UTS UNIVERSITY OF TECHNOLOGY SYDNEY

EEG-Based Stroke Rehabilitation: Enhancing Motor Imagery and Movement Classification

by Htoo Wai Aung

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

under the supervision of Dr. Jiao Jiao Li and Prof. Steven Su

University of Technology Sydney
Faculty of Engineering and IT

June, 2025

Certificate of Authorship / Originality

I, Htoo Wai Aung, declare that this thesis is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Biomedical Engineering at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis. This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Signature: Production Note: Signature removed prior to publication.

Date: June 13, 2025

Abstract

Stroke affects millions worldwide, leading to severe motor and cognitive impairments. Effective rehabilitation is essential but labor-intensive. Robotic exoskeletons integrated with Brain-Computer Interfaces (BCIs) using Electroencephalography (EEG) enable user-driven rehabilitation, reducing therapist workload. However, real-time EEG classification remains challenging due to signal complexity.

This thesis develops EEG_GLT-Net, a spectral Graph Neural Network (GNN) for real-time classification of EEG Motor Imagery (MI) signals at single time points ($\frac{1}{160}$ s). It introduces the EEG Graph Lottery Ticket (EEG_GLT) method, which dynamically constructs adjacency matrices without prior knowledge of EEG channel relationships, improving accuracy and efficiency. Evaluation on PhysioNet shows superior performance over state-of-the-art (SOTA) methods.

Beyond stroke rehabilitation, EEG_GLT is applied to economic forecasting, demonstrating its adaptability. Additionally, EEG Synergistic Gated Network (EEG_SGNet), a CNN-GNN hybrid, enhances window-based EEG classification, validated on BCIC_iv-2a and HGD datasets. Lastly, EEG_RL-Net, a reinforcement learning model, optimises classification by selectively skipping uncertain time points, improving computational efficiency.

These contributions advance EEG-based rehabilitation, enabling intelligent, adaptive systems that enhance stroke recovery and broader neurorehabilitation applications.

Dedication

To my family, for their sacrifices and unwavering belief in me throughout this journey.

Acknowledgements

This PhD journey would not have been possible without the support and guidance of numerous individuals and institutions, to whom I am deeply grateful.

First and foremost, I wish to express my profound gratitude to my research supervisors: Prof. Steven Su, Dr. Jiao Jiao Li, and A/Prof. Adrian Bishop. Their collective expertise and unwavering support have been invaluable throughout this research endeavour. I am especially indebted to Prof. Su and Dr. Li, whose guidance has been instrumental in shaping this research and my development as a researcher. Their combined wisdom, patience, and dedication have made this journey not only possible, but truly rewarding. I am eternally grateful for the opportunity to have worked under their supervision.

Additionally, I extend my sincere thanks to the Commonwealth Scientific and Industrial Research Organisation (CSIRO) supervisory team: Dr. Zili Zhu, Dr. Geoffrey Lee, Dr. Ying Guo, and Dr. Jiaming Li. Their insights and expertise were crucial in this research journey, providing invaluable guidance and support.

I am grateful for the generous financial support that made this research possible. The UTS Research Excellence Scholarship, along with the Australian Government's Research Training Program, provided crucial funding, covering my tuition fees and providing a stipend that allowed me to fully dedicate myself to this work. Additionally, I am thankful to CSIRO Data61 for their research top-up scholarship. This research was further enabled by the UTS Interactive High Performance Computing facility, which provided access to state-of-the-art computing resources and valuable training in machine learning model development and testing.

Finally, I owe a tremendous debt of gratitude to my family and friends. Their unwavering love, encouragement, and understanding throughout these challenging years have been my constant source of strength. While I have attempted to acknowledge all those who have supported me, any omissions are unintentional. This work is dedicated to all who have believed in me.

Htoo Wai Aung
June 13, 2025
Sydney, Australia

Publications

[1] EEG_GLT-Net: Optimising EEG graphs for real-time motor imagery signals classification (Published at Biomedical Signal Processing and Control Journal)

[2] A Real-Time Framework for EEG Signal Decoding with Graph Neural Networks and Reinforcement Learning (Published at IEEE Transactions on Neural Networks and Learning Systems Journal)

[3] Enhancing EEG Motor Imagery Time Point Signal Classification through Reinforcement Learning and Graph Neural Networks (Published at ICMLA 2024 Conference Proceedings)

[4] Revealing Economic Relationship Structures through Causal Inference: A Data-Driven Approach (At under review stage by Journal of Economic Dynamics and Control)

[5] Uncover Temporal Causal Relationships of Major Economic Variables through Peter-Clarke Algorithms (At under review stage by Journal of Computational Economics)

[6] Economic Graph Lottery Ticket: A GNN based Economic Forecasting Model (Published at AJCAI 2024 Conference Proceedings)

[7] Graph Neural Network Framework Based Economic Forecasting (Published at ISF 2024 Conference Presentation)

Statement of Contribution

[1] EEG_GLT-Net: Optimising EEG graphs for real-time motor imagery signals classification (Published at Biomedical Signal Processing and Control Journal)

Statement of Contribution of Others

Jiao Jiao Li, Bin Shi, Yang An, and Steven W. Su. EEG_GLT-Net: Optimising EEG graphs for real-time motor imagery signals classification. Biomedical Signal Processing and Control, vol. 104, p. 107 458, 2025. (**Published**)

I attest that Research Higher Degree candidate Htoo Wai Aung was the first author of this publication.

Production Note:
Signature removed prior to publication.

Jiao Jiao Li

Production Note:
Signature removed prior to publication.

Bin Shi

Production Note:
Signature removed prior to publication.

Yang An

Production Note:
Signature removed prior to publication.

Steven W. Su

[2] A Real-Time Framework for EEG Signal Decoding with Graph Neural Networks and Reinforcement Learning (Published at IEEE Transactions on Neural Networks and Learning Systems Journal)

Statement of Contribution of Others

Jiao Jiao Li, Yang An, and Steven W. Su. **A Real-Time Framework for EEG Signal Decoding with Graph Neural Networks and Reinforcement Learning**. IEEE Transactions on Neural Networks and Learning Systems, TNNLS.2025.3558171 (**Published**)

I attest that Research Higher Degree candidate Htoo Wai Aung was the first author of this publication.

Production Note:
Signature removed prior to publication.

Jiao Jiao Li

Production Note:
Signature removed prior to publication.

Yang An

Production Note:
Signature removed prior to publication.

Steven W. Su

[3] Enhancing EEG Motor Imagery Time Point Signal Classification through Reinforcement Learning and Graph Neural Networks (ICMLA 2024 Conference Proceedings)

Statement of Contribution of Others

Jiao Jiao Li, Yang An, and Steven W. Su. **Enhancing EEG Motor Imagery Time Point Signal Classification through Reinforcement Learning and Graph Neural Networks**. 23rd International Conference on Machine Learning and Applications. (**Published**)

I attest that Research Higher Degree candidate Htoo Wai Aung was the first author of this publication.

Production Note:
Signature removed prior to publication.

Jiao Jiao Li

Production Note:
Signature removed prior to publication.

Yang An

Production Note:
Signature removed prior to publication.

Steven W. Su

[4] Revealing Economic Relationship Structures through Causal Inference: A Data-Driven Approach (Under review stage by Journal of Economic Dynamics and Control)

Statement of Contribution of Others

Ying Guo, Zili Zhu, Jiaming Li, and Geoffrey Lee. **Revealing Economic Relationship Structures through Causal Inference: A Data-Driven Approach**. Journal of Economic Dynamics and Control, JEDC-D-24-00650 (**Submitted**)

I attest that Research Higher Degree candidate Htoo Wai Aung was the first author of this publication.

Production Note:
Signature removed prior to publication.

Ying Guo

Production Note:
Signature removed prior to publication.

Zili Zhu

Production Note:
Signature removed prior to publication.

Jiaming Li

Production Note:
Signature removed prior to publication.

Geoffrey Lee

[5] Uncover Temporal Causal Relationships of Major Economic Variables through Time-variant PC Algorithms (Under review stage by Journal of Computational Economics)

Statement of Contribution of Others

Ying Guo, Zili Zhu, Jiaming Li, and Geoffrey Lee. **Uncover Temporal Causal Relationships of Major Economic Variables through Time-variant PC Algorithms**. Journal of Computational Economics, CSEM-D-25-00729 (**Submitted**)

I attest that Research Higher Degree candidate Htoo Wai Aung was the first author of this publication.

Production Note:
Signature removed prior to publication.

Ying Guo

Production Note:
Signature removed prior to publication.

Zili Zhu

Production Note:
Signature removed prior to publication.

Jiaming Li

Production Note:
Signature removed prior to publication.

Geoffrey Lee

[6] Economic Graph Lottery Ticket: A GNN based Economic Forecasting Model (AJCAI 2024 Conference Proceedings)

Statement of Contribution of Others

Ying Guo, Jiaming Li, and Geoffrey Lee, and Zili Zhu. **Economic Graph Lottery Ticket: A GNN based Economic Forecasting Model**. 37th Australasian Joint Conference on Artificial Intelligence, AI 2024. (**Published**)

I attest that Research Higher Degree candidate Htoo Wai Aung was the first author of this publication.

Production Note:
Signature removed prior to publication.

Ying Guo

Production Note:
Signature removed prior to publication.

Jiaming Li

Production Note:
Signature removed prior to publication.

Geoffrey Lee

Production Note:
Signature removed prior to publication.

