

Neural Synergy in Compact Biomedical IoT Devices: Spintronic Pathways for MEG-EEG Integration and Portability

by

Mohamed Elshafei

B.Sc. in Biomedical Engineering, Mansoura University, Egypt 2019



A THESIS
SUBMITTED TO THE DEPARTMENT OF COMPUTER SCIENCE
AND THE FACULTY OF GRADUATE STUDIES
OF LAKEHEAD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
**MASTER OF SCIENCE (SPECIALIZATION IN ARTIFICIAL
INTELLIGENCE)**

© Copyright 2021 by Mohamed Elshafei
Lakehead University
Thunder Bay, ON, Canada

Supervisory Committee

Dr. Zubair Fadlullah

Supervisor

(Research Chair, Thunder Bay Regional Health Research Institute

External Adjunct Professor, Department of Computer Science, Lakehead University, Thunder Bay, Ontario, Canada.

Associate Professor, Department of Computer Science, University of Western Ontario, London, Ontario, Canada.)

Dr. Mostafa Fouda

Co-Supervisor

(Associate Professor, Department of Electrical and Computer Engineering, Idaho State University, USA.)

Dr. Garima Bajwa

Internal Examiner

(Assistant Professor, Department of Computer Science, Lakehead University, Thunder Bay, Ontario, Canada.)

Dr. Ahmed Elnakib

External Examiner

(Associate Professor, Department of Electrical and Computer Engineering, School of Engineering, Penn State Erie-The Behrend College, Erie, PA, USA.)

ABSTRACT

The human brain, a marvel of nature, consists of intricate neural networks that have fascinated and perplexed the scientific community for generations. As scientists and researchers globally endeavour to unravel the mysteries of bioelectrical activities that form the basis of our cognitive functions and experiences, our research emerges at the nexus of biology and cutting-edge technology. Specifically, we spotlight the remarkable capabilities of magnetoencephalography (MEG) and electroencephalography (EEG). These potent neuroimaging tools, celebrated for their unparalleled spatial and temporal precision, are synergistically combined in our study. We aim to map MEG innovatively signals onto their EEG replicas, employing avant-garde spintronic devices, with a particular emphasis on Magnetic Tunnel Junctions (MTJ).

Drawing inspiration from the wonders of nature, such as the awe-inspiring magnetoreception abilities exhibited by homing pigeons, our exploration is driven by a desire to harness and amplify similar bio-magnetic potentials latent within the human brain. Methodically segmented into distinct chapters, our research unfolds a series of groundbreaking contributions:

- **First:** Venturing into uncharted territories, we introduce the integration of artificial intelligence (AI) into MEG/EEG mapping. This novel approach propels our understanding of cerebral activities into exciting new directions.
- **Second:** Taking a deeper plunge, we explore the temporal features of the M/EEG signals as a further step toward optimizing the mapping solution. This insight paves the way for envisioning a future where bulky equipment gives way to compact, efficient neuroimaging devices.
- **Third:** In a display of our team’s innovative spirit, we simulate the MTJ sensor’s noise and utilize the AI approaches to denoise that noised dataset.
- **Finally:** Transitioning to the tangible realm of practical implementations, our focus gravitates towards the intricate facets of hardware design. We particularly emphasize the potential of quantizing the BiLSTM model to optimize and revolutionize the biomagnetic sensing ecosystem.

In essence, this research serves as a confluence of biology and state-of-the-art technological advancements. It meticulously lays a solid foundation for future explorations, from early detection and prediction of neurological disorders to pioneering strides in federated learning. Our fervent hope is that through our findings, we not only expand the horizons of current knowledge but also inspire a renewed perspective on how we understand and harness the myriad activities of the human brain.

ACKNOWLEDGEMENTS

In the name of Allah, the Most Gracious, the Most Merciful,

All praise is due to Allah for His divine guidance, unwavering support, and boundless blessings illuminating my academic journey. My heart swells with gratitude as I humbly acknowledge the immense favour of Allah, the Most Compassionate and Most Merciful.

Allah's grace has been the guiding light that has led me through the labyrinth of knowledge and understanding, and for this, I am profoundly grateful.

I extend my heartfelt appreciation to my esteemed academic mentor, Dr. Zubair Fadllulah, whose profound wisdom, patient guidance, and unyielding support have been instrumental in shaping the success of this thesis. His mentorship has enriched my academic understanding and nurtured my character as a scholar.

I am equally grateful to my co-supervisor, Dr. Mostafa Fouda, whose expertise and tireless dedication have further enriched this academic journey. His guidance and scholarly insights have contributed significantly to the depth and quality of this thesis.

I extend my deepest gratitude to my diligent examiners, whose valuable insights and critical evaluation have elevated the quality of this work. Your rigorous examination has been instrumental in refining the scholarly contribution of this thesis.

To my beloved family, your unwavering faith in my abilities and unwavering support has been the bedrock upon which I have built my academic pursuits. Your encouragement and sacrifices have fueled my determination.

To my dear wife, your boundless patience, unwavering encouragement, and steadfast belief in my abilities have been my most excellent source of strength and motivation. Your presence has been a constant reminder of the importance of love and support in pursuing knowledge.

To my friends, your camaraderie and encouragement have strengthened me throughout this academic odyssey. Your friendship and solidarity have lightened the burdens of this journey.

In conclusion, I stand on the shoulders of those who have supported and guided me throughout this academic endeavour. I offer my deepest thanks and appreciation to Allah, to each of you.

PUBLICATIONS

Parts of this thesis have been submitted for peer-review, published or accepted for publication:

- **Optimizing MEG-EEG Mapping in Resource-Constrained Non-Intrusive Bio-Magnetic Sensing Systems: A Data-Driven Approach** has been **accepted** in The 11th International Conference on Information and Communication Technology (ICoICT) 2023 (part of Chapter 3)
- **Data-Centric Enhancement of MEG-EEG Translations in low-Resource Bio-Magnetic Systems** is under revision (received review comments in first round), IEEE ACCESS.(part of Chapter 3)
- **Optimizing Temporal Pattern Detection Models for MEG-EEG Mapping in MTJ-based Portable Sensing Systems** is submitted, IEEE ACCESS. (part of Chapter 4)
- **Simulating and De-noising Magnetic Tunnel Junction Sensor Signals for MEG to EEG Mapping** is submitted, IEEE ACCESS. (part of Chapter 5)
- **Quantizing BiLSTM for Efficient Neuromorphic Hardware Deployment: An Exploration in Mapping MEG to EEG Signals** to be submitted (part of Chapter 6)

Contents

Supervisory Committee	ii
Abstract	iii
Acknowledgements	iv
Publications	v
Table of Contents	vi
List of Tables	ix
List of Figures	x
1 Introduction, Motivation, Contributions and Structure of Thesis	1
2 Preliminary ideas of bio-magnetic signals, spintronic sensors, AI-model-based sensory data acquisition challenges, and research gap identification	6
2.1 Introduction	7
2.2 Biology overview	8
2.2.1 Action Potential	8
2.2.2 Bio-magnetic Signals	10
2.2.2.1 Magnetoencephalography (MEG)	10
2.2.2.2 Magnetocardiography (MCG)	11
2.3 Spintronic sensors	12
2.4 AI-Based Models	15
2.4.1 Linear Regression (LR)	15
2.4.2 K-Nearest Neighbors (KNN) Regressor	15
2.4.3 Decision Tree (DT)	17
2.4.4 Deep Neural Network (DNN)	18
2.4.5 Convolutional Neural Network (CNN)	19
2.4.6 Gated Recurrent Unit (GRU)	20

2.4.7	Long Short-Term Memory (LSTM)	21
2.4.8	Bidirectional LSTM (BiLSTM)	22
2.5	Research Gap	23
3	Optimizing MEG-EEG Mapping in Resource-Constrained Non-Intrusive Bio-Magnetic Sensing Systems: A Data-Driven Approach	24
3.1	Introduction	25
3.2	Related work	26
3.3	Motivation and Problem Description	27
3.3.1	High Cost	28
3.3.2	Complexity	29
3.3.3	Weight and Mobility	29
3.4	Proposed MEG-EEG Mapping	30
3.4.1	Data Preprocessing and Channel Selection	30
3.4.2	Candidate Models for MEG-EEG Mapping	31
3.4.2.1	Linear Regression (LR)	31
3.4.2.2	K-Nearest Neighbors (KNN) Regressor	33
3.4.2.3	Decision Tree (DT)	33
3.4.2.4	Deep Neural Network (DNN)	33
3.4.2.5	Convolutional Neural Network (CNN)	34
3.4.3	Postprocessing	35
3.5	Performance Evaluation	36
3.6	Conclusion	38
4	Improved MEG-EEG Mapping Via Combined Spatial-Temporal Pattern Detection	40
4.1	Introduction	40
4.2	Related work	42
4.3	Considered System Design and Problem Description	44
4.4	Envisioned RNN Models Temporal Pattern Detector for MEG-EEG Mapping	45
4.4.1	Candidate RNN Models	45
4.4.1.1	GRU	45
4.4.1.2	LSTM	46
4.4.1.3	Bi-directional LSTM	47
4.4.2	Proposed Recurrent Layer based Convolution: Fusion of spatial and temporal detectors for MEG-EEG mapping	47
4.5	Data Preparation and Performance Evaluation	50
4.5.1	Data Preprocessing and Channel Selection	50
4.5.2	Evaluation metrics	52

4.6	Conclusion	53
5	De-noising Noisy Signals in MTJ Systems Via AI-assisted MEG-EEG Mapping	55
5.1	Introduction	55
5.2	Related work	57
5.3	MTJ Sensors and Noise Characteristics in MEG Signal Acquisition	58
5.4	Performance evaluation	59
	5.4.0.1 Noise simulation	59
	5.4.0.2 Findings and Analysis	61
5.5	Conclusion	63
6	Quantizing BiLSTM for Efficient Neuromorphic Hardware Deployment: An Exploration in Mapping MEG to EEG Signals	68
6.1	Introduction	69
6.2	Integration of Model Choices, Neuromorphic Principles, and Quantization Techniques	70
	6.2.1 From Linear Regression to Deep Learning: A Deliberate Choice	70
	6.2.2 Neuromorphic Computing	71
	6.2.3 Quantization of BiLSTM	72
	6.2.3.1 Reduced Memory Footprint	72
	6.2.3.2 Enhanced Processing Speed	73
	6.2.3.3 Energy Efficiency	73
6.3	Related work	74
6.4	Adaptive Precision Quantization for BiLSTM Networks	75
6.5	Performance Evaluation Through Hyperparameter Variation	77
6.6	Conclusion	78
7	Conclusion and Future Work on Integrating denoised and mapped signal for disease prediction and fed learning	83
7.1	Future Directions	84
	Bibliography	87

List of Tables

Table 3.1	Performance metrics and prediction times for different models.	37
Table 4.1	Performance metrics and prediction times for different models.	54
Table 5.1	Performance metrics and prediction times for different models.	63
Table 6.1	Performance Metrics RMSE and MAE for Different Models.	77

List of Figures

Figure 1.1	Organization of all the chapters of the thesis.	4
Figure 2.1	Action Potential Stages: 1) Stimulus initiates the threshold event. 2) Depolarization sees sodium ions entering the neuron. 3) During Repolarization, potassium ions exit, restoring internal negativity. 4) The Refractory Period ensures one-way signal travel by temporarily resisting new potentials.	9
Figure 2.2	Key Elements of Spintronic Sensor Operation: The upper part showcases the central spin valve structure, comprising layers of magnetic and non-magnetic materials, the lower section elucidates the Giant GMR effect	13
Figure 3.1	Traditional brain activities monitoring vs. our envisioned AI-based MEG monitoring system for brain activity monitoring, including a schematic of the MTJ sensor and the outline of a lightweight AI pipeline. Note that the core focus of our research in this chapter is on the MEG to EEG (MEG-EEG) mapping component of the pipeline.	28
Figure 3.2	Filtered signals of channel 41 and time range 4-5 seconds: Actual vs. predicted signals by different models.	32
(a)	Actual vs predicted signal performance with the LR model for channel 41.	32
(b)	Actual vs predicted signal performance with the KNN model for channel 41	32
(c)	Actual vs predicted signal performance with the DNN model for channel 41.	32
(d)	Actual vs predicted signal performance with the CNN model for channel 41	32
(e)	Actual vs predicted signal performance with the DT model for channel 41	32
Figure 3.3	Fix caption.	35
(a)	DNN model architecture	35

	(b) Training vs validation loss curves for the proposed DNN model	35
Figure 3.4	comparison of Training and Validation Loss Curves and Model Archi- tectures for the Proposed DNN and CNN Models	35
	(a) CNN model architecture	35
	(b) Training vs validation loss curves for the proposed CNN model	35
Figure 3.5	Performance comparison of MAE and RMSE for different models. . .	36
Figure 3.6	The first 20 EEG channels filtered based on the LR model's predictions	37
Figure 3.7	Actual vs. predicted signals for randomly selected EEG channels. . .	38
Figure 4.1	Comparison of machine learning and deep learning models for MEG to EEG mapping. (a) Spatial pattern models: Linear regression, KNN regressor, Decision tree regressor, DNN, and CNN, which focus on learning the spatial patterns of the signals. (b) Temporal pattern models: RNN-based models, including GRU, LSTM, and BiLSTM, are specifically designed to capture the temporal patterns in the sig- nals. (c) Integration of spatial and temporal patterns: RCNN model seamlessly couples the temporal and spatial pattern detectors to en- hance the mapping accuracy and highlight the importance of consid- ering both dependencies in the MEG-EEG mapping	44
Figure 4.2	Performance comparison of MAE and RMSE for different models. . .	50
Figure 4.3	Comparison of the model architectures. The four subfigures, (a), (b), (c), and (d), represent the structures of the proposed GRU, LSTM, BiLSTM, and RCNN models, respectively.	51
Figure 4.4	The first 20 predicted EEG channels based on the RCNN model . . .	51
Figure 5.1	Bar plot showing the average power levels (in dB) across different frequency bands of interest, namely Alpha, Gamma, Beta, Theta, and Delta in our clean MEG signal	60
Figure 5.2	Performance comparison of RMSE and MAE for different models. . .	63
Figure 5.3	Comparison between the Power Spectral Density (PSD) of Prediction Models and the PSD of Clean EEG Signals. The PSD is calculated as the average of all channels.	64
	(a) Actual vs. predicted signal performance with the LR model for channel 40.	65
	(b) Clean EEG vs. predicted signal performance with the KNN model for channel 40.	65
	(c) Clean EEG vs. predicted signal performance with the DNN model for channel 40.	65

(d)	Clean EEG vs. predicted signal performance with the CNN model for channel 40.	65
(e)	Clean EEG vs. predicted signal performance with the LSTM model for channel 40.	65
(f)	Clean EEG vs. predicted signal performance with the GRU model for channel 40.	65
Figure 5.4	Filtered signals of channel 40 within the time range of 20-500 seconds: a comparison between clean and denoised predicted signals using various models.	66
(g)	Clean EEG vs. predicted signal performance with the BiLSTM model for channel 40.	66
(h)	Clean EEG vs. predicted signal performance with the RNN&CNN model for channel 40.	66
(i)	Clean EEG vs. predicted signal performance with the RCNN model for channel 40.	66

Figure 5.5	Visual comparison of the predicted signals generated by various proposed AI-based denoising models for Channel 40, showcasing their cumulative performance from the best at the bottom to the worst at the top. The clean EEG signal is also included for reference, with the time range set from sample 20 to 700. The ordered arrangement of the models offers a clear depiction of their relative accuracy in capturing the underlying EEG dynamics.	67
------------	---	----

Figure 6.1	Analysis of RMSE and MAE based on the variation of different hyperparameters. The line plots illustrate the trend of RMSE and MAE values as a function of each hyperparameter. The comparative analysis provides insights into the sensitivity of the model's performance to changes in the hyperparameter configurations	79
------------	---	----

(a)	Model 1: Actual vs quantized signal performance with the $num_sets = 4, set_size = 8, input_bit_width = 32, weight_bit_width = 32, output_bit_width = 32$	81
(b)	Model 2: Actual vs quantized signal performance with the $num_sets = 4, set_size = 8, input_bit_width = 32, weight_bit_width = 32, output_bit_width = 16$	81
(c)	Model 3: Actual vs quantized signal performance with the $num_sets = 4, set_size = 6, input_bit_width = 16, weight_bit_width = 32, output_bit_width = 32$	81

(d)	Model 4: Actual vs quantized signal performance with the $num_sets = 4$, $set_size = 6$, $input_bit_width = 32$, $weight_bit_width = 32$, $output_bit_width = 16$	81
(e)	Model 5: Actual vs quantized signal performance with the $num_sets = 6$, $set_size = 6$, $input_bit_width = 16$, $weight_bit_width = 16$, $output_bit_width = 32$	81
(f)	Model 6: Actual vs quantized signal performance with the $num_sets = 6$, $set_size = 8$, $input_bit_width = 16$, $weight_bit_width = 32$, $output_bit_width = 32$	81
Figure 6.2	Performance comparison of actual vs Filtered normalized quantized signals using different configuration settings.	82
(g)	Model 7: Actual vs quantized signal performance with the $num_sets = 6$, $set_size = 6$, $input_bit_width = 32$, $weight_bit_width = 32$, $output_bit_width = 16$	82
(h)	Model 8: Actual vs quantized signal performance with the $num_sets = 6$, $set_size = 8$, $input_bit_width = 16$, $weight_bit_width = 32$, $output_bit_width = 16$	82

Chapter 1

Introduction, Motivation, Contributions and Structure of Thesis

The human brain, a miraculous conglomerate of billions of neurons, intrigues and challenges the scientific community. It is not just the sheer number of neurons but their intricate interconnections, leading to the manifestation of thoughts, emotions, and conscious experiences, that bewilder our understanding. The nuanced bioelectrical activities underlying these phenomena are both the bedrock of our cognitive essence and the subject of intense research scrutiny.

Amidst the spectrum of neuroimaging tools, magnetoencephalography (MEG) and electroencephalography (EEG) are pillars that stand tall, each offering unique vantage points to peer into the brain's electric symphony. MEG's prowess in capturing the magnetic fields produced by neuronal activity provides unparalleled spatial resolution. On the other hand, EEG, recording the electrical potentials as they dance on the scalp, offers exceptional temporal clarity. The harmonization of these techniques through mapping MEG signals to EEG paves the way for a richer, more comprehensive interpretation of neural activities.

This research endeavours to bridge these two realms at the intersection of biology and technology. The vision is not merely to comprehend but to transcend to harness the potential of Artificial Intelligence (AI) and merge it seamlessly with our understanding of brain activities. This research's heartbeat is the ambitious aim to map MEG signals to EEG counterparts, leveraging the might of spintronic devices, with Magnetic Tunnel Junctions (MTJ) standing as emblematic pillars. Imagine a world where bulky, cumbersome MEG machines are relics of the past, replaced by sleek, smart bands adorning wrists. These bands, driven by the innovations presented in this thesis, promise to continuously monitor and assess brain activities.

Our expedition through this research is not merely academic. It is a journey filled with purpose and promise. The vision is noble: empowering individuals with tools that make life easier and profoundly more meaningful. Every challenge faced, and every solution devised, has been a step toward laying a foundation - a robust platform upon which future innovations in brain monitoring can thrive. The following paragraphs delve into the heart of our inspiration and motivation behind this endeavour.

Nature has always been a reservoir of inspiration for humans. From birds' flight to jellyfish luminescence, it is a testament to the myriad ways life has evolved to navigate its surroundings and communicate. A particularly intriguing adaptation is found in the seemingly ordinary homing pigeon, which holds clues that resonate profoundly with the essence of our research.

Researchers have unearthed fascinating insights into the navigational prowess of pigeons. Embedded in their beaks are tiny fragments of magnetite, a naturally magnetized rock. This magnetite acts as nature's GPS, providing the pigeon with positional information concerning Earth's magnetic poles. Beyond the beak, the pigeon's eyes harbour specialized cells believed to allow them to perceive magnetic fields visually. These dual mechanisms empower pigeons to traverse vast, landmark-devoid terrains, such as oceans, with an uncanny sense of direction [1].

Historically, humans have harnessed the pigeon's impeccable navigational skills since the times of Babylon, utilizing them to deliver messages across challenging terrains, the mystique surrounding how pigeons consistently found their way home baffled scientists for ages. The findings in Science [2] unveil this enigma, emphasizing that pigeons utilize magnetoreception. This inherent ability to detect Earth's magnetic fields allows them to accurately pinpoint their location, as every geographical point on Earth has a distinct magnetic intensity and inclination.

Intriguingly, humans, too, have traces of magnetite nestled within the bones of our noses. While it is debatable whether we leverage Earth's magnetic fields for navigation, it is a tantalizing hint that perhaps we are not so different from our avian counterparts.

Drawing inspiration from these natural marvels, our research delves deep into the domain of bio-magnetic signals. Suppose pigeons, with their innate magnetoreception abilities, can navigate vast distances. How can we, with the boon of modern technology and AI, not harness the magnetic signals within our brains for greater purposes? If the biological systems of pigeons can denoise such noisy signals to identify their way, we can do similar. This research's objective to map MEG signals to EEG, leveraging spintronic devices like MTJ, feels like a poetic ode to nature's magnetic navigators. By synthesizing the wonders of biology and the potential of technology, we aspire to bring solutions that could reshape how we perceive and harness our brain's activities, akin to how pigeons tap into the Earth's magnetic symphony to find their way home.

Building upon this deep-seated inspiration and aspiration, we have made specific, tangible contributions to biomagnetic signal processing and mapping. Let us delve into our research journey's key advancements and innovations birthed from our research journey.

- **Untapped Potential Exploration:** We initiated an investigation into the area of MEG/EEG mapping using AI, a field that has largely remained untouched. This provides a new avenue for future research, promising richer insights into brain activities.
- **Harnessing Spintronics:** Capitalizing on the capabilities of spintronic MTJ sensors, we demonstrated their utility in acquiring MEG data, paving the way for more compact and efficient data acquisition tools.
- **Robust AI-based Mapping:** We have successfully developed robust and precise AI models explicitly tailored for MEG to EEG mapping, setting a new benchmark in signal translation accuracy.
- **Pattern-Driven Mapping Enhancement:** We unveiled a groundbreaking approach that synchronizes spatial and temporal patterns, leading to a remarkable improvement in MEG-EEG mapping accuracy.
- **Dataset Crafting:** We introduced a simulated dataset to recognize the absence of MTJ-based MEG signal datasets. This dataset not only mimics the inherent noise found in natural environments but also is a valuable resource for researchers and practitioners.
- **Advanced De-noising Techniques:** Our research introduces rigorous models to de-noise the simulated dataset effectively, ensuring that signal mapping maintains the highest fidelity possible.
- **Edge of Hardware Implementation:** We have extended our research to touch the frontiers of hardware implementation. A key highlight is the quantization of the BiLSTM model, promising a more efficient and portable bio-magnetic sensing ecosystem.
- **Identification of Challenges:** An exhaustive study elucidated the nuanced challenges associated with spintronic sensory and bio-magnetic signal acquisition, providing a more precise roadmap for future endeavours.
- **Future Outlook:** We have set the stage for the next era of bio-magnetic research, where refined signals can be pivotal in predicting neurological diseases and advancing federated learning.

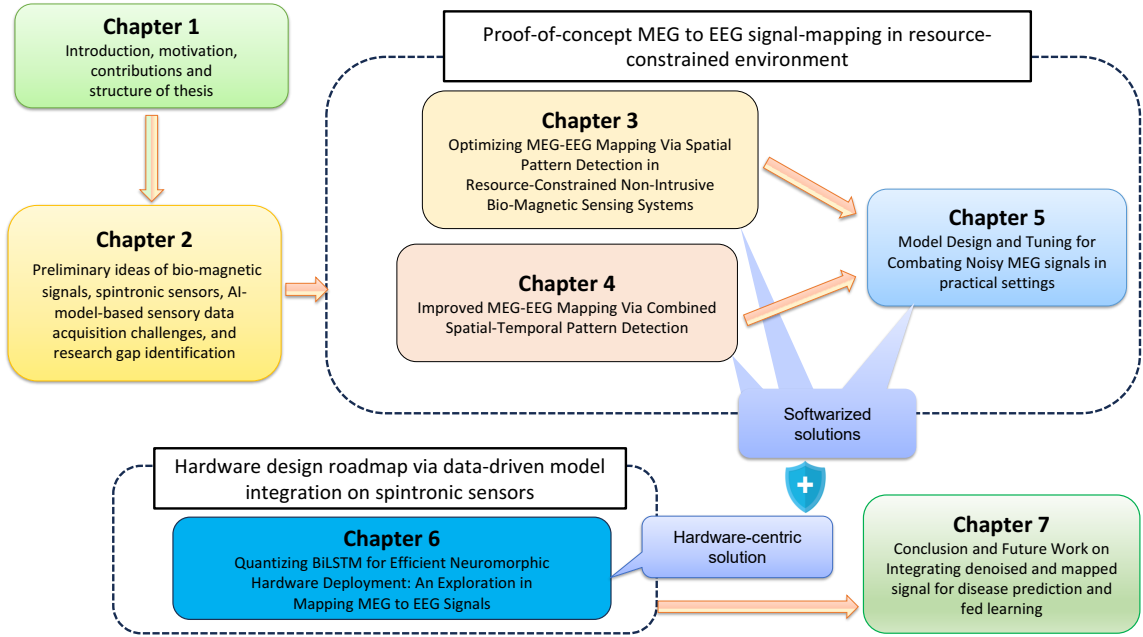


Figure 1.1: Organization of all the chapters of the thesis.

In Figure 1.1, we offer readers a panoramic view of our thesis structure, meticulously detailing its evolution from one chapter to the next. This illustrative roadmap demystifies the complexities inherent in our research journey and acts as a compass, guiding readers toward a deeper understanding. By presenting this organized breakdown, we aim to weave a cohesive narrative that underscores the synergies and interconnections among various research facets.

Chapter 2 sets the stage, presenting the cornerstone concepts upon which subsequent chapters are built. Readers are acquainted with the foundational knowledge of bio-magnetic signals in this chapter. We navigate the multifaceted landscape of spintronic sensory data acquisition, demystifying the associated challenges. Additionally, in our quest for comprehensive research, we shed light on the gaps that have yet to be bridged within this domain.

Moving forward, Chapters 3 through 5 are the heart of our exploration. These chapters are unified by a singular vision: to establish a "proof of concept for MEG to EEG signal mapping in resource-constrained environments." Nevertheless, they each offer distinct insights, all seamlessly integrated within the broader framework of software-centric solutions. Chapter 3 commences this deep dive, emphasizing a data-centric paradigm and introducing strategies tailored for efficient MEG-EEG mapping, especially in environments where resources are limited. Following suit, Chapter 4 presents a more intricate perspective, introducing a combined spatial-temporal pattern detection methodology. This approach, nuanced in its design, promises enhanced accuracy in MEG-EEG signal correlation. Chapter 5 encapsulates our endeavours in signal quality enhancement. Here, we chronicle our

strategies for simulating and refining Magnetic Tunnel Junction sensor signals, fortifying the entire mapping process of MEG to EEG.

With Chapter 6, we pivot our focus, anchoring our discussions on the tangible implications of our research in hardware. Within these pages, we delve into the futuristic potential of neuromorphic computing. A spotlight is cast on the quantization of BiLSTM, elucidating its capacity to usher in transformative shifts within the bio-magnetic sensing hardware paradigm.

Concluding our research quest, Chapter 7 provides a holistic retrospective, amalgamating insights from preceding chapters and contemplating prospective trajectories. As we reflect upon the research tapestry we have woven, we also cast our gaze forward, speculating on the broader applications of our findings. This includes the potent capability of utilizing the refined signals for diagnostic purposes, such as predicting neurological anomalies and venturing into the dynamic world of federated learning. Through this chapter, we aspire to leave readers with both a sense of closure and excitement for the untapped potential that lies ahead.

Chapter 2

Preliminary ideas of bio-magnetic signals, spintronic sensors, AI-model-based sensory data acquisition challenges, and research gap identification

2.1	Introduction	7
2.2	Biology overview	8
2.2.1	Action Potential	8
2.2.2	Bio-magnetic Signals	10
2.2.2.1	Magnetoencephalography (MEG)	10
2.2.2.2	Magnetocardiography (MCG)	11
2.3	Spintronic sensors	12
2.4	AI-Based Models	15
2.4.1	Linear Regression (LR)	15
2.4.2	K-Nearest Neighbors (KNN) Regressor	15
2.4.3	Decision Tree (DT)	17
2.4.4	Deep Neural Network (DNN)	18
2.4.5	Convolutional Neural Network (CNN)	19
2.4.6	Gated Recurrent Unit (GRU)	20
2.4.7	Long Short-Term Memory (LSTM)	21
2.4.8	Bidirectional LSTM (BiLSTM)	22

2.5	Research Gap	23
-----	------------------------	----

2.1 Introduction

Bio-magnetic signals, unique characteristics and significance in physiological and neurological processes have emerged as a valuable resource for understanding complex biological phenomena. In this chapter, we delve into the world of bio-magnetic signals, spintronic sensors, AI-based models, and the identification of research gaps. By exploring these topics, we aim to pave the way for optimized biomagnetic signal detection, advanced sensing technologies, and innovative data acquisition and analysis approaches. We pull the strings together to achieve a higher purpose that has the potential to make the individual's life easier.

First, we provide an overview of bio-magnetic signals and their relevance in unravelling physiological and neurological processes. These signals, including Magnetoencephalography (MEG), Magnetocardiography (MCG), and others, offer unique insights into brain activity, cardiac function, and other biological phenomena. By understanding the characteristics and applications of these biomagnetic signals, we can explore their potential in diverse fields such as neuroscience, cardiology, and diagnostics.

Next, we introduce spintronic sensors and their importance in bio-magnetic signal detection. These sensors provide an innovative approach to sensing bio-magnetic fields by leveraging the principles of spintronics. We discuss the fundamental principles of spintronics and highlight how spintronic sensors can revolutionize the detection and analysis of biomagnetic signals. Furthermore, we explore their advantages, including high sensitivity and low power consumption, and highlight recent advancements in this rapidly evolving field. Additionally, we address the challenges of acquiring sensory data from spintronic sensors and the need for tailored approaches to ensure reliable and accurate signal acquisition.

Integrating artificial intelligence (AI) models in bio-magnetic signal analysis is another exciting research avenue. This chapter examines the potential and limitations of AI-based models, including deep learning and machine learning used in this work, in extracting meaningful information from bio-magnetic signals. We explore the capabilities of these models in enhancing signal analysis, identifying patterns, and improving regression accuracy. However, we also acknowledge the challenges in developing robust AI models for bio-magnetic signal analysis, such as data availability, model interpretability, and generalizability.

Lastly, we identify research gaps in bio-magnetic signal detection, spintronic sensors, and AI-based data acquisition. We uncover limitations and shortcomings that call for further research and innovation by critically evaluating the current approaches, methodologies, and technologies. Addressing these research gaps becomes essential to advance the field,

improve signal detection techniques, enhance sensor technologies, and refine AI-based data acquisition methodologies.

