteristics, such as using identical montages and adult patient data, ensuring consistency and comparability in analysis. Integrating these complementary datasets enhances the comprehensiveness and reliability of the AI model for seizure detection. The EPILEPSIAE dataset consists of 30 patients, of whom 19 are male and 11 are female. The RPAH dataset from one of Australia's major hospitals reliably maintained one of the largest datasets from adult epilepsy patients nationwide. This work uses nine years (2011–2019) of data, testing nearly 14,590 hours of EEG data from 192 patients over 1,006 sessions, each averag-

ing around 15 hours of recording. Out of 212 patients, 20 were excluded for reasons including excessive seizures (more than 11 seizures/24 hours), missing electrode data, or seizures confirmed only by video.

The Australian dataset is about 16 times larger than the US training dataset, with longer inter-ictal periods and more background data, making false-positive evaluation highly robust. The distribution of the RPAH dataset across three domains—seizure type and frequency, age and gender, and seizure occurrence within a 24-hour cycle—highlights essential patterns. Notably, seizure occurrences are derived from intermittent monitoring, providing insights into the likely timing of seizures. Ethical approval was obtained to access this clinical data.

7.2.2 The KAN-EEG Structure

Characteristics of KANs

KANs are founded on the Kolmogorov-Arnold representation theorem, which posits that any multivariate continuous function can be broken down into a finite composition of continuous univariate functions and addition operations Liu et al. (2025). This foundational principle allows it to substitute traditional linear weights with spline-parametrised univariate functions. Moreover, it employs adaptive univariate activation functions along the network edges, enhancing flexibility and precision. These functions adapt to the data, leading to more accurate approximations. Spline functions enable dynamic adaptation to the data, providing refined representations that effectively capture smooth transitions. Additionally, KANs require fewer parameters than MLPs, improving computational efficiency and model interpretability. The learnable functions can also be visualised for better understanding. The operation of a KAN layer can be described by the Equ. (7.1):

$$\hat{x}_i = \sum_{j=1}^{n} \varphi_{ij}(x_j) \tag{7.1}$$

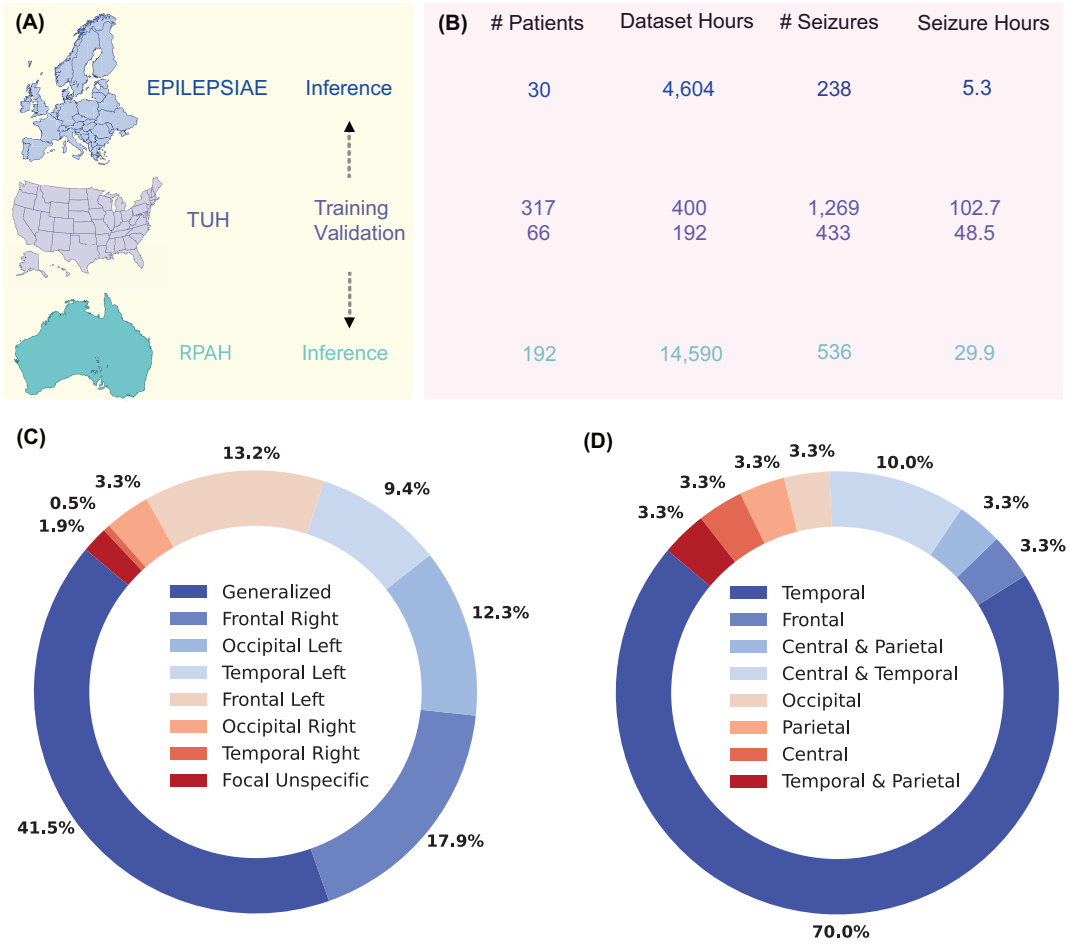

Figure 7.2 – Summary of the datasets being used. The TUH dataset was used as the training and validation dataset and subsequently for inference in both the RPAH and EPILEPSIAE datasets (A). Testing across all datasets incorporates extensive background and seizure-session data, facilitating comprehensive analysis of model efficacy across diverse clinical settings (B). Types of Seizures Across the RPAH Dataset (C) and EPILEPSIAE Dataset (D).

where $\hat{x}_i$ represents the activation value at node i, and φ_{ij} denotes the learnable activation function on the edge connecting node j to node i.

Mathematical Formulation of KANs

KANs utilise the Kolmogorov-Arnold representation theorem to break down a high-dimensional function into a sum of univariate functions. This decomposition is given by the Equ. (7.2):

$$f(x) = \sum_{q=1}^{2n+1} \Phi_q \left(\sum_{p=1}^{n} \varphi_{q,p}(x_p) \right) \tag{7.2}$$

In this formulation, Φ_q and $\varphi_{q,p}$ are univariate functions parameterised as splines, while x_p represents the input features. The inner functions $\varphi_{q,p}(x_p)$ transform the input features into intermediate representations, which are then aggregated and processed by the outer functions Φ_q. This structured approach enables KANs to effectively capture compositional structures and univariate functions, providing a robust framework for function approximation Liu et al. (2025). An overall representation of the KAN structure is seen in Fig. 7.1.

7.2.3 Pre-Processing

We utilised two signal processing techniques, ICA and STFT, to address the challenges associated with raw EEG data. Initially, the EEG signals were divided into 12-second segments, and the ICA algorithm was applied to decompose the signals into 19 independent components using BSS. ICA separates EEG signals into statistically independent components, as represented in Equ. (7.3),

$$T \approx MA^T \ , \tag{7.3}$$

where T contains the EEG data, M contains the time information, and A contains the weights for topographic maps. Pearson correlation was used to identify independent sources strongly associated with eye movement, detected from the 'FP1' and 'FP2' EEG channels. These sources related to eye movement were removed, leaving EEG

signals free of such artifacts. Subsequently, the STFT was applied to the cleaned EEG signals. This involved using a window length of 250 samples (equivalent to 1 second), with 50% overlap, and eliminating the DC component of the transform. As a result, the data dimensions were $(N \times 23 \times 125)$, where N denotes the number of electrodes, 23 denotes the time index, and 125 denotes the frequencies. Data preprocessing is performed separately from the KAN model, ensuring the data is adequately prepared before input into the KAN model for further training/inference, as depicted in Fig. 7.3.

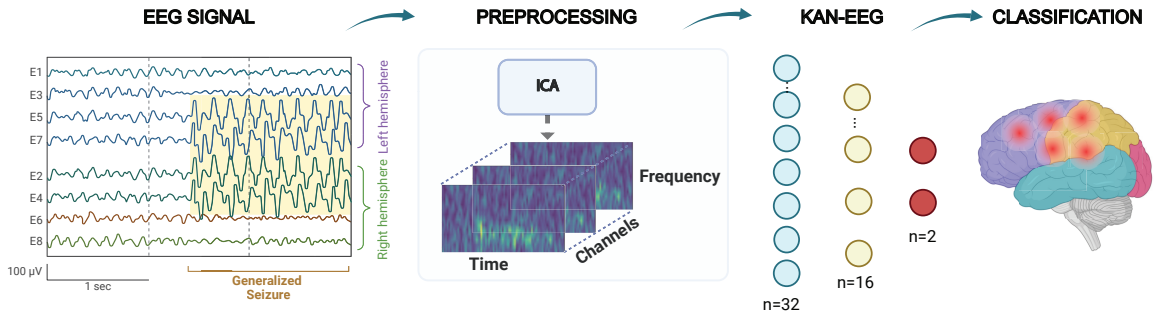

Figure 7.3 – KAN-EEG Seizure System. We analysed different kinds of seizures utilising the TUH dataset for training. We then preprocessed our model utilising Independent Component Analysis ICA and STFT. Subsequently, we incorporate a shallow KAN algorithm, yielding efficient results.

7.3 Experiments and Results

7.3.1 Training and Validation in Sample

Our model was trained and validated using the TUH dataset. We achieved an impressive AUROC score of 0.89 in Fig. 7.4(B). It's important to highlight that our model was trained on 400 hours of data, considerably fewer than the 752 hours used for the ConvLSTM model and the 910 hours for the Transformer model. These preliminary results highlight the exceptional performance of our approach, which not only rivals but also exceeds that of contemporary methods, despite using a smaller

training dataset and a comparable, and even larger, test dataset. Our model demonstrates superior efficacy in accurately detecting epileptic seizures, underscoring its potential for clinical applications. The structure used to obtain the outlined results is highlighted in Table 7.2, and we use this as our baseline for our following assessment.

The training process was performed for 100 epochs, and the trend across different metrics is shown in Fig. 7.4(A). Among these, the discernible reduction in the training loss is evident, gradually converging towards optimal thresholds. Noteworthy is the observation that metrics such as Recall, Precision, and AUROC manifest a discernible tendency to reach the convergence phase approximately around the 20th epoch. Given the nature of a small architecture, training losses converge to an average of 0.2-0.3.

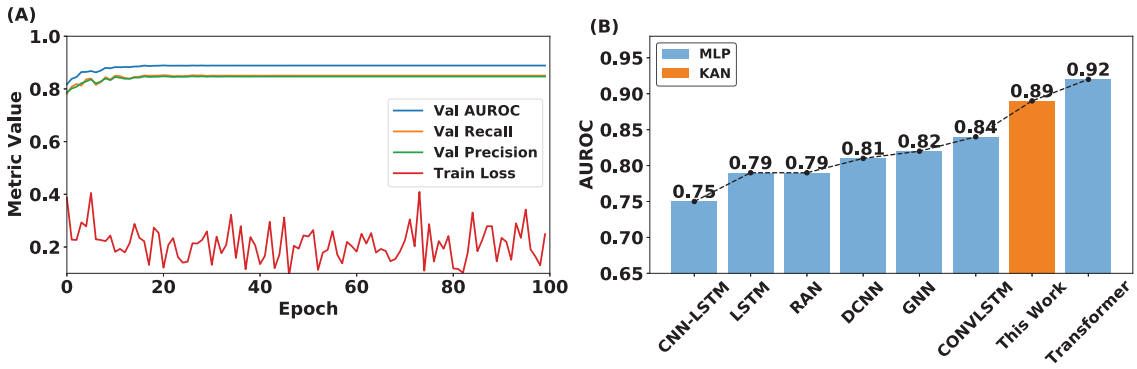

Figure 7.4 – In-Sample results on the TUH dataset. Our model was trained on 400 hours of data, significantly less than the data used to train other models, such as ConvLSTM and Transformer (752 and 910 hours, respectively) (A). A Comparison of a KAN-EEG (Architecture **I-764-256-O** [see Table 7.2]) training with other MLP based models Ahmedt-Aristizabal et al. (2020); Ma et al. (2023); Tang et al. (2022); Wang et al. (2017); Yang et al. (2022b) demonstrates better performance with less training data.