Zili Zhu

[7] Graph Neural Network Framework Based Economic Forecasting (ISF 2024 Conference Presentation)

Statement of Contribution of Others

Ying Guo, Jiaming Li, and Geoffrey Lee, and Zili Zhu. **Graph Neural Network Framework Based Economic Forecasting**. 44th International Symposium on Forecasting, Dijon, France. (**Published**)

I attest that Research Higher Degree candidate Htoo Wai Aung was the first author of this publication.

Production Note:
Signature removed prior to publication.

Ying Guo

Production Note:
Signature removed prior to publication.

Jiaming Li

Production Note:
Signature removed prior to publication.

Geoffrey Lee

Production Note:
Signature removed prior to publication.

Zili Zhu

Contents

List of Abbreviations **xxii**

1 Introduction **1**
1.1 Research Hypothesis . . . 2
1.2 Aims and Scope of Thesis . . . 3
1.3 Novelty, Contribution, and Significance . . . 4
1.4 Thesis Overview . . . 5

2 Literature Review **8**
2.1 Stroke . . . 8
2.2 Risk Factors for Stroke . . . 9
2.3 Burden of Stroke . . . 10
2.4 Rehabilitation for Stroke Patients . . . 10
2.5 Human Brain System . . . 14
2.6 EEG Signals . . . 16
2.7 EEG Electrodes Positioning . . . 17
2.8 EEG Signal Amplification . . . 18
2.9 EEG Signal Classification . . . 20
2.9.1 Convolutional Neural Network . . . 21
2.9.2 Graph Convolutional Neural Network . . . 23
2.9.3 Reinforcement Learning . . . 24

3 EEG_GLT-Net: Optimising EEG Graphs for Real-time Motor Imagery Signals Classification **26**
3.1 Introduction . . . 26
3.2 Related Work . . . 27
3.3 Methodology . . . 30
3.3.1 Overview . . . 30
3.3.2 Dataset Description . . . 30
3.3.3 Data Pre-processing and Feature Extraction . . . 31
3.3.4 Graph Preliminary . . . 31

3.3.5 EEG Graph Lottery Ticket (EEG_GLT) . . . 34
3.3.6 General Model Architecture . . . 34
3.3.7 Model Setting . . . 36
3.4 Results and Discussion . . . 40
3.4.1 Geodesic vs PCC Adjacency Matrix Construction Method . . . 40
3.4.2 EEG_GLT Method vs PCC Method in Adjacency Matrix Construction . 44
3.4.3 Model Setting vs Adjacency Matrix Construction Methods . . . 46
3.4.4 MACs Saving using EEG_GLT Method . . . 46
3.4.5 Comparison with Current State-of-the-Art Models . . . 46
3.5 Limitations and Future Works . . . 53
3.6 Conclusion . . . 54
3.7 Significance and Novelty . . . 55

4 Economic Graph Lottery Ticket: A GNN based Economic Forecasting Model 56
4.1 Introduction . . . 56
4.2 Economic Graph Lottery Ticket Algorithm . . . 57
4.3 Real Data Study: Implementation and Evaluation of the EGLT Approach . . . 59
4.3.1 Model Initialisation and Hyper-parameter Settings . . . 60
4.3.2 EGLT Model Implementation . . . 61
4.3.3 Prediction Analysis . . . 63
4.4 Conclusion . . . 67
4.5 Significance and Novelty . . . 68

5 EEG_SGNet: A Parallel Architecture for Advanced EEG Signal Decoding 69
5.1 Introduction . . . 69
5.2 Methodology . . . 72
5.2.1 Overview . . . 72
5.2.2 Dataset Description . . . 73
5.2.3 Dataset Preprocessing . . . 74
5.2.4 Feature Extraction . . . 75
5.2.5 Graph Preliminary . . . 76
5.2.6 Model Setting . . . 78
5.3 Results and Discussion . . . 82
5.3.1 PCC vs EEG_GLT Adjacency Construction Method . . . 82
5.3.2 Impact of CNNs and GNNs Layers . . . 84
5.3.3 Comparison with Current State-of-the-Art Models . . . 91
5.4 Limitations and Future Works . . . 92
5.5 Conclusion . . . 93
5.6 Significance and Novelty . . . 93

6 A Real-time Framework for EEG Signal Decoding with Graph Neural Net-

works and Reinforcement Learning **95**
6.1 Introduction . . . 95
6.2 Methodology . . . 97
6.2.1 Overview . . . 97
6.2.2 Dataset Description and Pre-processing . . . 98
6.2.3 Graph Feature Extraction . . . 100
6.2.4 Problem Redefinition . . . 103
6.2.5 Data Preprocessing and Data Splitting . . . 106
6.2.6 Dueling Deep Q-Learning . . . 107
6.2.7 EEG_RL Algorithm . . . 108
6.2.8 Model Setting and Evaluation Metrics . . . 111
6.3 Results and Discussion . . . 113
6.3.1 EEG_RL-Net vs EEG_GLT-Net . . . 113
6.3.2 Study of Changing r_{right} Values . . . 115
6.3.3 Study of Changing r_{wrong} Values . . . 118
6.3.4 Effects of Episode Length Variation and Optimisation on Classification Performance . . . 122
6.3.5 Comparison with Current State-of-the-Art Models . . . 125
6.4 Limitations and Future Works . . . 127
6.5 Conclusion . . . 127
6.6 Significance and Novelty . . . 128

7 Conclusions and Future Work **129**
7.1 Conclusion . . . 129
7.2 Future Work . . . 130
7.2.1 Rehabilitation Applications . . . 131
7.2.2 Broader Applications Beyond Rehabilitation . . . 132

List of Figures

1.1 Thesis Overview 6

2.1 Ischaemic Stroke [14] 8
2.2 Haemorrhagic Stroke [14] 9
2.3 Lower Limb Exoskeleton Robot [27] 11
2.4 Upper Limb Exoskeleton Robot [28] 12
2.5 General Concept of Controlling Lower Limb Assistive Devices using EEG [29] . 12
2.6 Dorsal View of Human Brain [32] 14
2.7 Sectional View of Human Brain [37] 14
2.8 Measuring EEG Signals on Human Scalp [44] 16
2.9 10-20 EEG Electrodes Positioning System [49] 17
2.10 10-10 EEG Electrodes Positioning System [52] 18
2.11 Ag/AgCl EEG Electrodes [53] 19
2.12 32 Channels EEG Cap with Amplifier [54] 19
2.13 Typical CNN Architecture on Image Classification [57] 21
2.14 Typical GNN Architecture on EEG-based Seizure Detection [69] 23
2.15 Interaction of RL Agent to its Environment [79] 24

3.1 EEG_GLT-Net Model: (a) Overall architecture (classifying EEG MI of one time point $\frac{1}{160}s$ of signals from 64 EEG electrodes). Note that EEG Graph adjacency matrix can be $A^{Geodesic}$, A^{PCC} or A^{EEG_GLT}, (b) Components inside the spectral graph convolution block, (c) Chebyshev spectral graph convolution 30
3.2 EEG graph (m_g) pruning using Algorithm 1: At each N_{ep} iteration, the bottom p_g% are pruned, reducing density from 100% until the lowest density s_g%. Solid lines indicate remaining edges, while red-dashed lines depict removed edges . . . 34
3.3 Geodesic Distance Adjacency Matrix ($A^{Geodesic}$) 38
3.4 PCC Adjacency Matrix (A^{PCC}) of Subject S_6 and S_{14} 39
3.5 Representations of $m_{g\_EEG\_GLT}$ for Subject S_6 at 13.39% Density. (a) Adjacency Matrix - Model A (Accuracy: 78.13%) (b) Graph - Model A (c) Adjacency Matrix - Model E (Accuracy: 73.55%) (d) Graph - Model E 39
3.6 Comparison of Model Accuracy Across Different Adjacency Matrix Construction Methods 43

3.7 Performance Accuracy Across Different m_g Densities Using Different Models for Subject S_1 . . . 48
3.8 Performance Accuracy Across Different m_g Densities Using Different Models for Subject S_3 . . . 48
3.9 Performance Accuracy Across Different m_g Densities Using Different Models for Subject S_6 . . . 49
3.10 Performance Accuracy Across Different m_g Densities Using Different Models for Subject S_{12} . . . 49
3.11 Performance Accuracy Across Different m_g Densities Using Different Models for Subject S_{14} . . . 50
3.12 Performance Accuracy Across Different m_g Densities Using Different Models for Subject S_{15} . . . 50
3.13 EEG_GLT Adjacency matrix mask (m_g) of Subject S_6 at different densities using Model A. The m_g density at 20.49% produces the highest accuracy of 78.13% . . 51
3.14 EEG_GLT Adjacency matrix mask (m_g) of Subject S_{14} at different densities using Model A. The m_g density at 13.39% produces the highest accuracy of 79.06% . . 52

4.1 Flowchart of a multiple layer EGLT approach . . . 58
4.2 The implemented two-layer EGLT model for eight economic variables . . . 60
4.3 RMSE losses at different adjacency matrix densities, κ^s . . . 61
4.4 The optimal adjacency matrix of the EGLT model, m_{g_EGLT}, is found at 21.88% density with Necessary Edges only. (a) Heatmap of the Adjacency Matrix. (b) Directed Nodes-Edges Graph representation of the Adjacency Matrix. . . . 62
4.5 CPI Prediction Result . . . 63
4.6 Wage Prediction Result . . . 64
4.7 Short-term Interest Rate Prediction Result . . . 64
4.8 Long-term Interest Rate Prediction Result . . . 65
4.9 Domestic Bond Index Prediction Result . . . 65
4.10 International Bond Index Prediction Result . . . 66
4.11 Unemployment Rate Prediction Result . . . 66
4.12 Housing Price Index Prediction Result . . . 67