Through the exploration of bio-magnetic signals, spintronic sensors, AI-based models, and the identification of research gaps, this chapter sets the stage for subsequent discussions. The following sections delve deeper into each topic, examining their intricacies, potentials, and challenges. By doing so, we aim to contribute to developing optimized techniques for biomagnetic signal detection, foster advancements in sensing technologies, and pave the way for innovative approaches to data acquisition and analysis.

2.2 Biology overview

2.2.1 Action Potential

To comprehend the generation of bio-magnetic signals, it is crucial to understand the concept of the action potential, which serves as the fundamental electrical signal within the body. The action potential is the brief electrical signals that play a fundamental role in neuronal communication and are essential for the functioning of the nervous system [3].

Action potentials are generated due to a rapid change in the membrane potential of excitable cells, such as neurons. This change in membrane potential is triggered by the motion of ions across the cell membrane, especially sodium and potassium ions. The process begins with a stimulus that depolarizes the cell membrane, opening voltage-gated sodium channels and the influx of sodium ions. This influx of positive charge causes a rapid rise in the membrane potential, resulting in depolarization.

An action potential is triggered once the membrane potential reaches a certain threshold. This triggers a sequence of events, as shown in Figure 2.1 [4], including the rapid opening of voltage-gated sodium channels, permitting a massive influx of sodium ions. The influx of positive charge causes a rapid spike in the membrane potential, generating the rising phase of the action potential. Following the depolarization phase, the membrane potential rapidly repolarizes. This happens due to the opening of voltage-gated potassium channels and the efflux of potassium ions. The repolarization phase returns the membrane potential to its resting state, briefly hyperpolarizing the cell before reaching equilibrium. The propagation of action potentials along neurons allows for transmitting electrical signals over long distances, enabling communication between different nervous system regions. These electrical impulses generate corresponding magnetic fields, creating the basis for the bio-magnetic effect [5].

The magnetic field associated with a bioelectrical signal is perpendicular to the electrical field, resulting in a unique magnetic field pattern for each type of bioelectrical signal [6]. Thus, the bio-magnetic signals are an intrinsic consequence of the electrical activity occur-

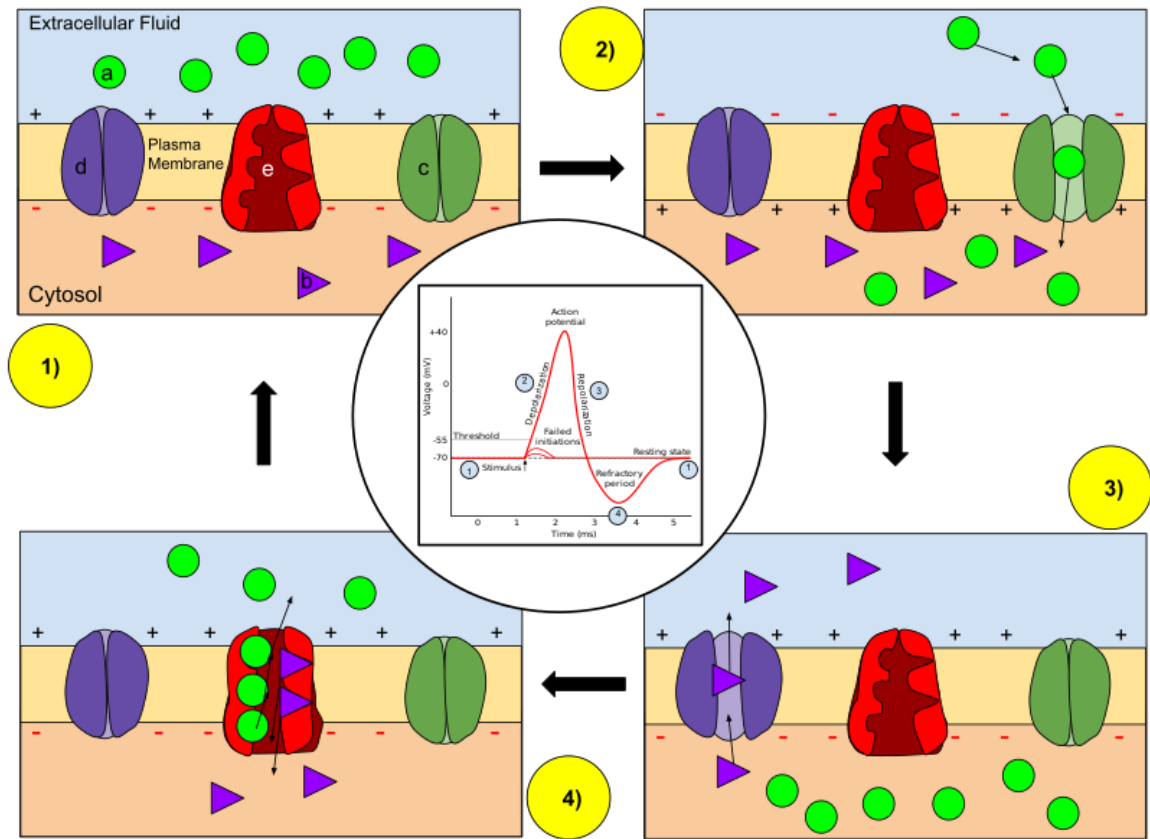


Figure 2.1: Action Potential Stages: 1) Stimulus initiates the threshold event. 2) Depolarization sees sodium ions entering the neuron. 3) During Repolarization, potassium ions exit, restoring internal negativity. 4) The Refractory Period ensures one-way signal travel by temporarily resisting new potentials.

ring within the body. The coherent relationship between the electrical and magnetic fields allows bio-magnetic signals to serve as valuable indicators of underlying physiological and neurological processes.

In neuroscience, for example, Magnetoencephalography provides insights into the spatiotemporal dynamics of brain activity, enabling the mapping of neuronal responses and studying cognitive processes. Similarly, in cardiology, Magnetocardiography offers a non-invasive means of assessing cardiac function and diagnosing various heart diseases [7]. By unravelling the intricate connection between electrical and magnetic fields, we gain a deeper understanding of the bio-magnetic signals and their potential applications in diverse fields. The characterization and analysis of these signals hold promise for advancing our understanding of physiological processes, aiding in disease diagnosis, and guiding therapeutic interventions.

2.2.2 Bio-magnetic Signals

The human body generates various magnetic fields and signals crucial in diagnosing, understanding, and analyzing internal body activities and interactions. These signals provide valuable insights into specific physiological processes and contribute to our understanding of the human body and its functions. Several techniques are employed to measure and analyze these biomagnetic signals, including Magnetoenterography (MEnG), Magnetooculography (MOG), and Magnetomyography (MMG).

Magnetoenterography (MEnG) is a technique used to measure the magnetic fields affected by the electrical activity of the gastrointestinal tract. By detecting and analyzing these magnetic fields, MEnG provides insights into the motility and function of the digestive system. It finds applications in studying gastrointestinal disorders and deepening our understanding of the physiological processes involved in digestion [8].

Similarly, Magnetooculography (MOG) measures the magnetic fields generated by eye movements. By capturing these magnetic fields, MOG allows for studying eye movement patterns and provides valuable insights into visual perception, oculomotor control, and vestibular function. MOG is employed in disciplines such as ophthalmology, neurology, and research on eye movement disorders [9].

Magnetomyography (MMG) is another technique that measures the magnetic fields generated by the electrical activity of skeletal muscles. By detecting and analyzing these magnetic fields, MMG provides insights into muscle contraction, neuromuscular diseases, and motor control. It can be effectively used to study muscle activity patterns and assess muscle function [10].

In addition to MEnG, MOG, and MMG, other biomagnetic signals offer a wide range of applications across various research and clinical practice areas. Each signal provides unique and valuable insights into specific physiological processes, contributing to our understanding of the human body and its intricate functions. While MEG and MCG are necessary biomagnetic signals, their significance in this research warrants a separate discussion. While MEG and MCG are necessary biomagnetic signals, their significance in this research warrants a separate discussion.

2.2.2.1 Magnetoencephalography (MEG)

Magnetoencephalography is a non-invasive neuroimaging technique that measures the magnetic fields generated by neuronal activity in the brain. MEG offers high spatial resolution, making it a valuable tool for studying the dynamics of brain activity and cognitive processes [11]. These magnetic fields are fragile and require specialized equipment and susceptible sensors to detect them accurately.

MEG exhibits several key characteristics that make it a powerful neuroimaging tech-

nique. Firstly, it offers high sensitivity, allowing it to detect even minute magnetic fields generated by neuronal activity. This high sensitivity enables researchers to gain insights into the brain's functioning with exceptional precision [12]. Additionally, MEG provides outstanding spatial resolution, enabling researchers to localize brain activity with millimetre-level precision. This precise localization helps identify specific brain regions involved in various cognitive processes, such as language processing, memory, attention, and sensory perception.

MEG also excels in its temporal resolution, capturing the rapid changes in brain activity within milliseconds. This temporal resolution is crucial for studying dynamic brain operations and understanding the precise timing of neuronal activity. MEG also allows for estimating the sources of neural activity in the brain [13]. By combining the measured magnetic fields with mathematical modelling techniques, researchers can reconstruct the underlying neuronal sources, providing insights into the brain's functional organization.

In clinical settings, MEG has found applications such as pre-surgical mapping of brain function in patients with epilepsy or brain tumours. It assists in identifying critical brain functional areas and mapping the epileptic focus, aiding in surgical planning and minimizing the risk of postoperative deficits.

MEG acquisition systems comprise a helmet-shaped array of susceptible magnetic sensors called Superconducting Quantum Interference Devices (SQUIDS). These systems are carefully shielded from outer magnetic interference to guarantee accurate measurements of the weak magnetic fields generated by the brain. The acquired signals are processed and interpreted using advanced computational techniques to extract influential information about brain activity. The data obtained from MEG provides researchers with valuable insights into the functioning of the brain and its role in cognitive processes [14].

2.2.2.2 Magnetocardiography (MCG)

Magnetocardiography is a non-intrusive approach that measures the magnetic fields induced by the heart's electrical activity. By detecting and analyzing these magnetic fields, MCG provides valuable insights into cardiac function and assists in diagnosing various heart conditions.

MCG exhibits several key characteristics that make it a powerful tool for investigating the heart's electrical activity. Like MEG, MCG offers high sensitivity, allowing it to detect the weak magnetic fields produced by the electrical currents generated during the cardiac cycle. This high sensitivity enables precise measurement of cardiac magnetic fields and provides valuable information about the electrical dynamics of the heart [15]. MCG also provides excellent temporal resolution, allowing for the capture of rapid changes in the magnetic fields associated with the heart's electrical activity. This temporal resolution enables researchers to analyze the various phases of the cardiac cycle, such as depolarization

and repolarization, with high precision.

MCG offers several benefits and has a range of applications in cardiology. Firstly, MCG provides insights into cardiac function and helps assess the heart's health. It can assist in diagnosing various heart conditions, including arrhythmias, ischemic heart disease, and congenital heart defects [16]. Furthermore, MCG aids in localizing cardiac abnormalities by identifying regions of abnormal electrical activity within the heart. This information can be crucial for planning interventions, such as ablation procedures, or guiding the placement of implantable devices, such as pacemakers. MCG also offers advantages in specific patient populations, such as infants, where traditional electrocardiography (ECG) may be challenging due to their small size or lack of cooperation. MCG provides a non-invasive alternative to assess cardiac function in these cases [17].

MCG acquisition systems consist of a specialized array of susceptible magnetic sensors similar to those used in MEG systems. These sensors are strategically placed around the chest area to capture the magnetic fields generated by the heart's electrical activity. The acquired MCG signals are processed and analyzed using sophisticated algorithms to extract information about the electrical dynamics of the heart. The data obtained from MCG provides cardiologists and researchers with valuable insights into cardiac function and aids in diagnosing and managing various heart conditions [18].

2.3 Spintronic sensors

Spintronic sensors have emerged as a groundbreaking technology in biomagnetic signal detection, offering tremendous potential for revolutionizing how we detect and analyze biomagnetic signals. By harnessing the principles of spintronics, these sensors provide unique advantages such as high sensitivity and low power consumption, making them highly attractive for bio-magnetic signal detection and analysis applications.

Spintronics, an abbreviation for spin electronics, explores the intrinsic spin property of electrons and their charge to develop new electronic devices. Unlike traditional electronics that solely rely on electron charge, spintronics utilizes the spin of electrons to encode and process information [19]. This fundamental principle enables advanced sensors to detect and measure magnetic fields accurately.

The fundamental component of spintronic sensors is the spin valve structure, which typically consists of thin layers of magnetic and non-magnetic materials. The magnetic layers are composed of materials with different magnetic properties, such as ferromagnetic or antiferromagnetic materials, the blue layers in Figure 2.2. The non-magnetic layers acts as a spacer between the magnetic layers, the orange layer. When a voltage is applied to the spin valve structure, an electric current of electrons flows through the structure. These electrons possess both charge and spin. The spin orientation of electrons can be manipulated

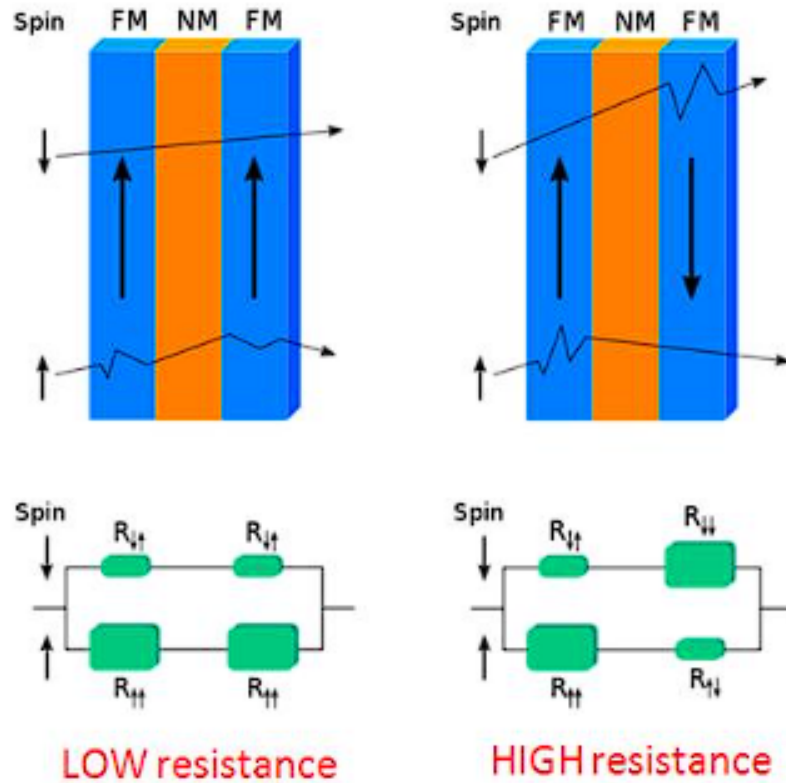


Figure 2.2: Key Elements of Spintronic Sensor Operation: The upper part showcases the central spin valve structure, comprising layers of magnetic and non-magnetic materials, the lower section elucidates the Giant GMR effect

by controlling the magnetic fields within the structure [20].

The spin-dependent phenomena utilized in spintronic sensors are the Giant Magnetoresistance (GMR) effect and the Tunnel Magnetoresistance (TMR) effect. Giant Magnetoresistance (GMR) Effect: The GMR effect appears when the electrical resistance of a spin valve structure changes due to the relative alignment between the magnetic layers. When the magnetic moments of the adjacent layers are parallel, electrons with aligned spins experience less resistance than when the magnetic moments are antiparallel, as illustrated in Figure 2.2 [21]. This change in resistance is detected as a change in the electrical signal, providing information about the magnetic field or the presence of magnetic materials [22].

Tunnel Magnetoresistance (TMR) Effect: The TMR effect is based on the phenomenon of electron tunnelling through a thin insulating barrier between two magnetic layers. When the magnetization of the two layers is parallel, electrons tunnel more efficiently, resulting in a higher electrical current. Conversely, electron tunnelling is inhibited when the mag-

netization is antiparallel, leading to a lower current. By measuring the electrical current through the tunnel junction, the relative orientation of the magnetic layers can be determined, delivering information about the applied magnetic field [23]. The output signal from a spintronic sensor is typically processed and amplified to provide valuable data about the measured magnetic field. This data can then be further analyzed and interpreted to understand the underlying phenomena or to extract relevant information about biological processes, such as biomagnetic signals.

The extraordinary sensitivity of spintronic sensors makes them significant in biomagnetic signal detection. Their high sensitivity allows the detection of even extremely weak magnetic fields, capturing subtle bio-magnetic signals produced by physiological processes. This sensitivity creates new possibilities for studying various biological phenomena, including brain activity, cardiac function, and muscle contractions [24]. Additionally, spintronic sensors exhibit low power consumption, making them energy-efficient and well-suited for portable and wearable applications. Their low power requirements extend the battery life of devices and minimize potential impacts on the biological systems under investigation [25].

In addition, their compatibility with integrated circuits allows for seamless integration into existing measurement systems and devices. This compatibility enables incorporating spintronic sensors into conventional electronic circuits, simplifying their implementation and enhancing their versatility. Moreover, spintronic sensors exhibit a fast response time, enabling the capture of rapid changes in magnetic fields. This fast response time makes them well-suited for real-time monitoring of dynamic processes and detecting transient biomagnetic signals.

Furthermore, spintronic sensors can be scaled to nanoscale dimensions, potentially developing miniaturized and susceptible sensing devices. This scalability allows for integrating spintronic sensors into compact and portable platforms, expanding their application in diverse settings.

While spintronic sensors offer significant advantages, there are challenges to overcome for their widespread adoption in bio-magnetic signal detection. These challenges include optimizing sensor performance, addressing noise sources, improving signal-to-noise ratios, and ensuring reliable and accurate data acquisition [26].

Future research efforts in spintronics aim to enhance the sensitivity, stability, and integration of spintronic sensors. This involves exploring new materials, engineering novel device structures, and developing advanced signal-processing techniques. Additionally, integrating spintronic sensors with other emerging technologies, such as artificial intelligence and machine learning, holds potential for further bio-magnetic signal detection and analysis advancements.

2.4 AI-Based Models

This section presents a concise overview of the AI-based techniques utilized in this study, highlighting their algorithms, strengths, and challenges. These techniques, including deep learning and machine learning, have demonstrated remarkable capabilities in extracting meaningful information from biomagnetic signals and enhancing signal analysis.

2.4.1 Linear Regression (LR)

Linear regression is a widely used machine learning algorithm that maps the relationship between a dependent variable and one or more independent variables. The algorithm aims to find the best-fitting linear equation representing the variables' linear relationship. The general form of a linear regression equation can be expressed as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \quad (2.1)$$

Here, Y represents the dependent variable, β_0 is the intercept term, β_1 to β_n are the coefficients corresponding to the independent variables X_1 to X_n , and X_1 to X_n are the independent variables.

Strengths of linear regression include its simplicity, interpretability, and computational efficiency. It provides a straightforward understanding of the relationship between the variables, as the coefficients indicate the magnitude and direction of each independent variable's impact on the dependent variable [27]. Linear regression also functions well when the relationship between the variables is approximately linear, making it suitable for many practical applications.

However, linear regression has certain limitations. One major weakness is its assumption of linearity in the relationship between the variables. If the actual relationship is non-linear, linear regression may not accurately capture the underlying patterns and may result in poor predictions [28]. Additionally, linear regression is sensitive to outliers, as they can disproportionately influence the estimated coefficients and affect the model's performance. Another area for improvement is the reliance on the independence of the errors; violating this assumption can lead to biased or inefficient estimates [29].

2.4.2 K-Nearest Neighbors (KNN) Regressor

K-Nearest Neighbors Regressor is a machine-learning algorithm for regression tasks. It predicts the value of a continuous target variable based on the values of its neighbouring data points. The algorithm finds the K nearest data points to a given query point and averages their target values to make predictions.

First, the number of neighbours K is chosen, indicating the nearest data points to be considered in the prediction process. This value of K can be determined based on the nature of the problem and the available data. Next, the algorithm computes the distance between the query point (the point we want to make a prediction) and all data points in the training set. A distance metric, such as the Euclidean distance, is commonly used to measure the closeness between points.

After calculating the distances, the K data points with the shortest distances to the query point are selected. These data points, also known as the nearest neighbours, are considered to be most similar to the query point. To make a prediction, the algorithm calculates the average or weighted average of the target values (e.g., the output variable) of the selected K neighbours. This aggregated value serves as the predicted target value for the query point.

By following this straightforward process, the KNN Regressor algorithm can provide predictions for new, unseen data points based on their proximity to existing data points in the training set. It leverages the notion that similar data points tend to have similar target values, allowing it to make predictions based on the nearest neighbours [30].

The K-Nearest Neighbors (KNN) Regressor has several strengths, making it a popular choice in machine learning. First, it offers simplicity in both understanding and implementation. The basic idea behind KNN is intuitive. It predicts a data point's value based on its tightest neighbours' values. This simplicity makes it accessible to both beginners and experienced practitioners.

Second, KNN is a non-parametric algorithm, which means it does not make any assumptions about the underlying data distribution [31]. This flexibility allows it to handle various data types, including numeric and categorical variables. This makes KNN a versatile choice for a wide range of applications.

Furthermore, KNN excels in localized learning. It captures local patterns in the data and adapts to different local structures. By considering the neighbours closest to a given data point, KNN can effectively capture local relationships and make accurate predictions within those regions.

The K-Nearest Neighbors (KNN) Regressor algorithm has some weaknesses that should be considered. First, the computational complexity of KNN increases as the dataset grows. Since the algorithm must calculate distances between the query point and all training data points, the computational cost can become significant, particularly for large datasets. As a result, the time required for prediction can be considerable [32].

Second, KNN Regressor requires storing the entire training dataset in memory. This can become memory-intensive, especially for datasets with many instances or high-dimensional features. As the dataset size increases, the memory usage of the algorithm also increases, potentially limiting its scalability.

Another weakness of the KNN Regressor is its sensitivity to noisy or irrelevant features in the data. Outliers or irrelevant features can significantly impact the distances between data points and influence the predictions. Therefore, it is crucial to preprocess the data and remove any noise or irrelevant features to improve the algorithm's robustness [33].

Determining the optimal value of K is another challenge in KNN Regressor. The choice of K affects the bias-variance trade-off in the algorithm. A small value of K may lead to overfitting, where the model captures noise or local variations too closely, resulting in poor generalization. On the other hand, a considerable value of K may result in underfitting, where the model oversimplifies the relationships in the data. Selecting the appropriate value of K requires careful tuning and validation to strike the right balance.

2.4.3 Decision Tree (DT)

The Decision Tree Regressor is a robust machine-learning algorithm used for regression tasks. It builds a tree-like model to predict continuous numerical values. It recursively partitions the data based on the values of input features, with each internal node representing a feature test and each leaf node representing a predicted numerical value. The algorithm selects the best feature and threshold to split the data, aiming to minimize the target variable's variance or mean squared error. Predictions are made by traversing the tree from the root to a leaf, following the appropriate branches based on the feature values of the predicted instance [34]. This process allows the model to adapt to complex decision boundaries and capture non-linear relationships between features and the target variable.

The strength of the Decision Tree Regressor lies in several key aspects. First, it provides interpretability, allowing an easy understanding of the decision-making process. The structure of the decision tree, with its if-else conditions at each node, enables a clear interpretation of how the features are used to make predictions. This transparency is valuable in domains where interpretability and explanation of predictions are crucial [35].

Second, Decision Trees can capture non-linear relationships between features and the target variable. By recursively splitting the data based on feature thresholds, the algorithm adapts to complex decision boundaries and captures interactions among different features. This flexibility enables the model to handle non-linear patterns in the data effectively. Furthermore, Decision Trees can handle numerical and categorical features, making them suitable for various datasets. They can automatically handle feature interactions and identify relevant features without extensive data preprocessing.

However, Decision Tree Regressors also have some limitations. One drawback is their tendency to overfit the training data, especially when the tree depth is not properly controlled. Overfitting occurs when the model becomes too complex and captures noise or peculiarities specific to the training data, resulting in poor generalization to new data. Regularization techniques, such as setting a maximum depth or pruning the tree, can help

mitigate overfitting and improve model performance [36].

Another limitation is the instability of Decision Trees to small changes in the data. A slight variation in the training set can lead to a significantly different decision tree. This sensitivity to data fluctuations can be addressed using ensemble methods like Random Forests or boosting algorithms combining multiple Decision Trees to improve stability and generalization.

2.4.4 Deep Neural Network (DNN)

The Deep Neural Network (DNN) model, also known as a deep learning model, is an artificial neural network consisting of multiple hidden layers between the input and output layers. It is designed to learn hierarchical representations of data by successively extracting more abstract features from the input.

The DNN model passes input data through multiple layers of interconnected nodes called neurons or units. Each neuron receives weighted inputs from the previous layer and applies a non-linear activation function to produce an output. The outputs from one layer are inputs to the next layer until the final layer, producing the model's prediction [37].

The computation for the weighted sum of inputs to a neuron, denoted as z , can be represented as:

$$z = w_1x_1 + w_2x_2 + \dots + w_nx_n + b \quad (2.2)$$

where w_1, w_2, \dots, w_n are the weights associated with the input features x_1, x_2, \dots, x_n , and b is the bias term [38].

The output of the neuron, denoted as a , is obtained by applying an activation function, typically a non-linear function, to the weighted sum:

$$a = \text{activation}(z) \quad (2.3)$$

Common activation functions used in neural networks include the sigmoid function, ReLU (Rectified Linear Unit), and tanh (hyperbolic tangent) function.

For a DNN with multiple hidden layers, the outputs from the previous layer serve as inputs to the next layer. The computations are performed layer by layer, propagating the information forward until reaching the output layer, where the final prediction is made.

DNN offers several advantages. First, DNNs can learn complex patterns, allowing them to handle intricate and non-linear relationships in data. This makes them well-suited for tasks such as image and speech recognition. Second, DNNs excel at automatic feature extraction, eliminating the need for manual feature engineering. They can automatically learn and extract relevant features from raw data, simplifying the modelling process. Third, DNNs are scalable and can handle large datasets with high-dimensional inputs. This scalability makes them suitable for big data applications [39].

However, the DNN model also has some drawbacks. One challenge is the extensive training data requirements. DNNs typically require a substantial amount of labelled training data to achieve good performance. Insufficient training data can lead to overfitting, where the model becomes too specialized to the training set and performs poorly on unseen data. Another consideration is the computational intensity of training deep neural networks. Complex tasks and large networks can be computationally demanding, requiring significant computational resources. This requirement can limit the applicability of DNNs in resource-constrained environments.

Additionally, DNNs are often considered black box models due to their lack of interpretability. Understanding and explaining the internal workings of DNNs can be challenging, which can be a drawback in domains where interpretability is crucial. Finally, DNNs are vulnerable to overfitting, especially when the model capacity is high relative to the available training data. Regularization techniques and careful model selection are necessary to mitigate overfitting and ensure generalization to unseen data [40].

2.4.5 Convolutional Neural Network (CNN)

The Convolutional Neural Network (CNN) model is a deep learning architecture that is particularly effective in processing structured grid-like data, such as images. It is designed to automatically learn and extract hierarchical representations of features directly from the input data. CNNs leverage convolution, a mathematical operation involving applying a filter (also known as a kernel) to the input data to extract local patterns and features. The CNN model works by using convolutional layers, pooling layers, and fully connected layers. The convolutional layers consist of multiple filters convolved with the input data, resulting in feature maps that capture different aspects of the input. These feature maps are created by applying convolutional operations, which involve element-wise multiplication and summation of the filter weights and the corresponding input values [41]. Mathematically, the convolution operation can be represented as:

$$(f * g)(i, j) = \sum_m \sum_n f(m, n) \cdot g(i - m, j - n) \quad (2.4)$$

where f represents the input data, g represents the filter, and (i, j) represents the spatial coordinates of the resulting feature map [42]. Pooling layers are then used to downsample the feature maps, reducing their spatial dimensions while retaining the most relevant information. Common pooling operations include max pooling, average pooling, and global pooling. Finally, fully connected layers are employed to classify or regress the learned features. These layers connect every neuron from the previous layer to every neuron in the current layer, allowing for high-level feature combinations and making the final predictions.

The advantages of CNN models include their ability to automatically learn hierarchi-

cal representations of features, which makes them highly effective for image and pattern recognition tasks. CNNs can capture the data’s local patterns and spatial dependencies, allowing for robust and accurate predictions. Additionally, CNNs can learn and generalize from many training examples, making them suitable for complex and large-scale datasets.

However, CNNs also have some drawbacks. One challenge is the high computational requirements, especially for training large networks on resource-constrained devices. Training CNNs typically involves many parameters and requires substantial computational resources. Additionally, CNNs may be prone to overfitting if the training dataset is limited, and regularization techniques are necessary to mitigate this issue [43]. Furthermore, the interpretability of CNNs can be challenging due to their complex and layered architectures, making it difficult to understand the exact reasoning behind their predictions.

2.4.6 Gated Recurrent Unit (GRU)

The Gated Recurrent Unit model is a recurrent neural network (RNN) architecture widely used for sequential data processing, such as natural language processing and time series analysis. It addresses some of the limitations of traditional RNNs, such as the vanishing gradient problem and the inability to capture long-term dependencies in sequences.

The GRU model works by utilizing gating mechanisms to control the flow of information within the network. It consists of update and reset gates, which determine how much of the previous hidden state to retain and how much new information to incorporate. This gating mechanism allows the GRU to update and propagate information over time selectively [44].

Mathematically, the GRU model can be described as follows:

$$\text{Update gate: } z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (2.5)$$

$$\text{Reset gate: } r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (2.6)$$

$$\text{Candidate hidden state: } h'_t = \tanh(W \cdot [r_t \cdot h_{t-1}, x_t]) \quad (2.7)$$

$$\text{Hidden state update: } h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot h'_t \quad (2.8)$$

Here, x_t represents the input at time step t , h_{t-1} represents the previous hidden state, z_t is the update gate, r_t is the reset gate, h'_t is the candidate hidden state, σ is the activation function, and h_t is the updated hidden state at time step t . W_z , W_r , and W are learnable weight matrices [45].

The advantages of the GRU model include its ability to capture long-term dependencies in sequences while avoiding the vanishing gradient problem. It achieves this by utilizing gating mechanisms that allow for selective information propagation and update. The GRU model also has a simpler architecture than other RNN variants, such as the Long Short-Term Memory (LSTM) model, making it computationally efficient and easier to train [46].

However, the GRU model also has some drawbacks. One limitation is its reduced memory compared to the LSTM model. The GRU has fewer gates and memory cells, which may limit its ability to capture complex long-term dependencies in some sequences. Additionally, like other deep learning models, the GRU model requires much-labelled training data to perform well. Limited training data can lead to overfitting or inadequate model generalization [47].