7.3.2 Robustness

The findings presented in Table 7.2 offer compelling insights into the robustness of the KAN-EEG architecture in seizure detection. By systematically reducing the model's complexity by decreasing both the number of layers and the number of hidden neurons, we sought to evaluate the architecture's adaptability and efficiency across vary-

Table 7.2 – **Model assessment robustness** in-sample validation using TUH dataset by decreasing layers and neurons leads to efficiency.

Model Structure	Layers	Hidden Neurons	Neuron Reduction	Model Size Reduction	AUROC	AUPRC	Precision	Recall	F1-Score
I-764-256-O$^{\circ}$ (Large)	2	1020	**Baseline**	**Baseline**	88.9	77.4	85.0	84.7	85.1
I-32-32-O*	2	64	6.27 %	4.5 %	86.9	73.5	83.2	83.0	83.3
I-32-16-O‡ (Small)	2	48	4.71 %	4.38 %	86.7	73.1	82.9	83.1	82.8
I-256-O	1	256	25 %	29.8 %	85.3	70.1	81.3	81.2	80.9

I: Input Data. We utilised 23 time and 125 frequency bins, followed by 19 EEG channels.
O: Output Neurons. We utilised 2 output neurons to represent seizure and non-seizure states.
$\diamond$ Out-of-sample generalisations are shown in Tables 7.3 and 7.4.
$\star$ Out-of-sample generalisations are shown in Tables 7.3 and 7.4.
‡ Out-of-sample generalisations are shown in Tables 7.3 and 7.4.

ing structural configurations. Our analysis reveals that, despite the reduction in model complexity, the KAN architecture consistently maintains high performance across key evaluation metrics, as reflected in the consistently high AUROC values across all models. For instance, the largest tested model has an AUROC of 0.89, while the smallest has an AUROC of 0.85. This indicates the model's ability to effectively differentiate between seizure and non-seizure states, even when architectural complexity is significantly reduced.

Moreover, while marginal decreases in metrics such as precision and recall are observed with model simplification, the overall performance, as quantified by the F1-Score, remains robust. This suggests that our architecture not only preserves its ability to detect seizures accurately but also maintains a balanced trade-off between precision and recall. It is crucial for real-world applications where false positives and negatives have significant implications. These results underscore the resilience of the proposed model to structural modifications, highlighting its adaptability and efficiency in resource-constrained environments. The model's ability to maintain high performance even with reduced computational complexity holds promising implications for practical deployment.

7.3.3 Compact Networks Leads to Out-of-sample Continental generalisation

Can the KAN-EEG generalise beyond in-sample testing? Out-of-sample generalisation is crucial in seizure detection models, as it assesses the model's ability to perform

effectively on unseen data that was not included during training. This is essential for establishing greater robustness and reliability of the model in real-world clinical settings, where variability in patient populations, seizure types, and environmental factors can significantly impact performance. To evaluate generalisation performance, we tested our initial model with 2 layers (764 and 256 neurons, respectively) on two distinct out-of-sample datasets. The RPAH dataset results are presented in Table 7.3, and the EPILEPSIAE dataset results are shown in Table 7.4. The results shown demonstrate some out-of-sample testing ability, albeit not efficiently, in both the RPAH and EPILEPSIAE datasets with AUROC results of 0.60 and 0.55, respectively. The observed results primarily stem from an architectural design showing indications of overfitting on the training dataset. This assertion suggests that the model has become overly specialised in capturing specific information inherent in the training/validation data, consequently compromising its ability to generalise in different datasets. As a result, addressing overfitting is crucial for enhancing the model's robustness and applicability in real-world scenarios, including unseen-world datasets with varying seizure patterns that can be specific to each patient.

To address the challenges of inefficient generalisation, we modified the model to have two layers with 32 and 16 neurons, respectively. Additionally, we conducted an additional evaluation to confirm the robustness of our findings by testing in a model with two layers, each with 32 neurons. As these architectures showed robustness in in-sample testing, with an AUROC of 0.87, we applied them to the generalisation test using training weights when the AUROC metric was not increasing.

As discussed previously, the KAN-EEG model with two layers (764 and 256 neurons) had a lower AUROC than the LTC-FPTT and ConvLSTM models. Interestingly, the KAN-EEG models with two layers (32 and 16 neurons, and 32 and 32 neurons) demonstrated comparable or superior AUROC values relative to these models. For the RPAH dataset, we demonstrated that the proposed compact networks achieve AUROC scores of 0.85 and 0.83, respectively, outperforming existing models. For the EPILEPSIAE dataset, although we did not surpass previous studies, our results were closer to 0.78 and 0.75, respectively.

This suggests that overfitting may occur in models with more complex neuron structures, adversely affecting their generalisation performance. These findings are significant as they indicate the potential for future deployable on-edge KAN algorithms, where memory size compression and robustness are crucial. These findings suggest the potential to replace MLP-based seizure-detection models with the KAN-EEG algorithm. Our framework remains robust and demonstrates promise for practical implementation in real-world applications. It is essential to note that our data was not trained on the full TUH dataset, indicating that there is even more room for improvement to achieve better results.

7.3.4 Computation Time in Training

We utilised a Tesla P100 GPU with 16 GiB of memory to compare the training performance of two neural network configurations of the KAN-EEG model: a larger network with 764-256 neurons and a smaller network with 32-16 neurons per layer respectively. Using a batch size of 64, the training time per epoch for the larger network averaged 4 minutes and 12 seconds, while the smaller network completed an epoch in 1 minute and 1 second. Regarding GPU memory consumption, the more extensive network required approximately 8.8 GiB, whereas the smaller network used only 1.6 GiB. These results underscore the efficiency of this neural network architecture, which is crucial for real-time applications. In contrast, SNNs often require significantly more training time, especially when analysing extended sequences with many timesteps. The KAN-EEG model's fast training performance makes it well-suited for real-time scenarios.

7.4 Discussion

We pioneer the application of shallow KANs to EEG data, leveraging this model's unique structure and capabilities inspired by the Kolmogorov-Arnold representation theorem. The proposed architecture not only achieves a higher accuracy in the in-

Table 7.3 – Out-of-sample generalisation results AUROC on the RPAH dataset.

Year	LTC-FPTT Herbozo Contreras et al. (2024a)	ConvLSTM Yang et al. (2022b)	KAN-EEG ◇	KAN-EEG ⋆	KAN-EEG †
2011	0.88	0.90	0.75	0.95	0.93
2012	0.92	0.91	0.64	0.90	0.88
2013	0.92	0.89	0.62	0.89	0.88
2014	0.78	0.73	0.61	0.80	0.78
2015	0.79	0.82	0.56	0.91	0.89
2016	0.81	0.85	0.55	0.86	0.84
2017	0.80	0.68	0.59	0.72	0.71
2018	0.80	0.72	0.56	0.80	0.79
2019	0.80	0.78	0.62	0.81	0.79
Overall	**0.83**	**0.82**	**0.60**	**0.85**	**0.83**

◇ Architecture **I-764-256-O** (see Table 7.2): 2 Layers (764 and 256 neurons).
⋆ Architecture **I-32-32-O** (see Table 7.2): 2 Layers (32 and 32 neurons).
‡ Architecture **I-32-16-O** (see Table 7.2): 2 Layers (32 and 16 neurons).

Table 7.4 – Out-of-sample generalisation results AUROC on the EPILEPSIAE dataset.

Pat-ID	LTC-FPTT Herbozo Contreras et al. (2024a)	ConvLSTM Yang et al. (2022b)	KAN-EEG ◇	KAN-EEG ⋆	KAN-EEG †
1	0.88	0.88	0.63	0.84	0.79
2	0.82	0.88	0.49	0.58	0.55
3	0.84	0.85	0.60	0.85	0.77
4	0.82	0.89	0.31	0.71	0.58
5	0.79	0.72	0.58	0.73	0.72
6	0.78	0.69	0.42	0.64	0.60
7	0.78	0.81	0.53	0.75	0.76
8	0.79	0.81	0.63	0.88	0.82
9	0.81	0.92	0.65	0.90	0.88
10	0.81	0.68	0.43	0.57	0.49
11	0.81	0.81	0.82	0.96	0.95
12	0.81	0.74	0.48	0.75	0.59
13	0.83	0.86	0.33	0.78	0.63
14	0.83	0.85	0.65	0.84	0.82
15	0.83	0.71	0.58	0.81	0.82
16	0.83	0.85	0.77	0.93	0.91
17	0.83	0.83	0.44	0.65	0.63
18	0.83	0.82	0.52	0.82	0.68
19	0.84	0.91	0.74	0.88	0.87
20	0.84	0.68	0.34	0.71	0.65
21	0.84	0.83	0.47	0.83	0.71
22	0.84	0.95	0.62	0.74	0.70
23	0.85	0.84	0.60	0.89	0.84
24	0.84	0.83	0.76	0.85	0.82
25	0.84	0.62	0.56	0.75	0.73
26	0.84	0.84	0.38	0.67	0.57
27	0.84	0.86	0.60	0.71	0.70
28	0.84	0.91	0.70	0.88	0.88
29	0.84	0.72	0.51	0.79	0.76
30	0.85	0.88	0.43	0.75	0.62
Overall	**0.83**	**0.81**	**0.55**	**0.78**	**0.73**

◇ Architecture **I-764-256-O** (see Table 7.2): 2 Layers (764 and 256 neurons).
⋆ Architecture **I-32-32-O** (see Table 7.2): 2 Layers (32 and 32 neurons).
‡ Architecture **I-32-16-O** (see Table 7.2): 2 Layers (32 and 16 neurons).

sample dataset but can also be generalized across different scalp-EEG, as smaller KANs can outperform much larger MLPs in terms of data fitting and partial differential equation solving, owing to faster neural scaling laws Liu et al. (2025). By applying KANs to EEG data, we expect to unlock new potentials in seizure prediction models, given its ability to learn compositional structures and optimise univariate functions Liu et al. (2025). This novel application could yield more accurate and interpretable predictions, advancing the field of neural network-based seizure prediction. The model's resilience against catastrophic forgetting will also ensure stable and continuous learning, a critical requirement for medical applications involving long-term EEG monitoring. To this extent, our next steps will include deploying this architecture on memristor devices. Recent studies have demonstrated the potential of incorporating KAN network with neuromorphic hardware. Given that memristors are a key component in developing neuromorphic platforms, this approach offers exciting opportunities for deploying such models on non-conventional hardware Peng et al. (2024). Furthermore, approaches to incorporate KAN to neuromorphic frameworks would be to introduce learnable activation functions by replacing the soma of the LIF-based SNN. Alternatively, hybrid ANN-SNN architectures could be explored, where the KAN component serves as the artificial neural network backbone while the spiking component provides event-driven, neuromorphic processing."

7.4.1 Scope and Model Limitations

This research serves as a proof of concept to evaluate whether the KAN model, which utilises a learnable activation function, can form a viable AI-based approach for seizure detection. However, further exploration into the optimal architectural design for improved performance is still needed. Although we have demonstrated that shallow networks tend to have better generalisation and robustness, as described in Table 7.2. The current scope is limited to seizure detection, but we will incorporate insights from these architectures for seizure prediction. Deployability has not been considered within this investigation. However, based on the architecture model, its aim can be tested in future studies. One notable drawback we found in our work

was a linear increase in memory consumption with increasing input data. Therefore, future studies should focus on addressing these challenges.

The data used in this model is in the frequency domain, which utilised a pre-processing step to transform the raw signals into a format suitable for analysis. This additional pre-processing imposes limitations, especially in real-time applications, as it can introduce latency and increase computational demands, making the model less practical for real-world scenarios that require rapid or continuous data processing. Consequently, this limitation affects the model's usability in time-sensitive or resource-constrained environments, such as portable or embedded systems.