5.1 Overall Architecture of EEG_SGNet Model . . . 72
5.2 Electrode positions of BCIC_iv-2a dataset[145]. Left: Electrode montage corresponding to the international 10-20 system. Right: Electrode montage of the three monopolar EOG channels . . . 73
5.3 Boxplot of Average Accuracy Across Subjects for Different Adjacency Construction Methods Across Model Settings on the BCIC_iv-2a Dataset . . . 83
5.4 Maximum Accuracy Achieved with Each Adjacency Construction Method Across Different Model Settings for Subjects S_1 to S_5 on the BCIC_iv-2a Dataset . . . 83

5.5 Maximum Accuracy Achieved with Each Adjacency Construction Method Across Different Model Settings for Subjects S_6 to S_9 on the BCIC_iv-2a Dataset 84
5.6 Boxplot of Average Accuracy Across Subjects for Different Adjacency Construction Methods Across Model Settings on the HGD Dataset 85
5.7 Maximum Accuracy Achieved with Each Adjacency Construction Method Across Different Model Settings for Subjects S_1 to S_7 on the HGD Dataset 86
5.8 Maximum Accuracy Achieved with Each Adjacency Construction Method Across Different Model Settings for Subjects S_8 to S_{14} on the HGD Dataset 86

6.1 (a) EEG_GLT-Net model [144]: Overall architecture (classifying EEG MI of one time point $\frac{1}{160}s$ of signals from 64 EEG electrodes), (b) EEG_RL-Net model (our proposed): Incorporation of an optimal pre-trained EEG_GCN Block at a 13.39% m_g density from the EEG_GLT-Net [144], coupled with an EEG_RL Block 98
6.2 Electrode positions of PhysioNet dataset[127] 99
6.3 Agent interaction with the EEG_RL Environment. The agent transitions to the next state within the current episode after performing the skipping action ($a = 0$), receiving a reward of r_{skip}. The agent transitions to the terminal state after performing a classifying action ($a \neq 0$), receiving a reward of r_{right} for a correct classification or r_{wrong} for an incorrect classification 104
6.4 Conversion of EEG MI time points into states using the pre-trained EEG_GCN Block, grouped into episodes comprising 20 states each. 106
6.5 EEG_RL Block: Featuring the Dueling DQN, this component predicts the q-values linked to various actions 109
6.6 Performance Mean Accuracy Comparison between EEG_GLT-Net (PCC Adj), EEG_GLT-Net (m_{g_GLT} Adj) and EEG_RL-Net Methods for Subjects S_1 - S_{10} . 113
6.7 Performance Mean Accuracy Comparison between EEG_GLT-Net (PCC Adj), EEG_GLT-Net (m_{g_GLT} Adj) and EEG_RL-Net Methods for Subjects S_{11} - S_{20} . 114
6.8 Overall Performance Mean Accuracy Comparison between EEG_GLT-Net (PCC Adj), EEG_GLT-Net (m_{g_GLT} Adj) and EEG_RL-Net Methods 114
6.9 Performance Mean Accuracy Comparison Varying r_{right} Values for Subjects S_1 - S_{10} 116
6.10 Performance Mean Accuracy Comparison Varying r_{right} Values for Subjects S_{11} - S_{20} 116
6.11 Overall Performance Mean Accuracy Comparison Varying r_{right} Values 117
6.12 Overall Mean Classification Time (Time Points) Comparing Varying r_{right} Values 118
6.13 Performance Mean Accuracy Comparison Varying r_{wrong} Values for Subjects S_1 - S_{10} 119
6.14 Performance Mean Accuracy Comparison Varying r_{wrong} Values for Subjects S_{11} - S_{20} 119
6.15 Overall Performance Mean Accuracy Comparison Varying r_{wrong} Values 120
6.16 Overall Mean Classification Time (Time Points) Comparing Varying r_{wrong} Values 121

6.17 Overall Mean Classification Time (Time Points) Comparing Varying Horizon (H) Values 122
6.18 Performance Mean Accuracy Comparison between EEG_GLT-Net (PCC Adj), EEG_GLT-Net (m_{g_GLT} Adj) and EEG_RL-Net Methods at Optimal Reward Settings for Subjects S_1 - S_{10} 123
6.19 Performance Mean Accuracy Comparison between EEG_GLT-Net (PCC Adj), EEG_GLT-Net (m_{g_GLT} Adj) and EEG_RL-Net Methods at Optimal Reward Settings for Subjects S_{11} - S_{20} 123
6.20 Overall Performance Mean Accuracy Comparison between EEG_GLT-Net (PCC Adj), EEG_GLT-Net (m_{g_GLT} Adj) and EEG_RL-Net Methods at Optimal Reward Settings 124

List of Tables

2.1 Generalised EEG Signal Processing Methods 21

3.1 Generalised Architecture of EEG_GLT-Net Model 36
3.2 Model Settings of EEG_GLT-Net 37
3.3 Hyperparameter Settings for Training EEG_GLT-Net 38
3.4 Accuracy Comparison of Different Adjacency Matrix Construction Methods for Subjects S_1 to S_{10} 40
3.5 Accuracy Comparison of Different Adjacency Matrix Construction Methods for Subjects S_{11} to S_{20} 41
3.6 Accuracy Comparison Across Different Methods of Adjacency Matrix Construction for Each Model 42
3.7 Optimal EEG_GLT Adjacency Matrix ($m_{g\_EEG\_GLT}$) Density of Each Subject Across Models 45
3.8 MACs Savings (%) for Each Subject: PCC's Best Model Accuracy vs. EEG_GLT Accuracy from Models with Adjacency Matrix Densities Just Surpassing PCC's Best Accuracy 47
3.9 Performance Comparisons with State-of-the-Art Models 53

4.1 Implementation Details of the 2-layer EGLT Model 60
4.2 Comparison of the RMSE losses for the SUPA and the EGLT algorithm across various economic variables. 63

5.1 Generalised Architecture of EEG_SGNet Model 80
5.2 Model Settings of EEG_SGNet 81
5.3 Hyperparameter Settings for Training EEG_SGNet 81
5.4 Maximum Accuracy Achieved by Various Methods for Each Subject on the BCIC_iv-2a Dataset 82
5.5 Maximum Accuracy Achieved by Various Methods for Each Subject on the HGD Dataset 85
5.6 Average Accuracy and F1 Score Across Subjects for Models A to E with Different Adjacency Matrix Construction Methods on the BCIC_iv-2a Dataset 87

5.7 Average Accuracy and F1 Score Across Subjects for Models F to I with Different Adjacency Matrix Construction Methods on the BCIC_iv-2a Dataset 88
5.8 Average Accuracy and F1 Score Across Subjects for Models A to E with Different Adjacency Matrix Construction Methods on the HGD Dataset 89
5.9 Average Accuracy and F1 Score Across Subjects for Models F to I with Different Adjacency Matrix Construction Methods on the HGD Dataset 90
5.10 Performance Comparison of EEG_SGNet Model with State-of-the-Art Models on BCIC_iv-2a Dataset 91
5.11 Performance Comparison of EEG_SGNet Model with State-of-the-Art Models on HGD Dataset 92

6.1 Details of EEG_GLT-Net [144] Model 102
6.2 Hyperparameter Settings for Training EEG_GLT-Net [144] 103
6.3 Details of EEG_RL-Net Model 110
6.4 Hyperparameter Settings for Training EEG_RL-Net 111
6.5 Accuracy Assessment: EEG_RL-Net versus EEG_GLT-Net 115
6.6 Impact of Varying r_{right} Values on Accuracy and Classification Time 117
6.7 Impact of Varying r_{wrong} Values on Accuracy and Classification Time 120
6.8 Impact of Varying Episode Lengths (H) Values on Accuracy, F1 Score and Classification Time 122
6.9 Subject-wise Classification Accuracy and Time with Optimal Reward Settings and Episode Lengths 124
6.10 Performance Comparison of EEG_RL-Net Model with State-of-the-Art Models on PhysioNet Dataset 125
6.11 Performance Comparison of EEG_RL-Net Model with State-of-the-Art Models on HGD Dataset 126

List of Abbreviations

10-10	International 10-10 System (EEG electrode positioning)
10-20	International 10-20 System (EEG electrode positioning)
10-5	International 10-5 System (EEG electrode positioning)
ADC	Analog-to-Digital Converter
Ag/AgCl	Silver/Silver Chloride
AI	Artificial Intelligence
AIHW	Australian Institute of Health and Welfare
ANNs	Artificial Neural Networks
ASX200	Australian Securities Exchange 200 Index
BCIC_IV-2a	BCI Competition IV Dataset 2a
BCI	Brain-Computer Interface
BN	Batch Normalisation
BRIs	Brain-Robot Interfaces
CNN	Convolutional Neural Network
CP3, CP4, CPz	Centroparietal (EEG electrodes)
CPI	Consumer Price Index
CSIRO	Commonwealth Scientific and Industrial Research Organisation
CSP	Common Spatial Pattern
CT	Computerised Tomography
DALY	Disability-Adjusted Life Years
DQN	Deep Q-Network
ECG	Electrocardiogram
ECoG	Electrocorticography
EEG	Electroencephalography
EEG_GLT-Net	EEG Graph Lottery Ticket Network
EEG_RL-Net	EEG Reinforcement Learning Network
EEG_SGNet	EEG Synergistic Gated Network
EGLT	Economic Graph Lottery Ticket
ELU	Exponential Linear Unit
EMG	Electromyography
EOG	Electrooculography