2.4.7 Long Short-Term Memory (LSTM)

The Long Short-Term Memory model is a recurrent neural network (RNN) designed to overcome the vanishing gradient problem and capture long-term dependencies in sequential data. It is widely used in various fields, including natural language processing, speech recognition, and time series analysis.

At its core, the LSTM model consists of memory cells that store information over time and gates that regulate the flow of information within the cells. These gates include the input gate, forget gate, and output gate, which control the information flow and selectively update the cell state [48].

Mathematically, the LSTM model can be described as follows:

$$\text{Input gate: } i_t = \sigma(W_i \cdot [h_{t-1}, x_t]) \quad (2.9)$$

$$\text{Forget gate: } f_t = \sigma(W_f \cdot [h_{t-1}, x_t]) \quad (2.10)$$

$$\text{Cell state update: } \tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t]) \quad (2.11)$$

$$\text{Cell state: } C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (2.12)$$

$$\text{Output gate: } o_t = \sigma(W_o \cdot [h_{t-1}, x_t]) \quad (2.13)$$

$$\text{Hidden state: } h_t = o_t \cdot \tanh(C_t) \quad (2.14)$$

Here, x_t represents the input at time step t , h_{t-1} represents the previous hidden state, i_t is the input gate, f_t is the forget gate, \tilde{C}_t is the candidate cell state, C_t is the updated cell state, o_t is the output gate, and h_t is the hidden state at time step t . W_i , W_f , W_C , and W_o are learnable weight matrices [49].

The advantages of the LSTM model include its ability to capture long-term dependencies, handle variable-length sequences, and alleviate the vanishing gradient problem. By incorporating the memory cells and gates, LSTMs can store and retrieve relevant information over extended intervals, making them well-suited for tasks involving long-range dependencies [50].

However, the LSTM model also has some drawbacks. One drawback is its increased computational complexity compared to simpler models like standard RNNs. The additional gates and memory cells contribute to more significant parameters and more intensive

computations, requiring more resources for training and inference. The LSTM model may also be susceptible to overfitting, especially when trained on small datasets. Regularization techniques and careful model selection are necessary to mitigate this issue [47].

2.4.8 Bidirectional LSTM (BiLSTM)

The Bidirectional LSTM model is a variant of the LSTM model that incorporates information from both past and future contexts by using two separate LSTM layers: one that processes the input sequence in the forward direction and another that processes it backward. This allows the model to capture dependencies in both temporal directions, enabling a more comprehensive understanding of the input sequence.

Typically, the BiLSTM can be described as follows. Given an input sequence, the forward LSTM layer processes the sequence from left to right, while the backward LSTM layer processes it from right to left. At each time step, both layers compute their respective hidden states based on the input at that time step and the hidden states from the previous time step. The final hidden state concatenates the forward and backward hidden states at each time step [51].

The BiLSTM model offers several advantages. Firstly, it enhances contextual understanding by incorporating information from both past and future contexts. This allows the model to capture a more comprehensive understanding of the input sequence, which is beneficial for tasks where the context in both directions is essential, such as part-of-speech tagging or named entity recognition. Secondly, the BiLSTM model is more robust to variations in the input sequence. Considering both preceding and succeeding elements, it can handle sequences with varying lengths and better capture long-range dependencies [52].

Another advantage of the BiLSTM model is its flexibility in architecture. It can be easily integrated into various architectures for different sequence-based tasks, including sequence labelling, machine translation, sentiment analysis, and more. Its versatility makes it popular in natural language processing and sequential data analysis domains.

However, the BiLSTM model also has some drawbacks. Firstly, its bidirectional nature increases computational complexity. Processing the input sequence in both directions requires more computation than unidirectional models like the LSTM. This can result in more extended training and inference times, especially for large datasets. Secondly, the BiLSTM model requires higher memory requirements. It needs to store the hidden states from both the forward and backward directions, which increases memory usage compared to unidirectional models. This can be a concern in memory-constrained environments or when dealing with large-scale models [53].

2.5 Research Gap

Based on the provided overview of the different components used in the thesis, including bio-magnetic signals (specifically MEG signals), spintronic sensors (specifically MTJ sensors), and AI-based models (including various machine and deep learning algorithms), there is a tremendous opportunity to elevate the project to new heights. The goal is to develop a portable, intelligent, and user-friendly MEG device that integrates these cutting-edge technologies.

While previous research has made significant strides in exploring MEG and EEG signals separately, leveraging deep learning techniques within these modalities remains a critical research gap in their combined exploration. This gap underscores the need to investigate the joint mapping of MEG and EEG signals within a unified framework. By undertaking this ambitious endeavour, the potential exists to revolutionize the field of biomedical devices and propel our understanding of neuroscience forward.

By embracing the synergistic nature of these signals and harnessing the power of deep learning and machine learning, this research holds promise in bridging the existing gap and realizing the substantial benefits and applications of an integrated approach. The upcoming chapters of the thesis will delve further into the intricacies of this mapping, emphasizing its importance and exploring its wide-ranging implications in bio-magnetic signal analysis and neuroscience.

Ultimately, this research aims to venture into uncharted territory by seamlessly combining MEG and EEG signals. This innovative approach not only propels advancements in biomedical devices but also unravels novel insights into the intricate workings of the human brain. By pushing the boundaries of interdisciplinary research, this project has the potential to reshape our understanding of neuroscience and pave the way for groundbreaking discoveries.

Chapter 3

Optimizing MEG-EEG Mapping in Resource-Constrained Non-Intrusive Bio-Magnetic Sensing Systems: A Data-Driven Approach

3.1	Introduction	25
3.2	Related work	26
3.3	Motivation and Problem Description	27
3.3.1	High Cost	28
3.3.2	Complexity	29
3.3.3	Weight and Mobility	29
3.4	Proposed MEG-EEG Mapping	30
3.4.1	Data Preprocessing and Channel Selection	30
3.4.2	Candidate Models for MEG-EEG Mapping	31
3.4.2.1	Linear Regression (LR)	31
3.4.2.2	K-Nearest Neighbors (KNN) Regressor	33
3.4.2.3	Decision Tree (DT)	33
3.4.2.4	Deep Neural Network (DNN)	33
3.4.2.5	Convolutional Neural Network (CNN)	34
3.4.3	Postprocessing	35
3.5	Performance Evaluation	36

3.6 Conclusion	38
--------------------------	----

3.1 Introduction

Improvements in brain imaging technologies have revolutionized our knowledge of the human brain and its intricate functioning. Among those techniques are magnetoencephalography (MEG) and electroencephalography (EEG). They are central tools for measuring high temporal and spatial resolution of neural activity [54]. However, the massive and expensive nature of the MEG devices and the complexity of EEG recording limit their practicality in real-life settings. In this chapter, we present an unconventional yet efficient data-driven approach that seeks to overcome these barriers by introducing the concept of a portable MEG device enabled by the outstanding capacities of recently emerging, ultra-sensitive Magnetic Tunnel Junction (MTJ) sensors fabricated by prominent research groups and industry stakeholders across the world including TDK, MDK, and so forth [55, 56].

While MEG and EEG work in tandem in abnormality diagnosis, each modality exhibits its unique advantage. For instance, EEG is superior to long-term video EEG recordings and can precisely detect temporal resolution [57]. Several diseases can be detected effectively by EEG more than MEG alone; for example, sleep abnormalities, such as sleep apnea, narcolepsy, and parasomnia, and abnormalities in sleep patterns (e.g., disruptions in sleep stages, sleep spindles, or REM sleep behavior disorder) [58]. There are other certain neurodegenerative disorders, such as dementia with Lewy body (DLB), that are difficult to differentiate from Alzheimer’s disease. Diagnosis of such diseases is difficult and expensive with available imaging and bio-molecular clinical tests [59]. We are motivated by MTJ/TMR-based bio-magnetic sensing systems which can achieve this much-desired functionality.

Our work, in this chapter, considers an ultra-sensitive MTJ/TMR bio-magnetic sensor and focuses on developing a robust mapping model that utilizes machine learning and deep learning algorithms to produce the EEG signals using the MEG channel records. By harnessing the potential of MTJ sensors and developing innovative signal processing methods, we pave the way towards a new era of portable and wearable brain imaging, poised to revolutionize healthcare, neurology, and brain-computer interfaces (BCIs).

To accomplish this mission, we pay attention to the mapping model by employing a dataset of MEG and EEG signals simultaneously recorded from a subject conducting various sensory tasks. The data are preprocessed by applying independent component analysis (ICA) to remove artifacts associated with eye movements (EOG) and cardiac activity (ECG) [60]. Furthermore, we perform bandpass filtering on the EEG and MEG data to extract relevant frequency details. Next, we separate the channels into EEG and MEG

channels. Then, we split the data into epochs based on event triggers, allowing us to analyze brain activity associated with specific stimuli. We then reshape the MEG and EEG data to create suitable input structures for our machine-learning models. The challenge of this problem is the variance in features between MEG and EEG records. We develop several models for MEG/EEG mapping and compare the results to suggest the best performance prototype depending on the mean square error (MSE), mean absolute error (MAE), and root mean square error (RMSE) evaluation matrices used in the regression tasks.

Our research opens up new prospects for developing portable MEG devices by successfully mapping MEG signals to related EEG signals. Integrating MTJ sensors in such devices enables real-time brain activity tracking in various settings, including clinical environments and everyday life [61].

The remainder of the chapter is organized as follows. The related research work is surveyed in Section 5.2. Then, we discuss the dimensions of the problem in Section 5.3. Next, the solution of the mapping problem is presented in Section 4.4 followed by the performance evaluation of our proposal in Section 5.4 Finally, we provide concluding remarks in Section 6.6.

3.2 Related work

Biomedical signal mapping has become an attractive topic for researchers, especially with the observed revolution in the powerful tools and algorithms in the machine and deep learning domains. This review [62] monitors the magnitude of the complex data preserved rapidly in the different healthcare sectors: hospitals, clinics, medical equipment and medical research. This big data becomes more valuable with the advancement of artificial intelligence/ machine and deep learning tools in a way that enhances healthcare delivery. Their focus on identifying the research gaps related to healthcare and artificial intelligence helps us investigate this area of biosignal mapping to utilize the power of such technologies.

We are also inspired by another review whose interest is understanding the relationship between biosignals and psychological stress. This work [63] has analyzed the consequences of stress on the human body, using various biosignal measures (EEG, electrocardiography (ECG), electrodermal activity (EDA), electromyography (EMG), respiratory rate, speech patterns, skin temperature, pupil size, and eye activity) to capture the physiological and physical responses associated with stressors. This type of mapping remains limited unless further investigation arises to unleash a deeper understanding of the multitude of biosignal features. This chapter uses deep learning models to investigate precisely MEG and EEG signal mapping.

In the context of our ongoing research, the exploration of mapping between Magneto-myography (MMG) and Electromyography (EMG) holds tremendous potential in advancing

our understanding of human physiological processes and enhancing the accuracy and efficacy of our own work [64]. They applied a recently proposed electronic Rotating Neuron Reservoir (eRNR) model using the self-collected MMG data to predict the EMG from the corresponding noisy MMG signal. Their model successfully maps the MMG signal to EMG with adequate normalised root mean square error ($mse = 0.3894$).

In the domain of mapping between different types of biosignals, a remarkable study by [65] concentrated on mapping cardiac dysfunctions using Electro/Magnetocardiogram (E/MCG) signals. Their research aimed to solve the inverse problem of estimating activities at the source level by constructing a spatial matrix based on vectors and utilizing a Bayesian approach. They incorporate the updates derived from Vectorcardiography signals to enhance the accuracy of the inverse problem solution. Similarly, the [66] authors used the ECG signals obtained by the exact position of MCG records to validate the detected signals by their invented wearable multichannel MCG system based on a spin exchange relaxation-free regime (SERF) magnetometer array. They applied the independent component analysis (ICA) and empirical mode decomposition (EMD) techniques to denoise the multichannel MCG records. They ended up with a robust system that can effectively capture the MEG signals through this wearable device.

Several endeavours study the relationships between the different biosignals through AI algorithms. Among those attempts to map such signals, the work in [67] may be regarded as pioneering. A deep learning model was developed in that work to accurately map the magnetocardiography (MCG) to the electrocardiography (ECG). They designed a model comprising a one-dimensional convolution layer, Gated Recurrent Unit (GRU) layer, and a fully-connected neural layer to learn the relationship between the MCG and ECG. Consequently, their developed model was shown to predict the ECG signal given the MCG signal acquired by the MTJ sensor. In addition, researchers in [68] mapped the noisy MCG to obtain an AI-filtered MCG signal by adopting and comparing several deep learning approaches [69]. As mentioned earlier, the extension of the research is to investigate the non-linear MEG/EEG mapping to enrich the biomedical signal domain, which has yet to be studied.

3.3 Motivation and Problem Description

There is no doubt about the qualitative shift in brain imaging brought about by the MEG technique, which we have benefited from until now; even our research that provides a proof of concept aims to replace these devices with others that do not suffer from what these devices suffer from. This section highlights some of the MEG/EEG device's challenges, including high cost, complexity, and portability, which necessitate developing alternative solutions, such as a portable and affordable MEG device based on ultra-sensitive MTJ

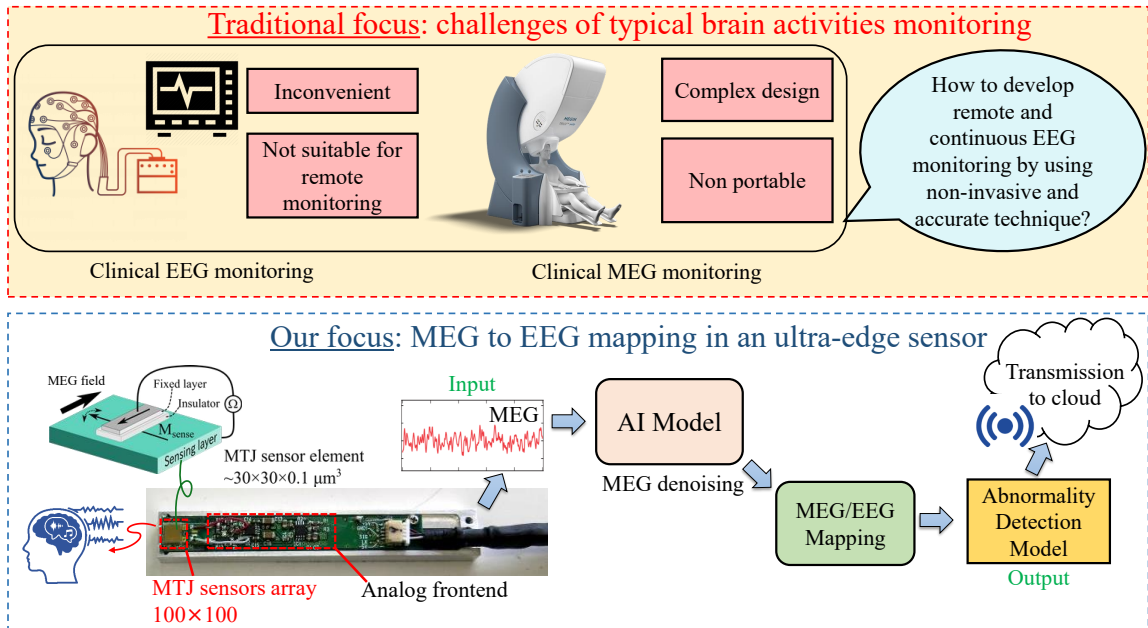


Figure 3.1: Traditional brain activities monitoring vs. our envisioned AI-based MEG monitoring system for brain activity monitoring, including a schematic of the MTJ sensor and the outline of a lightweight AI pipeline. Note that the core focus of our research in this chapter is on the MEG to EEG (MEG-EEG) mapping component of the pipeline.

sensors.

3.3.1 High Cost

Conventional MEG machines are known for their high operational cost. The sophisticated hardware and advanced electronics used in constructing MEG scanners contribute to their high price tags. The machine cost is around 2.5 million Euros plus the running expenses, which can be about 70,000 Euros per year for cooling with helium [70]. This cost poses a significant barrier, particularly for smaller research institutions, clinical settings, and resource-limited regions. The financial burden of acquiring and maintaining traditional MEG systems often bounds their availability and hampers the progress of vital neuroscientific research.

Our MEG/EEG mapping work is meaningful since we seek to avoid the constraints brought about by the high operational cost of MEG devices by creating precise and trustworthy mapping methodologies. The ability to map MEG signals to corresponding EEG signals provides a cost-effective avenue for accessing valuable information about brain activity. Researchers and clinicians can gain insights into neural dynamics and brain connectivity without expensive MEG equipment by utilizing EEG systems in conjunction with our mapping techniques. This reduces the financial burden and expands the accessibility of

advanced neuroscientific research and clinical applications to a broader range of institutions and regions.

3.3.2 Complexity

MEG devices typically consist of intricate systems of superconducting materials, cryogenic cooling mechanisms, and complex data acquisition systems. The design of this machine necessitates a shielded room to prevent interaction with the other magnetic fields, in addition to superconducting quantum interference devices (SQUIDs), which require a massive cooling system to provide connectivity [71]. These features require specialized technical expertise for installation, function, and maintenance. The complexity of traditional MEG systems makes them less accessible to researchers and clinicians without extensive training in MEG technology. Additionally, the complexity introduces portability and ease of use challenges, limiting their application in various settings beyond dedicated research laboratories.

MEG/EEG mapping is a feasible solution to overcome the complexity associated with traditional MEG devices. By mapping MEG signals to their corresponding EEG counterparts, we effectively bypass the need for intricate superconducting materials, cryogenic cooling systems, and specialized data acquisition mechanisms. The mapping process enables researchers and clinicians to leverage the accessibility and simplicity of EEG technology while still harnessing the advantages of MEG by simplifying the experimental setup and data acquisition process. This approach not only reduces the financial burden of acquiring and maintaining expensive MEG systems but also eliminates the need for shielded rooms and complex cooling setups.

3.3.3 Weight and Mobility

The weight of the MEG machines is the primary limitation to portability due to the requirements needed to acquire the data and eliminate any magnetic interactions. Despite this, our proposed technique uses deep learning and machine learning models to perform this task accurately. The lightest MEG device that exists nowadays is a few hundred kilograms. Therefore, any attempt to put that vast device into an IoT device and be available to people 24 hours is worth it.

In this chapter, we address the aforementioned challenges by conceptualizing a novel idea of a portable MEG device. We build our theory on the sensitivity of emerging MTJ sensors, which have been demonstrated to capture extremely weak bio-magnetic fields that are associated with the neural activity of humans. The captured signal passes through an AI-based filtering model to purify the signal from the sensor-added noises similar in the spirit in the coauthor's earlier work with denoising MCG signals [67]. Our research focuses on developing a mapping model between the EEG signals from the pure MEG activities. We

aim to construct a data-driven model to understand the intricate, non-linear relationship between MEG and EEG signals, which can be later widely used in many applications, such as neurology, healthcare, and brain-computer interfaces, as depicted in Figure 3.1

3.4 Proposed MEG-EEG Mapping

In this section, we present our envisioned MEG-EEG mapping technique. First, we describe the data preprocessing and channel selection process. Next, we perform a systematic investigation of the relevant data-driven models for MEG-EEG mapping. Then, we delineate how to post-process the mapped ECG signal for acceptable resolution for visualization and clinical interpretability purpose.

3.4.1 Data Preprocessing and Channel Selection

This study investigates the MEG/EEG mapping problem using the “sample” dataset available in the MNE Python library. The following steps and considerations guided our research.

First, we loaded the dataset and carefully examined its characteristics. We identified the ECG and EOG channels deemed inappropriate for further analysis. Thus, we excluded it from our subsequent computations. Next, a key step in our approach consisted in the utilization of Independent Component Analysis (ICA). This technique, known for its ability to separate mixed signals into their underlying sources, was chosen to address the challenge of isolating independent components associated with specific brain activity patterns [72].

After applying ICA, we focused on artifact removal to enhance the data quality. Specifically, we targeted the removal of ECG and muscle artifacts by identifying and eliminating the corresponding components. Since the dataset combined MEG and EEG signals, we partitioned the data into two files to facilitate the subsequent training process.

The raw data encompassed various essential details. The measurement date was December 03, 2002, at 19:01:10 GMT. The MEG team experimented, and the participant’s identity remained unknown. The dataset comprised 146 digitized points, including 204 Gradiometers, 102 Magnetometers, 9 Stimulus channels, 60 EEG channels, and 1 EOG channel. We noted the presence of bad channels, namely MEG 2443 and EEG 053. Furthermore, EOG channel 061 was identified as an EOG channel, while ECG channels were unavailable. The sampling frequency was 600.61 Hz, and a highpass filter at 0.10 Hz and a lowpass filter at 172.18 Hz were applied [73].

Following artifact removal, we applied bandpass filtering to the MEG and EEG signals. For the MEG data, we employed a bandpass filter ranging from 1 Hz to 100 Hz to contain all the relevant frequencies, as suggested in the relevant literature [74], aiming to remove unwanted noise and concentrate on the desired frequency range. Similarly, the EEG data underwent bandpass filtering within 0.5 Hz to 55 Hz [75].

We divided the filtered data into training and testing sets to facilitate model training and evaluation, adopting a 75-25 ratio. Regarding our analysis, we employed five distinct candidate models to address the MEG/EEG mapping task that are presented in the following subsection.

3.4.2 Candidate Models for MEG-EEG Mapping

From hereon, we describe our considered candidate data-driven models to map MEG to EEG signals that include linear regression, K-nearest neighbours (KNN), decision trees, artificial neural network (ANN), and convolutional neural network (CNN) models to capture the complex relationships between the MEG and the EEG signals.

3.4.2.1 Linear Regression (LR)

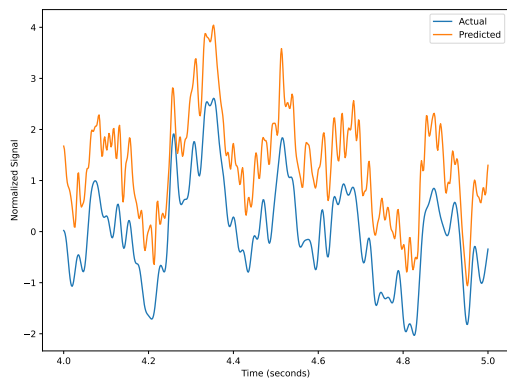
We applied linear regression, a widely-used supervised learning algorithm, to designate a linear relationship between the input MEG and output EEG channels [76]. The dependent variable represents the EEG signal, while the independent variables represent the MEG signals. The dependent variable, the EEG signal, is the target variable that we want to predict or map using the MEG signals. It consists of samples measured across 59 channels, which capture the electrical activity of the brain at specific locations. On the other hand, the independent variables are the MEG signals, which serve as predictors in the linear regression model. These MEG signals are measured across 305 channels and represent the magnetic fields generated by the electrical activity of the brain.

The purpose of the linear regression model is to establish a relationship between the MEG signals (independent variables) and the corresponding EEG signal (dependent variable). The model learns the coefficients for each MEG channel that best predict the corresponding EEG signal across the 59 channels. This model minimizes the mean squared error between the predicted and actual values. The mapping process can be mathematically represented as follows:

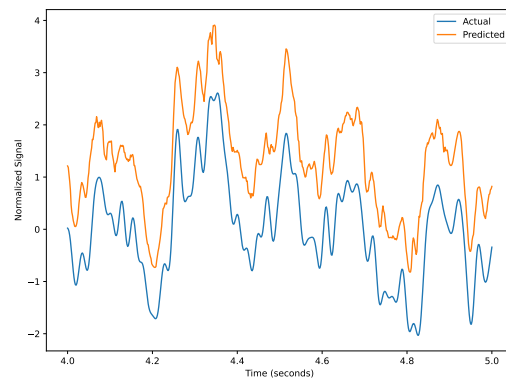
$$\mathbf{y}_{eeg} = \mathbf{X}_{meg}\mathbf{W} + \mathbf{b} \quad (3.1)$$

where \mathbf{y}_{eeg} is the predicted EEG signals, a 2D array of dimensions (number of samples, number of EEG channels). \mathbf{X}_{meg} represents the MEG input signals, a 2D array of dimensions (number of samples, number of MEG channels). \mathbf{W} is the weight matrix, a 2D array of dimensions (305, 59), representing the weights for mapping MEG to EEG. \mathbf{b} is the bias vector, a 1D array of length 59, representing the biases for each EEG channel [77].

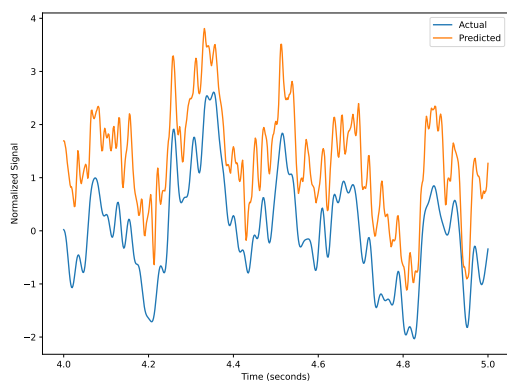
In this formulation, the model learns the weight matrix \mathbf{W} and bias vector \mathbf{b} during training to map the MEG input signals to the corresponding EEG signals.



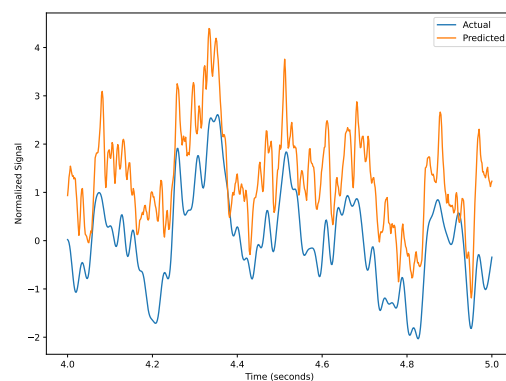
(a) Actual vs predicted signal performance with the LR model for channel 41.



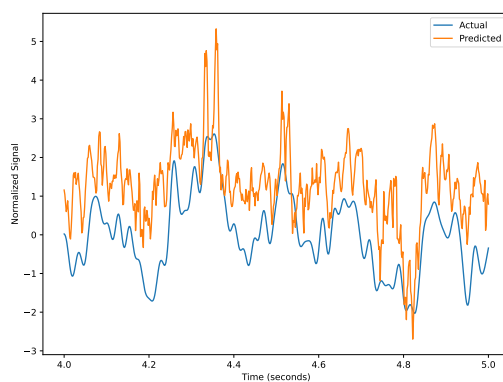
(b) Actual vs predicted signal performance with the KNN model for channel 41.



(c) Actual vs predicted signal performance with the DNN model for channel 41.



(d) Actual vs predicted signal performance with the CNN model for channel 41.



(e) Actual vs predicted signal performance with the DT model for channel 41.

Figure 3.2: Filtered signals of channel 41 and time range 4-5 seconds: Actual vs. predicted signals by different models.

3.4.2.2 K-Nearest Neighbors (KNN) Regressor

KNN regressor is a non-parametric algorithm representing a non-linear approach to estimating the input/output relationship, especially those that parametric methods cannot recognize. It calculates the distance between the new data point and all the training data points. Then, the KNN regressor considers the closest k neighbors to take the average of their target values and accordingly calculate the output.

The anticipated EEG signal, \mathbf{y}_{eeg} , is computed as the average of the target values of these K neighbours. Mathematically, the KNN Regressor can be represented as:

$$\mathbf{y}_{eeg} = \frac{1}{K} \sum_{n=1}^K \mathbf{y}_{meg}^{(n)} \quad (3.2)$$

where $\mathbf{y}_{meg}^{(n)}$ represents the target values of the K nearest neighbors, and K is the number of neighbors considered for prediction, we choose $k = 5$ neighbours empirically [78].

3.4.2.3 Decision Tree (DT)

Decision trees are adaptable machine learning models that partition the feature space into branches based on a hierarchy of if-else conditions. By recursively splitting the data based on the most independent features, the decision tree regressor predicts the output value by averaging the target values of the samples within each leaf node [79]. We construct our tree model with up to 15 split levels before reaching the output leaf node. This design enhances the capability of the model to capture the complexity within the MEG and the corresponding EEG features.

$$\hat{y}_{eeg} = f(\mathbf{x}_{meg}) = \sum_{j=1}^m C_j \cdot \mathbb{I}(\mathbf{x}_{meg} \in \mathbf{R}_j) \quad (3.3)$$

In this equation, [80], \hat{y}_{eeg} denotes the predicted EEG signal, \mathbf{x}_{meg} is the input MEG signal with 305 channels, \mathbf{C}_i is the constant weight associated with the i -th leaf node, \mathbf{R}_i represents the region defined by the conditions at the i -th internal node, and $\mathbb{I}(\cdot)$ is the indicator function that returns 1 if the condition is true and 0 otherwise. The equation sums up the constant values of the leaf nodes for which the input MEG signal falls into the corresponding regions.

3.4.2.4 Deep Neural Network (DNN)

Since the MEG-EEG mapping deals with a non-linear relationship, neural network structures may be regarded as a natural candidate model to formulate such a relationship. Therefore, next, we utilized an artificial neural network with deep structures, also referred

to as a Deep Neural Network (DNN), which is a versatile and powerful deep learning approach. The DNN model, as shown in Figure 3.3(a), begins with a dense layer of 128 units, which involves the Rectified Linear Unit (ReLU) activation function to introduce non-linearity. This layer receives input data corresponding to 305 channels. One additional dense layer with 256 and 128 units, respectively, and ReLU activation are added to capture complex patterns in the data. To prevent overfitting, dropout layers are incorporated after the second dense layer, randomly dropping out 50% of the units during training. The final dense layer, comprising 59 units and linear activation, provides the regression output for the EEG signal.

During the training process, the DNN model was compiled using the Adam optimizer and the mean squared error (MSE) loss function, commonly employed in regression tasks. The training-validation split of the dataset generates the training loss vs validation loss curve to evaluate the model’s performance and pinpoint any potential overfitting or underfitting issues. In our experiment, we trained the DNN model for 100 epochs using an early stopping technique with a patience value of 5. If the validation loss did not improve for five consecutive epochs, the training process would be stopped early to prevent overfitting. Interestingly, the model showed promising performance and achieved convergence within relatively few epochs. Specifically, the training stopped at epoch number 24, as shown in 3.3(b), indicating that the model had already captured the underlying patterns in the data and that further training was unnecessary. This early stopping mechanism helped us save computational resources and avoid potential overfitting issues.

3.4.2.5 Convolutional Neural Network (CNN)

CNNs are deep learning models that excel at processing grid-like data, such as images or sequences [81]. We deploy a CNN model consisting of 5 cascaded layers as shown in Figure 3.4(a). The temporal and spatial convolutions are performed to capture local patterns in the data using the ReLU activation function in the first and third layers. The second and the fourth layers are max pooling layers. They are crucial in extracting the most relevant features from the signal and abstracting the underlying patterns. Then we apply a flattened layer that reshapes the output into a one-dimensional vector to bridge the previous convolutional layers and the fully connected dense layer to get a proper outcome that fits our problem. Our CNN model accepts a 3D input of MEG samples, the number of MEG channels, and one, and produces 2D output with EEG samples and EEG channels. We use Adam as an adaptive optimizer and MSE as an evaluation matrix and loss function.