7.5 Conclusion

This work has introduced the KANs as a novel approach for epileptic seizure detection using EEG signals. By leveraging its unique architecture and learning properties, which feature learnable edge activation functions rather than fixed node activations, we have demonstrated the potential to achieve comparable accuracy and higher accuracy in both in-sample and out-of-sample seizure detection compared to traditional MLP models. Our results indicate that KANs are a viable alternative to MLP-based models for seizure detection applications, requiring smaller network sizes while maintaining robust performance. The significant advancements presented in this work highlight the potential of KANs to improve clinical outcomes for patients with epilepsy by providing a more accurate, efficient, and interpretable tool for seizure detection, thereby enhancing decision-making, treatment planning, and overall disease management.

Chapter 8

Concluding Remarks

8.1 Thesis Contributions

The central objective of this thesis is to establish foundational principles for *neuromorphic neuromodulation*, with a specific focus on artificial intelligence as the control layer for responsive neurotherapies. Using epilepsy as a clinically relevant case study, this work examines how neuromorphic AI can enable fully embedded closed-loop neurostimulation under strict constraints in latency, power, memory, privacy, and long-term autonomy.

Building on this foundation, the thesis presents a unified software pipeline for seizure analysis that enables robust on-device operation, demonstrating strong out-of-sample generalisation across large-scale, heterogeneous EEG datasets while reducing reliance on extensive optimisation and manual annotation. A major architectural advance is the introduction of liquid-time-constant dynamics and dendrite-inspired computation into spiking neural networks, culminating in a unified framework. Across multiple clinical datasets, this approach improves AUROC and related metrics by approximately 5 to 10% over state-of-the-art baselines, while reducing memory consumption to the kilobyte scale, lowering energy requirements by three to five orders of magnitude, and decreasing both training and inference complexity, establishing a clear pathway toward efficient neuromorphic intelligence.

In parallel, this thesis explores compact and learnable function representations that deliver strong performance at significantly reduced computational cost, offering a practical pathway for integration into neuromorphic systems where memory, power, and latency are primary constraints.

Finally, this work addresses autonomous patient-specific adaptation on the edge by introducing neuromorphic learning mechanisms that support online and few-shot learning under severe resource constraints, enabling personalised, self-responsive closed-loop detection and prediction systems without reliance on cloud-based computation.

Together, these contributions define a clear pathway from conceptual foundations to deployable neuromorphic intelligence, spanning generalisable seizure analysis, low-power architectures, compact function learning, and on-device adaptation for closed-loop neurostimulation.

8.2 Future Research Directions

This thesis establishes neuromorphic computing as a viable foundation for next-generation seizure detection, prediction, and closed-loop neuromodulation. Building on the models, frameworks, and system concepts introduced throughout this work, several important directions for future research emerge, spanning hardware translation, adaptive intelligence, biologically grounded modelling, and clinical deployment.

Neuromorphic Systems for Real-World Deployment

A critical next step is the transition from algorithmic validation to the end-to-end realisation of a neuromorphic system. While this thesis demonstrates hardware-compatible models and validates feasibility on commercial platforms such as Raspberry Pi and AKIDA, future work should establish a dedicated deployment pathway for the proposed architectures, moving beyond proof-of-concept toward purpose-built systems for clinical neurotechnology. Long-term evaluation under continuous operation, addressing stability, fault tolerance, and safety, will be essential for real-world translation.

An important direction in this context is the progression from FPGA-based prototyping to application-specific integrated circuit ASIC implementations. FPGA platforms enable rapid architectural exploration but primarily serve as an intermediate step toward custom silicon. Advancing to ASIC design is essential to realise the power, area, and latency advantages required for clinically viable implantable and wearable systems. Future work should therefore prioritise hardware–algorithm co-design that integrates the Liquid-Dendrite and neuromorphic learning frameworks with digital and mixed-signal architectures, enabling efficient mapping of neuronal dynamics, synaptic plasticity, and online learning rules into silicon.

The benefits observed in this thesis should not be attributed solely to either the hardware platform or the spiking/neuro-inspired algorithms in isolation. Rather, they arise from different levels of hardware-algorithm interaction across the thesis. In chapter 3-5, the models are primarily evaluated on conventional digital platforms, including GPU-based experimentation and Raspberry Pi deployment. In these cases, the observed benefits are mainly algorithmic, arising from compact spiking and neuro-inspired model structures that reduce model complexity, parameter count, and computational demand. However, these platforms do not fully exploit the event-driven nature of spike-based computation. In contrast, chapter 6 more directly demonstrates hardware-algorithm co-design, as the proposed approach is mapped onto a low-power neuromorphic chip capable of exploiting sparse/event-driven execution.

Continual and Personalised Learning on the Fly

This thesis introduces adaptive learning paradigms for seizure detection and prediction based on few-shot edge learning under strict resource constraints, as exemplified by deployment on neuromorphic platforms such as Akida. Building on this foundation, future research should extend these ideas toward lifelong, personalised neuromorphic learning, addressing fundamental challenges including catastrophic forgetting, stability–plasticity trade-offs, and robust adaptation under highly imbalanced seizure data distributions.

Neural Encoding for Neuromorphic Time-Series Processing

An important and largely underexplored opportunity lies in the design of neural encoding mechanisms for time-series data. Much of the neuromorphic literature has focused on encoding strategies developed for vision tasks, where spatial structure dominates. In contrast, neural signals such as EEG exhibit highly non-stationary and non-deterministic temporal dynamics that require fundamentally different representational approaches. Future work should therefore investigate encoding schemes tailored to continuous-time signals, with particular emphasis on domains such as audio as an intermediate test-bed, where rich temporal structure and multi-hertz dynamics provide a natural bridge toward more complex neural data. Developing encoding strategies specifically for time-series modalities may unlock significant gains in both performance and energy efficiency for neuromorphic seizure analysis.

Toward Closed-Loop Neuromorphic Neuromodulation

A long-term goal of this research is the realisation of fully autonomous closed-loop neuromodulation. Future work should move beyond isolated detection or prediction modules toward unified neuromorphic frameworks that integrate sensing, learning, decision-making, and stimulation. Such systems have the potential to support personalised, adaptive therapies that operate entirely on-device, minimising latency, preserving data privacy, and enabling long-term scalability in clinical settings. Achieving this vision will require sustained interdisciplinary collaboration across neuroscience, hardware design, and clinical practice, ensuring that algorithmic advances remain aligned with translational and regulatory realities.

List of References

Abdelhameed, A. M., Daoud, H. G., and Bayoumi, M. (2018). Epileptic seizure detection using deep convolutional autoencoder. In *IEEE International Workshop on Signal Processing Systems (SiPS)*, pages 223–228, Cape Town, South Africa.

Afsaneh, E. and Zarei Ghobadi, M. (2023). The role of neuromodulation-related technologies in neurology for the next 10 years. *Brain-Apparatus Communication: A Journal of Bacomics*, 2(1):2147405.

Aguilar, I., Bersani-Veroni, T., Herbozo Contreras, L. F., Love, T., Nikpour, A., Querlioz, D., and Kavehei, O. (2025). Continuous metaplastic training on brain signals. *npj Unconventional Computing*, 2(1):9.

Ahmedt-Aristizabal, D., Fernando, T., Denman, S., Petersson, L., Aburn, M. J., and Fookes, C. (2020). Neural memory networks for seizure type classification. In *42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 569–575, Montreal, QC, Canada.

Akopyan, F., Sawada, J., Cassidy, A., Alvarez-Icaza, R., Arthur, J., Merolla, P., Imam, N., Nakamura, Y., Datta, P., Nam, G.-J., et al. (2015). Truenorth: Design and tool flow of a 65 mw 1 million neuron programmable neurosynaptic chip. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 34(10):1537–1557.

Akrout, M. (2019). On the adversarial robustness of neural networks without weight transport. In *Real Neurons & Hidden Units: Future directions at the intersection of neuroscience and artificial intelligence @ NeurIPS 2019*, Vancouver, Canada.

Ali, O., Saif-ur Rehman, M., Dyck, S., Glasmachers, T., Iossifidis, I., and Klaes, C. (2022). Enhancing the decoding accuracy of EEG signals by the introduction of anchored-STFT and adversarial data augmentation method. *Scientific Reports*, 12(1):4245.

Alotaibi, S. M., ur Rahman, A., Basheer, M. I., and Khan, M. A. (2021). Ensemble machine learning based identification of pediatric epilepsy. *Computers, Materials & Continua*, 68(1):149–165.

Amann, J., Blasimme, A., Vayena, E., Frey, D., and Madai, V. I. (2020). Explainability for artificial intelligence in healthcare: A multidisciplinary perspective. *BMC Medical Informatics and Decision Making*, 20(1):1–9.

Ariav, G., Polsky, A., and Schiller, J. (2003). Submillisecond precision of the input-output transformation function mediated by fast sodium dendritic spikes in basal dendrites of CA1 pyramidal neurons. *Journal of Neuroscience*, 23(21):7750–7758.

Ashourvan, A., Pequito, S., Khambhati, A. N., Mikhail, F., Baldassano, S. N., Davis, K. A., Lucas, T. H., Vettel, J. M., Litt, B., Pappas, G. J., and Others (2020). Model-based design for seizure control by stimulation. *Journal of Neural Engineering*, 17(2):26009.

Avcu, M. T., Zhang, Z., and Chan, D. W. S. (2019). Seizure detection using least EEG channels by deep convolutional neural network. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1120–1124, Brighton, UK.

Balzekas, I., Sladky, V., Nejedly, P., Brinkmann, B. H., Crepeau, D., Mivalt, F., Gregg, N. M., Pal Attia, T., Marks, V. S., Wheeler, L., et al. (2021). Invasive electrophysiology for circuit discovery and study of comorbid psychiatric disorders in patients with epilepsy: Challenges, opportunities, and novel technologies. *Frontiers in Human Neuroscience*, 15:702605.

Banerjee, P. N., Filippi, D., and Hauser, W. A. (2009). The descriptive epidemiology of epilepsy - A review. *Epilepsy Research*, 85:31–45.

Bartels, J., Gallou, O., Ito, H., Cook, M., Sarnthein, J., Indiveri, G., and Ghosh, S. (2025). Event driven neural network on a mixed signal neuromorphic processor for EEG based epileptic seizure detection. *Scientific Reports*, 15(1):15965.

Benjamin, B. V. and et al. (2014). Neurogrid: A mixed-analog-digital multichip system for large-scale neural simulations. *Proceedings of the IEEE*, 102(5):699–716.

Bergey, G. K., Morrell, M. J., Mizrahi, E. M., Goldman, A., King-Stephens, D., Nair, D., Srinivasan, S., Jobst, B., Gross, R. E., Shields, D. C., and Others (2015). Long-term treatment with responsive brain stimulation in adults with refractory partial seizures. *Neurology*, 84(8):810–817.

Beveridge, M. and Pereira, L. (2022). Interpretable spatiotemporal forecasting of arctic sea ice concentration at seasonal lead times. In *NeurIPS 2022 Workshop on Tackling Climate Change with Machine Learning*.

Birjandtalab, J., Heydarzadeh, M., and Nourani, M. (2017). Automated EEG-based epileptic seizure detection using deep neural networks. In *IEEE International Conference on Healthcare Informatics (ICHI)*, pages 552–555.

Blouw, P., Choo, X., Hunsberger, E., and Eliasmith, C. (2019). Benchmarking keyword spotting efficiency on neuromorphic hardware. In *Proceedings of the 7th Annual Neuro-Inspired Computational Elements Workshop*, New York, NY, USA.

Bohnstingl, T., Scherr, F., Pehle, C., Meier, K., and Maass, W. (2019). Neuromorphic hardware learns to learn. *Frontiers in Neuroscience*, 13:483.