ERP	Event-Related Potential
fMRI	functional Magnetic Resonance Imaging
fNIRS	functional Near-infrared Spectroscopy
FC	Fully Connected Layer
FN	Number of features (in model, also False Negative)
FP	False Positive
Fp1, Fp2, Fpz	Frontal pole (EEG electrodes)
GCN	Graph Convolutional Network
GDP	Gross Domestic Product
GLT	Graph Lottery Ticket
GNN	Graph Neural Network
HCI	Human-computer Interaction
HGD	High-Gamma Movement Dataset
HPI	Housing Price Index
ICH	Intracerebral Hemorrhage
LDA	Linear Discriminant Analysis
LSTM	Long Short-term Memory
MACs	Multiply-Accumulate Operations
MEG	Magnetoencephalography
MI	Motor Imagery
ML	Machine Learning
MLP	Multi-Layer Perceptron
MSE	Mean Squared Error
MRI	Magnetic Resonance Imaging
NHS	National Health Survey
O1, O2, Oz	Occipital (EEG electrodes)
P3, P4, Pz	Parietal (EEG electrodes)
PCC	Pearson Correlation Coefficient
PhysioNet	PhysioNet EEG Motor Imagery Dataset
PO3, PO4, POz	Parieto-Occipital (EEG electrodes)
PLV	Phase-locking Value
PSD	Power Spectral Density
r_{right}	Positive reward for correct classification
r_{wrong}	Negative reward for wrong classification
ReLU	Rectified Linear Unit
RL	Reinforcement Learning
RMSE	Root Mean Squared Error
SAH	Subarachnoid Hemorrhage
SNR	Signal to Noise Ratio
SOTA	State-of-the-Art

SVM	Support Vector Machine
T3, T4, T7, T8	Temporal (EEG electrodes)
TN	True Negative
TP	True Positive
UGS	Unified GNN Sparsification Technique
UTS	University of Technology Sydney
WSO	World Stroke Organisation

Chapter 1

Introduction

Stroke attacks a human control centre which is the brain that can change one's life instantly. Stroke is a classically characterised as a neurological deficit attributed to the Central Nervous System (CNS) due to vascular cause including intracerebral haemorrhage (ICH), cerebral infraction and subarachnoid haemorrhage (SAH) [1]. In simple terms, human brain cells die when there is a lack of nutrients and oxygen that are normally carried by arteries via blood. Blocked or busting of arteries causes stroke that kills human brain cells at a rate of 1.9 million every 60 seconds [2]. Depending on the location of the brain strikes and the severity, every stroke differs. When a stroke happens, it damages the human body functions such as sensations and movement, cognitive ability, eating and swallowing, speech, perception of surroundings, emotional control, and sexual ability [3].

Stroke impacts on day-to-day activities of humans. Therefore, the increment of developing better quality and quantity of rehabilitation is imperative for stroke patients. Clinical treatment only is not enough to restore a stroke patient's physical, psychological and neurological abilities. Thus, rehabilitation is required to complement the clinical treatment in the restoration of stroke patients' abilities [4].

The rehabilitation treatment is a labour-intensive task, especially for the lower limb recovery treatment that requires a minimum of three therapists to manually support a patient's legs and torso to do the training [5]. An exoskeleton is a mechanical wearable device that produces torque by actuators and is transferred to the human body to assist in the user's movement intention. Thus, it reduces the amount of therapist involvement in training. Per [6], a robotic exoskeleton can be divided into categories based on their serving purposes like enhancing human mechanical performances, assisting people with movement disorders and rehabilitating stroke patients. Also, based on the part of the human body involved, the exoskeleton can be divided into three based on the part of the human body involved such as an upper limb, a lower limb and specific joints like the elbow, knee, ankle joints, shoulder, etc [6].

A Brain-computer Interface (BCI) is a system that is able to acquire brain signals and analyse

them. Then, the analysed signals are converted into commands which are relayed to output devices to perform desired actions [7]. In the BCI system, the acquired brain signals are trained using machine learning algorithms with quality labels for desired actions. Machine learning is a subset of Artificial Intelligence (AI), which imitates the way humans learn via the use of data and algorithms without being explicitly programmed [8], [9].

It's important to focus on the development of state-of-the-art control system techniques and algorithms of exoskeletons to improve the accuracy and performance for users. Human electrophysiological signals measured from the human body can control the exoskeleton robots using the user's motion intention [10]. Electroencephalography (EEG) is one of the most popular methods to acquire brain signals from human scalps via electrodes to analyse some brain diseases because the method is non-invasive and convenient [11]. It can also be used in controlling exoskeleton robots.

1.1 Research Hypothesis

This thesis is based on the hypothesis that:

> Advancing EEG-based stroke rehabilitation through optimised graph learning, parallel deep learning architectures, and reinforcement learning can enhance motor imagery and movement classification accuracy while reducing computational complexity, enabling real-time and adaptive rehabilitation strategies.

To validate this hypothesis, the following chapter-specific hypotheses are proposed:

- An optimised adjacency matrix for EEG-based graph learning can improve motor imagery classification accuracy while reducing computational requirements, enabling real-time applications in stroke rehabilitation.

- The EEG_GLT-based adjacency matrix construction method can generalise beyond EEG applications, improving the prediction accuracy of economic forecasting models.

- A parallel deep learning framework combining convolutional and graph-based learning with EEG_GLT-based adjacency matrix construction method improves EEG-based motor imagery and movement classification, contributing to more effective stroke rehabilitation strategies.

- Reinforcement learning-based decision-making in EEG signal decoding can improve classification accuracy and response efficiency, enhancing real-time adaptability for stroke rehabilitation.

1.2 Aims and Scope of Thesis

The primary focus of this PhD study is to develop effective and efficient predictive models for classifying EEG signals based on user intentions, with potential applications in rehabilitation facilities. The following aims are set to be accomplished by the end of this PhD study:

- To develop a Graph Neural Network (GNN) based model for classifying real-time EEG Motor Imagery (MI) signals, enhancing the accuracy of user intention detection.
- To evaluate the performance of the model using single time-point EEG MI signals ($\frac{1}{160}$ seconds) from publicly available datasets.
- To design an algorithm for constructing adjacency matrices tailored to individual subjects and model settings, without requiring prior knowledge of relationships between EEG channels, thereby improving both performance accuracy and computational efficiency.
- To create a parallel architecture combining Convolutional Neural Network (CNN) and GNN models to classify window-based EEG signals, where temporal information is critical.
- To develop a Reinforcement Learning (RL) based model for classifying real-time EEG MI signals, where the RL agent can decide whether to classify or skip a time point to achieve higher performance accuracy.
- To investigate optimal reward settings for the RL agent to achieve high classification accuracy with fast inference speeds.
- To evaluate the performance of the models using publicly available datasets, including both an EEG MI dataset and an EEG High-Gamma movement (HGD) dataset.

During the fulfillment of the above aims, the following research questions were raised:

- What is the influence of the spectral GNN filter order on the accuracy of EEG MI time-point classification?
- How can adjacency matrices be constructed to minimise computational requirements without compromising performance accuracy?
- What are the effects of different adjacency matrix construction methods, under the same model settings, on performance accuracy and computational efficiency?
- What are the optimal adjacency matrix densities for each subject?
- What features are extracted from window-based EEG signals for the CNN and GNN pathways in the parallel model architecture?
- How does the importance coefficient of the pathway (α) influence performance in the parallel model architecture?

- How does increasing the reward magnitude for correct decisions made by the RL agent affect overall performance accuracy and inference speed?
- How does increasing the penalty magnitude for incorrect decisions made by the RL agent affect overall performance accuracy and inference speed?
- What is the optimal number of time points per episode for the RL agent to achieve optimal performance?
- What are the optimal reward settings and model configurations for the RL agent to achieve the best performance?

1.3 Novelty, Contribution, and Significance

This PhD research proposes novel methods and models to estimate user movement intentions via motor imagery and physical movement based on EEG signals, with the goal of improving BCI human-machine interactions. These advancements focus on enhancing both performance accuracy and computational efficiency for applications such as rehabilitation robot control. The key novel aspects and significant contributions of this research are as follows:

A. **Development of a spectral GNN-based model for real-time EEG MI signal classification:**
   - Proposed the EEG_GLT-Net, a spectral GNN-based model capable of classifying EEG MI signals at single time points ($\frac{1}{160}$ seconds).
   - Designed optimal EEG_GLT-Net configurations for individual subjects, including the number of GNN layers, model parameters, and GNN filter orders.

B. **Introduction of a state-of-the-art adjacency matrix construction method:**
   - Proposed the EEG_GLT method, an innovative approach for constructing adjacency matrices without requiring prior knowledge of relationships between EEG channels.
   - Demonstrated the method's effectiveness in improving performance accuracy and computational efficiency when integrated with GNN-based models.

C. **Creation of a parallel CNN and GNN-based model for window-based EEG MI and movement signal classification:**
   - Proposed the EEG_SGNet, a novel model combining the strengths of CNN and GNN architectures to classify window-based EEG MI and movement signals, achieving enhanced performance accuracy.