We employ early stopping with a patience 5 to ensure optimal performance while avoiding overfitting, like the DNN model. In our case, the model stops training at epoch number 41. The training and validation loss curves are depicted in Figure 3.4(b), showing the progression and convergence of the model during training. The decreasing trend of both curves

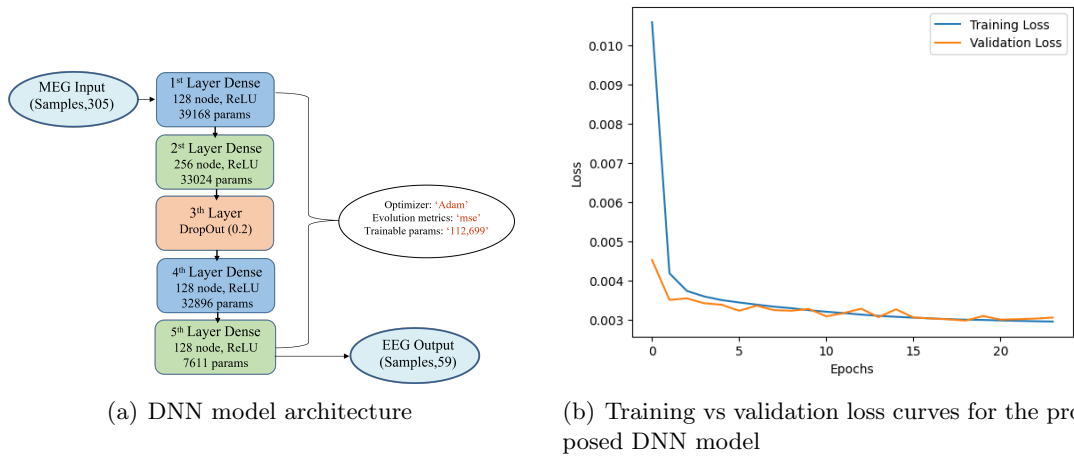


Figure 3.3: Fix caption.

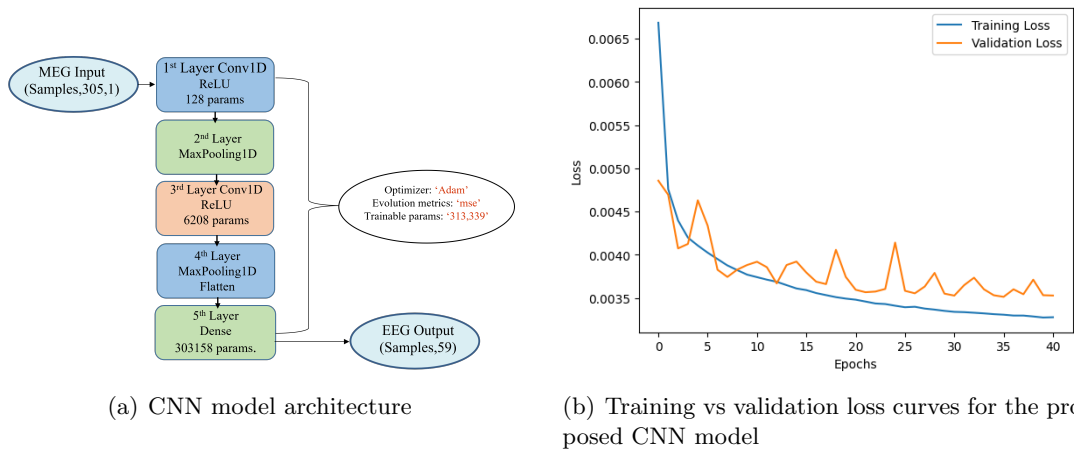


Figure 3.4: comparison of Training and Validation Loss Curves and Model Architectures for the Proposed DNN and CNN Models

indicates that the model learns from the training data and generalizes well to the validation data.

3.4.3 Postprocessing

After obtaining the predicted EEG signals from each model described in section 3.4.2, we applied a moving average filter to smooth the signals and reduce high-frequency noise. The moving average filter takes the input signal and window size as parameters. The function applies a moving average operation to the signal using a window of ones divided by the window size. The window of size five smooths out the high-frequency noise in the predicted signals, resulting in cleaner and more interpretable waveforms. By involving the moving

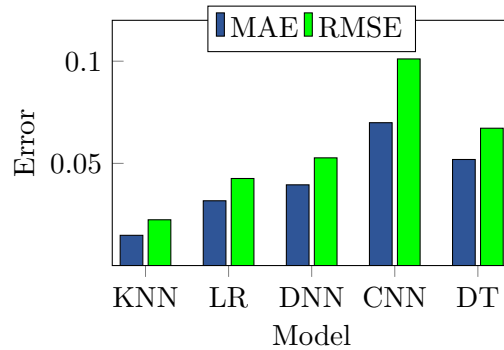


Figure 3.5: Performance comparison of MAE and RMSE for different models.

average filter in the predicted EEG signals, we improve the overall signal quality and reduce the impact of short-term fluctuations, thereby enhancing the interpretability and visual clarity of the predicted EEG waveforms [82]. This postprocessing step prepares the signals for further investigation and evaluation of their performance.

Figure 3.2 illustrates the typical EEG test signal along with the prediction results of different models after the post-processing step for channel number 41. Specifically, Figure 3.2(d) represents the prediction results using the Linear Regression (LR) model, Figure 3.2(b) shows the results obtained from the K-Nearest Neighbors (KNN) model, Figure 3.2(d) displays the predictions from the Convolutional Neural Network (CNN) model, Figure 3.2(e) depicts the predictions made by the Decision Tree (DT) model, and Figure 3.2(c) showcases the results obtained using the Deep Neural Network (DNN) model. To facilitate visual interpretation, we focus on a specific time interval of 1 second, specifically between the second 4 to 5. This selected timeframe allows us to observe the performance of each model in capturing the EEG signal dynamics and comparing it to the ground truth signal. By examining these figures, we can gain insights into how well each model predicts the EEG signal for channel number 41, thereby assessing their effectiveness in MEG/EEG signal mapping.

3.5 Performance Evaluation

The evaluation measures used in this chapter to assess the proposed models' effectiveness for the MEG/EEG mapping task include root mean square error (RMSE), mean absolute error (MAE), and the prediction time [83]. Figure 4.2 visually represents the performance comparison among different models in terms of MAE and RMSE. The x-axis represents the models, including KNN, LR, DNN, CNN, and DT, while the y-axis represents the error values.

Table 6.1 indicates a detailed summary of each model's performance metrics and pre-

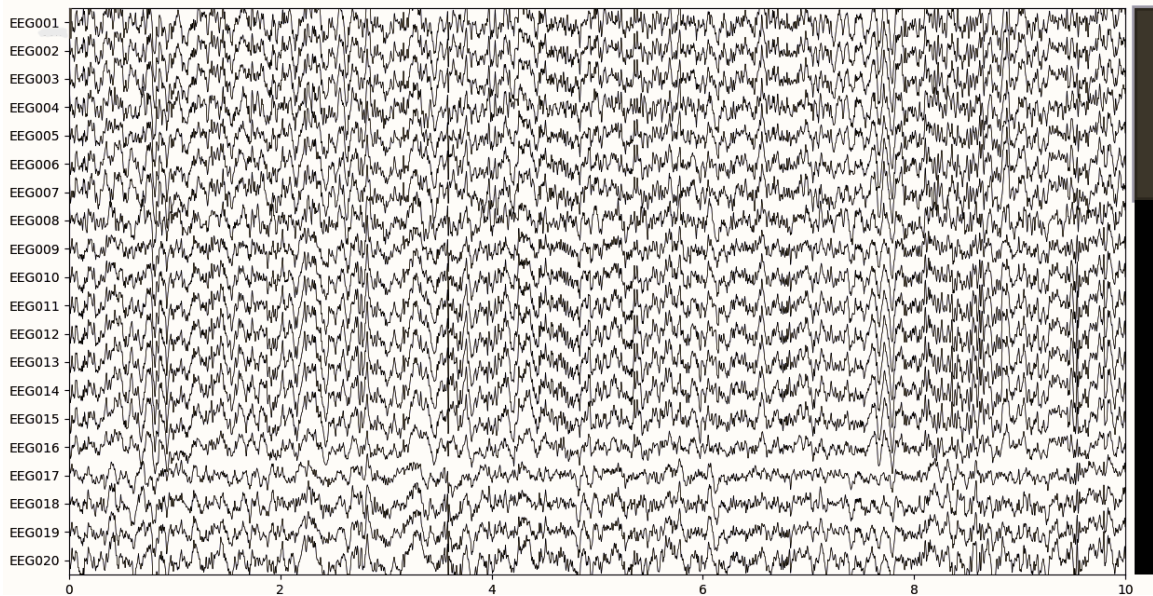


Figure 3.6: The first 20 EEG channels filtered based on the LR model's predictions

Table 3.1: Performance metrics and prediction times for different models.

Model	MAE	RMSE	Prediction Time (s)
KNN	0.0148	0.0224	29.0600
LR	0.0317	0.0426	0.1255
DNN	0.0395	0.0527	7.1060
CNN	0.0699	0.1011	6.2057
DT	0.0519	0.0672	0.0911

diction times. The MAE and RMSE values are listed for each model, indicating the error level between the predicted and actual values. Additionally, the prediction times for each model are included, offering insights into the computational efficiency of the algorithms. These findings contribute to evaluating and selecting the most fitting (i.e., most appropriate) model for the MEG-EEG mapping task that could be viable for resource-constrained IoT (Internet-of-Things) capability incorporation with MTJ sensing systems.

The KNN regressor exhibited the best performance in terms of minimization error. However, it is computationally expensive since it requires the longest prediction time among all models. On the other hand, the DT model predicted the test samples instantly. However, it did not demonstrate our expected performance. From the optimization point of view, the LR model achieved high performance in a reasonable time, as depicted in Figure 5.5.

Figure 3.7 represents the predicted EEG signal after applying the average moving filter

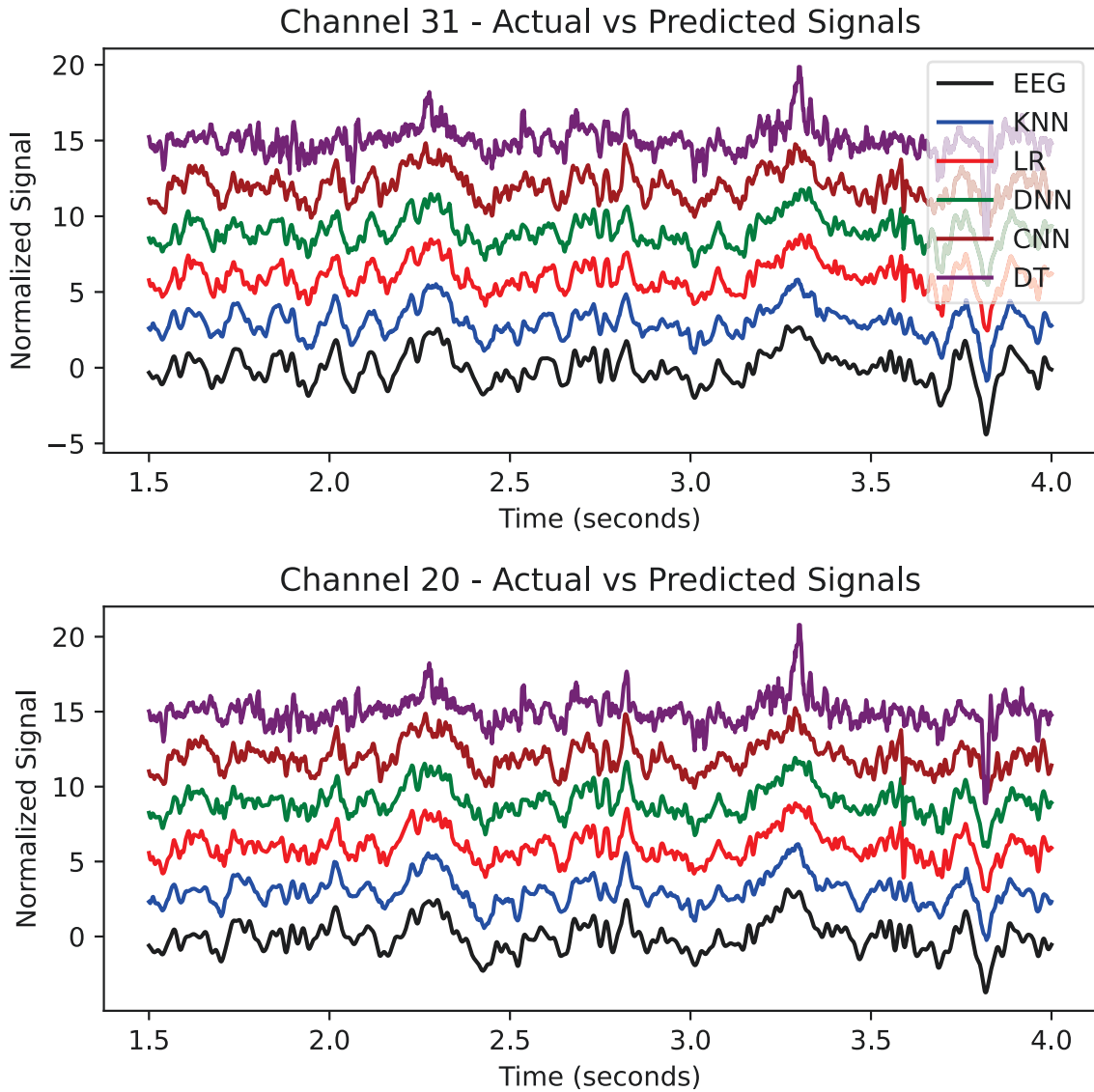


Figure 3.7: Actual vs. predicted signals for randomly selected EEG channels.

with the actual EEG record for two random channels from the 59 EEG channels within the selected period, i.e., 1.5-4 seconds. We normalized the signal amplitude for the sake of better visualization. The results in the figure demonstrate how implemented models are morphologically similar to the standard version of the EEG signal.

3.6 Conclusion

In this chapter, we proposed the concept of integrating the lightweight IoT-based biomagnetic sensors with a softwarized mapping model that transfers the captured MEG signals

into EEG signals that can be used to detect the EEG abnormality. This proof-of-concept is a baseline to improve the functionality of emerging wearable devices, powered by the MTJ sensing technology, for spontaneously monitoring and tracking brain health. We focused in this research on the MEG to EEG mapping component, and will thoroughly investigate the other elements to discover an optimized system that may provide users with non-intrusive brain function monitoring at home with accurate, clinical-grade performance with portability.

As candidate data-driven solutions, we investigated various machine learning models, including KNN, LR, DNN, CNN, and DT, to evaluate their respective performances in terms of MAE, RMSE, and prediction time. The results indicated that the KNN model outperformed other models in terms of accuracy while incurring a high computation time, while LR showcased a good balance between accuracy and computational efficiency. On the other hand, the DNN and CNN models exhibited a promising potential for capturing complex patterns in the EEG signals that could be leveraged for fast MEG to EEG mapping and thereby real-time disease inference on board resource-constrained MTJ-based IoT devices.

Chapter 4

Improved MEG-EEG Mapping Via Combined Spatial-Temporal Pattern Detection

4.1	Introduction	40
4.2	Related work	42
4.3	Considered System Design and Problem Description	44
4.4	Envisioned RNN Models Temporal Pattern Detector for MEG-EEG Mapping	45
4.4.1	Candidate RNN Models	45
4.4.1.1	GRU	45
4.4.1.2	LSTM	46
4.4.1.3	Bi-directional LSTM	47
4.4.2	Proposed Recurrent Layer based Convolution: Fusion of spatial and temporal detectors for MEG-EEG mapping	47
4.5	Data Preparation and Performance Evaluation	50
4.5.1	Data Preprocessing and Channel Selection	50
4.5.2	Evaluation metrics	52
4.6	Conclusion	53

4.1 Introduction

Magnetoencephalography (MEG) and electroencephalography (EEG) are powerful modalities for recording brain activity and have revolutionized the field of neuroscience. These

non-invasive techniques provide invaluable insights into the complex workings of the human brain and have been instrumental in understanding various cognitive processes, neural dynamics, and brain disorders. MEG excels in capturing the fine-grained spatial details of neural activity, while EEG provides excellent temporal resolution [84]. The complementary nature of MEG and EEG signals makes their combined analysis crucial for a comprehensive understanding of brain function.

However, traditional MEG and EEG systems present challenges that impede widespread adoption and utilization [85], [86]. The setup and operation of these devices often require specialized infrastructure, making them accessible only to a limited number of well-equipped research centers. The complexity and weight of the equipment pose practical limitations, especially for studies involving special populations or prolonged recording sessions. Moreover, the high costs associated with MEG and EEG devices can further limit their usage in various research and clinical settings. These challenges necessitate exploring alternative methods for acquiring MEG signals and mapping them to corresponding EEG signals.

In recent years, spintronic sensors, particularly Magnetic Tunnel Junctions (MTJs), have emerged promising alternatives for measuring brain activity. These sensors utilize the unique properties of magnetic materials to detect minute changes in magnetic fields generated by neural activity. The simplicity, lightweight nature, and integration capabilities of MTJs make them ideal for portable and wearable applications, offering the potential to overcome the limitations of traditional MEG and EEG devices [87]. Using MTJs presents an opportunity for more accessible and user-friendly MEG signal acquisition, enabling wider adoption of this modality in various research and clinical settings.

Mapping MEG signals to corresponding EEG signals is paramount, as these modalities provide significant information about brain activity. MEG offers precise localization of brain regions involved in various cognitive processes [88]. On the other hand, EEG enables the analysis of brain rhythms and the dynamics of neural activity over time [89]. One modality may better capture and understand certain brain disorders and cognitive processes. Therefore, integrating MEG and EEG signals through accurate mapping can enhance our ability to detect and diagnose neurological disorders more accurately, monitor cognitive processes in real-time, and develop targeted treatment strategies tailored to individual patients.

Advanced deep learning techniques are employed to address the temporal dependencies present in MEG and EEG signals. Recurrent Neural Network (RNN)-based models, including Gated Recurrent Units (GRU), Long Short-Term Memory (LSTM), and Bidirectional LSTM (BiLSTM), have proven effective in capturing temporal patterns in sequential data. Additionally, there has been a growing interest in combining Convolutional Neural Networks (CNNs) and RNNs to capture spatial and temporal patterns simultaneously. One such model is the Recurrent Convolutional Neural Network (RCNN), which incorporates Recurrent Convolutional Layer (RCL) to process the input data using both CNN and RNN

operations. The RCNN model offers a unique approach to considering temporal and spatial patterns in MEG and EEG signals.

Our contribution in this chapter is as follows:

1. Mapping MEG and EEG signals

- Develop a comprehensive framework for mapping MEG signals acquired using MTJ sensors to corresponding EEG signals.

2. Considering temporal patterns using RNN-based models

- Utilize advanced RNN-based models, including GRU, LSTM, and BiLSTM, to capture the temporal dependencies within the MEG and EEG signals.

3. Enhancing spatial and temporal resolution through RCL-based model

- Incorporate RCNN to consider the spatial and temporal features together of the mapped MEG and EEG signals.

This proposed framework combines the advantages of spintronic sensors, such as MTJs, and the power of RNN-based models to overcome the limitations of traditional MEG and EEG systems. The aim is to enable accurate mapping between these modalities and develop more accessible, user-friendly, and comprehensive approaches for studying brain activity. The study holds great potential for applications in neuroscience research, brain-computer interfaces, and clinical diagnostics.

The remainder of the chapter is organized as follows. The related research work is surveyed in Section 5.2. Then, we discuss the dimensions of the problem and the design considerations in Section 5.3. Next, the solution of the mapping problem is presented in Section 4.4 followed by the performance evaluation of our proposal in Section 5.4. Finally, we provide concluding remarks in Section 6.6.

4.2 Related work

Few studies have investigated the mapping of different physiological signals using deep learning techniques, demonstrating the potential of these models in capturing complex relationships and improving prediction accuracy. One example is mapping magnetocardiography (MCG) to electrocardiography (ECG) signals. These studies have successfully employed deep learning models to establish accurate mappings between MCG and ECG signals, showcasing these techniques' effectiveness in physiological signal mapping [61], [69], [68].

In the previous chapter, we focused on the importance of mapping magnetoencephalography (MEG) to electroencephalography (EEG) signals and applied various machine learning and deep learning techniques. We aimed to explore the potential of these models in capturing the complex relationships between MEG and EEG signals. The techniques employed

included K-nearest neighbours (KNN), decision trees (DT), linear regression (LR), deep neural networks (DNN), and convolutional neural networks (CNN).

To evaluate the performance of these models, we employed metrics such as root mean squared error (RMSE) and mean absolute error (MAE), which provided insights into the prediction error of each model. Additionally, we considered the time required for prediction, as it was crucial to choose an optimal model suitable for real-time applications, particularly in an Internet of Things (IoT) system.

Before applying the machine learning and deep learning techniques, we performed pre-processing on the MEG and EEG data to enhance its quality and prepare it for analysis. This preprocessing involved noise removal, artifact removal, and feature extraction. By transforming the raw data into a suitable format, we ensured that the models received high-quality input data for mapping purposes.

Our experiments also incorporated post-processing techniques to refine the output signals further. Specifically, we applied an average filter to smooth the output signals generated by all the models. This post-processing step aimed to enhance the overall quality of the predicted EEG signals and improve their coherence and interpretability.

Through our comprehensive evaluation, we aimed to identify the optimal model for MEG-EEG mapping, considering both prediction accuracy and computational efficiency. Our results demonstrated the linear regression model as an optimal solution that highlights the potential of machine learning and deep learning techniques in accurately mapping MEG signals to corresponding EEG signals.

However, despite these achievements, there are still challenges and opportunities for further research. For instance, exploring more advanced deep learning architectures, such as recurrent neural networks (RNNs) or hybrid models combining CNNs and RNNs, may improve the modelling of temporal dependencies within MEG and EEG signals. Additionally, investigating the potential of transfer learning and domain adaptation techniques could enhance the generalizability of the mapping models across different subjects and experimental settings [90].

In this paper [91], researchers have recognized the impact of incorporating both spatial and temporal features of EEG signals for tasks such as imagined speech decoding. A novel model called hybrid-scale spatial-temporal dilated convolution network (HS-STDCN) was proposed to capture the spatial-temporal reliances and long-range contextual cues in EEG-based imagined speech recognition. The HS-STDCN integrated feature learning from temporal and spatial information into a unified end-to-end model, utilizing hybrid-scale temporal and depthwise spatial convolution layers. By leveraging the spatial-temporal representation of the input EEG data, the model employed dilated convolution layers to learn discriminative features for accurate classification.

In this paper, we build upon our previous work by proposing novel approaches to improve

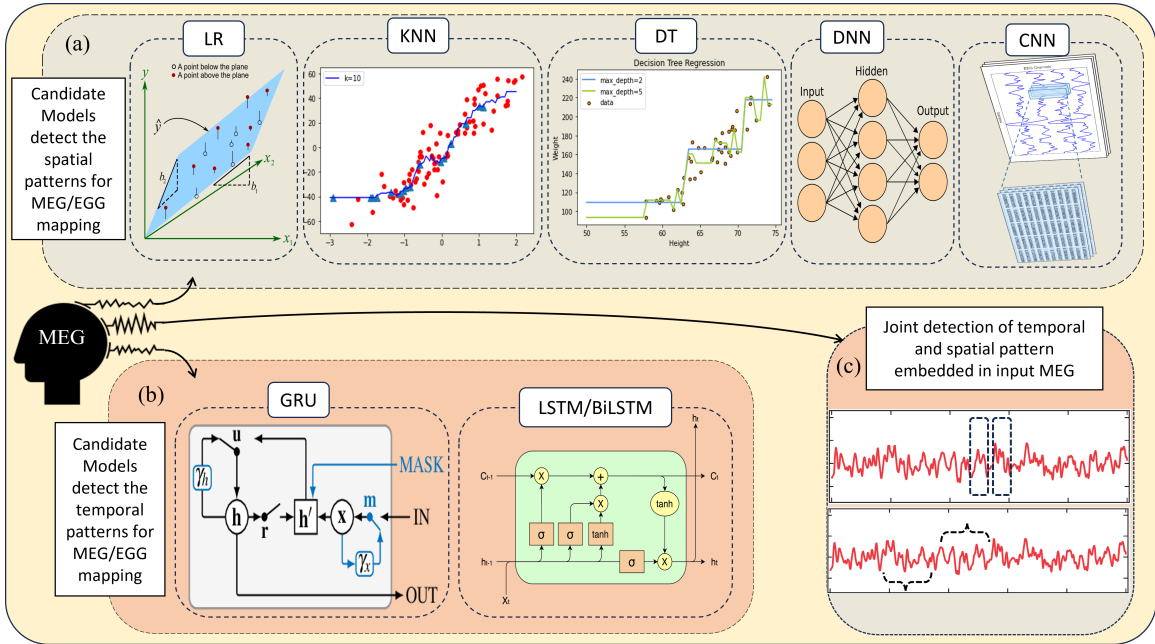


Figure 4.1: Comparison of machine learning and deep learning models for MEG to EEG mapping. (a) Spatial pattern models: Linear regression, KNN regressor, Decision tree regressor, DNN, and CNN, which focus on learning the spatial patterns of the signals. (b) Temporal pattern models: RNN-based models, including GRU, LSTM, and BiLSTM, are specifically designed to capture the temporal patterns in the signals. (c) Integration of spatial and temporal patterns: RCNN model seamlessly couples the temporal and spatial pattern detectors to enhance the mapping accuracy and highlight the importance of considering both dependencies in the MEG-EEG mapping

the accuracy and efficiency of MEG-EEG mapping. We address the limitations identified in our earlier studies and introduce new methodologies to achieve superior performance. We aim to provide a robust and reliable mapping framework for MEG and EEG signals with potential applications in neuroscience research, clinical diagnostics, and brain-computer interfaces by integrating advanced machine learning and deep learning techniques, preprocessing strategies, and post-processing refinements.

4.3 Considered System Design and Problem Description

The nature of MEG and EEG signals reveals a unique characteristic in sequential patterns, which holds the key to understanding the complex relationships between these modalities. In our study, we leverage this sequential nature by employing RNN-based models to capture the temporal dependencies in MEG and EEG signals [92]. By doing so, we aim to unlock new possibilities in neuroscience and brain-computer interfaces (BCIs).

Our proposed system design is driven by the objective of mapping MEG signals to their

corresponding EEG signals, with a focus on considering the temporal dependencies inherent in both modalities [93]. To evaluate the feasibility and effectiveness of this approach, we take a step-by-step process, gradually advancing from more straightforward machine learning techniques to more sophisticated deep learning methods.

Initially, we investigate the application of classical machine learning techniques, exploring their potential in mapping MEG and EEG signals. This initial exploration provides worthwhile insights and serves as a foundation for assessing the feasibility of the mapping problem. Encouraged by promising results and reasonable regression accuracy, we delve deeper into deep learning, specifically Recurrent Neural Networks (RNNs), as represented in fig 4.1.

The utilization of RNNs is driven by the distinctive nature of MEG and EEG signals, which exhibit temporal dependencies and complex patterns. RNNs, by their ability to recognize sequential information, are a natural fit for modelling and mapping these signals [94]. Our approach embraces this integration between the temporal nature of MEG and EEG signals and the power of RNNs, enabling us to explore and optimize our goals further.

Our ongoing efforts focus on optimizing and refining our solution for enriching the neuroscience field and improving patients' quality of life involves continuous discovery and optimization. We strive to integrate lightweight sensors, such as spintronic sensors, with the magic of deep learning techniques without putting computing and energy burdens on the MTJ-IoT system. This integration holds the potential to track and analyze brain activities in a more accessible and efficient manner, enabling daily monitoring and facilitating personalized interventions.

4.4 Envisioned RNN Models Temporal Pattern Detector for MEG-EEG Mapping

The following sections present a set of data-driven models we have considered for mapping MEG to EEG signals. These models encompass various RNN-based techniques such as GRU, LSTM, BiLSTM, CNN-RNN merging, and RCNN. We aim to capture the intricate connections between the MEG and EEG signals by utilizing these models. Subsequently, we discuss the post-processing steps involved in refining the mapped ECG signal to ensure satisfactory resolution for effective visualization and clinical interpretability.

4.4.1 Candidate RNN Models

4.4.1.1 GRU

The Gated Recurrent Unit (GRU) model has demonstrated remarkable potential in mapping magnetoencephalography (MEG) signals to corresponding electroencephalography (EEG)

signals, unravelling the intricate relationships between these two modalities. In our proposed GRU model, we incorporate two GRU layers with 128 and 64 units, respectively, as illustrated in Figure 4.3. This design allows the model to effectively capture the complex dynamics inherent in MEG and EEG signals, facilitating an accurate mapping process.

By leveraging the GRU layers in our architecture, we empower the model to retain critical information over longer sequences and effectively model the temporal dependencies within the MEG and EEG data [95]. These layers capture the nuanced dynamics and patterns, enabling an accurate and comprehensive mapping from MEG to EEG signals.

We introduce dropout layers with a dropout rate of 0.2 after each GRU layer to enhance the model’s generalization and mitigate overfitting. The incorporation of dropout regularization aids in mitigating the model’s reliance on specific features or patterns in the data, improving its robustness and enabling better generalization to unseen MEG and EEG samples. Adding dropout layers ensures the model captures MEG and EEG signals’ underlying patterns and dynamics more effectively.

Furthermore, the final layer of our GRU model is a dense layer comprising 59 units, representing the output dimensions of the EEG signals. This layer acts as a flexible mapping component, facilitating the transformation of the learned representations from previous layers into the target output space. By compiling the model with the mean squared error (MSE) loss function and optimizing it with the Adam optimizer, we aim to minimize the discrepancy between the predicted and actual EEG signals. Training the GRU model for 100 epochs with a batch size of 128 allows the model to converge to an accurate MEG-to-EEG mapping, enabling comprehensive insights and analysis of brain activity.

4.4.1.2 LSTM

To develop precise models for mapping MEG signals to corresponding EEG signals, we explore the potential of Long Short-Term Memory (LSTM) units, a specialized recurrent neural network (RNN) developed to capture and process temporal reliances in sequential data [96]. Our proposed model architecture incorporates two LSTM layers, each consisting of 128 and 64 units. These LSTM layers are vital components for modelling the hierarchical and temporal characteristics of the MEG-to-EEG mapping.

Our LSTM model’s input shape (samples/10, 10, 305) represents a sequence of 10-time steps with 305 channels corresponding to the MEG signal channels. By employing multiple LSTM layers, we aspire to capture the intricate dynamics and temporal patterns inherent in the MEG and EEG signals, encouraging an exact mapping process.