Bouaziz, B., Chaari, L., Batatia, H., and Quintero-Rincón, A. (2019). Epileptic seizure detection using a convolutional neural network. In *Digital Health Approach for Predictive, Preventive, Personalised and Participatory Medicine*, pages 79–86. Springer.

Brice, J. and Mclellan, L. (1980). Suppression of intention tremor by contingent deep-brain stimulation. *The Lancet*, 315(8180):1221–1222.

Broccard, F. D., Joshi, S., Wang, J., and Cauwenberghs, G. (2017). Neuromorphic neural interfaces: from neurophysiological inspiration to biohybrid coupling with nervous systems. *Journal of Neural Engineering*, 14(4):041002.

Buccelli, S., Bornat, Y., Colombi, I., Ambroise, M., Martines, L., Pasquale, V., Bisio, M., Tessadori, J., Nowak, P., Grassia, F., et al. (2019). A neuromorphic prosthesis to restore communication in neuronal networks. *iScience*, 19:402–414.

Burelo, K., Ramantani, G., Indiveri, G., and Sarnthein, J. (2022a). A neuromorphic spiking neural network detects epileptic high frequency oscillations in the scalp EEG. *Scientific Reports*, 12(1):1798.

Burelo, K., Sharifshazileh, M., Indiveri, G., and Sarnthein, J. (2022b). Automatic detection of high-frequency oscillations with neuromorphic spiking neural networks. *Frontiers in Neuroscience*, 16:861480.

Carrette, S., Boon, P., Sprengers, M., Raedt, R., and Vonck, K. (2015). Responsive neurostimulation in epilepsy. *Expert Review of Neurotherapeutics*, 15(12):1445–1454.

Cartiglia, M., Rubino, A., Narayanan, S., Frenkel, C., Haessig, G., Indiveri, G., and Payvand, M. (2022). Stochastic dendrites enable online learning in mixed-signal neuromorphic processing systems. In *IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 476–480, Austin, TX, USA.

Chahine, M., Hasani, R., Kao, P., Ray, A., Shubert, R., Lechner, M., Amini, A., and Rus, D. (2023). Robust flight navigation out of distribution with liquid neural networks. *Science Robotics*, 8(77):eadc8892.

Chen, G., Peng, P., Li, G., and Tian, Y. (2023). Training full spike neural networks via auxiliary accumulation pathway. *arXiv preprint arXiv:2301.11929*.

Chen, W., Kirkby, L., Kotzev, M., Song, P., Gilron, R., and Pepin, B. (2021). The role of large-scale data infrastructure in developing next-generation deep brain stimulation therapies. *Frontiers in Human Neuroscience*, 15:717401.

Chen, X., Ji, J., Ji, T., and Li, P. (2018). Cost-sensitive deep active learning for epileptic seizure detection. In *Proceedings of the ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, pages 226–235, New York, NY, USA.

Chiappalone, M., Cota, V. R., Carè, M., Di Florio, M., Beaubois, R., Buccelli, S., Barban, F., Brofiga, M., Averna, A., Bonacini, F., et al. (2022). Neuromorphic-based neuroprostheses for brain rewiring: State-of-the-art and perspectives in neuroengineering. *Brain Sciences*, 12(11):1578.

Chicca, E., Stefanini, F., Bartolozzi, C., and Indiveri, G. (2014). Neuromorphic electronic circuits for building autonomous cognitive systems. *Proceedings of the IEEE*, 102(9):1367–1388.

Choi, G., Park, C., Kim, J., Cho, K., Kim, T.-J., Bae, H., Min, K., Jung, K.-Y., and Chong, J. (2019). A novel multi-scale 3D CNN with deep neural network for epileptic seizure detection. In *IEEE International Conference on Consumer Electronics (ICCE)*, pages 1–2, Las Vegas, NV, USA.

Christensen, D. V., Dittmann, R., Linares-Barranco, B., Sebastian, A., Le Gallo, M., Redaelli, A., Slesazeck, S., Mikolajick, T., Spiga, S., Menzel, S., et al. (2022). 2022 roadmap on neuromorphic computing and engineering. *Neuromorphic Computing and Engineering*, 2(2):022501.

Chua, A., Jordan, M. I., and Muller, R. (2022). SOUL: An energy-efficient unsupervised online learning seizure detection classifier. *IEEE Journal of Solid-State Circuits*, 57(8):2532–2544.

Clément, C. (2019). *Brain-Computer Interface Technologies*. Springer.

Co., L. B. (2021). Ultra-thin LiPo Battery.

Contreras, L. F. H., Yu, L., Huang, Z., Aguilar, I., Nikpour, A., and Kavehei, O. (2025). Spiking neural networks with liquid time-constant dynamics and dendritic branches for efficient time-domain edge-based seizure detection. *Machine Learning: Health*, 1(1):015001.

Cook, M. J., O'Brien, T. J., Berkovic, S. F., Murphy, M., Morokoff, A., Fabinyi, G., D'Souza, W., Yerra, R., Archer, J., Litewka, L., and Others (2013). 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*, 12(6):563–571.

Corradi, F. and Indiveri, G. (2015). A neuromorphic event-based neural recording system for smart brain-machine-interfaces. *IEEE Transactions on Biomedical Circuits and Systems*, 9(5):699–709.

Costa, F., Schaft, E. V., Huiskamp, G., Aarnoutse, E. J., van't Klooster, M. A., Krayenbühl, N., Ramantani, G., Zijlmans, M., Indiveri, G., and Sarnthein, J. (2024). Robust compression and detection of epileptiform patterns in ecog using a real-time spiking neural network hardware framework. *Nature Communications*, 15(1):3255.

Covert, I. C., Krishnan, B., Najm, I., Zhan, J., Shore, M., Hixson, J., and Po, M. J. (2019). Temporal graph convolutional networks for automatic seizure detection. In *Machine Learning for Healthcare Conference*, pages 160–180, Michigan, USA. PMLR.

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2(4):303–314.

Dammers, J., Schiek, M., Boers, F., Silex, C., Zvyagintsev, M., Pietrzyk, U., and Mathiak, K. (2008). Integration of amplitude and phase statistics for complete artifact removal in independent components of neuromagnetic recordings. *IEEE Transactions on Biomedical Engineering*, 55(10):2353–2362.

Davenport, T. and Kalakota, R. (2019). The potential for artificial intelligence in healthcare. *Future Healthcare Journal*, 6(2):94.

Davies, M. and other authors (2018). Loihi: a neuromorphic manycore processor with on-chip learning. *IEEE Micro*, 38:82–99.

De Lange, M., Aljundi, R., Masana, M., Parisot, S., Jia, X., Leonardis, A., Slabaugh, G., and Tuytelaars, T. (2021). A continual learning survey: Defying forgetting in classification tasks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(7):3366–3385.

Degnan, B., Marr, B., and Hasler, J. (2015). Assessing trends in performance per watt for signal processing applications. *IEEE Transactions on Very Large Scale Integration Systems*, 24:58–66.

Delgado, J. M., Roberts, W. W., and Miller, N. E. (1954). Learning motivated by electrical stimulation of the brain. *American Journal of Physiology-Legacy Content*, 179(3):587–593.

Dellaferrera, G. and Kreiman, G. (2022). Error-driven input modulation: Solving the credit assignment problem without a backward pass. In *International Conference on Machine Learning*, pages 4937–4955, Maryland, USA. PMLR.

Demler, M. (2019). Brainchip Akida is a fast learner, spiking-neural-network processor identifies patterns in unlabeled data. *Microprocessor Report*.

Deng, L., Wang, G., Li, G., Li, S., Liang, L., Zhu, M., Wu, Y., Yang, Z., Zou, Z., Pei, J., et al. (2020). Tianjic: A unified and scalable chip bridging spike-based and continuous neural computation. *IEEE Journal of Solid-State Circuits*, 55(8):2228–2246.

Denison, T., Consoer, K., Santa, W., Avestruz, A.-T., Cooley, J., and Kelly, A. (2007). A 2μW 100 nV/$\sqrt{\text{hz}}$ chopper-stabilized instrumentation amplifier for chronic measurement of neural field potentials. *IEEE Journal of Solid-State Circuits*, 42(12):2934–2945.

Denison, T. and Morrell, M. J. (2022). Neuromodulation in 2035: The neurology future forecasting series. *Neurology*, 98(2):65–72.

Denison, T. J. and Santa, W. A. (2017). Seizure prediction. https://patents.google.com/patent/US9788750B2. US Patent 9,788,750 B2.

Deuschl, G., Paschen, S., and Witt, K. (2013). Clinical outcome of deep brain stimulation for Parkinson's disease. *Handbook of Clinical Neurology*, 116:107–128.

Do Valle, B. G., Cash, S. S., and Sodini, C. G. (2016). Low-power, 8-channel EEG recorder and seizure detector ASIC for a subdermal implantable system. *IEEE Transactions on Biomedical Circuits and Systems*, 10(6):1058–1067.

Donati, E. and Indiveri, G. (2023). Neuromorphic bioelectronic medicine for nervous system interfaces: from neural computational primitives to medical applications. *Progress in Biomedical Engineering*, 5(1):013002.

Donati, E., Payvand, M., Risi, N., Krause, R., and Indiveri, G. (2019). Discrimination of EMG signals using a neuromorphic implementation of a spiking neural network. *IEEE Transactions on Biomedical Circuits and Systems*, 13(5):795–803.

Donati, E. and Valle, G. (2024). Neuromorphic hardware for somatosensory neuroprostheses. *Nature Communications*, 15(1):556.

Dümpelmann, M. (2019). Early seizure detection for closed loop direct neurostimulation devices in epilepsy. *Journal of Neural Engineering*, 16(4):041001.

Duun-Henriksen, J., Baud, M., Richardson, M. P., Cook, M., Kouvas, G., Heasman, J. M., Friedman, D., Peltola, J., Zibrandtsen, I. C., and Kjaer, T. W. (2020). A new era in electroencephalographic monitoring? Subscalp devices for ultra–long-term recordings. *Epilepsia*, 61(9):1805–1817.

Eissa, S., Corradi, F., de Putter, F., Stuijk, S., and Corporaal, H. (2023). QMTS: Fixed-point quantization for multiple-timescale spiking neural networks. In *International Conference on Artificial Neural Networks*, pages 407–419, Crete, Greece. Springer.

Emami, A., Kunii, N., Matsuo, T., Shinozaki, T., Kawai, K., and Takahashi, H. (2019). Autoencoding of long-term scalp electroencephalogram to detect epileptic seizure for diagnosis support system. *Computers in Biology and Medicine*, 110:227–233.

Esser, S. K., Merolla, P. A., Arthur, J. V., Cassidy, A. S., Appuswamy, R., Andreopoulos, A., Berg, D. J., McKinstry, J. L., Melano, T., Barch, D. R., et al. (2016). From the cover: Convolutional networks for fast, energy-efficient neuromorphic computing. *Proceedings of the National Academy of Sciences of the United States of America*, 113(41):11441.

Famm, K., Litt, B., Tracey, K. J., Boyden, E. S., and Slaoui, M. (2013). A jump-start for electroceuticals. *Nature*, 496(7444):159–161.

Fang, Z., Leung, H., and Choy, C. S. (2018). Spatial temporal GRU convnets for vision-based real time epileptic seizure detection. In *15th IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 1026–1029, Washington, D.C., United States.

Fares, H., Ronchini, M., Zamani, M., Farkhani, H., and Moradi, F. (2022). In the realm of hybrid brain: Human brain and AI. *arXiv preprint arXiv:2210.01461*.

Fisher, R. S., Acevedo, C., Arzimanoglou, A., Bogacz, A., Cross, J. H., Elger, C. E., Engel Jr, J., Forsgren, L., French, J. A., Glynn, M., et al. (2014). ILAE official report: A practical clinical definition of epilepsy. *Epilepsia*, 55(4):475–482.

Frenkel, C., Bol, D., and Indiveri, G. (2023). Bottom-up and top-down approaches for the design of neuromorphic processing systems: Tradeoffs and synergies between natural and artificial intelligence. *Proceedings of the IEEE*, 111(6):623–652.