D. **Development of an RL-based model for real-time EEG MI signal classification:**
   - Proposed the EEG_RL-Net, a novel framework where an RL agent classifies single time-point EEG MI signals based on extracted graph features and has the ability to

skip classification when sufficient confidence is not reached.

- Demonstrated the effectiveness of the EEG_RL-Net in achieving high accuracy and efficiency by optimising reward settings to enable the RL agent to make correct decisions within minimal time.

By integrating high performance, computational efficiency, and versatility, this research addresses critical challenges in EEG signal processing and graph-based learning. It demonstrates significant advancements in real-time EEG classification, enabling precise time-point-level analysis for applications like rehabilitation robotics. The domain-agnostic adjacency matrix construction method further enhances adaptability, paving the way for scalable graph-based systems. These contributions not only advance neurorehabilitation technologies but also open new research directions in dynamic graph learning, time-sensitive signal classification, multimodal physiological data integration, and domain-agnostic applications in areas such as economic forecasting, climate modeling, and network optimisation. This work establishes a foundation for innovation across healthcare, engineering, and beyond.

1.4 Thesis Overview

Chapter 2 provides a comprehensive literature review, outlining the key interdisciplinary areas relevant to this study. It begins with an overview of stroke, highlighting its global prevalence, associated risk factors, and the challenges faced in rehabilitation. The fundamentals of EEG are then introduced, covering its role in capturing brain activity, the principles of EEG acquisition, and commonly used signal processing techniques. The chapter also examines BCIs, emphasising their significance in enabling communication between the brain and external devices, particularly in rehabilitation contexts. A review of advancements in rehabilitation robotics follows, with a focus on the integration of BCIs and EEG-driven systems to support patients with motor impairments. Additionally, machine learning approaches such as CNNs, GNNs, and RL are discussed as the methodological foundation for this research.

Chapter 3 introduces the EEG_GLT-Net model, designed to classify real-time EEG MI signals. This model is based on a spectral GNN and can classify EEG MI signals at a single time point of $\frac{1}{160}$ seconds. EEG_GLT-Net has potential applications in rehabilitation facilities, where real-time classification of EEG MI signals is essential. As part of this project, an innovative method for constructing adjacency matrices, termed EEG_GLT, will be presented. This method does not rely on prior knowledge of the relationships between EEG channels and dynamically optimises these relationships, making it adaptable to both subject-specific data and GNN models. The publicly available PhysioNet EEG MI dataset will be used for evaluation, focusing on data from 20 subjects. The EEG_GLT method will be compared against other adjacency matrix construction approaches, including the Geodesic method and the Pearson Correlation Coefficient (PCC). Additionally, six model frameworks, varying in the number of GCN filters, filter polynomial orders, and fully connected (FC) hidden nodes, will be evaluated. Finally, the

performance of EEG_GLT-Net will be benchmarked against seven state-of-the-art methods.

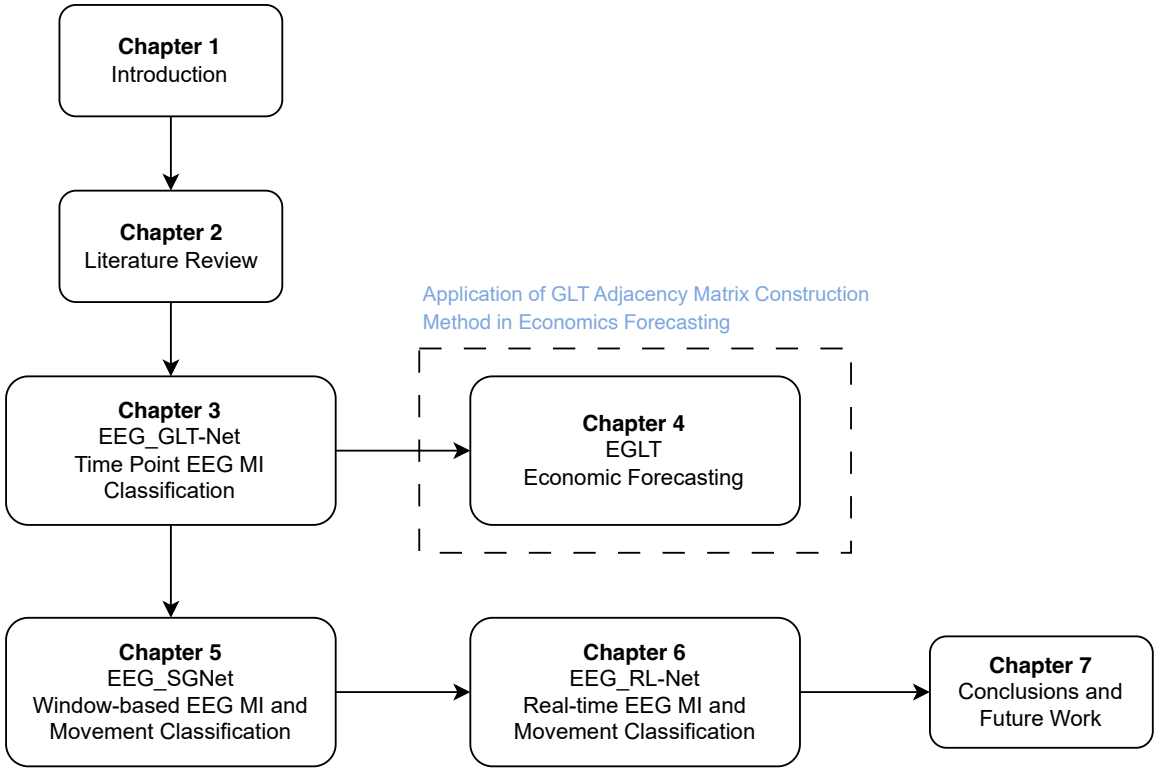


Figure 1.1: Thesis Overview

The EEG_GLT adjacency matrix construction method is not limited to EEG MI signal classification; it also has applications in economic forecasting. Chapter 4 explores its use in a GNN framework to forecast eight economic variables without requiring prior knowledge of their interrelationships. These variables include the consumer price index (CPI), wage index, short-term interest rate, long-term interest rate, domestic bond index, international bond index, unemployment rate, and housing price index, based on 20 years of data spanning from June 2000 to July 2020. The model predicts the next quarter's values for all variables using data from the previous three quarters while simultaneously determining the optimal relationships between the variables. Finally, the predictions for each variable are compared in detail with the outputs of the state-of-the-art SUPA model, which was developed using industry expertise.

Chapter 5 introduces EEG_SGNet, a novel framework for classifying EEG signals. Unlike the previously introduced models, EEG_SGNet is designed to classify window-based EEG signals, where temporal information plays a critical role. The framework consists of parallel CNN and GNN pathways. Various configurations, including the number and sizes of filters, as well as the number of CNN and GNN layers, are explored to achieve optimal results. The project utilises two datasets: the BCIC_iv-2a dataset, an EEG MI dataset, and the HGD dataset, an EEG movement dataset. This dual-dataset approach demonstrates that EEG_SGNet is effective for

both motor imagery and movement datasets, highlighting its versatility. Furthermore, it shows the model's potential applicability in rehabilitation robotics, where high classification accuracy is essential. Finally, EEG_SGNet's performance is benchmarked against ten state-of-the-art models for the BCIC_iv-2a dataset and eight state-of-the-art models for the HGD dataset.

Chapter 6 introduces the EEG_RL-Net model, a novel framework for classifying EEG MI signals. The model integrates a GNN module to extract graph features from single time points and a DQN-based RL framework for classification. By enabling the RL agent to skip classification at uncertain time points, the model enhances both accuracy and efficiency. Various configurations, such as reward strategies and episode lengths, are tested to optimise performance for each subject. The model's evaluation is conducted using the publicly available PhysioNet and HGD dataset, and its performance is benchmarked against state-of-the-art models, including EEG_GLT-Net and EEG_SGNet, demonstrating its advancements in real-time EEG signal classification.

Chapter 2

Literature Review

2.1 Stroke

Stroke is one of the most common diseases that about every one in four persons is affected over their lifetimes. Also, it is the third leading cause of disability and the second cause of death globally [12]. Internal carotid arteries and vertebral arteries are two sources that are supplying blood to the brain. The anterior and middle cerebral arteries are two major branches of internal carotid arteries [13]. Fundamentally, two types of strokes are Ischaemic stroke and Haemorrhagic stroke.

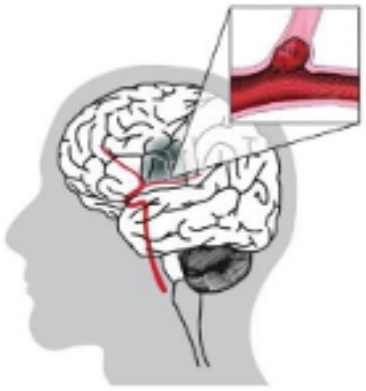


Figure 2.1: Ischaemic Stroke [14]

An Ischaemic stroke happens when there is a blockage anywhere along the arteries to the brain. The latter type, Haemorrhagic stroke is due to the breaks or bursts in the brain. Ischaemic stroke is more common than Haemorrhagic stroke which accounts for about 60 - 90% of the stroke. Only 10 - 40% of strokes are Haemorrhagic stroke type [15], [16]. Since the management of Ischaemic stroke is different from Haemorrhage, reliable differentiation of the type of stroke is imperative. The clinical symptoms and signs are not sufficient to reliably distinguish stroke subtype and mechanism. Thus, brain scanning technologies such

as Computerised Tomography (CT), Magnetic Resonance Imaging (MRI), etc. are crucial to complement the clinical examination [17].