To ensure better generalization and mitigate overfitting, we introduce dropout layers with a rate of 0.2 after each LSTM layer. Dropout layers play a vital role in regularization by randomly deactivating a proportion of the neurons during training. This regularization technique reduces the model’s dependence on specific features or patterns in the data, en-

hancing its ability to generalize to unseen MEG and EEG samples. Including dropout layers in our LSTM model promote robustness and enables the model to capture the underlying patterns and dynamics in the MEG, and EEG signals more effectively.

The final layer of our LSTM model is a dense layer comprising 59 units, aligning with the output dimensions of the EEG signals. This dense layer is a flexible mapping component, enabling the model to accurately convert the learned representations from the preceding layers into the target output space.

Training the LSTM model involves utilizing the mean squared error (MSE) loss function, quantifying the discrepancy between the predicted and actual EEG signals. The Adam optimizer, known for its adaptive learning rate and efficient convergence properties, optimizes the model. Training the model over 100 epochs with a batch size of 128 facilitates effective learning from many MEG-EEG pairs, allowing the model to refine its mapping capabilities and converge to an accurate representation of the MEG-to-EEG mapping.

4.4.1.3 Bi-directional LSTM

In addition to the GRU and LSTM models, we explore the effectiveness of Bidirectional Long Short-Term Memory (BiLSTM) units in mapping MEG signals to corresponding EEG signals. Like the GRU and LSTM models, the BiLSTM model captures temporal dependencies and patterns inherent in the MEG and EEG data.

Our proposed BiLSTM model incorporates two BiLSTM layers with 128 units each. This architecture allows the model to capture both forward and backward temporal dependencies simultaneously, providing a comprehensive understanding of the relationship between MEG and EEG signals [97].

To enhance generalization and mitigate overfitting, dropout layers with a dropout rate 0.2 are introduced after each BiLSTM layer, similar to the GRU and LSTM models. The BiLSTM model shares the same evaluation metric and loss function as the GRU and LSTM models. We choose to Optimize the model with the Adam optimizer and train it for 100 epochs with a batch size of 128, aiming to converge to an accurate mapping from MEG to EEG signals.

Figure 4.3 illustrates the Comparison of the model architectures. The four subfigures, (a), (b), (c), and (d), represent the structures of the GRU, LSTM, BiLSTM, and RCNN models, respectively.

4.4.2 Proposed Recurrent Layer based Convolution: Fusion of spatial and temporal detectors for MEG-EEG mapping

After extensively exploring the application of machine and deep learning models that specifically focus on capturing spatial patterns and delving into models that effectively utilize tem-

poral sequences, we recognized the importance of integrating both patterns in the mapping task. By harnessing the strengths of both approaches, we aimed to develop a comprehensive model that leverages the power of convolutional and recurrent layers sequentially. Consequently, we constructed a deep learning model that, in series, combines convolutional and recurrent layers, enabling the extraction of spatial and temporal dependencies [98]. This adding approach facilitates a more comprehensive and accurate mapping process between MEG and EEG signals, taking advantage of the spatial and temporal aspects inherent in the data.

Inspired by the abovementioned motivation, we designed a novel deep learning model to effectively capture the complex connections between MEG and EEG signals. The model architecture combines convolutional and recurrent layers, allowing for the extraction of spatial and temporal dependencies.

The model begins with a reshape layer, which reshapes the input data to match the desired dimensions. This is followed by a series of convolutional layers, which apply convolutional operations across the temporal dimension of the data. The LeakyReLU activation function introduces non-linearity and enhances the model’s ability to capture complex patterns. Incorporating the LeakyReLU activation function is vital in precluding the dying ReLU problem and allows the network to continue learning even when some neurons become inactive. By introducing a slight negative slope of 0.2, LeakyReLU enables the model to capture and propagate gradients more effectively. This can lead to improved learning and convergence during the training step.

For the overfitting issue, Dropout layers are applied after each Conv1D layer, randomly deactivating a portion of the neurons during training. This regularization technique mitigates the risk of relying too heavily on specific features or patterns in the data.

After the convolutional layers, a Flatten layer converts the multidimensional output into a flat representation. This is followed by a series of recurrent layers, starting with a GRU layer and then an LSTM layer. These recurrent layers enable the model to capture the data’s long-term dependencies and temporal patterns.

To further mitigate overfitting, dropout layers are applied after each recurrent layer. Finally, a dense layer with the appropriate output channels is added to produce the predicted EEG signals. During training, the model is fed with the reshaped MEG and corresponding EEG data. The training continues for several epochs with a defined batch size. The model’s performance is evaluated using MSE loss metric on the entire dataset.

By utilizing this model architecture and training methodology, we aim to achieve an accurate and comprehensive mapping between MEG and EEG signals, capturing the spatial and temporal dependencies inherent in the data.

In our pursuit of a more comprehensive understanding and a deeper exploration of the relationships between spatial and temporal patterns, we initially combined the CNN and

RNN models in series. This allowed us to capture the individual strengths of each approach and obtain valuable insights into the data.

As our research progressed, we contemplated the potential of seamlessly integrating spatial and temporal information within a single layer. This realization motivated us to explore the concept of recurrent convolution layers (RCL) in the development of our RCNN model. The RCLs enable the integration of recurrent connections within a convolutional layer, allowing the network to evolve over time even when the input remains static. These connections ensure that each unit in the network is influenced by its neighboring units, facilitating the incorporation of contextual information. By embracing the RCL concept [99], we have achieved a robust architecture that effectively combines the strengths of both CNN and RNN models in a unified framework.

By leveraging the RCL, we were able to build a more advanced and sophisticated model that effectively harnesses the spatial and temporal information within the MEG and EEG signals. This integration not only improved the performance of our model but also provided a more detailed and accurate representation of the underlying brain activity.

Our RCNN model is designed to capture both spatial and temporal dependencies in the MEG and EEG signals. It utilizes a custom RCL, combining convolutional and recurrent operations within a single layer.

The RCNN model starts with an input layer of shape (4, 305), representing a sequence of 4-time steps with 305 channels corresponding to the MEG signal channels. The input data then goes through a series of operations.

First, a 1D convolutional layer with 128 filters and a kernel size of 11 is applied to extract spatial features from the input signals. The LeakyReLU activation function with an alpha value of 0.2 introduces the non-linearity. Batch normalization is performed to improve training stability, followed by dropout regularization with a rate of 0.2 to prevent overfitting. Next, the RCL is employed, which consists of a convolutional layer and a Simple RNN layer. This layer further captures temporal dependencies in the data. The convolutional layer with 64 filters and a kernel size of 5 performs convolutional operations, followed by the Simple RNN layer with the same number of filters. Again, batch normalization and dropout regularization are applied after each layer.

To model the temporal dependencies more comprehensively, a GRU layer with 128 units and the LeakyReLU activation function is employed. This layer captures the intricate dynamics and patterns inherent in the MEG and EEG signals, facilitating an accurate mapping process. Following the GRU layer, additional batch normalization and dropout regularization are applied to improve the model's generalization and robustness. The output is flattened and passed through a dense layer with 128 units and the LeakyReLU activation function. Batch normalization and dropout regularization are applied again. Finally, a dense layer with the output shape of (4*59) is employed to obtain the desired output shape

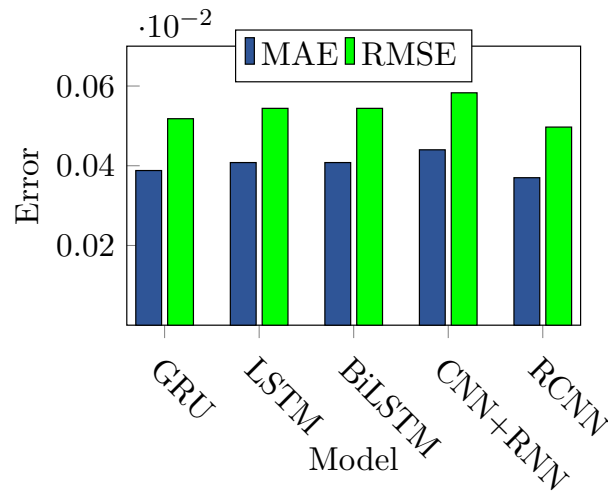


Figure 4.2: Performance comparison of MAE and RMSE for different models.

for the EEG signals. The output is then reshaped to (4, 59) to match the original time steps and EEG channel dimensions.

The RCNN model combines convolutional and recurrent operations within a unified architecture, effectively capturing both spatial and temporal dependencies in the MEG and EEG signals. By leveraging the capabilities of both operations, the model can achieve a more comprehensive understanding of the complex relationships within the data.

4.5 Data Preparation and Performance Evaluation

4.5.1 Data Preprocessing and Channel Selection

This section outlines our proposed MEG-EEG mapping technique, starting with the data preprocessing and channel selection process. We systematically investigate data-driven models for effective MEG-EEG mapping and provide post-processing steps for improved visualization and clinical interpretability.

Our study utilizes the "sample" dataset available in the MNE Python library. We examine the dataset's characteristics and make informed decisions regarding the channels to include in our analysis. Channels associated with ECG and EOG signals, deemed unsuitable for further analysis, are excluded. To isolate independent components related to specific brain activity patterns, we employ Independent Component Analysis (ICA) to separate mixed signals into their underlying sources [72].

After applying ICA, our focus shifts to artifact removal to enhance the quality of the data. We specifically target the removal of ECG and muscle artifacts by identifying and eliminating the corresponding components. As the dataset combines MEG and EEG signals,

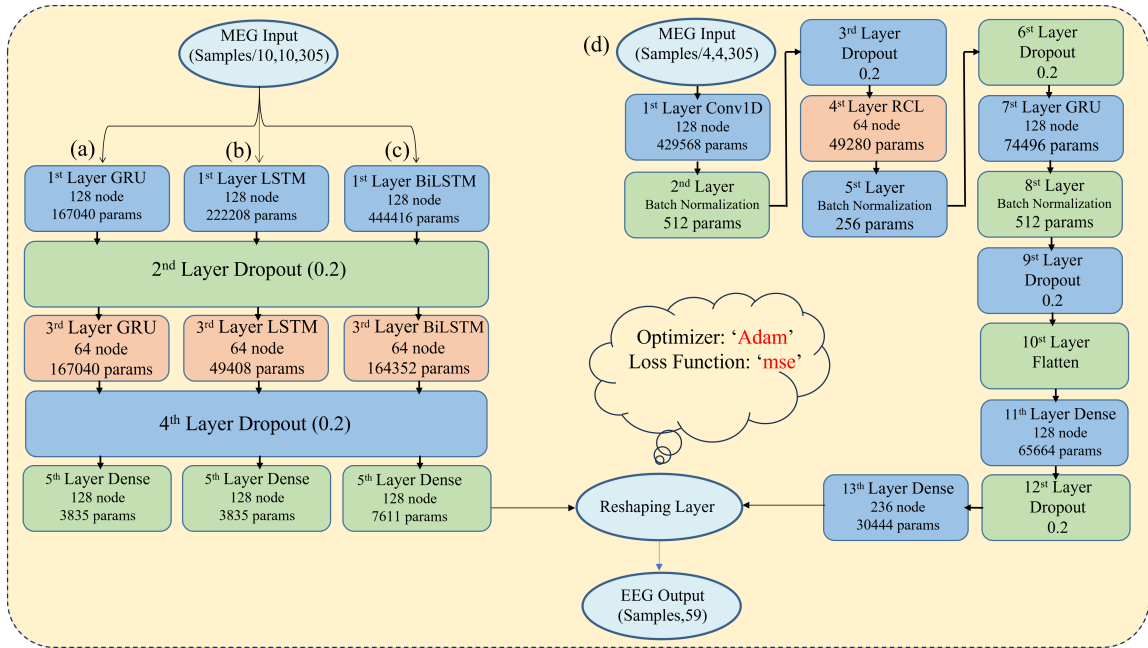


Figure 4.3: Comparison of the model architectures. The four subfigures, (a), (b), (c), and (d), represent the structures of the proposed GRU, LSTM, BiLSTM, and RCNN models, respectively.

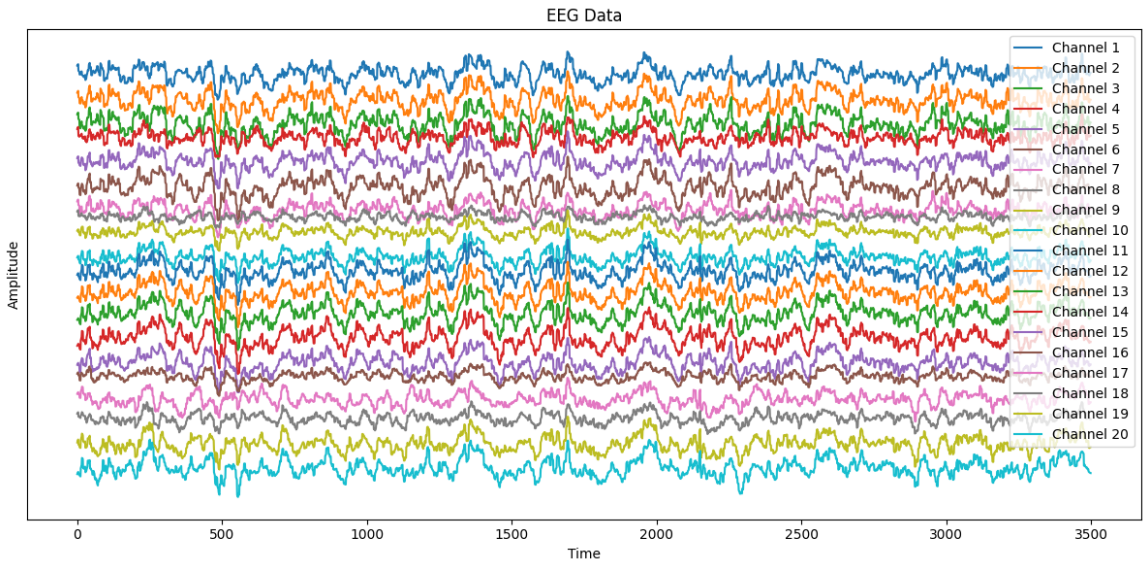


Figure 4.4: The first 20 predicted EEG channels based on the RCNN model

we partition the data into two files to facilitate the subsequent training process.

The raw data contains essential details, including the measurement date, MEG team information, and participant identity. It comprises various digitized points, including Gradiometers, Magnetometers, Stimulus channels, EEG channels, and an EOG channel. We

identify and note the presence of bad channels, such as MEG 2443 and EEG 053. Additionally, we determine EOG channel 061 as an EOG channel, while the ECG channels are unavailable. The sampling frequency is 600.61 Hz, and the data undergoes highpass and lowpass filtering at specific frequency ranges [73].

Following artifact removal, we apply bandpass filtering to the MEG and EEG signals. We utilize a bandpass filter ranging from 1 Hz to 100 Hz for the MEG data to retain relevant frequencies. Similarly, the EEG data undergoes bandpass filtering within the 0.5 Hz to 55 Hz range [75].

We divide the filtered data into training and testing sets to facilitate model training and evaluation, following a 75-25 split. With the preprocessed data in place, we explore and compare five candidate models designed explicitly for the MEG-EEG mapping task, as detailed in the following subsection.

4.5.2 Evaluation metrics

Before evaluating our models, we employed a moving average filter to enhance the signal quality and reduce high-frequency noise. The moving average filter takes the input signal and window size as parameters. It performs a moving average operation on the signal using a window consisting of ones divided by the window size. To smooth out high-frequency noise in the predicted signals, we utilized a window size of five. This process resulted in cleaner and more interpretable waveforms. By incorporating the moving average filter into the predicted EEG signals, we enhanced the overall signal quality and mitigated the impact of short-term fluctuations. Consequently, the interpretability and visual clarity of the predicted EEG waveforms were improved [82]. This postprocessing step was crucial in preparing the signals for further investigation and evaluation of their performance.

In evaluating regression models, the Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) serve as valuable metrics for assessing performance. RMSE emphasizes significant errors more by squaring the differences between predicted and actual values without considering their direction, while MAE treats all errors equally. RMSE estimates the standard deviation of prediction errors, while MAE represents the average magnitude of errors [100]. The equations for RMSE and MAE are calculated as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\text{predicted}_i - \text{true}_i)^2} \quad (4.1)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\text{predicted}_i - \text{true}_i| \quad (4.2)$$

where n represents the number of the data points or samples used for evaluation.

These metrics quantitatively estimate the accuracy and reliability of predictions, aiding

in selecting the optimal solution for real-world applications. A lower RMSE or MAE value indicates improved accuracy, with values closer to zero indicating higher precision. These metrics offer valuable insights into model performance by assessing the magnitude and direction of errors.

In addition to RMSE and MAE, prediction time is an essential metric for evaluation. It measures the computational time required for the model to generate predictions. Minimizing prediction time is crucial in real-world applications, ensuring efficient and timely data processing.

By considering RMSE, MAE, and prediction time as evaluation metrics, we comprehensively understand the model’s accuracy, precision, and efficiency. These metrics are pivotal in selecting the optimal solution that aligns with the requirements of practical use cases.

The table 6.1 presents the performance metrics and prediction times for different models. The findings show that the GRU model achieved an MAE of 0.0388 and an RMSE of 0.0518, with a prediction time of 1.9108 seconds. The LSTM model yielded an MAE of 0.0408 and an RMSE of 0.0544, with a prediction time of 2.3665 seconds. Similarly, the BiLSTM model also achieved an MAE of 0.0408 and an RMSE of 0.0544, with a slightly higher prediction time of 2.4893 seconds. In comparison, the CNN+RNN model resulted in an MAE of 0.0440 and an RMSE of 0.0583, with a prediction time of 4.5880 seconds. Finally, the RCNN model exhibited the lowest MAE of 0.0370 and the lowest RMSE of 0.0497 but had a longer prediction time of 6.0500 seconds.

These findings indicate that the RCNN model outperformed the other models regarding MAE and RMSE, achieving the lowest error values as shown in fig 4.2. However, it also had the longest prediction time among the models. The GRU model showed the best balance between performance and prediction time, with relatively low errors and a faster prediction time than the other models. Moreover, it is worth noting that the models achieved convergence at different rates. The GRU model required 77 iterations to converge, while the LSTM model converged in 68 iterations. Notably, the BiLSTM model exhibited faster convergence, reaching convergence within 45 iterations when employing the early stop method with a patience of 10.

Overall, the results demonstrate the effectiveness of the proposed models for the given task, with varying trade-offs between accuracy and prediction time. Researchers can use these findings to select the optimal model that best suits their requirements regarding accuracy and efficiency.

4.6 Conclusion

In conclusion, our research addressed the challenges associated with the high cost, limited accessibility, invasiveness, and complexity of traditional MEG and EEG modalities for di-

Table 4.1: Performance metrics and prediction times for different models.

Model	MAE	RMSE	Prediction Time (s)
GRU	0.0388	0.0518	1.9108
LSTM	0.0408	0.0544	2.3665
BiLSTM	0.0408	0.0544	2.4893
CNN+RNN	0.0440	0.0583	4.5880
RCNN	0.0370	0.0497	6.0500

agnosing various diseases and abnormal brain activities. To overcome these challenges, we proposed an innovative approach using MTJ-based sensors and leveraging the power of deep learning techniques in a way to develop user-friendly and daily-based monitoring systems.

We focused on extracting and utilizing the temporal nature of MEG and EEG signals to create robust models capable of accurately mapping brain activity. We explored several RNN-based models, including GRU, LSTM, and biLSTM. Additionally, we investigated models that seamlessly integrated RNN and CNN to capture both spatial and temporal patterns within the signals.

After thorough experimentation, we discovered that the RCNN model demonstrated superior performance compared to all other models. It excelled in accurately capturing and analyzing the complex patterns present in the signals. However, considering efficiency and time consumption, the GRU model emerged as the optimal choice. It balanced accuracy and computational efficiency, making it a valuable solution for real-world applications.

By successfully developing and evaluating these models, we have paved the way for more accessible, cost-effective, and non-intrusive monitoring of brain activity. This research contributes to advancing the field of neural diagnostics and lays the foundation for future advancements in brain-computer interfaces and daily monitoring systems.

Chapter 5

De-noising Noisy Signals in MTJ Systems Via AI-assisted MEG-EEG Mapping

5.1	Introduction	55
5.2	Related work	57
5.3	MTJ Sensors and Noise Characteristics in MEG Signal Acquisition	58
5.4	Performance evaluation	59
	5.4.0.1 Noise simulation	59
	5.4.0.2 Findings and Analysis	61
5.5	Conclusion	63

5.1 Introduction

Brain activities play a pivotal role in disease diagnosis and overall healthcare management. Consequently, ongoing advancements in brain signal capture are crucial and valuable. While existing modalities like Magnetoencephalography (MEG) and Electroencephalography (EEG) have proven effective in capturing brain signals, addressing challenges such as portability, complexity, and high costs remains a prominent research focus. To overcome these limitations, researchers have been exploring innovative approaches, and one promising attempt involves harnessing the power of spintronic devices, specifically Magnetoresistive Tunnel Junction (MTJ) sensors, for capturing the intricate magnetic fields associated with brain activity.

This cutting-edge sensor demonstrates significant promise in detecting magnetic field signals. Its compact form allows for seamless integration on-chip with silicon logic circuits, facilitating the development of portable lab/logic-in-sensor diagnostic devices. By its logic-in-sensor architecture, the sensor effectively addresses critical obstacles, encompassing communication delay, overhead, and power consumption. As a result, this groundbreaking technology enables prolonged and accurate monitoring of a patient’s brain activities while allowing them to maintain their daily activities undisturbed.

Although MTJ sensors offer numerous advantages, they also introduce multiple types of noise into the captured signals. Separating and removing this noise with traditional signal processing methods becomes exceptionally challenging, particularly when the noise band overlaps the signal frequency range. This difficulty hinders the accurate extraction of the target brain signals from the noise, necessitating novel solutions to enhance brain signal fidelity.

Fortunately, the remarkable advancements in AI, deep learning, and machine learning have opened doors to transformative possibilities in this realm. This chapter aims to contribute to biomedical signal processing and neuroscience to better understand the relationship between the MEG and EEG by introducing promising machine and deep learning models. Specifically, we aim to develop models capable of accurately mapping noisy MEG signals to their corresponding denoised EEG signals. By leveraging the potential of AI, we envision a new era of brain activity monitoring that can revolutionize healthcare applications.

The core focus of this research is to explore and evaluate various AI-based models, including Recurrent Neural Networks (RNN), Convolutional Neural Networks (CNN), Deep Neural Networks (DNN), k-Nearest Neighbors (KNN) regressor, Linear Regression (LR), and Decision Trees (DT). We aim to optimize these models for the task of MEG signal denoising, thereby enhancing the accuracy and reliability of brain activity measurements.

Through our efforts, we aim to develop a portable MEG device, bridging the gap between the potential of spintronic sensors and the effective denoising capabilities of AI-based models. Ultimately, this work strives to foster significant advancements in biomedical signal processing, opening new avenues for brain activity monitoring and healthcare diagnostics.

The remainder of the chapter is organized as follows. The related research work is surveyed in Section 5.2. Then, we discuss the dimensions of the problem and the design considerations 5.3. Next, the solution of the mapping problem is presented in Section 4.4 followed by the performance evaluation of our proposal in Section 5.4 Finally, we provide concluding remarks in Section 6.6.

5.2 Related work

Several noteworthy studies have been conducted in the context of denoising biomedical signals, including the work by [68]. In their research, the authors focused on denoising noisy MCG (Magnetocardiography) signals captured by MTJ sensors and achieved precise extraction of pure MCG signals. They further correlated these denoised MCG signals with corresponding Electrocardiography (ECG) signals.

To accomplish this task, the authors employed a combination of convolutional signal processing techniques and AI-based models for comparison. Notably, their findings revealed the superiority of the proposed deep learning approach, incorporating Convolutional Layers, Gated Recurrent Unit (GRU) Layer, and Fully Connected Layer, over the conventional moving average technique. This significant outcome underscores the remarkable power of AI in denoising MEC signals, surpassing classical methods.

The susceptibility of Electroencephalography (EEG) signals to noise, mainly when recorded from scalp sensors, has motivated researchers In [101] to explore practical denoising approaches. The authors proposed a novel denoising method based on deep learning, employing a deep convolutional autoencoder to enhance brain-computer interface diagnosis and communication. While numerous EEG denoising methods have been suggested, many of these algorithms tend to be complex.

To evaluate the performance of their proposed approach, the researchers conducted experiments using two types of noise, eye blinks and jaw clenching. Performance assessment was accomplished using the peak signal-to-noise ratio (PSNR). The results unequivocally showcased the superiority of their deep learning-based approach over the conventional baseline bandpass filtering method. Remarkably, the confidence intervals consistently favoured their approach, demonstrating its effectiveness in reducing noise and enhancing the quality of EEG signals.

The growing importance of incorporating deep learning techniques into signal processing has been evident in recent research, exemplified by the study conducted by [102]. Traditional active noise control (ANC) methods, utilizing adaptive signal processing with the least mean square algorithm, are inherently limited in handling nonlinear distortions due to their linear nature. To address this challenge, this paper takes a new approach, treating ANC as a supervised learning problem, and introduces an innovative deep learning method known as "deep ANC."

The core principle underlying deep ANC revolves around harnessing the strength of deep learning to encode optimal control parameters tailored to various noise types and environmental conditions. Through the training of a convolutional recurrent network (CRN) to estimate the real and imaginary spectrograms of the cancelling signal from the reference signal, the deep ANC technique enables effective elimination or attenuation of primary noise

in the ANC system, using the corresponding anti-noise. Ensuring robustness and versatility, large-scale multi-condition training is employed to facilitate excellent generalization across diverse noise scenarios.

The deep ANC method proves its adaptability in active noise cancellation, regardless of whether the reference signal contains noise or noisy speech. Moreover, a delay-compensated strategy is introduced to mitigate potential latency challenges in ANC systems. Empirical evidence from experiments highlights the efficacy of deep ANC in achieving wideband noise reduction, impressively demonstrating its ability to generalize even to untrained noise types. Furthermore, the proposed approach effectively establishes ANC within a quiet zone and exhibits resilience in the face of varying reference signals.

5.3 MTJ Sensors and Noise Characteristics in MEG Signal Acquisition

In this section, we focus on providing an in-depth exploration of the Magnetic Tunnel Junction (MTJ) sensors used in capturing MEG signals. We discuss their fundamental principles and functionalities, emphasizing their advantages and relevance in the context of brain activity monitoring. Additionally, we delve into the various types of noise inherent to MTJ sensors and their potential impact on the acquired MEG signals.

MTJ sensors consist of two ferromagnetic metals (FMs) separated by an insulating tunnelling barrier, typically made of materials like magnesium oxide [67]. By applying an external magnetic field (H), the magnetization angles of the FMs change, resulting in a linear dependence of MTJ resistance on H . This fascinating property, known as the tunnelling magnetoresistance effect (TMR), enables straightforward measurement of MTJ sensors and seamless integration with the fabrication process of integrated circuits. Consequently, MTJ sensors and structures have become instrumental in driving advancements in spintronics research and finding innovative applications in information storage, as extensively reviewed in [103].

One of the primary obstacles sensors face is noise at the lower frequency range of the spectrum. In MTJ sensors, this noise exhibits a $1/f$ characteristic akin to numerous other systems [104]. This problem intensifies in the high-sensitivity region, as indicated by [105]. The relationship can describe the power spectral density (PSD) of low-frequency noise [106]:

$$S_v \propto \frac{\chi}{M_s V} \frac{1}{f^\beta} \quad (5.1)$$

In this context, the variables χ , M_s , V , f , and β represent specific properties of the sensor. χ relates to the sensor susceptibility, which is directly proportional to its sensitivity; M_s denotes the sensor's saturation magnetization; V represents the sensor volume; f

indicates the spectral frequency; and β corresponds to the exponent of the noise spectrum.

A critical challenge arises in this context when certain targeted brain activities and sensor-related noise oscillate within the same frequency band. This overlapping phenomenon makes the separation of these components notably tricky. Notably, research by [107] emphasizes the valuable information embedded in the low-frequency band of MEG signals. Understanding the dynamics of the brain, as well as detecting and characterizing various diseases, heavily relies on this low-frequency band. Fig 5.1 demonstrates the frequency bands of Alpha, Gamma, Beta, Theta, and Delta. The analysis provides valuable insights into the distribution of power in these frequency bands, shedding light on the characteristics of our clean MEG signals, and emphasizing the importance of the low-frequency levels.

In this research, our primary objective is to develop a practical and lightweight solution to address the challenge posed by $1/f$ noise in MEG signals while accurately providing the corresponding EEG signals. To achieve this, we propose integrating various deep learning and machine learning techniques, which hold significant promise in effectively reducing noise and enhancing the accuracy of denoising processes. By leveraging these advanced techniques, we aim to pave the way for a robust and efficient approach to separate the targeted brain activities from the interfering noise, facilitating a more comprehensive understanding of brain dynamics and enabling better disease detection and characterization.

5.4 Performance evaluation

This section serves a dual purpose, encompassing both the simulation of noisy MEG signals and the performance evaluation of AI-based models used for denoising. The simulation involves generating a dataset that faithfully emulates the noise characteristics associated with MTJ sensors. This realistic and diverse dataset plays a pivotal role in effectively training and evaluating denoising models.

By combining simulation and performance evaluation, we strive to push the boundaries of denoising capabilities and enhance the accuracy of brain signal restoration. The simulation process ensures that the denoising models are exposed to a wide range of noise scenarios, closely mirroring real-world conditions. Subsequently, the performance evaluation of these AI-based models provides valuable insights into their effectiveness, enabling us to identify the most promising candidates for practical applications.

5.4.0.1 Noise simulation

The simulation approach in this study ingeniously leverages pink noise, an essential component in accurately emulating the noise patterns observed in real-world MEG recordings captured by Magnetoencephalography with MTJ sensors. Pink noise is characterized by a $1/f$ characteristic, where its power decreases at a rate inversely proportional to frequency.