Frenkel, C., Legat, J.-D., and Bol, D. (2020). A 28-nm convolutional neuromorphic processor enabling online learning with spike-based retinas. In *International Symposium on Circuits and Systems (ISCAS)*, pages 1–5, Seville, Spain. IEEE.

Fukumori, K., Nguyen, H. T. T., Yoshida, N., and Tanaka, T. (2019). Fully data-driven convolutional filters with deep learning models for epileptic spike detection. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2772–2776, Brighton, UK.

Furber, S. B., Galluppi, F., Temple, S., and Plana, L. A. (2014). The SpiNNaker Project. *Proceedings of the IEEE*, 102(5):652–665.

Gallou, O., Bartels, J., Ghosh, S., Schindler, K., Sarnthein, J., and Indiveri, G. (2024). Online epileptic seizure detection in long-term iEEG recordings using mixed-signal neuromorphic circuits. In *IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pages 1–5, Xi'an, China.

Ganzer, P. D., Darrow, M. J., Meyers, E. C., Solorzano, B. R., Ruiz, A. D., Robertson, N. M., Adcock, K. S., James, J. T., Jeong, H. S., Becker, A. M., Goldberg, M. P., Pruitt, D. T., Hays, S. A., Kilgard, M. P., and Rennaker Robert L, I. I. (2018). Closed-loop neuromodulation restores network connectivity and motor control after spinal cord injury. *eLife*, 7:e32058.

Geirhos, R., Jacobsen, J.-H., Michaelis, C., Zemel, R., Brendel, W., Bethge, M., and Wichmann, F. A. (2020). Shortcut learning in deep neural networks. *Nature Machine Intelligence*, 2(11):665–673.

Geller, E. B., Skarpaas, T. L., Gross, R. E., Goodman, R. R., Barkley, G. L., Bazil, C. W., et al. (2017). Brain-responsive neurostimulation in patients with medically intractable mesial temporal lobe epilepsy. *Epilepsia*, 58(6):994–1004.

Geng, M., Zhou, W., Liu, G., Li, C., and Zhang, Y. (2020). Epileptic seizure detection based on stockwell transform and bidirectional long short-term memory. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(3):573–580.

Getty, N., Brettin, T., Jin, D., Stevens, R., and Xia, F. (2021). Deep medical image analysis with representation learning and neuromorphic computing. *Interface Focus*, 11(1).

Ghazali, S. M., Alizadeh, M., Mazloum, J., and Baleghi, Y. (2022). Modified binary salp swarm algorithm in EEG signal classification for epilepsy seizure detection. *Biomedical Signal Processing and Control*, 78:103858.

Ghorbani, R., Reinders, M. J., and Tax, D. M. (2024). Personalized anomaly detection in PPG data using representation learning and biometric identification. *Biomedical Signal Processing and Control*, 94:106216.

Giftakis, J. E., Wu, J., and Nelson, D. E. (2014). Seizure probability metrics. https://patents.google.com/patent/US8812098B2. US Patent 8,812,098 B2.

Goddard, G. V., McIntyre, D. C., and Leech, C. K. (1969). A permanent change in brain function resulting from daily electrical stimulation. *Experimental Neurology*, 25(3):295–330.

Golmohammadi, M., Harati Nejad Torbati, A. H., Lopez de Diego, S., Obeid, I., and Picone, J. (2019). Automatic analysis of EEGs using big data and hybrid deep learning architectures. *Frontiers in Human Neuroscience*, 13:76.

Gordon, B., Lesser, R. P., Rance, N. E., Hart Jr, J., Webber, R., Uematsu, S., and Fisher, R. S. (1990). Parameters for direct cortical electrical stimulation in the human: histopathologic confirmation. *Electroencephalography and Clinical Neurophysiology*, 75(5):371–377.

Greene, B. R., Boylan, G. B., Reilly, R. B., de Chazal, P., and Connolly, S. (2007). Combination of EEG and ECG for improved automatic neonatal seizure detection. *Clinical Neurophysiology*, 118(6):1348–1359.

Grübl, A., Billaudelle, S., Cramer, B., Karasenko, V., and Schemmel, J. (2020). Verification and design methods for the brainscales neuromorphic hardware system. *Journal of Signal Processing Systems*, 92:1277–1292.

Gunasekaran, S., Kembay, A., Ladret, H., Zhu, R.-J., Perrinet, L., Kavehei, O., and Eshraghian, J. (2025). A predictive approach to enhance time-series forecasting. *Nature Communications*, 16(1):8645.

Hannun, A. Y., Rajpurkar, P., Haghpanahi, M., Tison, G. H., Bourn, C., Turakhia, M. P., and Ng, A. Y. (2019). Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature Medicine*, 25(1):65–69.

Harrer, S., Kiral-Kornek, F. I., Mashford, B. S., Subhrajit, R., and Saha, S. (2020). Seizure detection, prediction and prevention using neurostimulation technology and deep neural network. https://patents.google.com/patent/US10596377B2. US Patent 10,596,377 B2.

Hasani, R., Lechner, M., Amini, A., Liebenwein, L., Ray, A., Tschaikowski, M., Teschl, G., and Rus, D. (2022). Closed-form continuous-time neural networks. *Nature Machine Intelligence*, pages 1–12.

Hasani, R., Lechner, M., Amini, A., Rus, D., and Grosu, R. (2021). Liquid time-constant networks. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(9):7657–7666.

Hasani, R., Lechner, M., Wang, T.-H., Chahine, M., Amini, A., and Rus, D. (2023). Liquid structural state-space models. In *The Eleventh International Conference on Learning Representations*, Kigali, Rwanda.

Haykin, S. (1998). *Neural Networks: A Comprehensive Foundation*. Prentice Hall PTR.

Heck, C. N., King-Stephens, D., Massey, A. D., Nair, D. R., Jobst, B. C., Barkley, G. L., et al. (2014). Two-year seizure reduction in adults with medically intractable partial onset epilepsy treated with responsive neurostimulation: Final results of the RNS System Pivotal trial. *Epilepsia*, 55(3):432–441.

Herbozo Contreras, L. F., Huang, Z., Yu, L., Nikpour, A., and Kavehei, O. (2024a). Biological plausible algorithm for seizure detection: Toward AI-enabled electroceuticals at the edge. *APL Machine Learning*, 2(2).

Herbozo Contreras, L. F., Truong, N. D., Eshraghian, J. K., Xu, Z., Huang, Z., Bersani-Veroni, T. V., Aguilar, I., Leung, W. H., Nikpour, A., and Kavehei, O. (2024b). Neuromorphic neuromodulation: Towards the next generation of closed-loop neurostimulation. *PNAS Nexus*, 3(11):488.

Herbozo Contreras, L. F., Yu, L., Zhang, Z., Huang, Z., Nikpour, A., and Kavehei, O. (2024c). Tiny dLIF: A dendritic spiking neural network enabling a time-domain energy-efficient seizure detection system. *Neuromorphic Computing and Engineering*.

Hinton, G. (2022). The forward-forward algorithm: Some preliminary investigations. *arXiv preprint arXiv:2212.13345*.

Hofmeister, M., Memedovich, A., Brown, S., Saini, M., Dowsett, L. E., Lorenzetti, D. L., McCarron, T. L., MacKean, G., and Clement, F. (2020). Effectiveness of Neurostimulation Technologies for the Management of Chronic Pain: A Systematic Review. *Neuromodulation: Technology at the Neural Interface*, 23(2):150–157.

Horowitz, M. (2014). 1.1 computing's energy problem (and what we can do about it). In *IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, pages 10–14, San Francisco, CA, USA.

Hossain, M. S., Amin, S. U., Alsulaiman, M., and Muhammad, G. (2019). Applying deep learning for epilepsy seizure detection and brain mapping visualization. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 15(1s):1–17.

Hsieh, Y.-Y., Lin, Y.-C., and Yang, C.-H. (2022). A 96.2-nJ/class neural signal processor with adaptable intelligence for seizure prediction. *IEEE Journal of Solid-State Circuits*, 58(1):167–176.

Hu, Y., Zheng, Q., Jiang, X., and Pan, G. (2023). Fast-SNN: fast spiking neural network by converting quantized ANN. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(12).

Huang, Z., Herbozo Contreras, L. F., Yu, L., Truong, N. D., Nikpour, A., and Kavehei, O. (2024a). S4D-ECG: A shallow state-of-the-art model for cardiac abnormality classification. *Cardiovascular Engineering and Technology*, pages 1–12.

Huang, Z., Leung, W. H., Yu, L., Herbozo Contreras, L. F., Zhang, Z., Truong, N. D., Nikpour, A., and Kavehei, O. (2024b). On-device edge-learning for cardiac abnormality detection using a bio-inspired and spiking shallow network. *APL Machine Learning*, 2(2).

HumanBrainProject (2023). Hardware—The human brain project. https://www.humanbrainproject.eu/en/science-development/focus-areas/neuromorphic-computing/hardware/. Accessed 21 June 2023.

Hutson, T., Pizarro, D., Pati, S., and Iasemidis, L. D. (2018). Predictability and resetting in a case of convulsive status epilepticus. *Frontiers in Neurology*, 9:172.

Iasemidis, L. D., Shiau, D.-S., Chaovalitwongse, W., Sackellares, J. C., Pardalos, P. M., Principe, J. C., Carney, P. R., Prasad, A., Veeramani, B., and Tsakalis, K. (2003). Adaptive epileptic seizure prediction system. *IEEE Transactions on Biomedical Engineering*, 50(5):616–627.

Ivanov, D., Chezhegov, A., Kiselev, M., Grunin, A., and Larionov, D. (2022). Neuromorphic artificial intelligence systems. *Frontiers in Neuroscience*, 16:1513.

Jabri, M. and Flower, B. (1992). Weight perturbation: An optimal architecture and learning technique for analog VLSI feedforward and recurrent multilayer networks. *IEEE Transactions on Neural Networks*, 3(1):154–157.

Jarosiewicz, B. and Morrell, M. (2021). The RNS System: brain-responsive neurostimulation for the treatment of epilepsy. *Expert Review of Medical Devices*, 18(2):129–138.

Jimenez-Shahed, J. (2021). Device profile of the Percept PC Deep Brain Stimulation system for the treatment of Parkinson's disease and related disorders. *Expert Review of Medical Devices*, 18(4):319–332.

Jing, J., Sun, H., Kim, J. A., Herlopian, A., Karakis, I., Ng, M., Halford, J. J., et al. (2020). Development of Expert-Level Automated Detection of Epileptiform Discharges During Electroencephalogram Interpretation. *JAMA Neurology*, 77(1):103–108.

Jobst, B. C., Kapur, R., Barkley, G. L., Bazil, C. W., Berg, M. J., Bergey, G. K., et al. (2017). Brain-responsive neurostimulation in patients with medically intractable seizures arising from eloquent and other neocortical areas. *Epilepsia*, 58(6):1005–1014.

Karuppiah Ramachandran, V. R., Alblas, H. J., Le, D. V., and Meratnia, N. (2018). Towards an online seizure advisory system - An adaptive seizure prediction framework using active learning heuristics. *Sensors*, 18(6).

Ker, M.-D., Lin, C.-Y., and Chen, W.-L. (2011). Stimulus driver for epilepsy seizure suppression with adaptive loading impedance. *Journal of Neural Engineering*, 8(6):066008.

Kerzenmacher, S. (2013). 6 - Biofuel cells as sustainable power sources for implantable systems. In Inmann, A. and Hodgins, D., editors, *Implantable Sensor Systems for Medical Applications*, pages 183–212. Woodhead Publishing.

Kim, S., Kang, S., Kim, J., Lee, D., Kim, S., Lee, J., Jang, K.-I., Oh, Y.-S., Rah, J.-C., Huh, M. S., et al. (2021). Closed-loop neuromodulation for Parkinson's disease: Current state and future directions. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, 7(4):209–223.