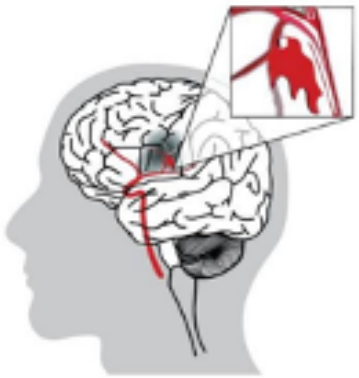


Figure 2.2: Haemorrhagic Stroke [14]

2.2 Risk Factors for Stroke

Stroke is becoming an increasingly common problem because of people's lifestyles (high-fat diet, lack of exercise, etc.). Per the National Health Survey (NHS) data [18], six risk factors associated with strokes are high cholesterol, overweight and obesity, smoking, high blood pressure, atrial fibrillation, and physical inactivity. The limit of high cholesterol is 5.5 mmol/L. Similarly, when a person's blood pressure exceeds 140/90 mmHg, the person has high blood pressure. If a person's Body Mass Index is between 25 and 30, the person is overweight. However, the person is obese once the BMI level exceeds 30. The factor of physical inactivity is that there was no physical activity in the week before having a stroke.

The report of the World Stroke Organisation (WSO) stated that 89% of global stroke deaths and disabilities resided in the low-to-middle-income countries [19]. In Australia, the native Aboriginals and Torres Strait Islanders are 1.7 times more probable to be hospitalised due to stroke than other ethnicities [20]. These are due to the higher rate of smoking, physical inactivity, and obesity compared to non-Aboriginal and Torres Strait Islander counterparts. In the rate of mortality due to stroke, per AIHW, Aboriginal and Torres Strait Islanders have a greater risk of about 1.3 times compared to non-Aboriginal and Torres Strait Islanders [21].

Public health initiatives aim to address these risk factors through education, screening, and lifestyle interventions. However, despite these efforts, the global incidence of stroke continues to rise [22]. While preventive measures are crucial, they have not been universally effective in reducing the overall burden of stroke. This highlights the importance of post-stroke rehabilitation as a complementary approach to addressing the long-term impacts of the disease. Rehabilitation efforts, such as those explored in this thesis, are particularly critical for improving recovery outcomes after stroke. By focusing on rehabilitation using EEG systems for both motor imagery and motor movement, this research aims to enhance survivors' recovery

potential, helping them regain functional independence. These innovations are not only vital for individuals but also play a role in alleviating the broader societal and economic burdens associated with stroke.

2.3 Burden of Stroke

Stroke imposes a significant burden on individuals, families, and healthcare systems worldwide. In Australia, over 445,087 people were living with stroke in 2020, with 27,428 new cases reported that year—equivalent to one stroke every 19 minutes [23]. Globally, there were 12.2 million new cases in 2019, equating to one stroke every 3 seconds, with 101 million people living with the effects of stroke—nearly double the 1990 figure of 54.7 million [19].

Stroke mortality remains high, with 8,703 deaths in Australia in 2020 and a projected 2.5-fold increase by 2050 [3], [23]. Worldwide, stroke claims 6.5 million lives annually and accounts for the loss of over 143 million Disability-adjusted Life Years (DALYs) each year. Beyond its health impact, stroke has immense economic consequences. The global cost of stroke in 2017 was estimated at USD 451 billion (0.36% of global GDP), while the Australian economic cost of stroke in 2020 was $6.2 billion, including $2.9 billion in productivity losses, $1.3 billion in healthcare costs, and $0.7 billion in caregiving expenses [3], [19].

In 2020, the federal government bore 40% of stroke-related costs, with individuals contributing 24%, followed by society (16%), employers (7%), and state governments (6%) [3]. The burden is expected to rise by 45.7% by 2050, increasing stroke incidence to one every 10 minutes. Welfare payments, including Carer Allowance and Disability Support Pension (accounting for 85% of welfare payments), highlight the societal cost, with $120.8 million allocated to support 6,595 Australians affected by stroke [24].

Given these challenges, improving stroke rehabilitation is critical. Enhancing recovery efficiency could significantly reduce disability, cut lifetime management costs, and alleviate strain on caregivers and public resources. Even a 10% improvement in rehabilitation efficiency could save over $600 million in Australia and billions globally. This thesis contributes by developing EEG-based rehabilitation systems to optimise recovery pathways, offering the potential to enhance survivors' quality of life, reduce societal costs, and enable more effective allocation of healthcare resources.

2.4 Rehabilitation for Stroke Patients

When a stroke happens, it damages the human body functions such as sensations and movement, cognitive ability, eating and swallowing, speech, perception of surroundings, emotional control, and sexual ability [3]. [25] suggested that better results in the chronic stages of recovery can be expected if recovery takes place early during the acute stage. Stroke rehabilitation involves therapies and activities to assist the human brain to relearn the skills that are affected

by stroke. The rehabilitation trainings are to restore dysfunctional neural pathways or create new brain pathways to perform the same tasks before the stroke using different parts of the brain [26].

Figure 2.3: Lower Limb Exoskeleton Robot [27]

Usually, the task of motor rehabilitation is practiced manually by therapists at rehabilitation centers or hospitals. Those tasks demand intensive work, especially for lower limb treatment. It requires three therapists' inputs to manually support the patient's torso and legs to perform training [5]. The experience of therapists also plays a crucial role in traditional manual rehabilitation, as it would require more resources in repetitive and high-intensity training. Since the last decade, wearable robots (exoskeleton robotic devices) have been developed to assist in labour-intensive rehabilitation systems.

The robotic exoskeleton can be fundamentally divided into three types based on their serving purpose [6]. The first type of exoskeleton robot is to enhance human mechanical performances such as lifting and handling heavy items that are physically challenging for humans. They are mainly used in areas such as manufacturing industries, military applications, or development sites. The second category of exoskeleton robots is to assist people with movement disorders because of muscle weakness, stroke, or injury. The third one is therapeutic exoskeletons which are primarily used in the area of rehabilitation.

Figure 2.4: Upper Limb Exoskeleton Robot [28]

Based on the part of the human body to cater for, the exoskeleton robot can be classified into three that are upper limb exoskeleton robot, lower limb exoskeleton robot, and specific joint exoskeleton robots such as the knee, ankle, elbow, shoulder, etc. In the upper limb exoskeleton robot, it can be further classified into two categories such as exoskeletons and prostheses [6]. The former type, exoskeletons or orthoses are the orthopaedic equipment to assist the person with disabilities and recover the functions of the affected limb. The latter type, prostheses are to replace a missing human body part. The lower limb exoskeleton robots are mechatronic system devices to assist in the recovery of standing or walking. The lower limb exoskeletons can be further classified into two groups. The first group of devices is to serve as human gait trainers to recover the human walking pattern. The second group is the signal joint exoskeleton robot.

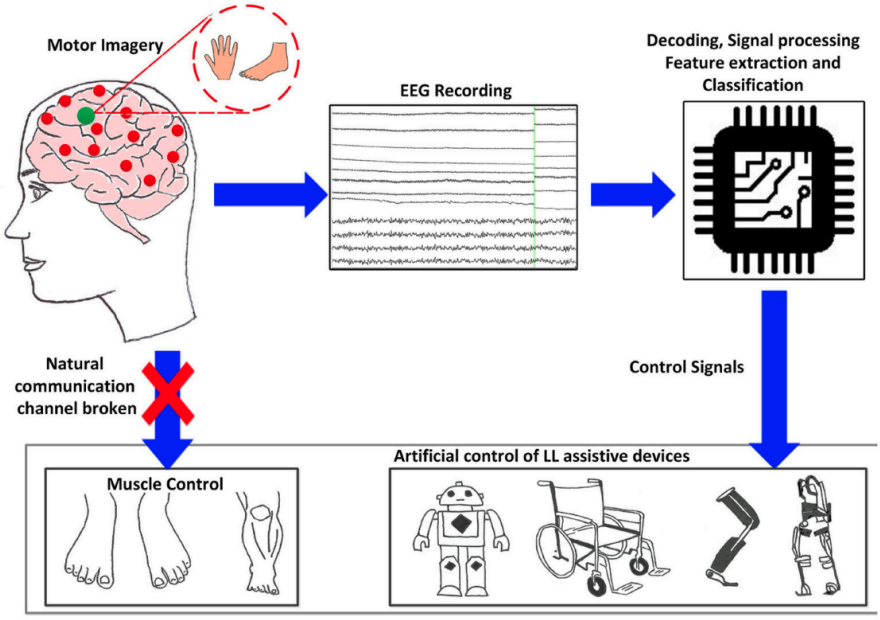


Figure 2.5: General Concept of Controlling Lower Limb Assistive Devices using EEG [29]

Alongside the advanced exoskeleton robot development, it's important to focus on the development of state-of-the-art control system techniques and algorithms of exoskeletons to improve the accuracy and performance for exoskeleton robot users. The development of a sophisticated interface between the human and wearable robot is also imperative as the result of the estimation of the user's movement intention mainly depends on it. Based on the human-robot interaction method, many types of control strategies have been developed. The electrophysiological signal is a physiological signal measured from the human body via an exoskeleton that shows the user's motion intention [10]. In addition, it greatly influences exoskeleton assistive improvement.