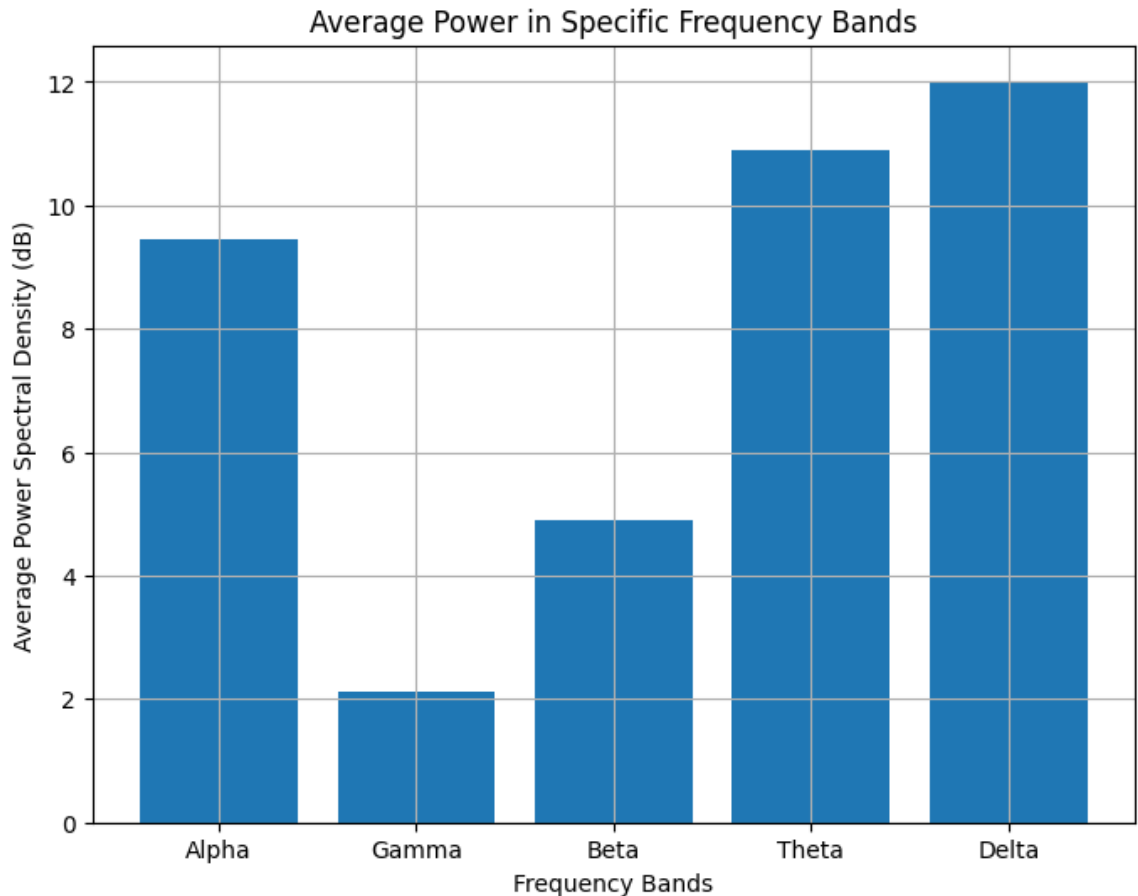


Figure 5.1: Bar plot showing the average power levels (in dB) across different frequency bands of interest, namely Alpha, Gamma, Beta, Theta, and Delta in our clean MEG signal

This unique property closely resembles the noise patterns commonly found in MEG signals, making it a suitable choice for simulating noise in our dataset.

The carefully chosen simulation parameters effectively create a dataset that precisely mimics the noise associated with MTJ sensors. We set the sampling frequency to 600 Hz, ensuring sufficient temporal resolution to capture brain activities accurately [108]. The power spectral density is set at 1×10^{-4} , representing the baseline white noise power spectral density. White noise, random noise with equal power across all frequencies, forms an integral part of the simulated dataset.

The simulated pink noise exhibits a characteristic power distribution governed by an exponent (β) of 1.5, resulting in the $1/f^\beta$ behaviour. This careful choice of β imparts the distinctive property to the pink noise, ensuring it closely resembles the observed noise patterns in MEG signals. The value of β plays a crucial role in controlling the slope of the pink noise power spectrum, determining the rate at which its power decreases with increasing frequency [109].

Furthermore, to accurately emulate the noise observed in MEG signals captured by Magnetoencephalography with MTJ sensors, a cutoff frequency of 10 Hz is introduced. This threshold effectively separates the $1/f^\beta$ region from the flat line region, where the pink noise dominates. As a result, the pink noise takes precedence in the lower frequency range, reflecting the prominent noise behaviour in this band [110]. Conversely, the white noise component becomes more pronounced at higher frequencies, as the effects of MTJ sensor noise predominantly impact the low-frequency band.

The precisely generated pink and baseline white noise is integrated with the clean MEG signals using Fourier transform techniques. This integration process ensures that the simulated pink and white noise accurately resembles the noise acquired by MTJ sensors during actual data recording. The result is a comprehensive dataset containing noisy MEG signals that realistically simulate the impact of MTJ sensor noise on signal quality.

By harnessing this diverse and realistically simulated dataset, various denoising models can be thoroughly trained and rigorously evaluated. The primary objective is to achieve high accuracy in effectively separating the targeted brain activities from the noise and accurately predicting the corresponding EEG signals. This novel approach aims to enhance our understanding of brain dynamics and significantly improve disease detection and characterization through more accurate brain signal analysis.

The successful combination of pink and white noise simulation, alongside advanced AI-based denoising techniques, opens new avenues for neuroscience. It brings us closer to realizing highly effective and non-invasive brain activity monitoring, which holds immense potential to advance our knowledge of brain functions and improve healthcare outcomes.

With robust denoising models and a reliable simulated dataset, this research paves the way for transformative applications in brain signal processing and healthcare diagnostics. The integration of pink noise simulation, white noise representation, and advanced AI methodologies brings transformative potential to the field of neuroscience. It significantly contributes to improving healthcare diagnostics and understanding the complexities of brain activity.

5.4.0.2 Findings and Analysis

In this section, we present the findings and analysis of the denoising models applied to the EEG signals. Table 6.1 showcases the performance metrics, including Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Prediction Time, for each model. We evaluated nine models, namely Linear Regression (LR), K-Nearest Neighbors (KNN) regressor, Deep Neural Network (DNN), Convolutional Neural Network (CNN), Gated Recurrent Unit (GRU), Long Short-Term Memory (LSTM), Bidirectional LSTM (BiLSTM), CNN combined with RNN (CNN+RNN), and Recurrent Convolutional Neural Network (RCNN). Figure 6.2 shows the performance of each model compared to the actual signal in

the range of 20-500 samples.

The table shows that the LR model achieves the lowest MAE of 0.0373, indicating its ability to approximate the clean EEG signals accurately. However, it is noteworthy that the LR model's RMSE is slightly higher at 0.0459 but still the best among the candidate models. On the other hand, the KNN model shows the highest MAE of 0.0737, indicating relatively more significant signal approximation errors than other models. The KNN model also exhibits the highest RMSE of 0.0923, displaying the enormous discrepancy from the actual signals.

Regarding computation efficiency, the LR model has the lowest prediction time of 0.126 seconds, making it the fastest model among all the models evaluated. Conversely, the KNN model requires significantly more time, with a prediction time of 30.973 seconds.

To gain deeper insights into the models' performance, we generated a bar plot in Figure 5.2 comparing the RMSE and MAE values across the different models. From the plot, we can visually compare the error levels of each model for better model selection. The results demonstrate that the BiLSTM and RCNN models are coming after LR in terms of both RMSE and MAE, offering a balance between accuracy and efficiency.

Moreover, fig 5.3 visually compares the Power Spectral Density (PSD) between the prediction models and the clean EEG signals. The PSD is calculated as the average of all channels, and the plot provides valuable insights into the distribution of power in different frequency bands, shedding light on the performance of the models, especially in the low-frequency band.

Besides, we present Figure 5.5, which visually compares the predicted signals generated by various proposed AI-based denoising models for Channel 40 after applying a moving average filter with window size 5 to smooth the curves. The models are arranged in descending order based on their cumulative performance, with the best model placed at the bottom and the worst at the top. The clean EEG signal is also included for reference, with the time range set from sample 20 to 700. The ordered arrangement of the models in Figure 5.5 clearly depicts their relative accuracy in capturing the underlying EEG dynamics. The gradual decrease in denoising errors from the bottom to the top demonstrates the effectiveness of each model, with the best models closely resembling the clean EEG signals.

These findings highlight the potential of the LR, BiLSTM and RCNN models in denoising EEG signals effectively and efficiently, making them promising candidates for further investigation and application in real-world scenarios. The insights gained from this study contribute to our understanding of the performance and suitability of various models for EEG signal denoising, facilitating future advancements in brain signal analysis and related fields.

Table 5.1: Performance metrics and prediction times for different models.

Model	MAE	RMSE	Prediction Time (s)
LR	0.0373	0.0459	0.126
KNN	0.0737	0.0923	30.973
DNN	0.0483	0.0588	2.564
CNN	0.0521	0.0633	5.384
GRU	0.0450	0.0551	2.085
LSTM	0.0476	0.0568	1.532
BiLSTM	0.0399	0.0488	1.143
CNN+RNN	0.0444	0.0550	3.308
RCNN	0.0419	0.0497	1.452

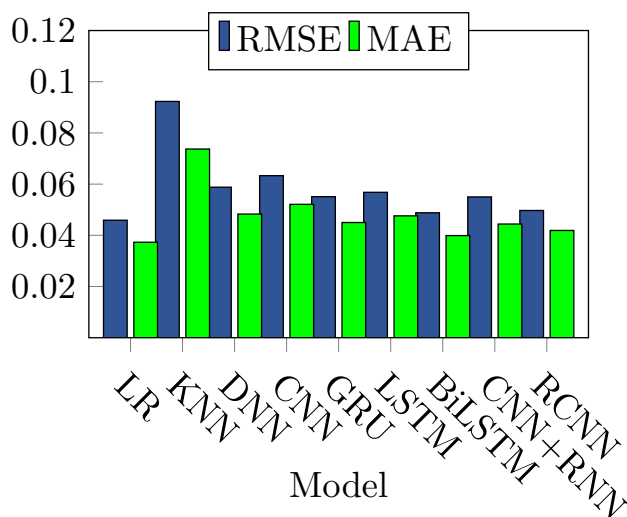


Figure 5.2: Performance comparison of RMSE and MAE for different models.

5.5 Conclusion

This chapter embarks on a trailblazing endeavour to confront the formidable challenges of denoising MEG signals captured by MTJ sensors. Armed with the power of AI-based models, we transcend the boundaries of brain signal fidelity, heralding a transformative era in healthcare diagnostics.

We underscore the profound importance of brain activity monitoring while exposing the limitations of traditional EEG and MEG methods. With stubborn determination, we venture into the realm of MTJ sensors, unearthing their potential to conquer portability, complexity, and cost challenges while capturing the enigmatic magnetic fields of brain ac-

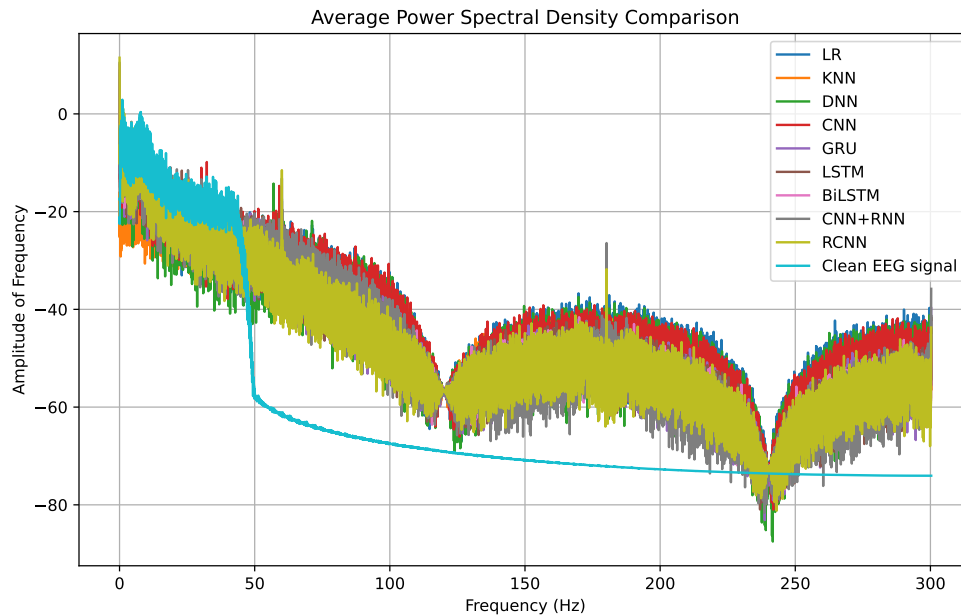


Figure 5.3: Comparison between the Power Spectral Density (PSD) of Prediction Models and the PSD of Clean EEG Signals. The PSD is calculated as the average of all channels.

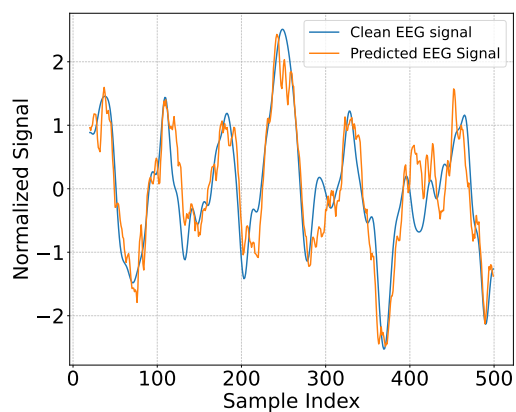
tivity. Yet, within this promise lies the irritating presence of various noise types, impeding our path to accurate signal analysis.

Through an interdisciplinary exploration, we harness the collective might of AI, deep learning, and machine learning to forge a cohort of innovative denoising models. RNN, CNN, DNN, KNN regressor, and LR models converge, unified in their mission to liberate MEG signals from noise shackles.

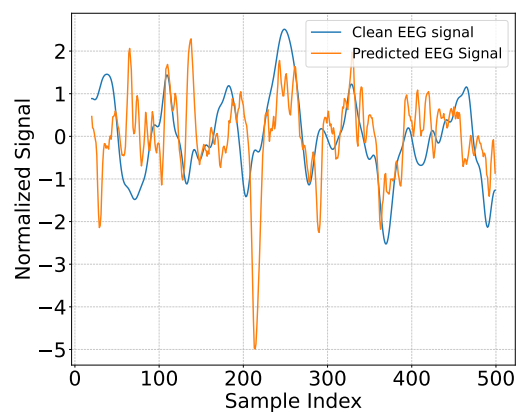
The stage is set, and the results unveil the triumph of our models. The LR model emerges as a beacon of promise, conquering the realm of approximation with the lowest MAE and unyielding RMSE values, hinting at the nature of the linear relationship between MEG and EEG. The BiLSTM and RCNN models join this vanguard, striking an exquisite balance between unwavering accuracy and unrivalled efficiency.

The symphony of success resonates as we unlock the latent potential of spintronic sensors while the cadence of AI-based denoising resounds through the halls of biomedical signal processing. The future unfurls before us, where portable MEG devices unfettered by noise empower us to unravel the mysteries of the mind.

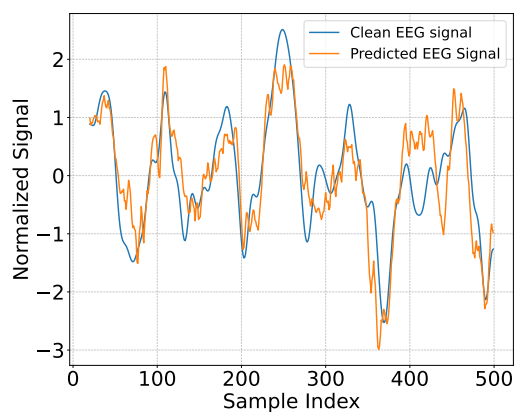
This work is poised to be a step in shaping the future of portable MEG devices and biomedical signal processing by unlocking the potential of spintronic sensors and harnessing the denoising capabilities of AI-based models. Integrating cutting-edge technology with neuroscience holds immense promise for transformative healthcare applications.



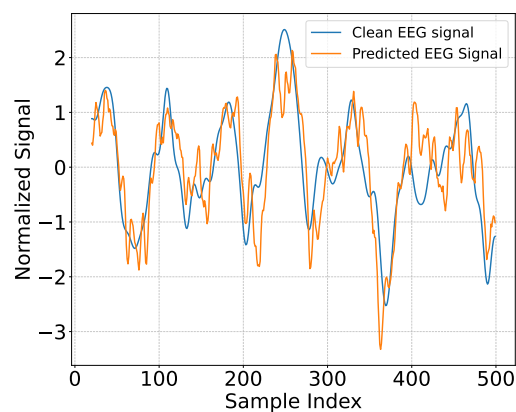
(a) Actual vs. predicted signal performance with the LR model for channel 40.



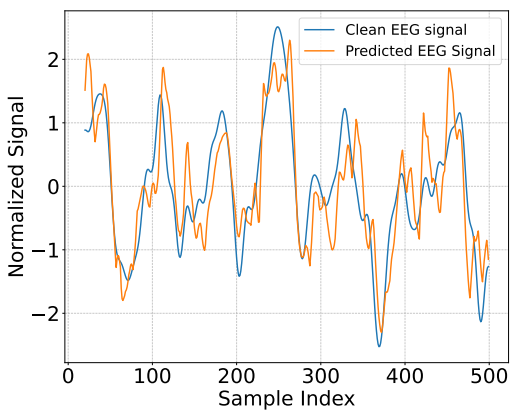
(b) Clean EEG vs. predicted signal performance with the KNN model for channel 40.



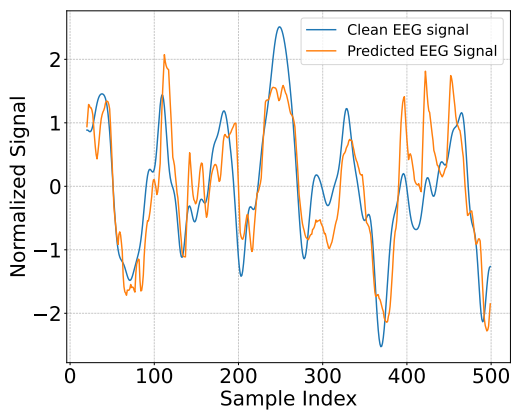
(c) Clean EEG vs. predicted signal performance with the DNN model for channel 40.



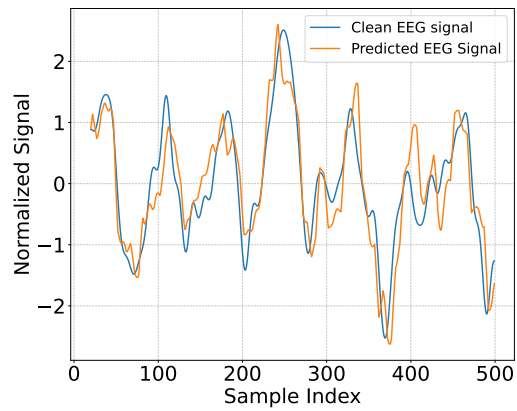
(d) Clean EEG vs. predicted signal performance with the CNN model for channel 40.



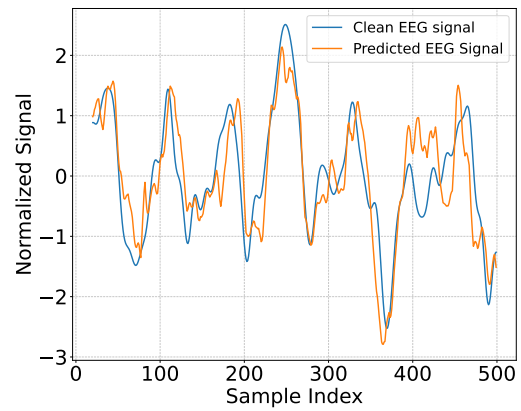
(e) Clean EEG vs. predicted signal performance with the LSTM model for channel 40.



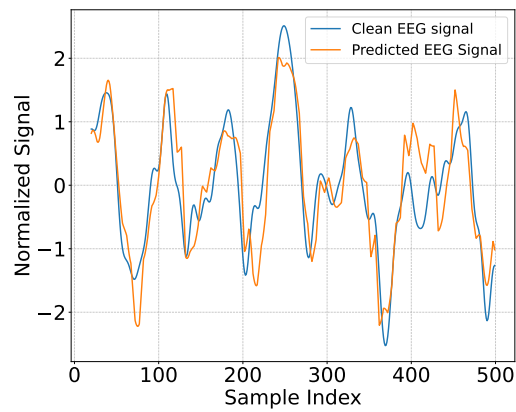
(f) Clean EEG vs. predicted signal performance with the GRU model for channel 40.



(g) Clean EEG vs. predicted signal performance with the BiLSTM model for channel 40.



(h) Clean EEG vs. predicted signal performance with the RNN&CNN model for channel 40.



(i) Clean EEG vs. predicted signal performance with the RCNN model for channel 40.

Figure 5.4: Filtered signals of channel 40 within the time range of 20-500 seconds: a comparison between clean and denoised predicted signals using various models.

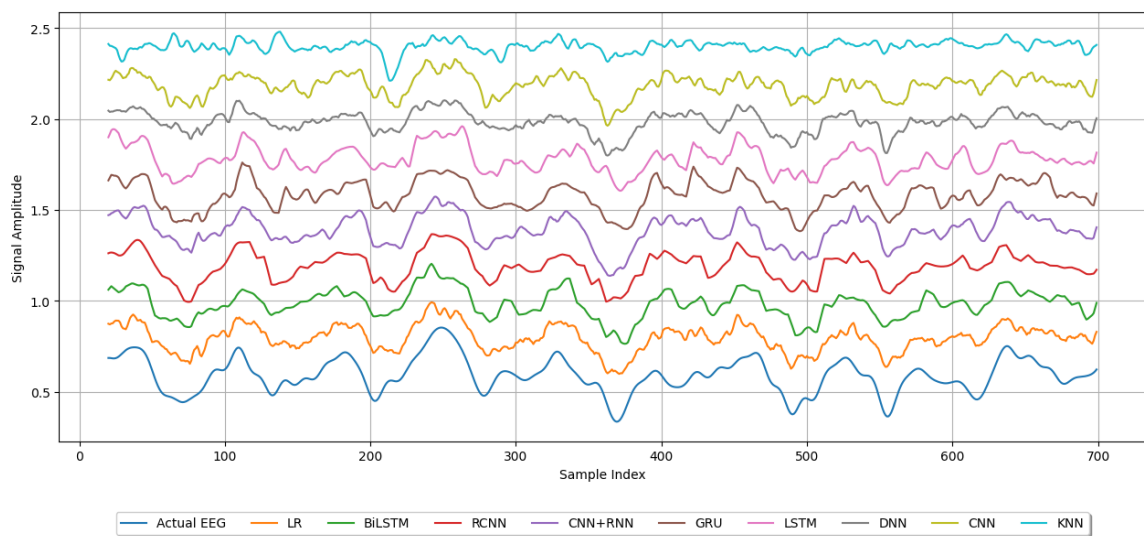


Figure 5.5: Visual comparison of the predicted signals generated by various proposed AI-based denoising models for Channel 40, showcasing their cumulative performance from the best at the bottom to the worst at the top. The clean EEG signal is also included for reference, with the time range set from sample 20 to 700. The ordered arrangement of the models offers a clear depiction of their relative accuracy in capturing the underlying EEG dynamics.

Chapter 6

Quantizing BiLSTM for Efficient Neuromorphic Hardware Deployment: An Exploration in Mapping MEG to EEG Signals

6.1	Introduction	69
6.2	Integration of Model Choices, Neuromorphic Principles, and Quantization Techniques	70
6.2.1	From Linear Regression to Deep Learning: A Deliberate Choice . . .	70
6.2.2	Neuromorphic Computing	71
6.2.3	Quantization of BiLSTM	72
6.2.3.1	Reduced Memory Footprint	72
6.2.3.2	Enhanced Processing Speed	73
6.2.3.3	Energy Efficiency	73
6.3	Related work	74
6.4	Adaptive Precision Quantization for BiLSTM Networks	75
6.5	Performance Evaluation Through Hyperparameter Variation	77
6.6	Conclusion	78

6.1 Introduction

In recent years, the fusion of deep learning techniques and neuroscience has opened new avenues for analyzing brain signals, particularly in Magnetoencephalography (MEG) and Electroencephalography (EEG). These non-invasive neuroimaging methodologies offer unprecedented perspectives into brain activities and are indispensable in advancing our comprehension of cognitive processes and neurological disorders. The efficacy of these techniques hinges on precise signal processing and formidable predictive models capable of decoding the intricate temporal and spatial intricacies inherent in brain signals.

Our preceding endeavours tackled pivotal challenges in MEG devices, pioneering the use of home-temperature sensors like the Magnetic Tunnel Junction (MTJ) as a robust alternative [111]. Our initial exploration delved into the transformation from MEG to EEG signals harnessing the prowess of artificial intelligence. Subsequently, we embraced the potency of Recurrent Neural Network (RNN) models, particularly the BiLSTM architecture, achieving commendable results in processing sequential data. Furthering our research, we examined the noise implications in MEG measurements, introducing state-of-the-art AI methodologies to fortify signal quality.

Building on this foundation, this chapter underscores the significance of adapting deep learning models, especially the BiLSTM, for seamless deployment on neuromorphic hardware platforms. We emphasize the pivotal role of the Multiply-Accumulate-Activate-Pool (MAAP) set, ingeniously tailored for efficient operational procedures, resonating with the paradigms of neuromorphic computing, renowned for energy conservation and parallel processing [112].

In this contemporary edge computing epoch, accentuating hardware implementation is imperative. The proficient adaptation of deep learning models, primarily through quantization, onto platforms such as Field-Programmable Gate Arrays (FPGAs) signifies monumental strides in real-time analysis, energy efficiency, and proximate AI data processing [113]. Consequently, this manuscript delineates the comprehensive methodology embraced in quantizing the BiLSTM model for optimal hardware compatibility. By demonstrating the practicability and merits of this approach, we envision a robust framework for IoT integration and energy-conscious MEG signal interpretation.

The chapter unfolds as follows: Section 6.2 delves into the seminal works pertinent to neuromorphic hardware implementations and BiLSTM. Section 6.3 provides an overview of existing methodologies and their relevance. Section 6.4 elucidates the nuances of the model optimization and quantization for the BiLSTM network. Section 6.5 offers a critical analysis of our results and discusses the effects of varying hyperparameters. Finally, Section 6.6 encapsulates the chapter, summarizing our contributions and outlining future research directions. At its core, this study aspires to seamlessly intertwine deep learning

advancements with pragmatic hardware solutions, heralding a renaissance in brain signal analytics and its transformative applications across neuroscience, healthcare, and beyond.

6.2 Integration of Model Choices, Neuromorphic Principles, and Quantization Techniques

This section comprehensively explores model choices, delving into our decision to transition from Linear Regression to the BiLSTM model. We then traverse the intricate landscape of quantization, elucidating its methodologies, significance, particularly for BiLSTM networks, and the inherent trade-offs and mitigations. As we progress, the alignment of quantization with neuromorphic computing principles is unveiled, spotlighting its pivotal role. Collectively, these discussions illuminate our strategic roadmap toward actualizing AI-driven brain signal analysis for real-time, real-world applications.

6.2.1 From Linear Regression to Deep Learning: A Deliberate Choice

In our comprehensive investigation into mapping MEG to EEG, we applied traditional machine learning and avant-garde deep learning models to noisy MEG data. The landscape of our findings from this exploration was both enlightening and instrumental in shaping our future endeavours. In our third major study, the humble Linear Regression (LR) model emerged as a dark horse, outshining many sophisticated models in terms of Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and prediction time. The Bi-directional Long Short-Term Memory (BiLSTM) network, though slightly trailing the LR model in RMSE and MAE metrics, was a close second but required marginally more time for prediction.

At a cursory glance, due to its superior performance metrics, the apparent choice for deployment in hardware solutions is the Linear Regression model. However, the decision is not so black and white on deeper introspection. One of the primary challenges with the LR model is the necessity for regular calibration [114]. Given the ever-fluctuating nature of the real-world environment, changes in temperature, pressure, and humidity, to name a few. It becomes evident that such external factors could have a pronounced impact on sensor readings and, consequently, the data acquired. Imagine a device equipped with four sensors; calibrating each sensor's curve to ensure accurate EEG prediction would be cumbersome and a logistical nightmare, especially in environments that demand quick, accurate responses.

On the other hand, the BiLSTM model serves as a beacon of stability amidst the chaos. Its inherent design and architecture bestow upon it the robustness we desperately seek. The weights and biases within the BiLSTM are constant, eliminating the need for continual calibration and making it readily usable across multiple sensors without the associated

recalibration hassle [115].

Moreover, the essence of deep learning, specifically the BiLSTM network, is its capability to generalize and adapt to varying data patterns, which might not be explicitly linear. While the LR model might demonstrate superior performance under controlled conditions, its limitations might become pronounced in real-world, unpredictable scenarios. The BiLSTM, being more versatile, offers a better promise of consistency across diverse situations.

Furthermore, our commitment to realizing a real-world solution continues after model selection. We have also embarked on the path of quantizing the BiLSTM model. This process, aimed at reducing the memory and computational requirements of the model, ensures that our chosen model is accurate and also efficient and deployable in resource-constrained hardware environments. The fact that the BiLSTM can be quantized, making it more amenable to hardware implementations, further cements its place as our choice model.

6.2.2 Neuromorphic Computing

Neuromorphic computing embodies an avant-garde confluence of neuroscience and artificial intelligence. Drawing inspiration from the architecture and mechanisms of the biological brain, it ventures to conceive systems that emulate brain-like computational processes [116]. This approach diverges from traditional von Neumann architectures, characterized by sequential information processing and distinct memory units. In stark contrast, neuromorphic systems echo the parallelism, adaptability, and event-driven essence innate to human neural networks.

The impetus for neuromorphic computing arises from traditional computational systems' constraints, especially when deciphering intricate and dynamic real-world data that mandate real-time processing with conservative power consumption [117]. By harnessing the computational prowess of the brain, neuromorphic systems offer salient advantages: unmatched energy efficiency, scalability, and the capacity for real-time complex cognition. Central to the neuromorphic ethos are Spiking Neural Networks (SNNs). These are structured on "spikes" or action potentials as their fundamental computational entities, mirroring biological neuronal communication. The event-driven core of SNNs facilitates efficient, sparse calculations, curbing memory needs and processing overheads [118].

Recent strides in hardware, exemplified by Field-Programmable Gate Arrays (FPGAs) and specialized neuromorphic chips, have catalyzed the tangible manifestation of neuromorphic systems. These platforms are tailor-made for the nuances of SNNs, proffering exceptional speed and power efficiencies over conventional counterparts [119]. Such systems have demonstrated their mettle in real-world challenges, ranging from brain-computer interfaces and neuromorphic vision systems to cutting-edge cognitive computing [120]. Embracing the brain's computational strategies, neuromorphic approaches promise to redefine

our understanding of brain functions and catalyze leaps in neuroscience.

The essence of this chapter is to elucidate the synergy between neuromorphic computing and the quest to map MEG signals to EEG. Specifically, the spotlight optimizes the BiLSTM model, tailored for MEG to EEG signal transformation via MTJ sensors. Venturing into the vistas of neuromorphic computing, primarily through the lens of the Multiply-Accumulate-Activate-Pool (MAAP) blueprint, we aspire to supercharge the efficiency and real-time prowess of deep learning algorithms in MEG data analytics. By weaving neuromorphic computing into the tapestry of real-time applications, we envisage groundbreaking revolutions in neuroscience, healthcare, and myriad other realms, forging an unprecedented alliance between biological and artificial intelligence.

6.2.3 Quantization of BiLSTM

Bi-directional Long Short-Term Memory networks (BiLSTMs), a particular category of Recurrent Neural Networks (RNNs), have been recognized for their remarkable capability to discern and represent both past and future context in sequential data, like that found in MEG and EEG signals. This prowess in temporal data understanding has driven the adoption of BiLSTM models in various applications, especially in neuroscience, where temporal dynamics are critical.