Kiourti, A. and Nikita, K. S. (2012). A review of implantable patch antennas for biomedical telemetry: Challenges and solutions [wireless corner]. *IEEE Antennas and Propagation Magazine*, 54(3):210–228.

Kiral-Kornek, I., Roy, S., Nurse, E., Mashford, B., Karoly, P., Carroll, T., Payne, D., Saha, S., Baldassano, S., O'Brien, T., et al. (2018). Epileptic seizure prediction using big data and deep learning: toward a mobile system. *EBioMedicine*, 27:103–111.

Kokoszka, M. A., Panov, F., Vega-Talbott, M. L., McGoldrick, P. E., Wolf, S. M., and Ghatan, S. (2018). Treatment of medically refractory seizures with responsive neurostimulation: 2 pediatric cases. *Journal of Neurosurgery: Pediatrics PED*, 21(4):421–427.

Kremen, V., Brinkmann, B. H., Kim, I., Guragain, H., Nasseri, M., Magee, A. L., Attia, T. P., Nejedly, P., Sladky, V., Nelson, N., et al. (2018). Integrating brain implants with local and distributed computing devices: a next generation epilepsy management system. *IEEE Journal of Translational Engineering in Health and Medicine*, 6:1–12.

Krucoff, M. O., Wozny, T. A., Lee, A. T., Rao, V. R., and Chang, E. F. (2021). Operative technique and lessons learned from surgical implantation of the neuropace responsive neurostimulation® system in 57 consecutive patients. *Operative Neurosurgery*, 20(2):E98–E109.

Kulkarni, S. R. and Rajendran, B. (2018). Spiking neural networks for handwritten digit recognition-supervised learning and network optimization. *Neural Networks*, 103:118–127.

Kupsch, A., Benecke, R., Müller, J., Trottenberg, T., Schneider, G.-H., Poewe, W., Eisner, W., Wolters, A., Müller, J.-U., Deuschl, G., et al. (2006). Pallidal deep-brain stimulation in primary generalized or segmental dystonia. *New England Journal of Medicine*, 355(19):1978–1990.

Laborieux, A., Ernoult, M., Hirtzlin, T., and Querlioz, D. (2021). Synaptic metaplasticity in binarized neural networks. *Nature Communications*, 12(1):2549.

Lagarde, S., Bonini, F., McGonigal, A., Chauvel, P., Gavaret, M., Scavarda, D., Carron, R., Régis, J., Aubert, S., Villeneuve, N., et al. (2016). Seizure-onset patterns in focal cortical dysplasia and neurodevelopmental tumors: relationship with surgical prognosis and neuropathologic subtypes. *Epilepsia*, 57(9):1426–1435.

Laxton, A. W., Tang-Wai, D. F., McAndrews, M. P., Zumsteg, D., Wennberg, R., Keren, R., Wherrett, J., Naglie, G., Hamani, C., Smith, G. S., et al. (2010). A phase I trial of deep brain stimulation of memory circuits in Alzheimer's disease. *Annals of Neurology*, 68(4):521–534.

Le, T. X., Le, T. T., Dinh, V. V., Tran, Q. L., Nguyen, L. T., and Nguyen, D. T. (2018). Deep learning for epileptic spike detection. *VNU Journal of Science: Computer Science and Communication Engineering*, 33(2):1–13.

Lechner, M., Hasani, R., Amini, A., Henzinger, T. A., Rus, D., and Grosu, R. (2020). Neural circuit policies enabling auditable autonomy. *Nature Machine Intelligence*, 2(10):642–652.

Lee, J. H., Delbruck, T., and Pfeiffer, M. (2016). Training deep spiking neural networks using backpropagation. *Frontiers in Neuroscience*, 10:508.

Levy, R., Deer, T. R., Poree, L., Rosen, S. M., Kapural, L., Amirdelfan, K., Soliday, N., Leitner, A., and Mekhail, N. (2019). Multicenter, randomized, double-blind study protocol using human spinal cord recording comparing safety, efficacy, and neurophysiological responses between patients being treated with evoked compound action potential–controlled closed–loop spinal cord stimulation or open–loop spinal cord stimulation (the evoke study). *Neuromodulation: Technology at the Neural Interface*, 22(3):317–326.

Liao, Q., Leibo, J., and Poggio, T. (2016). How important is weight symmetry in backpropagation? *Proceedings of the AAAI Conference on Artificial Intelligence*, 30(1).

Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., and Hinton, G. (2020). Backpropagation and the brain. *Nature Reviews Neuroscience*, 21(6):335–346.

Litt, B., D'Alessandro, A., Esteller, R., Echauz, J., and Vachtsevanos, G. (2003). Translating seizure detection, prediction and brain stimulation into implantable

devices for epilepsy. *First International IEEE EMBS Conference on Neural Engineering*, pages 485–488.

Liu, J., Zhu, Z., Zhou, Y., Wang, N., Dai, G., Liu, Q., Xiao, J., Xie, Y., Zhong, Z., Liu, H., et al. (2021). A reconfigurable biomedical AI processor with adaptive learning for versatile intelligent health monitoring. *IEEE International Solid-State Circuits Conference-(ISSCC)*, pages 62–64.

Liu, Z., Wang, Y., Vaidya, S., Ruehle, F., Halverson, J., Soljacic, M., Hou, T. Y., and Tegmark, M. (2025). KAN: Kolmogorov–Arnold Networks. *The Thirteenth International Conference on Learning Representations*.

Lo, M.-C. and Widge, A. S. (2017). Closed-loop neuromodulation systems: next-generation treatments for psychiatric illness. *International Review of Psychiatry*, 29(2):191–204.

Loeffler, A., Diaz-Alvarez, A., Zhu, R., Ganesh, N., Shine, J. M., Nakayama, T., and Kuncic, Z. (2023). Neuromorphic learning, working memory, and metaplasticity in nanowire networks. *Science Advances*, 9(16):eadg3289.

Long, Y., Li, J., Yang, F., Wang, J., and Wang, X. (2021). Wearable and implantable electroceuticals for therapeutic electrostimulations. *Advanced Science*, 8(8):2004023.

Lozano, A. M., Fosdick, L., Chakravarty, M. M., Leoutsakos, J.-M., Munro, C., Oh, E., Drake, K. E., Lyman, C. H., Rosenberg, P. B., Anderson, W. S., et al. (2016). A phase II study of fornix deep brain stimulation in mild Alzheimer's disease. *Journal of Alzheimer's Disease*, 54(2):777–787.

Lulic, D., Ahmadian, A., Baaj, A. A., Benbadis, S. R., and Vale, F. L. (2009). Vagus nerve stimulation. *Neurosurgical Focus FOC*, 27(3):E5.

Lundstrom, B. N., Brinkmann, B. H., and Worrell, G. A. (2021). Low frequency novel interictal EEG biomarker for localizing seizures and predicting outcomes. *Brain Communications*, 3(4):fcab231.

Ma, B. B., Fields, M. C., Knowlton, R. C., Chang, E. F., Szaflarski, J. P., Marcuse, L. V., and Rao, V. R. (2020). Responsive neurostimulation for regional neocortical epilepsy. *Epilepsia*, 61(1):96–106.

Ma, Y., Liu, C., Ma, M. S., Yang, Y., Truong, N. D., Kothur, K., Nikpour, A., and Kavehei, O. (2023). TSD: Transformers for seizure detection. *bioRxiv*, pages 2023–01.

Maeng, W.-Y., Tseng, W.-L., Li, S., Koo, J., and Hsueh, Y.-Y. (2022). Electroceuticals for peripheral nerve regeneration. *Biofabrication*, 14(4):042002.

Marković, D., Mizrahi, A., Querlioz, D., and Grollier, J. (2020). Physics for neuromorphic computing. *Nature Reviews Physics*, 2(9):499–510.

Martins, N. R., Angelica, A., Chakravarthy, K., Svidinenko, Y., Boehm, F. J., Opris, I., Lebedev, M. A., Swan, M., Garan, S. A., Rosenfeld, J. V., et al. (2019). Human brain/cloud interface. *Frontiers in Neuroscience*, page 112.

Martins, N. R., Erlhagen, W., and Freitas Jr, R. A. (2016). Human connectome mapping and monitoring using neuronanorobots. *Journal of Ethics and Emerging Technologies*, 26(1):1–25.

Mayberg, H. S., Lozano, A. M., Voon, V., McNeely, H. E., Seminowicz, D., Hamani, C., Schwalb, J. M., and Kennedy, S. H. (2005). Deep brain stimulation for treatment-resistant depression. *Neuron*, 45(5):651–660.

Mead, C. (1990). Neuromorphic electronic systems. *Proceedings of the IEEE*, 78(10):1629–1636.

Mead, C. (2020). How we created neuromorphic engineering. *Nature Electronics*, 3(7):434–435.

Mekhail, N., Levy, R. M., Deer, T. R., Kapural, L., Li, S., Amirdelfan, K., Hunter, C. W., Rosen, S. M., Costandi, S. J., Falowski, S. M., et al. (2020). Long-term safety and efficacy of closed-loop spinal cord stimulation to treat chronic back and leg pain (EVOKE): a double-blind, randomised, controlled trial. *The Lancet Neurology*, 19(2):123–134.

Mekhail, N. A., Levy, R. M., Deer, T. R., Kapural, L., Li, S., Amirdelfan, K., Pope, J. E., Hunter, C. W., Rosen, S. M., Costandi, S. J., et al. (2023). ECAP-controlled closed-loop versus open-loop scs for the treatment of chronic pain: 36-month results of the evoke blinded randomized clinical trial. *Regional Anesthesia & Pain Medicine*.

Mel, B. W. and Schiller, J. (2004). On the fight between excitation and inhibition: location is everything. *Science's STKE*, 2004(250):pe44–pe44.

Mermillod, M., Bugaiska, A., and Bonin, P. (2013). The stability-plasticity dilemma: Investigating the continuum from catastrophic forgetting to age-limited learning effects. *Frontiers in Psychology*, 4:504.

Merolla, P. A. and et al. (2014). A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science*, 345(6197):668–673.

Michmizos, K. P., Lindqvist, B., Wong, S., Hargreaves, E. L., Psychas, K., Mitsis, G. D., Danish, S. F., and Nikita, K. S. (2017). Computational neuromodulation: Future challenges for deep brain stimulation [life sciences]. *IEEE Signal Processing Magazine*, 34(2):114–119.

Minnekhanov, A. A., Emelyanov, A. V., Lapkin, D. A., Nikiruy, K. E., Shvetsov, B. S., Nesmelov, A. A., Rylkov, V. V., Demin, V. A., and Erokhin, V. V. (2019). Parylene based memristive devices with multilevel resistive switching for neuromorphic applications. *Scientific Reports*, 9(1):10800.

Mirowski, M., Mower, M. M., and Reid, P. R. (1980). The automatic implantable defibrillator. *American Heart Journal*, 100(6, Part 2):1089–1092.

Mivalt, F., Kremen, V., Sladky, V., Balzekas, I., Nejedly, P., Gregg, N. M., Lundstrom, B. N., Lepkova, K., Pridalova, T., Brinkmann, B. H., et al. (2022). Electrical brain stimulation and continuous behavioral state tracking in ambulatory humans. *Journal of Neural Engineering*, 19(1):016019.

Moradi, S., Qiao, N., Stefanini, F., and Indiveri, G. (2017). A scalable multicore architecture with heterogeneous memory structures for dynamic neuromorphic asynchronous processors (DYNAPs). *IEEE Transactions on Biomedical Circuits and Systems*, 12(1):106–122.

Morrell, M. (2006). Brain stimulation for epilepsy: can scheduled or responsive neurostimulation stop seizures? *Current Opinion in Neurology*, 19(2).

Mortazavi, A., Elliott, R.-J. S., Phan, T. N., Schreiber, J., Gaillard, W. D., and Oluigbo, C. O. (2021). Responsive neurostimulation for the treatment of medically refractory epilepsy in pediatric patients: strategies, outcomes, and technical considerations. *Journal of Neurosurgery: Pediatrics*, 28(1):54–61.