Among electrophysiological signals, electromyography (EMG) has been one of the most prominent types used in exoskeleton control as it can make a measurement that directly reflects a human movement and muscular actions. Also, electroencephalography (EEG) is another control technique that is non-invasive and convenient [11]. A Brain-computer Interface (BCI) is a system that is able to acquire brain signals and analyse them. Then, the analysed signals are converted into commands which are relayed to output devices to perform desired actions [7]. Thus, the acquired EEG signals are trained in the BCI system to translate desired actions using machine learning algorithms.

2.5 Human Brain System

In the human body, brain is the most important part as it plays a significant role in everyday life. It controls all parts of human body. It is a collection of neural networks that are interconnected with each other to decide the human behaviours. Generally, in the perspective of radiologists and medical doctors, the right side of the human body is controlled by the left side of human brain. Similarly, the left side of human body is associated with the right side of human brain [30], [31].

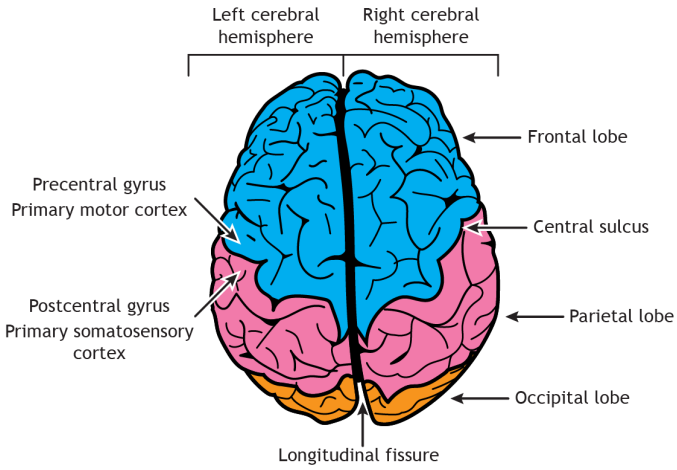


Figure 2.6: Dorsal View of Human Brain [32]

The human brain can be generally divided into the left and right hemispheres [33]. The left and right hemispheres are symmetrical in shape, and they can be further divided into four lobes that are temporal lobe, frontal lobe, occipital lobe, and parietal lobe. The frontal lobe is the largest lobe among the four lobes. It is located at the back of the forehead. It is mainly responsible for language and speech. Also, problem-solving, memory, decision-making, and other behaviours and emotions are concerned with the frontal lobe. The frontal lobe should not be injured/damaged as it has a big impact on human memory, emotions, and languages [34]–[36].

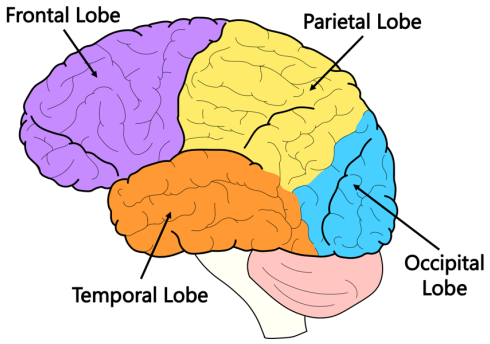


Figure 2.7: Sectional View of Human Brain [37]

The temporal lobe is situated on the sides of the human brain hemispheres behind the frontal lobe and under the parietal lobe. The temporal lobe is mainly responsible for various aspects of speech and hearing. Damaging the temporal lobe can be impairment of language, hearing, and sensory problems or loss permanently if the temporal lobe is seriously damaged. Another important lobe of the brain is the occipital lobe, which is located at the back of the human head. It is mainly responsible for the perception and processing of human visual information. Thus, damaging the occipital lobe can lead to visual and perception defects and loss of them if the injury is severe. The parietal lobe is located behind the frontal lobe of the human brain. The parietal lobe is located behind the frontal lobe of the human brain. The parietal lobe integrates different types of sensory information from different parts of the human body. Damaging or injuring the parietal lobe of the human brain can loss of the ability to locate parts of the human body [38], [39].

A particular part of the human brain (cranial nerve) controls the specific action and activities of the human body. The human eye moment function is controlled by oculomotor, trochlear, and abducens nerve while the human vision system is controlled by optic cranial nerves. The olfactory cranial nerve controls the function of the smell sensor. The tongue movement is responsible for hypoglossal cranial nerves while the hypoglossal cranial nerve controls the tongue movement [30]. The facial sensation is responsible for the trigeminal cranial nerve. Meanwhile, the facial cranial nerve controls human facial movement. Vagus cranial nerve controls the involuntary muscles of the human body while the accessory cranial nerve controls the voluntary neck muscle. Finally, the hearing and balancing of the human body are controlled by the vestibulocochlear cranial nerve. Therefore, understanding the cranial nerve and its associated function in the human body is important as it is helpful to place the EEG electrodes position.

2.6 EEG Signals

Today, EEG recording is one of the most common methods to acquire brain signals to analyse some brain diseases such as Alzheimer's, autism, and epilepsy. The method is non-evasive to the human body and has no side effects on patients [40]–[42]. The electric potential produced by brain sources can be recorded by an EEG device via electrodes on human scalps. An EEG device consists of electrodes, ADC (Analog to Digital Converter), conductive gel, and amplifiers to amplify the EEG signals. EEG electrodes can be made from different metals such as stainless steel, gold-plated silver, pure gold, tin, and Ag/AgCl. However, it should not mix EEG electrodes made from different metals in EEG recording because using two or more different types of electrodes will result in a high dc offset potential [43]. Among them, Ag/AgCl EEG electrode is the most common type because it maintains stable electrochemical potentials against biological tissues. Also, it has outstanding long-term electrical stability, has low dc offset variability and is free from potential allergenic compounds.

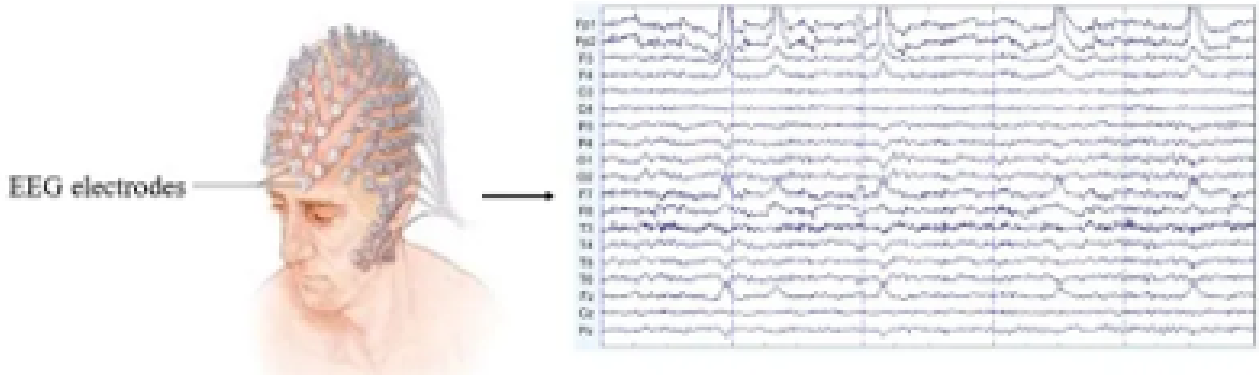


Figure 2.8: Measuring EEG Signals on Human Scalp [44]

A reusable disk is a type of EEG electrode that is placed on the scalp of the human head via a small amount of Ag-Cl conductive gel [45]. The cost of this type of electrode is low, and the life span of the electrode depends on the type of metal composition and the insulating materials. Another type of electrode is the EEG cap. Unlike the normal disk-type electrodes, the EEG cap has fewer chances of artifacts as the normal electrodes are easier to be fallen off from the human head scalp. The number of electrodes for the EEG cap and the material type can be chosen to suit for application. It's preferable to utilise in the application where multi-channel recording is required because individual electrode placing methods would not be practical at the exact placement of large quantities. The cap is made of fabric that has fixed electrode holders; so, placing electrodes does not require position measurements. However, it is important to note that troubleshooting of EEG cap is more difficult as tracing the faulty electrode is more difficult than the counterpart. It usually ends up in changing the whole cap [46].

2.7 EEG Electrodes Positioning

Electrodes should be placed systematically in EEG signals acquisition. Firstly, regions where electrodes on the lobes of the brain are differentiated with their initials. For instance, the electrodes to be placed on the temporal lobe region of the brain are started with 'T'. Similarly, the electrodes acquiring a signal from the Frontal lobe shall be labelled with 'F'. The electrodes placed on the Occipital lobe region are to be labelled with 'O', and the central lobe region of the brain should be 'C'. Finally, the midline region electrode is labelled with 'z' [46].

The second letter of electrodes indicates hemispherical side of the human brain. The even numbers are reserved for the electrodes to be placed on the right hemisphere of the human brain. Similarly, the odd numbers are reserved for the left hemisphere of the human brain. For instance, the C3 electrode identifies the electrode on the left hemisphere of the human brain at the Central lobe. In order to record special properties such as eye movements, respiration rate, skin conductance and heart rate, additional sensors are to be utilised [47].

In the positioning of electrodes, four anatomical essential landmarks are used. The first landmark is the nasion, which is located at the midline bony depression between the eyes. The second landmark is the inion. The third landmark is at the point 10% distance from the nasion toward inion. The final landmark is at the point that is 10% distance from the inion toward nasion. The traditional international 10-20 EEG electrode placement method is proposed by [48] to describe the location of scalp electrodes.