However, the complexity associated with BiLSTM networks, mainly due to their intricate weight matrices and recurrent architectures, poses challenges for deployment on conventional and neuromorphic hardware platforms [121]. These models must be optimized to harness the full potential of neuromorphic computing, primarily its real-time and energy-efficient processing capabilities. Enter quantization is a pivotal step toward achieving this goal.

Quantization pertains to constraining or limiting the number of distinct states that a system, such as a neural network model, can assume. In deep learning models, quantization typically involves reducing the number of bits representing the neural network's weights and biases [122]. The primary motivations for quantizing deep learning models for neuromorphic hardware deployment include the following:

6.2.3.1 Reduced Memory Footprint

The beauty of quantization lies in its ability to significantly reduce the memory requirements of neural models without substantially compromising performance. Neural networks, incredibly complex architectures like BiLSTM, come with large weight matrices that occupy substantial memory when represented using standard precision formats like 32-bit floating numbers. Quantization aims to represent these weights and biases using fewer bits, sometimes as low as 2 to 8 [123]. This drastic reduction implies that the model's overall size

shrinks considerably. Consequently, the model can be stored and run on devices with limited memory capacities. This is particularly essential for IoT devices, often embedded systems with tight memory constraints. Furthermore, neuromorphic chips, designed to mimic the parallel processing capabilities of the brain, can benefit immensely from these compressed models [124], as it allows them to run more models in parallel or allocate memory to other crucial tasks.

6.2.3.2 Enhanced Processing Speed

The computational performance of deep learning models is closely tied to the precision of their numerical representations. Higher precision often implies more computational overhead due to the intricacy of floating-point arithmetic. By reducing the bit width of weights and biases through quantization, the arithmetic operations become more straightforward and faster. For instance, operations on 8-bit integers are inherently swifter than those on 32-bit floating numbers [125]. This speed-up is even more pronounced on specialized hardware accelerators, which can perform multiple low-bit-width operations simultaneously. The cumulative effect is a model that responds rapidly to inputs and meets the stringent requirements of real-time applications, ensuring timely data analysis, inference, and decision-making.

6.2.3.3 Energy Efficiency

One of the pivotal concerns in today's electronics is energy consumption, especially as devices become more portable and edge-centric. Every bit of computation and data movement within a device consumes power. Quantization, by reducing the number of bits involved in computations, directly impacts the device's energy efficiency. Fewer bits entail fewer transistors switching on and off during computation, directly translating to lower power consumption [126]. This is a tremendous advantage for battery-operated devices, ensuring longer battery life and extended periods between charges. Moreover, as edge computing gains traction, where computations are performed closer to the data source, energy-efficient models become essential for sustainable and scalable growth. Whether a wearable health monitor or a remote environmental sensor, energy-efficient models ensure these devices remain operational for prolonged durations, enabling consistent and uninterrupted data analysis.

Despite the benefits, quantization introduces several challenges that can affect its overall purpose. One of the primary concerns stems from the reduction of numerical precision in model parameters. This efficiency-driven reduction can lead to the loss of vital, fine-grained information. In the intricate architecture of deep networks, even minute imprecisions originating from initial layers can compound and become magnified in the subsequent ones [127].

As a result, the model could stray significantly from its expected behaviour, resulting in a potential dip in overall accuracy.

Moreover, quantization can influence the training dynamics of models. Training with reduced precision, often termed quantization-aware training, can disrupt gradient propagation. This disruption could manifest as slower convergence or, in more extreme cases, push the model’s training into a state of divergence. The complexities do not end there. Making decisions on the granularity of quantization is a non-trivial task. There is a looming question over which components of the model to quantize, should it be limited to the weights, or should it extend to activations and gradients? Each decision here can have a marked impact on the model’s final performance and computational efficiency.

Lastly, the realm of hardware introduces its own set of challenges. Different hardware platforms come equipped with distinct requirements for low-precision arithmetic operations. Ensuring the compatibility of quantized models with intended devices becomes paramount. Moreover, due to the inherent constraints in their numerical representation, quantized models might need help with adaptability. This reduced resilience against changes in data distribution can be a significant drawback, especially in dynamic real-world scenarios where the data’s nature can evolve.

In the subsequent sections, we will provide a comprehensive breakdown of our methodology for quantizing the BiLSTM model, coupled with experimental results, demonstrating the optimized model’s effectiveness and feasibility on neuromorphic hardware platforms.

6.3 Related work

The relentless pursuit of efficient, compact, and swift neural network models suited for real-world scenarios has given rise to many techniques encompassing hardware innovations, unconventional devices, and architecture-centric modifications.

As an approach, spintronics leverages the intrinsic electron spin and its associated magnetic moment to create energy-efficient neural computations. A highlight in this domain is the Spin Hall Magnetic Tunnel Junction-based spintronic neurons, which are tailored for concurrent multi-functional neural computations, achieving an impressive 99% precision [128]. This is further augmented by investigations into Spin-Orbit Torque Magnetic Tunnel Junction (SOT-MTJ), mainly to optimize ReLU activation and max-pooling, yielding significantly improved energy and area efficiency [129].

Neuromorphic computing offers another paradigm that strives to be more closely aligned with biology, creating systems that emulate human brain modelling. Intel’s Loihi is a seminal example, demonstrating the potential of spiking neural networks with reduced latency and energy consumption, especially when deploying brain-inspired networks with nuanced features like synaptic plasticity and sparsity [130].

Diving deeper into neuromorphic systems, a fresh perspective emerges with a comprehensive review of the Hopfield algorithm’s model, which has recently been explored extensively in large-scale hardware projects. This paper [131] underscores the algorithm’s advancements and potential to shape future artificial intelligence projects. Hardware solutions are gaining traction for their ability to cater to specific neural operations. This is exemplified in recent studies proposing time-multiplexed architectures for the convolution layer in CNNs, boasting a staggering 97% reduction in power consumption [132]. Another innovation bridges the analog and digital realms, resulting in a mixed-signal architecture for CNNs that promises improved energy-delay metrics [133].

There is also an emergent focus on healthcare applications. A contactless sleep apnea detection method has been proposed, leveraging snoring signals and hybrid deep neural networks. This solution, adapted for real-time applications on embedded hardware platforms, promises early diagnosis and treatment for sleep apnea in domestic environments [134]. Hand gesture recognition, essential for prosthesis control, also witnessed breakthroughs with the advent of a real-time gesture recognition system that employs a convolutional neural network (CNN). This system utilizes a custom 32-channel HDsEMG electrode array, demonstrating a rapid response time and a 98.15% accuracy, paving the way for advancements in hand prostheses [135].

There has been a surge in breakthroughs in the vast arena of neural networks in healthcare. This is particularly evident when harnessing dedicated deep-learning accelerators and neuromorphic processors. From memristive devices to FPGAs, researchers are creating DL accelerators adept at handling diagnostics, pattern recognition, and signal processing [136]. The recent research trajectories in neural network methodologies, algorithms, and hardware implementations, especially within the healthcare and computer vision domains, reveal a thriving and rapidly evolving landscape [137].

6.4 Adaptive Precision Quantization for BiLSTM Networks

Adaptive Precision Quantization is a compelling technique to optimize deep neural networks, especially when considering deployment in real-world scenarios that require efficient, high-speed computations with memory constraints. This section delves into the meticulous quantization process implemented in the BiLSTM networks of our study.

Initially, the study utilized reshaped dataset tailored explicitly for our model. It consists of noisy MEG signals with 305 channels and pure EEG signals with 59 channels. This dataset has been divided into training and testing as a ratio of 0.75 to 0.25 and prepared to ensure data preprocessing is aptly aligned with the model’s requirements. Essential TensorFlow libraries and specific modules pertinent to model optimization are invoked. The TensorFlow Model Optimization Toolkit (‘tfmot’) is particularly critical, given its robust

suite of tools tailored for training quantization-aware models [138]. The quantization configuration is defined by the number of sets, set size, and bit widths for inputs, weights, and outputs. These parameters were pivotal in guiding the subsequent quantization of the BiLSTM model’s weights.

A bespoke customized function named `map-rnn-to-maap-sets` is designed to navigate the quantization process for each BiLSTM layer in the network. This function’s primary role is to extract the weights of a given BiLSTM layer, apply the defined quantization process, and restore these adjusted weights to the layer. By doing this for all BiLSTM layers, the BiLSTM model is retrofitted with quantized weights, aiming for a balance between performance and computational efficiency.

Once this weight adjustment is complete, a quantization-aware model is birthed. This model, a mirror image of the original BiLSTM but with an awareness of the quantization nuances, is compiled and trained. An early stopping mechanism was introduced during training to ensure convergence and avoid overfitting based on the validation loss.

After the rigorous training phase, evaluating the quantization-aware model’s performance using the test dataset is paramount. This step is essential to gauge the model’s accuracy and understand how well the quantized model generalizes to unseen data, particularly given the changes introduced during quantization. The test dataset, which represents previously unseen data, is passed through the quantization-aware model to do this. This resulted in a set of predictions. These predictions, however, are in a quantized format, a direct consequence of the quantization process. Quantized outputs are advantageous for computation and storage, but converting them back to their original or conventional format is crucial for interpretability and further processing.

Enter the `decode-predicted` function. This function is designed with a clear objective: translate the quantized outputs into a format that aligns more closely with the original data’s scale and nature. The decoding mechanism is achieved through scaling, rounding, and further scaling operations. The predicted outputs are initially scaled based on the previous configuration’s bit width and number of sets. This transformation aimed to bring the values from the quantized range to the original data’s range. Subsequently, these scaled values were rounded to the nearest integers, ensuring the predictions conformed to a precise, discrete set of values.

This quantization approach in the BiLSTM networks emphasizes the tangible benefits of marrying computational efficiency with performance, presenting a feasible pathway for deploying robust models in resource-constrained environments.

Table 6.1: Performance Metrics RMSE and MAE for Different Models.

Model	MAE	RMSE
Model 1	0.1843	0.2069
Model 2	0.1849	0.2075
Model 3	0.1847	0.2075
Model 4	0.1840	0.2067
Model 5	0.2380	0.2598
Model 6	0.2388	0.2605
Model 7	0.2389	0.2606
Model 8	0.2374	0.2591

6.5 Performance Evaluation Through Hyperparameter Variation

In our quest to ascertain the most favourable configurations for quantization, we delved into a systematic variation of critical hyperparameters. This experiment aimed to determine the influence of these configurations on the model’s performance. By varying parameters like the Number of Sets, Set Size, and Bit Widths (for input, weight, and output), we exposed our model to a spectrum of quantization scenarios.

Each unique configuration was put to the test, with its performance being meticulously evaluated. This iterative process offered a panoramic landscape view, revealing how different formats influenced model accuracy and efficiency. The overarching aim was to identify an ‘optimal’ design and understand the nuanced interplay between quantization parameters and their resultant effect on the model’s predictive capabilities.

In Table 6.1, we present the performance metrics of eight different model configurations, evaluated through both Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). Among these configurations, Model 4 exhibits the most promising accuracy with an RMSE of 0.2067, $num_sets = 4$, $set_size = 6$, $input_bit_width = 32$, $weight_bit_width = 32$, $output_bit_width = 16$, while Model 5 registers a relatively higher error with an RMSE of 0.2598, $num_sets = 6$, $set_size = 6$, $input_bit_width = 16$, $weight_bit_width = 16$, $output_bit_width = 32$. Notably, a perceptible performance gap exists when comparing the RMSE values of these quantized models to the original BiLSTM model’s 0.0488, as highlighted in our prior research.

This observed disparity in performance between the original and quantized models is intricately linked to the nature of quantization itself. The process inherently involves approximations, which could be the source of minor discrepancies. Furthermore, the complex-

ities and nuances inherent to the original BiLSTM model might not be wholly encapsulated post-quantization. While methodical, the quantization’s hyperparameter choices may only sometimes prove optimal across all model configurations. It’s also conceivable that some models become more sensitive to variations in data after undergoing quantization, resulting in escalated error rates. Despite these challenges, the benefits of quantized models in computational efficiency and memory conservation cannot be understated. The more optimization is given to fine-tuning step, the better the quantization, which is applicable.

Further illuminating our examination of the model’s performance post-quantization, Figure 6.2 graphically showcases the performance comparison of actual versus quantized signals for different configurations. Each subplot vividly demonstrates the disparities and congruences between the original and the quantized signals after applying a moving average filter of window 5 to smooth the output. The figures show the first 1000 samples for channel 4 for better visualization. For instance, in Model 1 (Figure 6.2(a)), the quantized signal is derived from a configuration with parameters such as *num_sets* = 4, *set_size* = 8, and so forth. Each subsequent subplot, from Model 2 (Figure 6.2(b)) through to Model 8 (Figure 6.2(h)), represents varied configuration settings. As one traverses through these visual plots, the intricate differences underscored by various hyperparameter combinations become more evident. Such visualizations are pivotal in comprehending the nuances of quantization on model performance and how subtle configuration changes can manifest in the model’s output. While some models appear to mimic the original signal with a commendable degree of accuracy, others highlight the challenges intrinsic to the quantization process.

Expanding on our investigation into the impact of quantization hyperparameters, Figure 6.1 delves into a granular analysis of RMSE and MAE based on the variation of different hyperparameters. The visual representation elucidates a clear trend in the model’s performance sensitivity concerning each hyperparameter. Predominantly, it is evident that the number of sets is the paramount hyperparameter that significantly influences the quantization’s efficacy. Its impact is more pronounced than the other parameters, suggesting its potential role as a critical determinant in the optimization of quantization processes. In contrast, variations in different hyperparameters appear to have a more subtle or marginal effect on the model’s performance. The discernment of such nuances emphasizes the importance of reasonable hyperparameter selection, especially when aiming to maximize the benefits of quantization without compromising on model accuracy.

6.6 Conclusion

This research journey, predicated on the symbiosis of deep learning techniques and neuroscience, has been a testament to the limitless potentialities that emerge when intricate brain signal analysis meets groundbreaking AI innovations. The landscape of Magnetoen-

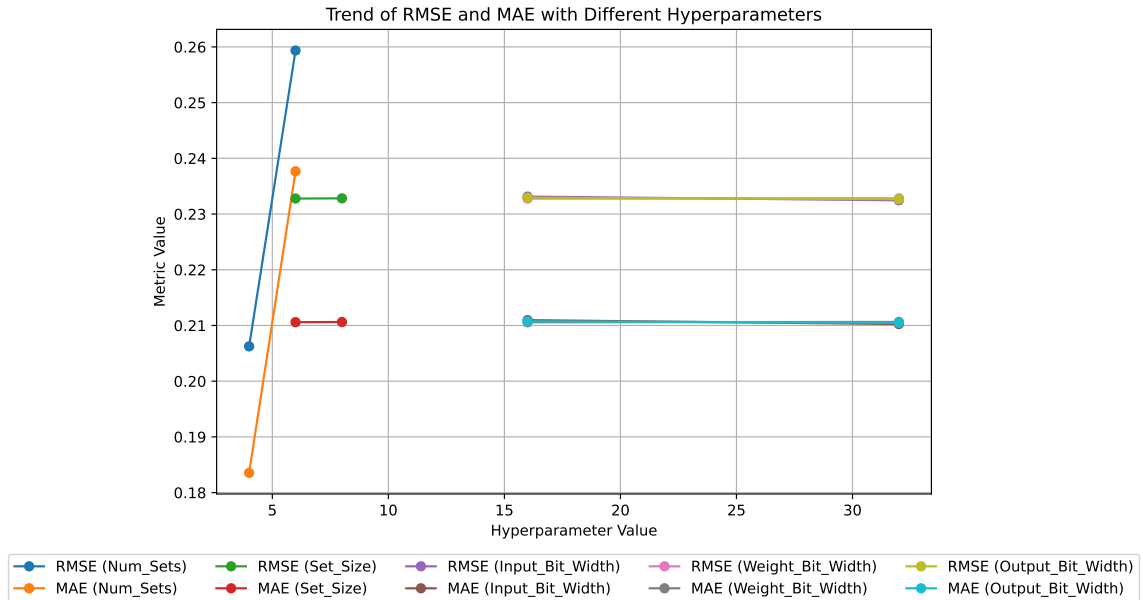


Figure 6.1: Analysis of RMSE and MAE based on the variation of different hyperparameters. The line plots illustrate the trend of RMSE and MAE values as a function of each hyperparameter. The comparative analysis provides insights into the sensitivity of the model’s performance to changes in the hyperparameter configurations

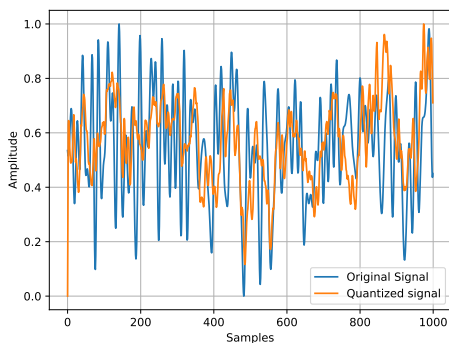
cephalography (MEG) and Electroencephalography (EEG) signal processing has been profoundly reshaped by our forays into employing Magnetic Tunnel Junction (MTJ) sensors and exploring the BiLSTM architecture. Our endeavours have highlighted the possibilities within these realms and underscored the profound impact of precise signal processing on advancing our understanding of cognitive paradigms.

The essence of this manuscript has been the emphasis on the meticulous quantization of the BiLSTM model, tailored to resonate harmoniously with neuromorphic hardware platforms. The detailed performance analysis presented underscores the robustness of our approach. Rather than presenting a full-fledged hardware implementation, we have laid down the foundational steps to ensure that the depth and intricacy of applied deep learning are faithfully represented when translated to a hardware-oriented context.

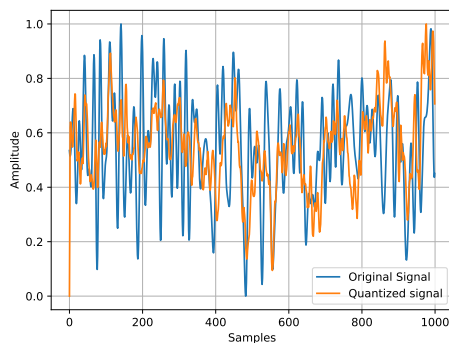
The world is at the cusp of an edge computing revolution, and our work stands as a beacon that emphasizes the indispensability of neuromorphic implementations for real-world applications. As we culminate this exploration, it is imperative to acknowledge the vast expanse of untapped potential that awaits. While our study has established a robust framework, the confluence of deep learning and neuromorphic hardware holds promises yet to be unravelled.

Looking forward, we remain committed to advancing this field. Our goal is to be at the forefront of brain signal research, using technology as our lens to explore the vast potential

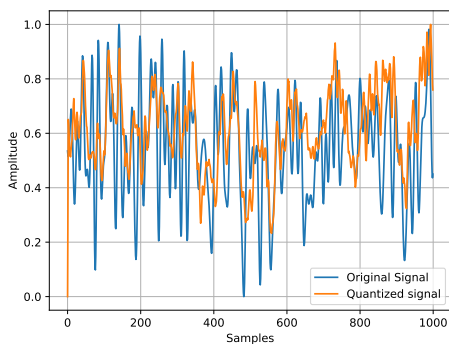
of the human mind. We dream of a world where our work serves as a catalyst, bridging technology and neuroscience to benefit sectors like healthcare and cognitive research.



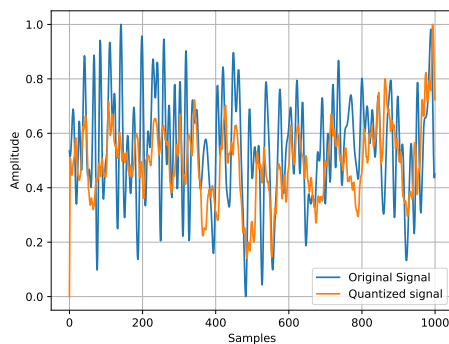
(a) Model 1: Actual vs quantized signal performance with the $num_sets = 4$, $set_size = 8$, $input_bit_width = 32$, $weight_bit_width = 32$, $output_bit_width = 32$



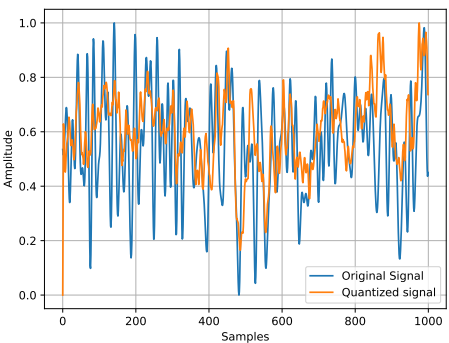
(b) Model 2: Actual vs quantized signal performance with the $num_sets = 4$, $set_size = 8$, $input_bit_width = 32$, $weight_bit_width = 32$, $output_bit_width = 16$



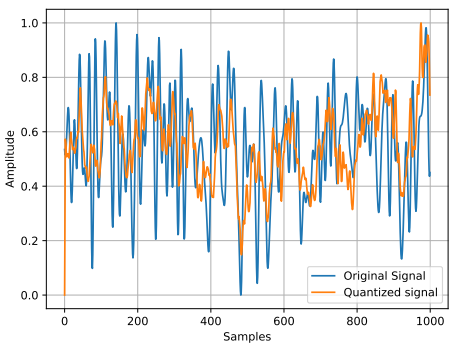
(c) Model 3: Actual vs quantized signal performance with the $num_sets = 4$, $set_size = 6$, $input_bit_width = 16$, $weight_bit_width = 32$, $output_bit_width = 32$



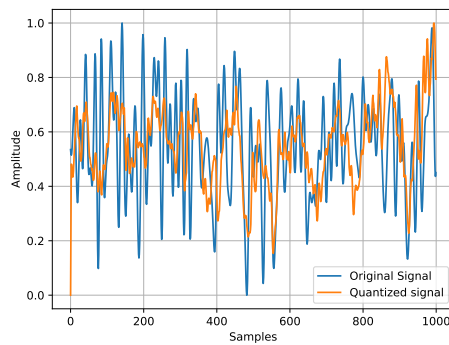
(d) Model 4: Actual vs quantized signal performance with the $num_sets = 4$, $set_size = 6$, $input_bit_width = 32$, $weight_bit_width = 32$, $output_bit_width = 16$



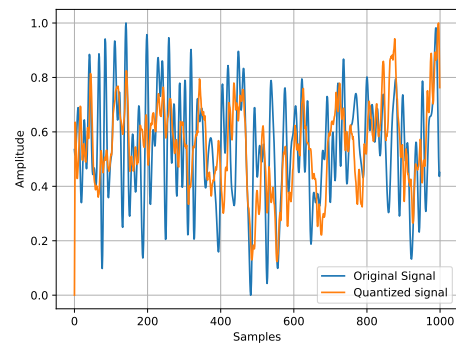
(e) Model 5: Actual vs quantized signal performance with the $num_sets = 6$, $set_size = 6$, $input_bit_width = 16$, $weight_bit_width = 16$, $output_bit_width = 32$



(f) Model 6: Actual vs quantized signal performance with the $num_sets = 6$, $set_size = 8$, $input_bit_width = 16$, $weight_bit_width = 32$, $output_bit_width = 32$



(g) Model 7: Actual vs quantized signal performance with the $num_sets = 6$, $set_size = 6$, $input_bit_width = 32$, $weight_bit_width = 32$, $output_bit_width = 16$



(h) Model 8: Actual vs quantized signal performance with the $num_sets = 6$, $set_size = 8$, $input_bit_width = 16$, $weight_bit_width = 32$, $output_bit_width = 16$

Figure 6.2: Performance comparison of actual vs Filtered normalized quantized signals using different configuration settings.

Chapter 7

Conclusion and Future Work on Integrating denoised and mapped signal for disease prediction and fed learning

The exquisite interplay between the human brain's complexity and the advancing frontier of modern technology is the central motif of this research. Embarking on a voyage through the intricacies of neuroscience intertwined with the possibilities of artificial intelligence, our efforts have heralded a transformative phase in the domain of Magnetoencephalography (MEG) and Electroencephalography (EEG) signal analysis. Embarking on this research, we ventured into the largely unexplored realms of Magnetic Tunnel Junction (MTJ) sensors, paired with the profound capacities of deep learning. This initial spark of curiosity and exploration has matured into a pioneering innovation, one that has the potential to significantly impact the future trajectory of brain health diagnostics and treatment methodologies.

At the onset of our research expedition, we presented a novel idea that held immense promise—a seamless integration of nimble, IoT-driven biomagnetic sensors and avant-garde AI frameworks. Our aim was ambitious yet clear: to adeptly capture MEG signals and transpose them into their EEG equivalents. This innovative amalgamation highlighted the latent potential of the next generation of wearable diagnostic devices. As our research suggests, these tools can usher in an era where monitoring brain functions from the comfort of one's home, without any compromise on quality or accuracy, becomes the norm. Diving into machine learning, we embarked on a meticulous assessment of diverse models. It became evident that while models like KNN and LR stood out for specific strengths, deep learning paradigms, especially DNN and CNN, exhibited a remarkable finesse in discerning the multifaceted patterns embedded within EEG signals.

As our exploration deepened, we grappled with age-old obstacles that have long cast shadows over conventional MEG and EEG diagnostic techniques. These challenges—ranging from prohibitive costs and limited accessibility to the invasive nature of some procedures and their inherent complexity—were formidable. However, by synergizing the potent capabilities of MTJ-based sensors with state-of-the-art deep learning methodologies, we envisioned the creation of monitoring systems that are user-friendly and seamlessly integrated into daily routines. Our scrupulous analysis of RNN models bore fruit when the RCNN model emerged as a clear front-runner, striking an optimal balance between processing efficiency and analytical precision. Concurrently, the GRU model also showed significant potential, marking itself as a viable alternative for specific applications.

Nevertheless, as with any pioneering endeavour, our path was challenged. A recurring nemesis in biomedical signal processing has been 'noise'—unwanted interference that can obscure actual signal data. This formidable adversary threatened to hinder our quest for capturing unadulterated MEG signals. However, fortified with an array of AI-driven denoising algorithms, we met this challenge with unwavering resolve. The results were illuminating, with models like LR, BiLSTM, and RCNN showcasing their robust capabilities in noise mitigation.

On another frontier, our efforts to quantize the BiLSTM model revealed its inherent compatibility with neuromorphic hardware platforms, signalling the dawn of an age dominated by edge computing. Instead of confining ourselves to presenting a pre-packaged solution, we took a broader perspective. We meticulously crafted a roadmap—a comprehensive blueprint that accentuates the harmonious marriage of deep learning techniques and neuromorphic hardware architectures.

To encapsulate, our research journey has been monumental in building bridges between the avant-garde realms of technology and the enigmatic world of neuroscience. A vision fuels us as we pause at this juncture, poised at the threshold of myriad unexplored possibilities. A vision of a tomorrow where technological advancements and neuroscience coalesce, heralding breakthroughs in sectors as pivotal as healthcare and cognitive studies. Guided by this vision, our mission remains steadfast—to continue our quest, leveraging technology as a lens, to delve deeper into the mysteries of the human psyche.

7.1 Future Directions

As delineated in our study, the nexus of neuroscience and avant-garde technology has established a robust platform for numerous forward-looking ventures. Gazing into the horizon, distinct and crucial paths beckon, ready to steer our subsequent initiatives and deepen our dedication to transforming the realm of brain health surveillance. Herein, we outline the salient pathways our research anticipates venturing into:

- **Transition to HDL Solutions:** Building upon our work on the quantized solution, the next logical stride lies in transitioning to Hardware Description Language (HDL) solutions. This will enable us to design and simulate the neuromorphic hardware platforms more intricately, offering a gateway to actual hardware implementation. Such a move would accelerate the transition from theory to tangible devices and refine our systems' performance and efficiency in real-world scenarios.
- **Early Detection of Abnormal Brain Activities:** With the foundational knowledge garnered from our research, there is an imperative to focus on developing AI-based models explicitly tailored for detecting abnormalities in brain activities. Such models leverage deep learning architectures to analyze and predict anomalies at their nascent stage, enabling timely interventions and potentially mitigating severe medical conditions.
- **Innovations in Real-time Monitoring:** Envisioning the future of non-intrusive brain health monitoring, we are inspired to design a smart band. This wearable would seamlessly integrate IoT capabilities with our AI-driven denoising and detection models, presenting a holistic solution that offers real-time insights into brain activity. Such a device would prioritize user comfort, ensuring continuous monitoring without disrupting daily activities.
- **Cloud Integration for Emergency Alerts:** As we advance our real-time monitoring solutions, there is an inherent need for safety mechanisms. We can implement advanced algorithms to identify any life-threatening anomalies by transmitting the monitored data to cloud infrastructures. Should any dangerous situation be detected, instant alerts could be dispatched to medical professionals and caregivers, facilitating rapid response and intervention.
- **Embracing Federated Learning:** Given the sensitive nature of brain data and the growing concerns surrounding data privacy, it becomes crucial to incorporate federated learning techniques. By processing data locally on IoT devices (like the envisioned smart band) and only sharing model updates rather than raw data, federated learning promises to uphold user privacy while benefiting from collective data intelligence. This decentralized approach to machine learning will ensure that our solutions are not only technologically advanced but also ethically responsible.

In conclusion, the path forward is illuminated with opportunities further to bridge the chasm between neuroscience and state-of-the-art technology. Our vision for the future is a world where the confluence of these domains leads to breakthroughs that transcend the traditional boundaries of healthcare, diagnostics, and cognitive research. With a clear

roadmap ahead, we remain steadfast in our mission to harness technology's potential to shed light on the enigma that is the human brain.