Mostafa, H., Müller, L. K., and Indiveri, G. (2015). An event-based architecture for solving constraint satisfaction problems. *Nature Communications*, 6(1):1–10.

Movassaghi, S., Abolhasan, M., Lipman, J., Smith, D., and Jamalipour, A. (2014). Wireless body area networks: A survey. *IEEE Communications Surveys & Tutorials*, 16(3):1658–1686.

Muneeb, A. and Kassiri, H. (2023). Energy-efficient spiking-cnn-based cross-patient seizure detection. In *IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 1–5.

Nejedly, P., Kremen, V., Sladky, V., Nasseri, M., Guragain, H., Klimes, P., Cimbalnik, J., Varatharajah, Y., Brinkmann, B. H., and Worrell, G. A. (2019). Deep-learning for seizure forecasting in canines with epilepsy. *Journal of Neural Engineering*, 16(3):036031.

Niedermeyer, E. and da Silva, F. L. (2005). *Electroencephalography: Basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins.

Pagkalos, M., Chavlis, S., and Poirazi, P. (2023). Introducing the dendrify framework for incorporating dendrites to spiking neural networks. *Nature Communications*, 14(1):131.

Park, J., Lee, J., and Jeon, D. (2019). A 65 nm 236.5 nJ/classification neuromorphic processor with 7.5% energy overhead on-chip learning using direct spike-only feedback. *IEEE International Solid-State Circuits Conference (ISSCC)*, pages 140–142.

Pedoeem, J., Bar Yosef, G., Abittan, S., and Keene, S. (2022). TABs: Transformer based seizure detection. In *Biomedical Sensing and Analysis: Signal Processing in Medicine and Biology*, pages 133–160. Springer.

Peh, W. Y., Thangavel, P., Yao, Y., Thomas, J., Tan, Y. L., and Dauwels, J. (2023). Multi-center assessment of CNN-transformer with belief matching loss for patient-independent seizure detection in scalp and intracranial EEG. *International Journal of Neural Systems*, 33(3):2350012.

Pehle, C., Billaudelle, S., Cramer, B., Kaiser, J., Schreiber, K., Stradmann, Y., Weis, J., Leibfried, A., Müller, E., and Schemmel, J. (2022). The BrainScaleS-2 accelerated neuromorphic system with hybrid plasticity. *Frontiers in Neuroscience*, 16.

Pei, J., Deng, L., Song, S., Zhao, M., Zhang, Y., Wu, S., Wang, G., Zou, Z., Wu, Z., He, W., et al. (2019). Towards artificial general intelligence with hybrid Tianjic chip architecture. *Nature*, 572(7767):106–111.

Penfield, W. and Jasper, H. (1954). *Epilepsy and the Functional Anatomy of the Human Brain*. Little, Brown and Company, Boston.

Peng, P., Song, Y., Yang, L., and Wei, H. (2022). Seizure prediction in EEG signals using STFT and domain adaptation. *Frontiers in Neuroscience*, 15:825434.

Peng, Y., Hooten, S., Vaerenbergh, T. V., Xiao, X., Fiorentino, M., and Beausoleil, R. G. (2024). Photonic KAN: a Kolmogorov-Arnold network inspired efficient photonic neuromorphic architecture. In *NeurIPS 2024 Workshop Machine Learning with new Compute Paradigms*.

Pepin, B. M. and Kotzev, M. T. (2021). Neuromodulation therapy monitoring and continuous therapy reprogramming. https://patents.google.com/patent/US20210052901A1. US Patent 2021/0052901 A1.

Perez-Nieves, N., Leung, V. C., Dragotti, P. L., and Goodman, D. F. (2021). Neural heterogeneity promotes robust learning. *Nature Communications*, 12(1):5791.

Pes, L., Luiken, R., Corradi, F., and Frenkel, C. (2024). Active dendrites enable efficient continual learning in time-to-first-spike neural networks. In *IEEE 6th International Conference on AI Circuits and Systems (AICAS)*, pages 41–45, Abu Dhabi, United Arab Emirates.

Piper, R. J., Richardson, R. M., Worrell, G., Carmichael, D. W., Baldeweg, T., Litt, B., Denison, T., and Tisdall, M. M. (2022). Towards network-guided neuromodulation for epilepsy. *Brain*, 145(10):3347–3362.

Pugh, J., Pycroft, L., Sandberg, A., Aziz, T., and Savulescu, J. (2018). Brainjacking in deep brain stimulation and autonomy. *Ethics and Information Technology*, 20:219–232.

Pycroft, L., Boccard, S. G., Owen, S. L., Stein, J. F., Fitzgerald, J. J., Green, A. L., and Aziz, T. Z. (2016). Brainjacking: Implant security issues in invasive neuromodulation. *World Neurosurgery*, 92:454–462.

Qian, C., Parramon, J., and Sanchez-Sinencio, E. (2011). A micropower low-noise neural recording front-end circuit for epileptic seizure detection. *IEEE Journal of Solid-State Circuits*, 46(6):1392–1405.

Qiu, X., Yan, F., and Liu, H. (2023). A difference attention ResNet-LSTM network for epileptic seizure detection using EEG signal. *Biomedical Signal Processing and Control*, 83:104652.

Rao, V. R. (2021). Chronic electroencephalography in epilepsy with a responsive neurostimulation device: current status and future prospects. *Expert Review of Medical Devices*, 18(11):1093–1105.

Rathore, H., Al-Ali, A., Mohamed, A., Du, X., and Guizani, M. (2017). DLRT: Deep learning approach for reliable diabetic treatment. In *IEEE Global Communications Conference (GLOBECOM)*, pages 1–6, Singapore.

Rathore, H., Wenzel, L., Al-Ali, A. K., Mohamed, A., Du, X., and Guizani, M. (2018). Multi-layer perceptron model on chip for secure diabetic treatment. *IEEE Access*, 6:44718–44730.

Rhew, H.-G., Jeong, J., Fredenburg, J. A., Dodani, S., Patil, P. G., and Flynn, M. P. (2014). A fully self-contained logarithmic closed-loop deep brain stimulation SoC with wireless telemetry and wireless power management. *IEEE Journal of Solid-State Circuits*, 49(10):2213–2227.

Ronchini, M., Rezaeiyan, Y., Zamani, M., Panuccio, G., and Moradi, F. (2023). NET-TEN: a silicon neuromorphic network for low-latency detection of seizures in local field potentials. *Journal of Neural Engineering*, 20(3):036002.

Ronchini, M., Zamani, M., Huynh, H. A., Rezaeiyan, Y., Panuccio, G., Farkhani, H., and Moradi, F. (2021). A CMOS-based neuromorphic device for seizure detection from LFP signals. *Journal of Physics D: Applied Physics*, 55(1):014001.

Roy, S., Kiral-Kornek, I., and Harrer, S. (2018). Deep learning enabled automatic abnormal EEG identification. In *40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2756–2759, Hawaii, United States.

Rueckert, U. (2020). Update on brain-inspired systems. In *NANO-CHIPS 2030: On-Chip AI for an Efficient Data-Driven World*, pages 387–403. Springer.

Saab, K., Dunnmon, J., Ré, C., Rubin, D., and Lee-Messer, C. (2020). Weak supervision as an efficient approach for automated seizure detection in electroencephalography. *npj Digital Medicine*, 3(1):1–12.

Sandamirskaya, Y. (2022). Rethinking computing hardware for robots. *Science Robotics*, 7(67):eabq3909.

Sangwan, V. K. and Hersam, M. C. (2020). Neuromorphic nanoelectronic materials. *Nature Nanotechnology*, 15(7):517–528.

Sawan, M., Salam, M. T., Le Lan, J., Kassab, A., Gélinas, S., Vannasing, P., Lesage, F., Lassonde, M., and Nguyen, D. K. (2013). Wireless recording systems: from noninvasive EEG-NIRS to invasive EEG devices. *IEEE Transactions on Biomedical Circuits and Systems*, 7(2):186–195.

Schachter, S. C. and Saper, C. B. (1998). Vagus nerve stimulation. *Epilepsia*, 39(7):677–686.

Schak, M. and Gepperth, A. (2019). A study on catastrophic forgetting in deep lstm networks. In *28th International Conference on Artificial Neural Networks (ICANN) on Deep Learning: , Munich, Germany, September 17–19, , Part II 28*, pages 714–728.

Scheid, B. H., Bernabei, J. M., Khambhati, A. N., Mouchtaris, S., Jeschke, J., Bassett, D. S., Becker, D., Davis, K. A., Lucas, T., Doyle, W., et al. (2022). Intracranial electroencephalographic biomarker predicts effective responsive neurostimulation for epilepsy prior to treatment. *Epilepsia*, 63(3):652–662.

Schemmel, J., Billaudelle, S., Dauer, P., and Weis, J. (2021). Accelerated analog neuromorphic computing. In *Analog Circuits for Machine Learning, Current/Voltage/Temperature Sensors, and High-speed Communication: Advances in Analog Circuit Design*, pages 83–102. Springer.

Schemmel, J., Brüderle, D., Grübl, A., Hock, M., Meier, K., and Millner, S. (2010). A wafer-scale neuromorphic hardware system for large-scale neural modeling. In *IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 1947–1950.

Scheuer, M. L., Wilson, S. B., Antony, A., Ghearing, G., Urban, A., and Bagić, A. I. (2021). Seizure Detection: Interreader Agreement and Detection Algorithm Assessments Using a Large Dataset. *Journal of Clinical Neurophysiology : Official Publication of the American Electroencephalographic Society*, 38(5):439–447.

Schliebs, S. and Kasabov, N. (2013). Evolving spiking neural network—a survey. *Evolving Systems*, 4:87–98.

Schuman, C. D., Kulkarni, S. R., Parsa, M., Mitchell, J. P., Date, P., and Kay, B. (2022). Opportunities for neuromorphic computing algorithms and applications. *Nature Computational Science*, 2(1):10–19.

Shah, V., Golmohammadi, M., Ziyabari, S., Von Weltin, E., Obeid, I., and Picone, J. (2017). Optimizing channel selection for seizure detection. In *IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, pages 1–5.

Sharifshazileh, M., Burelo, K., Sarnthein, J., and Indiveri, G. (2021). An electronic neuromorphic system for real-time detection of high frequency oscillations (HFO) in intracranial EEG. *Nature Communications*, 12(1):3095.

Shen, S., Wang, C., Huang, R., Zhong, Y., Guo, Q., Lu, Z., Zhang, J., and Leng, L. (2025). SpikingSSMs: Learning long sequences with sparse and parallel spiking state space models. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(19):20380–20388.

Shin, U., Ding, C., Zhu, B., Vyza, Y., Trouillet, A., Revol, E. C., Lacour, S. P., and Shoaran, M. (2022). NeuralTree: A 256-channel 0.227-μJ/class versatile neural activity classification and closed-loop neuromodulation SoC. *IEEE Journal of Solid-State Circuits*, 57(11):3243–3257.

Shoeb, A. H. (2009). *Application of machine learning to epileptic seizure onset detection and treatment*. PhD thesis, Massachusetts Institute of Technology.

Singhal, N. S., Numis, A. L., Lee, M. B., Chang, E. F., Sullivan, J. E., Auguste, K. I., and Rao, V. R. (2018). Responsive neurostimulation for treatment of pediatric drug-resistant epilepsy. *Epilepsy & Behavior Case Reports*, 10:21–24.

Skarpaas, T. L., Jarosiewicz, B., and Morrell, M. J. (2019). Brain-responsive neurostimulation for epilepsy (RNS® system). *Epilepsy Research*, 153:68–70.

Sladky, V., Nejedly, P., Mivalt, F., Brinkmann, B. H., Kim, I., St. Louis, E. K., Gregg, N. M., Lundstrom, B. N., Crowe, C. M., Attia, T. P., et al. (2022). Distributed brain co-processor for tracking spikes, seizures and behaviour during electrical brain stimulation. *Brain Communications*, 4(3):fcac115.