Bibliography

- [1] S. S. E. Center. (Date accessed) How do birds navigate? [Online]. Available: <https://ssec.si.edu/stemvisions-blog/how-do-birds-navigate>
- [2] Science. (Date accessed) Magnetic attraction: How do homing pigeons find their way home? [Online]. Available: <https://www.science.org/content/article/magnetic-attraction-home>
- [3] J. Zhang, “Basic neural units of the brain: neurons, synapses and action potential,” *arXiv preprint arXiv:1906.01703*, 2019.
- [4] Wikimedia Commons, “Membrane permeability of a neuron during an action potential,” 2021, [Online; accessed `{date-of-access}`]. [Online]. Available: https://commons.wikimedia.org/wiki/File:Membrane_Permeability_of_a_Neuron_During_an_Action_Potential.svg
- [5] W. Van der Kloot, “Calcium and neuromuscular transmission,” in *Calcium in drug action*. Springer, 1978, pp. 261–287.
- [6] J. Graham-Harper-Cater, A. Nogaret, H. Blakes, B. Metcalfe, and J. Taylor, “Resilience of neural electronics to high magnetic fields,” *Journal of Electronic Materials*, vol. 51, no. 8, pp. 4161–4168, 2022.
- [7] C. Demopoulos, “Magnetoencephalography.” 2023.
- [8] C. Sun, B. Zhao, T. Wu, J. Zhang, Y. Leng, H. Guo, and C. B. Group, “Monitoring the intestinal magnetic field with optically pumped atomic magnetometers,” *bioRxiv*, pp. 2022–11, 2022.
- [9] T. Katila, “the beginning of biomagnetism research and magnetoencephalography in finland in the 1970s,” *Fifty Years of Magnetoencephalography: Beginnings, Technical Advances, and Applications*, p. 8, 2020.
- [10] S. Zuo, K. Nazarpour, T. Böhnert, E. Paz, P. Freitas, R. Ferreira, and H. Heidari, “Integrated pico-tesla resolution magnetoresistive sensors for miniaturised magneto-

- myography,” in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2020, pp. 3415–3419.
- [11] D. Moreau, S.-C. Kao, and C.-H. Wang, “Probing brain dynamics correlates of motor expertise with mobile recordings,” 2023.
- [12] O. Feys and X. De Tiège, “From cryogenic to on-scalp magnetoencephalography for the evaluation of paediatric epilepsy,” *Developmental Medicine & Child Neurology*, 2023.
- [13] C. Zhao, L. Zhan, P. M. Thompson, and H. Huang, “Predicting spatio-temporal human brain response using fmri,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2022, pp. 336–345.
- [14] R. Rucco, F. Baselice, M. Ambrosanio, A. Vettoliere, P. Sorrentino, M. P. Riccio, C. Bravaccio, P. Silvestrini, and C. Granata, “Brain connectivity study by multi-channel system based on superconducting quantum magnetic sensors,” *Engineering Research Express*, vol. 2, no. 1, p. 015038, 2020.
- [15] K. Arai, A. Kuwahata, D. Nishitani, I. Fujisaki, R. Matsuki, Y. Nishio, Z. Xin, X. Cao, Y. Hatano, S. Onoda *et al.*, “Millimetre-scale magnetocardiography of living rats with thoracotomy,” *Communications Physics*, vol. 5, no. 1, p. 200, 2022.
- [16] M. K. Armstrong, M. G. Schultz, A. D. Hughes, D. S. Picone, and J. E. Sharman, “Physiological and clinical insights from reservoir-excess pressure analysis,” *Journal of human hypertension*, vol. 35, no. 9, pp. 758–768, 2021.
- [17] J. Kissane, J. A. Neutze, and H. Singh, *Radiology fundamentals: Introduction to imaging & technology*. Springer Nature, 2020.
- [18] D. Tyler, J. Wood, T. Sabir, C. McDonnell, A. S. M. Sayem, and N. Whittaker, “Wearable electronic textiles,” *Textile Progress*, vol. 51, no. 4, pp. 299–384, 2019.
- [19] P. Dey, J. N. Roy, P. Dey, and J. N. Roy, “Spintronics applications,” *Spintronics: Fundamentals and Applications*, pp. 223–267, 2021.
- [20] P. Kumar, A. Kumar, and D. Kaur, “Spin valve effect in sputtered fl-mos2 and ferromagnetic shape memory alloy based magnetic tunnel junction,” *Ceramics International*, vol. 47, no. 4, pp. 4587–4594, 2021.
- [21] Nanohub: Supporting documents. [Online]. Available: <https://nanohub.org/resources/24815/supportingdocs>

- [22] T. Kubota, Z. Wen, and K. Takashi, "Current-perpendicular-to-plane giant magnetoresistance effects using heusler alloys," *Journal of Magnetism and Magnetic Materials*, vol. 492, p. 165667, 2019.
- [23] Z. Zhou, K. Zhang, and Q. Leng, "Tunneling magnetoresistance (tmr) materials and devices for magnetic sensors," *Spintronics: Materials, Devices and Applications*, pp. 51–92, 2022.
- [24] C. Zheng, K. Zhu, S. C. De Freitas, J.-Y. Chang, J. E. Davies, P. Eames, P. P. Freitas, O. Kazakova, C. Kim, C.-W. Leung *et al.*, "Magnetoresistive sensor development roadmap (non-recording applications)," *IEEE Transactions on Magnetics*, vol. 55, no. 4, pp. 1–30, 2019.
- [25] S. Mostufa, P. Yari, B. Rezaei, K. Xu, and K. Wu, "Flexible magnetic field nanosensors for wearable electronics: A review," *ACS Applied Nano Materials*, 2023.
- [26] S. Zuo, H. Heidari, D. Farina, and K. Nazarpour, "Miniaturized magnetic sensors for implantable magnetomyography," *Advanced Materials Technologies*, vol. 5, no. 6, p. 2000185, 2020.
- [27] K. T. Bæk and K. P. Kepp, "Data set and fitting dependencies when estimating protein mutant stability: Toward simple, balanced, and interpretable models," *Journal of Computational Chemistry*, vol. 43, no. 8, pp. 504–518, 2022.
- [28] S. H. Krishna, O. J. Oyebode, A. Joshi, G. S. Reddy, P. Karthikeyan, and K. Ganesh, "The technical role of regression analysis in prediction and decision making," in *2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*. IEEE, 2023, pp. 183–187.
- [29] A. Weigard, S. Lane, K. Gates, and A. Beltz, "The influence of autoregressive relation strength and search strategy on directionality recovery in group iterative multiple model estimation." *Psychological Methods*, vol. 28, no. 2, p. 379, 2023.
- [30] A. Name. (Year) The basics: Knn for classification and regression. Accessed: Date. [Online]. Available: <https://towardsdatascience.com/the-basics-knn-for-classification-and-regression-c1e8a6c955>
- [31] G. E. Atteia, H. A. Mengash, and N. A. Samee, "Evaluation of using parametric and non-parametric machine learning algorithms for covid-19 forecasting," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 10, 2021.
- [32] K. Taunk, S. De, S. Verma, and A. Swetapadma, "A brief review of nearest neighbor algorithm for learning and classification," in *2019 international conference on intelligent computing and control systems (ICCS)*. IEEE, 2019, pp. 1255–1260.

- [33] R. Liu and S. Misra, "Machine learning assisted detection and localization of mechanical discontinuity," *International Journal of Fracture*, vol. 236, no. 2, pp. 219–234, 2022.
- [34] F. Hajjej, M. A. Alohali, M. Badr, and M. A. Rahman, "A comparison of decision tree algorithms in the assessment of biomedical data," *BioMed Research International*, vol. 2022, 2022.
- [35] G. Stiglic, P. Kocbek, N. Fijacko, M. Zitnik, K. Verbert, and L. Cilar, "Interpretability of machine learning-based prediction models in healthcare," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 10, no. 5, p. e1379, 2020.
- [36] R. Natras, B. Soja, and M. Schmidt, "Ensemble machine learning of random forest, adaboost and xgboost for vertical total electron content forecasting," *Remote Sensing*, vol. 14, no. 15, p. 3547, 2022.
- [37] V.-C. Nguyen, D.-T. Hoang, X.-T. Tran, M. Van, and H.-J. Kang, "A bearing fault diagnosis method using multi-branch deep neural network," *Machines*, vol. 9, no. 12, p. 345, 2021.
- [38] A. Ghatak and A. Ghatak, "Introduction to neural networks," *Deep learning with R*, pp. 23–63, 2019.
- [39] H. Robinson, S. Pawar, A. Rasheed, and O. San, "Physics guided neural networks for modelling of non-linear dynamics," *Neural Networks*, vol. 154, pp. 333–345, 2022.
- [40] M. Bordini, M. W. Rivolta, and R. Sassi, "Opening the black box: interpretability of machine learning algorithms in electrocardiography," *Philosophical Transactions of the Royal Society A*, vol. 379, no. 2212, p. 20200253, 2021.
- [41] F. Demir, A. M. Ismael, and A. Sengur, "Classification of lung sounds with cnn model using parallel pooling structure," *IEEE Access*, vol. 8, pp. 105 376–105 383, 2020.
- [42] L. Lin, C. Chen, and T. Xu, "Spatial-spectral hyperspectral image classification based on information measurement and cnn," *EURASIP Journal on Wireless Communications and Networking*, vol. 2020, pp. 1–16, 2020.
- [43] E. Tsalera, A. Papadakis, I. Voyiatzis, and M. Samarakou, "Cnn-based, contextualized, real-time fire detection in computational resource-constrained environments," *Energy Reports*, vol. 9, pp. 247–257, 2023.
- [44] F. M. Almasoudi, "Enhancing power grid resilience through real-time fault detection and remediation using advanced hybrid machine learning models," *Sustainability*, vol. 15, no. 10, p. 8348, 2023.

- [45] A. J. Dautel, W. K. Härdle, S. Lessmann, and H.-V. Seow, “Forex exchange rate forecasting using deep recurrent neural networks,” *Digital Finance*, vol. 2, pp. 69–96, 2020.
- [46] K. L. Tan, C. P. Lee, and K. M. Lim, “Roberta-gru: A hybrid deep learning model for enhanced sentiment analysis,” *Applied Sciences*, vol. 13, no. 6, p. 3915, 2023.
- [47] P. Kumari and D. Toshniwal, “Deep learning models for solar irradiance forecasting: A comprehensive review,” *Journal of Cleaner Production*, vol. 318, p. 128566, 2021.
- [48] F. Landi, L. Baraldi, M. Cornia, and R. Cucchiara, “Working memory connections for lstm,” *Neural Networks*, vol. 144, pp. 334–341, 2021.
- [49] G. Xu, T. Ren, Y. Chen, and W. Che, “A one-dimensional cnn-lstm model for epileptic seizure recognition using eeg signal analysis,” *Frontiers in neuroscience*, vol. 14, p. 578126, 2020.
- [50] J. P. Bharadiya, “Exploring the use of recurrent neural networks for time series forecasting,” *International Journal of Innovative Science and Research Technology*, vol. 8, no. 5, pp. 2023–2027, 2023.
- [51] S. Siami-Namini, N. Tavakoli, and A. S. Namin, “A comparative analysis of forecasting financial time series using arima, lstm, and bilstm,” *arXiv preprint arXiv:1911.09512*, 2019.
- [52] Z. Yang, J. Ma, H. Chen, J. Zhang, and Y. Chang, “Context-aware attentive multilevel feature fusion for named entity recognition,” *IEEE transactions on neural networks and learning systems*, 2022.
- [53] R. E. Nogales and M. E. Benalcázar, “Hand gesture recognition using automatic feature extraction and deep learning algorithms with memory,” *Big Data and Cognitive Computing*, vol. 7, no. 2, p. 102, 2023.
- [54] E. Gordon, “Brain imaging technologies: how, what, when and why?” *Australian & New Zealand Journal of Psychiatry*, vol. 33, no. 2, pp. 187–196, 1999.
- [55] D. Tech, “Introduction to TMR magnetic sensors,” <http://www.dowaytech.com/en/1776.html>, 2015, accessed: 12 3, 2015.
- [56] TDK, “TDK developing technologies,” <https://product.tdk.com/en/techlibrary/developing/bio-sensor/index.html>.
- [57] G. L. Barkley and C. Baumgartner, “MEG and EEG in epilepsy,” *Journal of clinical neurophysiology*, vol. 20, no. 3, pp. 163–178, 2003.

- [58] A. M. Alvi, S. Siuly, and H. Wang, “Neurological abnormality detection from electroencephalography data: a review,” *Artificial Intelligence Review*, vol. 55, no. 3, pp. 2275–2312, 2022.
- [59] J. J. Van der Zande *et al.*, “EEG characteristics of dementia with lewy bodies, alzheimer’s disease and mixed pathology,” *Frontiers in aging neuroscience*, vol. 10, p. 190, 2018.
- [60] Y. Zhu, T. Parviainen, E. Heinilä, L. Parkkonen, and A. Hyvärinen, “Unsupervised representation learning of spontaneous meg data with nonlinear ica,” *NeuroImage*, vol. 274, p. 120142, 2023.
- [61] S. Sakib, M. M. Fouda, M. Al-Mahdawi, A. Mohsen, M. Oogane, Y. Ando, and Z. M. Fadlullah, “Deep learning models for magnetic cardiography edge sensors implementing noise processing and diagnostics,” *IEEE Access*, vol. 10, pp. 2656–2668, 2021.
- [62] N. Mehta, A. Pandit, and S. Shukla, “Transforming healthcare with big data analytics and artificial intelligence: A systematic mapping study,” *Journal of biomedical informatics*, vol. 100, p. 103311, 2019.
- [63] G. Giannakakis *et al.*, “Review on psychological stress detection using biosignals,” *IEEE Transactions on Affective Computing*, vol. 13, no. 1, pp. 440–460, 2022.
- [64] Y. Ding, X. Liang, T. Middelman, J. Marquetand, and H. Heidari, “Mmg/emg mapping with reservoir computing,” in *2022 29th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, 2022, pp. 1–4.
- [65] V. Bhat, “Reconstruction of cardiac activities from vectorcardiography and magnetocardiography using bayesian approach with coherence mapping,” *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 9, no. 1, pp. 78–91, 2021.
- [66] Y. Yang *et al.*, “A new wearable multichannel magnetocardiogram system with a SERF atomic magnetometer array,” *Scientific Reports*, vol. 11, no. 1, p. 5564, 2021.
- [67] A. Mohsen, M. Al-Mahdawi, M. M. Fouda, M. Oogane, Y. Ando, and Z. M. Fadlullah, “AI aided noise processing of spintronic based IoT sensor for magnetocardiography application,” in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020.
- [68] S. Sakib, M. M. Fouda, M. Al-Mahdawi, A. Mohsen, M. Oogane, Y. Ando, and Z. M. Fadlullah, “Noise-removal from spectrally-similar signals using reservoir computing for MCG monitoring,” in *ICC 2021 - IEEE International Conference on Communications*, 2021.

- [69] S. Sakib, M. M. Fouda, and Z. M. Fadlullah, “Harnessing artificial intelligence for secure ECG analytics at the edge for cardiac arrhythmia classification,” in *Secure Edge Computing*, pp. 137–153. CRC Press, 2021.
- [70] H. Stefan and E. Trinka, “Magnetoencephalography (MEG): Past, current and future perspectives for improved differentiation and treatment of epilepsies,” *Seizure*, vol. 44, pp. 121–124, 2017.
- [71] M. J. Brookes, J. Leggett, M. Rea, R. M. Hill, N. Holmes, E. Boto, and R. Bowtell, “Magnetoencephalography with optically pumped magnetometers (OPM-MEG): the next generation of functional neuroimaging,” *Trends in Neurosciences*, vol. 45, no. 8, pp. 621–634, 2022.
- [72] G. Barbati, C. Porcaro, F. Zappasodi, P. M. Rossini, and F. Tecchio, “Optimization of an independent component analysis approach for artifact identification and removal in magnetoencephalographic signals,” *Clinical Neurophysiology*, vol. 115, no. 5, pp. 1220–1232, 2004.
- [73] A. Gramfort *et al.*, “MNE software for processing MEG and EEG data,” *Neuroimage*, vol. 86, pp. 446–460, 2014.
- [74] M. Schulz *et al.*, “An integrative MEG–fMRI study of the primary somatosensory cortex using cross-modal correspondence analysis,” *NeuroImage*, vol. 22, no. 1, pp. 120–133, 2004.
- [75] S. Vanhatalo, J. Voipio, and K. Kaila, “Full-band EEG (FbEEG): an emerging standard in electroencephalography,” *Clinical Neurophysiology*, vol. 116, no. 1, pp. 1–8, 2005.
- [76] J. M. Antelis, L. Montesano, A. Ramos-Murguialday, N. Birbaumer, and J. Minguetz, “On the usage of linear regression models to reconstruct limb kinematics from low frequency EEG signals,” *PloS one*, vol. 8, no. 4, article no. e61976, 2013.
- [77] D. Sabbagh, P. Ablin, G. Varoquaux, A. Gramfort, and D. A. Engemann, “Predictive regression modeling with MEG/EEG: from source power to signals and cognitive states,” *NeuroImage*, vol. 222, article no. 116893, 2020.
- [78] H. Huang, G. Han, F. Xiao, and R. Wang, “An online teaching video evaluation scheme based on EEG signals and machine learning,” *Wireless Communications and Mobile Computing*, vol. 2022, 2022.
- [79] S. Talebi, J. Waczak, B. A. Fernando, A. Sridhar, and D. J. Lary, “Data-driven EEG band discovery with decision trees,” *Sensors*, vol. 22, no. 8, article no. 3048, 2022.

- [80] E. Carrizosa, C. Molero-Rio, and D. Romero Morales, “Mathematical optimization in classification and regression trees,” *Top*, vol. 29, no. 1, pp. 5–33, 2021.
- [81] M. Jurczak, M. Kołodziej, and A. Majkowski, “Implementation of a convolutional neural network for eye blink artifacts removal from the electroencephalography signal,” *Frontiers in Neuroscience*, vol. 16, article no. 34, 2022.
- [82] A. Altameem *et al.*, “Performance analysis of machine learning algorithms for classifying hand motion-based EEG brain signals,” *Computer Systems Science and Engineering*, vol. 42, no. 3, pp. 1095–1107, 2022.
- [83] Y. Sadat-Nejad and S. Beheshti, “Efficient blinking component estimation in subspace-based eeg and meg analysis,” pp. 2028–2032, 2019.
- [84] X. Li, J. Chen, N. Shi, C. Yang, P. Gao, X. Chen, Y. Wang, S. Gao, and X. Gao, “A hybrid steady-state visual evoked response-based brain-computer interface with meg and eeg,” *Expert Systems with Applications*, vol. 223, p. 119736, 2023.
- [85] N. Coquelet, X. De Tiège, F. Destoky, L. Roshchupkina, M. Bourguignon, S. Goldman, P. Peigneux, and V. Wens, “Comparing meg and high-density eeg for intrinsic functional connectivity mapping,” *NeuroImage*, vol. 210, p. 116556, 2020.
- [86] R. M. Hill, E. Boto, M. Rea, N. Holmes, J. Leggett, L. A. Coles, M. Papastavrou, S. K. Everton, B. A. Hunt, D. Sims *et al.*, “Multi-channel whole-head opm-meg: Helmet design and a comparison with a conventional system,” *NeuroImage*, vol. 219, p. 116995, 2020.
- [87] S. E. Barnes, I. Opris, B. R. Noga, S. Huang, and F. Zuo, “Neural spintronics: Non-invasive augmentation of brain functions,” *Modern Approaches to Augmentation of Brain Function*, pp. 433–446, 2021.
- [88] B. Susan Philip, G. Prasad, and D. J. Hemanth, “A systematic review on artifact removal and classification techniques for enhanced meg-based bci systems,” *Brain-Computer Interfaces*, pp. 1–15, 2023.
- [89] C. E. Stevens Jr and D. L. Zabelina, “Creativity comes in waves: an eeg-focused exploration of the creative brain,” *Current Opinion in Behavioral Sciences*, vol. 27, pp. 154–162, 2019.
- [90] D. Walther, J. Viehweg, J. Haueisen, and P. Mäder, “A systematic comparison of deep learning methods for eeg time series analysis,” *Frontiers in Neuroinformatics*, vol. 17, p. 1067095, 2023.

- [91] F. Li, W. Chao, Y. Li, B. Fu, Y. Ji, H. Wu, and G. Shi, “Decoding imagined speech from eeg signals using hybrid-scale spatial-temporal dilated convolution network,” *Journal of Neural Engineering*, vol. 18, no. 4, p. 0460c4, 2021.
- [92] A. Sarkar, A. Singh, and R. Chakraborty, “A deep learning-based comparative study to track mental depression from eeg data,” *Neuroscience Informatics*, vol. 2, no. 4, p. 100039, 2022.
- [93] J. Sun, B. Wang, Y. Niu, Y. Tan, C. Fan, N. Zhang, J. Xue, J. Wei, and J. Xiang, “Complexity analysis of eeg, meg, and fmri in mild cognitive impairment and alzheimer’s disease: a review,” *Entropy*, vol. 22, no. 2, p. 239, 2020.
- [94] S. Hashemnia, L. Grasse, S. Soni, and M. S. Tata, “Human eeg and recurrent neural networks exhibit common temporal dynamics during speech recognition,” *Frontiers in Systems Neuroscience*, vol. 15, p. 617605, 2021.
- [95] L. Wang, M. Li, and L. Zhang, “Recognize enhanced temporal-spatial-spectral features with a parallel multi-branch cnn and gru,” *Medical & Biological Engineering & Computing*, pp. 1–20, 2023.
- [96] C. Dinh, J. G. Samuelsson, A. Hunold, M. S. Hämmäläinen, and S. Khan, “Contextual meg and eeg source estimates using spatiotemporal lstm networks,” *Frontiers in neuroscience*, vol. 15, p. 552666, 2021.
- [97] D. Dash, P. Ferrari, K. Berstis, and J. Wang, “Imagined, intended, and spoken speech envelope synthesis from neuromagnetic signals,” in *International Conference on Speech and Computer*. Springer, 2021, pp. 134–145.
- [98] X. Zhang, L. Yao, X. Wang, J. Monaghan, D. Mcalpine, and Y. Zhang, “A survey on deep learning-based non-invasive brain signals: recent advances and new frontiers,” *Journal of neural engineering*, vol. 18, no. 3, p. 031002, 2021.
- [99] L. Wiest, “Recurrent neural networks-combination of rnn and cnn-convolutional neural networks for image and video processing-tum wiki,” *Recurrent+ Neural+ Networks+-+ Combination+ of+ RNN+ and+ CNN*, 2017.
- [100] Y. Mohammadi and M. H. Moradi, “Prediction of depression severity scores based on functional connectivity and complexity of the eeg signal,” *Clinical EEG and Neuroscience*, vol. 52, no. 1, pp. 52–60, 2021.
- [101] N. M. N. Leite, E. T. Pereira, E. C. Gurjão, and L. R. Veloso, “Deep convolutional autoencoder for eeg noise filtering,” in *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2018, pp. 2605–2612.

- [102] H. Zhang and D. Wang, “Deep anc: A deep learning approach to active noise control,” *Neural Networks*, vol. 141, pp. 1–10, 2021.
- [103] A. Hirohata, K. Yamada, Y. Nakatani, I.-L. Prejbeanu, B. Diény, P. Pirro, and B. Hillebrands, “Review on spintronics: Principles and device applications,” *Journal of Magnetism and Magnetic Materials*, vol. 509, p. 166711, 2020.
- [104] F. Hooge, T. Kleinpenning, and L. K. Vandamme, “Experimental studies on 1/f noise,” *Reports on progress in Physics*, vol. 44, no. 5, p. 479, 1981.
- [105] P. Wisniowski, J. Almeida, and P. Freitas, “1/f magnetic noise dependence on free layer thickness in hysteresis free mgo magnetic tunnel junctions,” *IEEE Transactions on Magnetics*, vol. 44, no. 11, pp. 2551–2553, 2008.
- [106] Z. Lei, G. Li, W. F. Egelhoff, P. Lai, and P. W. Pong, “Review of noise sources in magnetic tunnel junction sensors,” *IEEE Transactions on Magnetics*, vol. 47, no. 3, pp. 602–612, 2011.
- [107] S. Waldert, H. Preissl, E. Demandt, C. Braun, N. Birbaumer, A. Aertsen, and C. Mehring, “Hand movement direction decoded from meg and eeg,” *Journal of neuroscience*, vol. 28, no. 4, pp. 1000–1008, 2008.
- [108] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen *et al.*, “Meg and eeg data analysis with mne-python,” *Frontiers in neuroscience*, p. 267, 2013.
- [109] W. He, T. Donoghue, P. F. Sowman, R. A. Seymour, J. Brock, S. Crain, B. Voytek, and A. Hillebrand, “Co-increasing neuronal noise and beta power in the developing brain,” *BioRxiv*, p. 839258, 2019.
- [110] M. Gerster, G. Waterstraat, V. Litvak, K. Lehnertz, A. Schnitzler, E. Florin, G. Curio, and V. Nikulin, “Separating neural oscillations from aperiodic 1/f activity: challenges and recommendations,” *Neuroinformatics*, vol. 20, no. 4, pp. 991–1012, 2022.
- [111] S. Sakib, M. M. Fouda, M. Al-Mahdawi, A. Mohsen, M. Oogane, Y. Ando, and Z. M. Fadlullah, “Deep learning models for magnetic cardiography edge sensors implementing noise processing and diagnostics,” *IEEE Access*, vol. 10, pp. 2656–2668, 2022.
- [112] N. K. Upadhyay, H. Jiang, Z. Wang, S. Asapu, Q. Xia, and J. Joshua Yang, “Emerging memory devices for neuromorphic computing,” *Advanced Materials Technologies*, vol. 4, no. 4, p. 1800589, 2019.
- [113] N. J. Aklo, “Simulation action potential signal based on fpga,” *University of Thi-Qar Journal for Engineering Sciences*, vol. 10, no. 1, pp. 163–167, 2019.

- [114] A. Almeida, M. Castel-Branco, and A. Falcão, “Linear regression for calibration lines revisited: weighting schemes for bioanalytical methods,” *Journal of Chromatography B*, vol. 774, no. 2, pp. 215–222, 2002.
- [115] M. Jiao, G. Wan, Y. Guo, D. Wang, H. Liu, J. Xiang, and F. Liu, “A graph fourier transform based bidirectional long short-term memory neural network for electrophysiological source imaging,” *Frontiers in Neuroscience*, vol. 16, p. 867466, 2022.
- [116] B. Sun, T. Guo, G. Zhou, S. Ranjan, Y. Jiao, L. Wei, Y. N. Zhou, and Y. A. Wu, “Synaptic devices based neuromorphic computing applications in artificial intelligence,” *Materials Today Physics*, vol. 18, p. 100393, 2021.
- [117] A. Calimera, E. Macii, and M. Poncino, “The human brain project and neuromorphic computing,” *Functional neurology*, vol. 28, no. 3, p. 191, 2013.
- [118] R.-J. Zhu, Q. Zhao, and J. K. Eshraghian, “Spikegpt: Generative pre-trained language model with spiking neural networks,” *arXiv preprint arXiv:2302.13939*, 2023.
- [119] B. Prashanth and M. R. Ahmed, “Fpga implementation of bio-inspired computing based deep learning model,” in *Advances in Distributed Computing and Machine Learning: Proceedings of ICADCML 2020*. Springer, 2020, pp. 237–245.
- [120] M. Pals, R. J. P. Belizón, N. Berberich, S. K. Ehrlich, J. Nassour, and G. Cheng, “Demonstrating the viability of mapping deep learning based eeg decoders to spiking networks on low-powered neuromorphic chips,” in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2021, pp. 6102–6105.
- [121] D. J. Mannion and A. J. Kenyon, “Artificial dendritic computation: The case for dendrites in neuromorphic circuits,” *arXiv preprint arXiv:2304.00951*, 2023.
- [122] T. Liang, J. Glossner, L. Wang, S. Shi, and X. Zhang, “Pruning and quantization for deep neural network acceleration: A survey,” *Neurocomputing*, vol. 461, pp. 370–403, 2021.
- [123] J. Grollier, D. Querlioz, K. Camsari, K. Everschor-Sitte, S. Fukami, and M. D. Stiles, “Neuromorphic spintronics,” *Nature electronics*, vol. 3, no. 7, pp. 360–370, 2020.
- [124] D. Marković, A. Mizrahi, D. Querlioz, and J. Grollier, “Physics for neuromorphic computing,” *Nature Reviews Physics*, vol. 2, no. 9, pp. 499–510, 2020.
- [125] M. Davies, “Lessons from loihi: Progress in neuromorphic computing,” in *2021 Symposium on VLSI Circuits*. IEEE, 2021, pp. 1–2.

- [126] M. Kang, Y. Lee, and M. Park, “Energy efficiency of machine learning in embedded systems using neuromorphic hardware,” *Electronics*, vol. 9, no. 7, p. 1069, 2020.
- [127] A. Gholami, S. Kim, Z. Dong, Z. Yao, M. W. Mahoney, and K. Keutzer, “A survey of quantization methods for efficient neural network inference,” in *Low-Power Computer Vision*. Chapman and Hall/CRC, 2022, pp. 291–326.
- [128] A. W. Stephan and S. J. Koester, “Spin hall MTJ devices for advanced neuromorphic functions,” vol. 67, no. 2. IEEE, 2020, pp. 487–492.
- [129] S. Soni, G. Verma, H. Nehete, and B. K. Kaushik, “Energy efficient spin-based implementation of neuromorphic functions in cnns,” *IEEE Open Journal of Nanotechnology*, vol. 4, pp. 102–108, 2023.
- [130] M. Davies, A. Wild, G. Orchard, Y. Sandamirskaya, G. A. F. Guerra, P. Joshi, P. Plank, and S. R. Risbud, “Advancing neuromorphic computing with loihi: A survey of results and outlook,” *Proceedings of the IEEE*, vol. 109, no. 5, pp. 911–934, 2021.
- [131] Z. Yu, A. M. Abdulghani, A. Zahid, H. Heidari, M. A. Imran, and Q. H. Abbasi, “An overview of neuromorphic computing for artificial intelligence enabled hardware-based hopfield neural network,” *IEEE Access*, vol. 8, pp. 67 085–67 099, 2020.
- [132] M. E. Nojehdeh, S. Parvin, and M. Altun, “Efficient hardware implementation of convolution layers using multiply-accumulate blocks,” in *2021 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*. IEEE, 2021, pp. 402–405.
- [133] Q. Lou, C. Pan, J. McGuinness, A. Horvath, A. Naeemi, M. Niemier, and X. S. Hu, “A mixed signal architecture for convolutional neural networks,” *ACM Journal on Emerging Technologies in Computing Systems (JETC)*, vol. 15, no. 2, pp. 1–26, 2019.
- [134] X. Lin, H. Cheng, Y. Lu, H. Luo, H. Li, Y. Qian, L. Zhou, L. Zhang, and M. Wang, “Contactless sleep apnea detection in snoring signals using hybrid deep neural networks targeted for embedded hardware platform with real-time applications,” *Biomedical Signal Processing and Control*, vol. 77, p. 103765, 2022.
- [135] S. Tam, M. Boukadoum, A. Campeau-Lecours, and B. Gosselin, “A fully embedded adaptive real-time hand gesture classifier leveraging hd-semg and deep learning,” *IEEE Transactions on Biomedical Circuits and Systems*, vol. 14, no. 2, pp. 232–243, 2020.
- [136] M. R. Azghadi, C. Lammie, J. K. Eshraghian, M. Payvand, E. Donati, B. Linares-Barranco, and G. Indiveri, “Hardware implementation of deep network accelerators

towards healthcare and biomedical applications,” *IEEE Transactions on Biomedical Circuits and Systems*, vol. 14, no. 6, pp. 1138–1159, 2020.

- [137] X. Feng, Y. Jiang, X. Yang, M. Du, and X. Li, “Computer vision algorithms and hardware implementations: A survey,” *Integration*, vol. 69, pp. 309–320, 2019.
- [138] M. Abadi *et al.*, “TensorFlow: Large-scale machine learning on heterogeneous systems,” 2015, software available from [tensorflow.org](https://www.tensorflow.org/). [Online]. Available: <https://www.tensorflow.org/>