Snyder, D. and Leyde, K. W. (2008). Methods and Systems for Characterizing and Generating a Patient-Specific Seizure Advisory System. https://patents.google.com/patent/US20080208074A1. US Patent 2008/0208074 A1, Abandoned.

Stanslaski, S., Herron, J., Chouinard, T., Bourget, D., Isaacson, B., Kremen, V., Opri, E., Drew, W., Brinkmann, B. H., Gunduz, A., et al. (2018). A chronically implantable neural coprocessor for investigating the treatment of neurological disorders. *IEEE Transactions on Biomedical Circuits and Systems*, 12(6):1230–1245.

Stirling, R. E., Maturana, M. I., Karoly, P. J., Nurse, E. S., McCutcheon, K., Grayden, D. B., Ringo, S. G., Heasman, J. M., Hoare, R. J., Lai, A., D'Souza, W., Seneviratne, U., Seiderer, L., McLean, K. J., Bulluss, K. J., Murphy, M., Brinkmann, B. H., Richardson, M. P., Freestone, D. R., and Cook, M. J. (2021). Seizure forecasting using a novel sub-scalp ultra-long term EEG monitoring system. *Frontiers in Neurology*, 12:713794.

Stöckl, C. and Maass, W. (2021). Optimized spiking neurons can classify images with high accuracy through temporal coding with two spikes. *Nature Machine Intelligence*, 3(3):230–238.

Sun, F. T. and Morrell, M. J. (2014). Closed-loop neurostimulation: the clinical experience. *Neurotherapeutics*, 11(3):553–563.

Sun, Y., Jin, W., Si, X., Zhang, X., Cao, J., Wang, L., Yin, S., and Ming, D. (2022). Continuous seizure detection based on transformer and long-term iEEG. *IEEE Journal of Biomedical and Health Informatics*, 26(11):5418–5427.

Surianarayanan, C., Lawrence, J. J., Chelliah, P. R., Prakash, E., and Hewage, C. (2023). Convergence of Artificial Intelligence and Neuroscience towards the Diagnosis of Neurological Disorders—A Scoping Review. *Sensors*, 23(6):3062.

Tanaka, G., Yamane, T., Héroux, J. B., Nakane, R., Kanazawa, N., Takeda, S., Numata, H., Nakano, D., and Hirose, A. (2019). Recent advances in physical reservoir computing: A review. *Neural Networks*, 115:100–123.

Tang, S., Dunnmon, J., Saab, K. K., Zhang, X., Huang, Q., Dubost, F., Rubin, D., and Lee-Messer, C. (2022). Self-supervised graph neural networks for improved electroencephalographic seizure analysis. In *International Conference on Learning Representations*.

Tang, S., Dunnmon, J. A., Liangqiong, Q., Saab, K. K., Baykaner, T., Lee-Messer, C., and Rubin, D. L. (2023). Modeling multivariate biosignals with graph neural networks and structured state space models. In *Conference on Health, Inference, and Learning*, pages 50–71. PMLR.

Tang, T., Yan, L., Park, J. H., Wu, H., Zhang, L., Lee, H. Y. B., and Yoo, J. (2020). EEG dust: A BCC-based wireless concurrent recording/transmitting concentric electrode. In *IEEE International Solid-State Circuits Conference (ISSCC)*, pages 516–518, San Francisco, California, USA.

Thakur, C. S. and et al. (2018). Large-scale neuromorphic spiking array processors: A quest to mimic the brain. *Frontiers in Neuroscience*, 12.

Tian, F., Yang, J., Zhao, S., and Sawan, M. (2021). A new neuromorphic computing approach for epileptic seizure prediction. In *IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 1–5.

Truong, N. D. (2020). *Epileptic Seizure Detection and Forecasting Ecosystems*. Phd thesis, University of Sydney. Available at https://ses.library.usyd.edu.au/handle/2123/21932.

Truong, N. D., Nguyen, A. D., Kuhlmann, L., Bonyadi, M. R., Yang, J., Ippolito, S., and Kavehei, O. (2018). Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram. *Neural Networks*, 105:104–111.

Turner, J., Page, A., Mohsenin, T., and Oates, T. (2014). Deep belief networks used on high resolution multichannel electroencephalography data for seizure detection. In *AAAI Spring Symposium Series*.

Valderrama, M., Nikolopoulos, S., Adam, C., Navarro, V., and Le Van Quyen, M. (2010). Patient-specific seizure prediction using a multi-feature and multi-modal EEG-ECG classification. *Proc. Conference on Medical and Biological Engineering and Computing*, pages 77–80.

Venkataramani, S., Roy, K., and Raghunathan, A. (2016). Efficient embedded learning for iot devices. In *21st IEEE Asia and South Pacific Design Automation Conference (ASP-DAC)*, pages 308–311, Macao, China.

Vidyaratne, L., Glandon, A., Alam, M., and Iftekharuddin, K. M. (2016). Deep recurrent neural network for seizure detection. In *IEEE International Joint Conference on Neural Networks (IJCNN)*, pages 1202–1207, Vancouver, Canada.

Wang, F., Jiang, M., Qian, C., Yang, S., Li, C., Zhang, H., Wang, X., and Tang, X. (2017). Residual attention network for image classification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3156–3164, Honolulu, HI, USA.

Wang, Y., Luo, H., Chen, Y., Jiao, Z., Sun, Q., Dong, L., Chen, X., Wang, X., and Zhang, H. (2021). A closed-loop neuromodulation chipset with 2-level classification achieving 1.5-Vpp CM interference tolerance, 35-dB stimulation artifact rejection in 0.5 ms and 97.8%-sensitivity seizure detection. *IEEE Transactions on Biomedical Circuits and Systems*, 15(4):802–819.

Weisdorf, S., Duun-Henriksen, J., Kjeldsen, M. J., Poulsen, F. R., Gangstad, S. W., and Kjær, T. W. (2019). Ultra-long-term subcutaneous home monitoring of epilepsy—490 days of EEG from nine patients. *Epilepsia*, 60(11):2204–2214.

Wen, T. and Zhang, Z. (2018). Deep convolution neural network and autoencoders-based unsupervised feature learning of EEG signals. *IEEE Access*, 6:25399–25410.

Wenger, N., Moraud, E. M., Raspopovic, S., Bonizzato, M., DiGiovanna, J., Musienko, P., Morari, M., Micera, S., and Courtine, G. (2014). Closed-loop neuromodulation of spinal sensorimotor circuits controls refined locomotion after complete spinal cord injury. *Science Translational Medicine*, 6(255):255ra133.

White, T., Blok, E., and Calhoun, V. D. (2022). Data sharing and privacy issues in neuroimaging research: Opportunities, obstacles, challenges, and monsters under the bed. *Human Brain Mapping*, 43(1):278–291.

Whittington, J. C. and Bogacz, R. (2019). Theories of error back-propagation in the brain. *Trends in Cognitive Sciences*, 23(3):235–250.

Wingeier, B. M. and Tcheng, T. K. (2016). Spatiotemporal pattern recognition for neurological event detection and prediction in an implantable device. US Patent 9,314,182.

Wu, H., Adler, S., Azagury, D. E., Bohon, C., Safer, D. L., Barbosa, D. A. N., Bhati, M. T., et al. (2020a). Brain-Responsive Neurostimulation for Loss of Control Eating: Early Feasibility Study. *Neurosurgery*, 87(6):1277–1288.

Wu, Q., Sun, Y., Yan, H., and Wu, X. (2020b). Ecg signal classification with binarized convolutional neural network. *Computers in Biology and Medicine*, 121:103800.

Wu, Y., Deng, L., Li, G., Zhu, J., Xie, Y., and Shi, L. (2019). Direct training for spiking neural networks: Faster, larger, better. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01):1311–1318.

Xiao, C., Wang, S., Iasemidis, L., Wong, S., and Chaovalitwongse, W. A. (2017). An adaptive pattern learning framework to personalize online seizure prediction. *IEEE Transactions on Big Data*.

Yamazaki, K., Vo-Ho, V.-K., Bulsara, D., and Le, N. (2022). Spiking neural networks and their applications: A review. *Brain Sciences*, 12(7):863.

Yan, Z., Zhou, J., and Wong, W.-F. (2021). Energy efficient ECG classification with spiking neural network. *Biomedical Signal Processing and Control*, 63:102170.

Yang, Y., Eshraghian, J. K., Truong, N. D., Nikpour, A., and Kavehei, O. (2023). Neuromorphic deep spiking neural networks for seizure detection. *Neuromorphic Computing and Engineering*, 3(1):014010.

Yang, Y., Truong, N. D., Eshraghian, J. K., Nikpour, A., and Kavehei, O. (2022a). Weak self-supervised learning for seizure forecasting: a feasibility study. *Royal Society Open Science*, 9(8):220374.

Yang, Y., Truong, N. D., Maher, C., Nikpour, A., and Kavehei, O. (2022b). Continental generalization of a human-in-the-loop AI system for clinical seizure recognition. *Expert Systems with Applications*, 207:118083.

Yin, B., Corradi, F., and Bohté, S. M. (2020). Effective and efficient computation with multiple-timescale spiking recurrent neural networks. In *International Conference on Neuromorphic Systems 2020*, pages 1–8.

Yin, B., Corradi, F., and Bohté, S. M. (2021). Accurate and efficient time-domain classification with adaptive spiking recurrent neural networks. *Nature Machine Intelligence*, 3(10):905–913.

Yin, B., Corradi, F., and Bohté, S. M. (2023). Accurate online training of dynamical spiking neural networks through forward propagation through time. *Nature Machine Intelligence*, pages 1–10.

Yin, M., Borton, D. A., Aceros, J., Patterson, W. R., and Nurmikko, A. V. (2013). A 100-Channel Hermetically Sealed Implantable Device for Chronic Wireless Neurosensing Applications. *IEEE Transactions on Biomedical Circuits and Systems*, 7(2):115–128.

Yoo, J. and Shoaran, M. (2021). Neural interface systems with on-device computing: Machine learning and neuromorphic architectures. *Current Opinion in Biotechnology*, 72:95–101.

Young, A. R., Dean, M. E., Plank, J. S., and Rose, G. S. (2019). A review of spiking neuromorphic hardware communication systems. *IEEE Access*, 7:135606–135620.

Yu, K.-H., Beam, A. L., and Kohane, I. S. (2018). Artificial intelligence in healthcare. *Nature Biomedical Engineering*, 2(10):719–731.

Yuan, Y., Xun, G., Jia, K., and Zhang, A. (2018). A multi-view deep learning framework for EEG seizure detection. *IEEE Journal of Biomedical and Health Informatics*, 23(1):83–94.

Zhang, M., Zhang, L., Tsai, C.-W., and Yoo, J. (2022). A patient-specific closed-loop epilepsy management SoC with one-shot learning and online tuning. *IEEE Journal of Solid-State Circuits*, 57(4):1049–1060.

Zhang, W., Gao, B., Tang, J., Yao, P., Yu, S., Chang, M.-F., Yoo, H.-J., Qian, H., and Wu, H. (2020). Neuro-inspired computing chips. *Nature Electronics*, 3(7):371–382.

Zhao, S., Yang, J., Xu, Y., and Sawan, M. (2020). Binary single-dimensional convolutional neural network for seizure prediction. In *IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 1–5. IEEE.

Zheng, H., Zheng, Z., Hu, R., Xiao, B., Wu, Y., Yu, F., Liu, X., Li, G., and Deng, L. (2024). Temporal dendritic heterogeneity incorporated with spiking neural networks for learning multi-timescale dynamics. *Nature Communications*, 15(1):277.

Zuo, R., Wei, J., Li, X., Li, C., Zhao, C., Ren, Z., Liang, Y., Geng, X., Jiang, C., Yang, X., et al. (2019). Automated detection of high-frequency oscillations in epilepsy based on a convolutional neural network. *Frontiers in Computational Neuroscience*, 13:6.