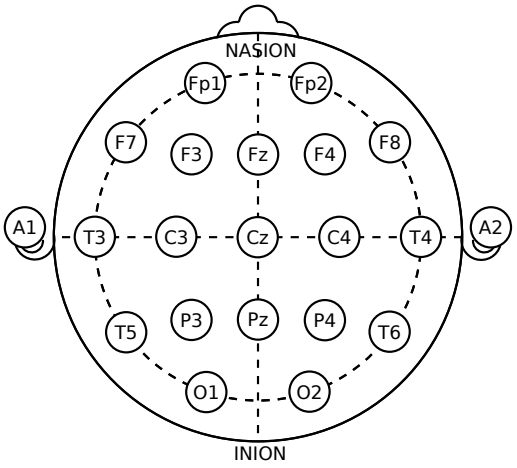


Figure 2.9: 10-20 EEG Electrodes Positioning System [49]

In the 10-20 system, the adjacent electrode distances are either 10% or 20% of the total distance between nasion and inion. The number of scalp positions is only up to 21. A 10-10 system is another EEG placement system proposed by [50] in 1985 to increase the number of electrode positions. In the 10-10 system, the number of electrode positions is increased to 74 locations. Then, the 10-5 system was introduced by [51] in which the number of electrode positions is increased up to 345 locations. The circumference of a standard human adult head is 58

cm. Thus, the typical inter-electrode distance in 10-20 system is between 53 mm and 74 mm. Similarly, the inter-electrode distance is between 28 mm and 38 mm in the 10-10 system, and between 22 mm and 31 mm in the 10-5 system.

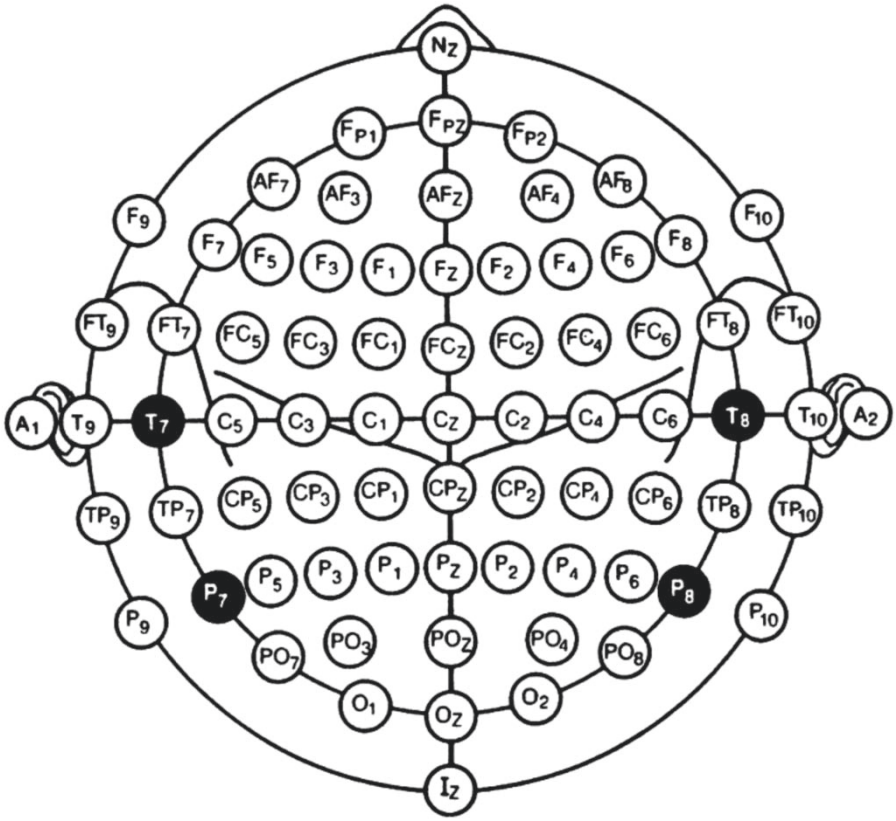


Figure 2.10: 10-10 EEG Electrodes Positioning System [52]

The electrode Cz, C3, and C4 mainly acquire the sensory and motor function of humans. Meanwhile, T4 and T6 acquire the emotional memory of humans. Verbal memory function signals are acquired at the other side of the hemisphere via T3 and T5 electrodes. The signals concerned with human cognitive function can be recorded at the Parietal lobe via Pz, P3 and P4. The O1, O2 and Oz electrodes can record the signal associated with visual processing stimuli. By using FP1, FP2 and, FPz electrodes, judgement and attention impulses can be obtained. Similarly, F3 and F4 electrodes are to acquire motor planning activities. The verbal and emotional expressions of a human signal can be obtained from the F7 and F8 EEG electrodes. The Fz electrode is positioned close to the motivational and intentional centres [47].

2.8 EEG Signal Amplification

The quality of the EEG amplifier plays an important role in EEG signal research because it's the location where the EEG signals are amplified while reducing or attenuating undesired components of the signals. Typically, there are three different types of signal components included in the EEG signals. They are biological signals, mains noise signals and electrode

offset signals. Among them, only biological signal is the only desired signal. The EEG biological signals are measured at the scalp electrode to record cortical and subcortical activity. Some artifacts are included in the biological signal which are endogenous artifacts and exogenous artifacts caused by sweat and scalp potentials, eye movement, etc. The amplitude of the biosignals can range from -2 mV to +2 mV. The dc offset of EEG signals result from the metallic type of EEG electrodes. The amount of dc offset can be varied several hundred millivolts [43]. Among them, Ag/AgCl is the most favoured due to the low offset value which is less than 100 mV and high temporal dc offset stability.

Figure 2.11: Ag/AgCl EEG Electrodes [53]

The mains noise is another undesired signal in the EEG recording. It is a sinusoidal artifact at the mains frequency is 50 Hz or 60 Hz. It is mainly due to the electrical devices powered by utility mains in the area of EEG recording. The magnitude of the mains noise depends on the type of mains-powered devices, particularly devices with a motor such as pumps, hair dryers, razors, etc. Standard procedures are followed in the EEG recording to reduce the presence of mains noise. One of the methods is bundling EEG electrodes or in a ribbon that greatly reduces the mains noise. In addition, rechargeable batteries are used in the EEG amplifier to further reduce the noise. Also, active shielding is simple and effective to achieve the amplifier's optimal Signal to Noise Ratio (SNR). Modern EEG amplifiers adopt a multichannel instrumentation method that is designed to amplify only the biological signal portion in the recording while the dc offset signal is unamplified and removes the mains noise [43].

Figure 2.12: 32 Channels EEG Cap with Amplifier [54]

The amplifier setting in the EEG signals recording is imperative to obtain quality data. The main parameters are sample rate, the gain, and characteristics of the low pass filter, high pass filter and notch filter. As per the signal sampling theorem, the sampling rate should be more

than two times higher than the highest frequency of interest. There is a trade-off between the sampling rate and the file size of the data. The gain of the amplifier is to be chosen as high as possible without reaching the saturation of the amplifier. The setting of highpass and lowpass filters depends on the frequency band of the spectral content of interest. The notch filter is a very steep special filter to filter out the narrow frequency band around the mains frequency.

2.9 EEG Signal Classification

EEG signal processing is critical for neurorehabilitation and brain-computer interface (BCI) systems, where high accuracy and the ability to process shorter time windows are key to practical and effective solutions. EEG signals are inherently noisy and dynamic, requiring algorithms capable of handling their complexity while delivering reliable and timely results. The desirable features of an effective algorithm include:

- **High Accuracy:** Ensure reliable classification across diverse subjects and conditions to support clinical applications.
- **Short Time Window Classification:** Enable accurate interpretation of data within shorter time windows, facilitating faster decision-making and real-time use.
- **Noise Robustness:** Handle artifacts such as eye movements (EOG), muscle activity (EMG), and environmental noise effectively to maintain signal integrity.
- **Temporal and Spatial Sensitivity:** Accurately capture dynamic changes over time and across electrode locations to enhance signal analysis.
- **Scalability and Efficiency:** Ensure generalisation across datasets and subjects, with computational efficiency suitable for real-time applications.

Modern EEG signal processing methods can be broadly categorised into three groups: traditional statistical methods, feature-based machine learning, and deep learning models which are summarised in Table 2.1.

Table 2.1: Generalised EEG Signal Processing Methods

Method	**Accuracy**	**Advantages**	**Disadvantages**
Statistical Approaches	Moderate	Simple, interpretable; effective for basic noise filtering.	Limited ability to capture complex, nonlinear; lower accuracy in dynamic tasks.
Feature-Based Machine Learning	High (with optimised features)	Incorporates domain knowledge; enables targeted feature extraction (e.g., CSP, PSD).	Performance depends on quality of feature engineering; time-intensive feature selection.
Deep Learning (e.g., CNNs, GNNs, etc.)	Very High	Learns complex spatial and temporal patterns automatically; high classification accuracy.	Computationally expensive; requires large datasets and careful hyperparameter tuning.

2.9.1 Convolutional Neural Network

Convolutional Neural Networks (CNNs) are a subset of Artificial Neural Networks (ANNs) designed to process structured data, such as images and sequential signals, by leveraging spatial hierarchies through convolutional layers. Inspired by the human visual cortex's role in object recognition, CNNs were initially developed for computer vision tasks [55]. However, their versatility has been demonstrated across various domains, including biomedical signal processing. CNNs excel at learning features directly from raw data, eliminating the need for extensive manual feature engineering [56].

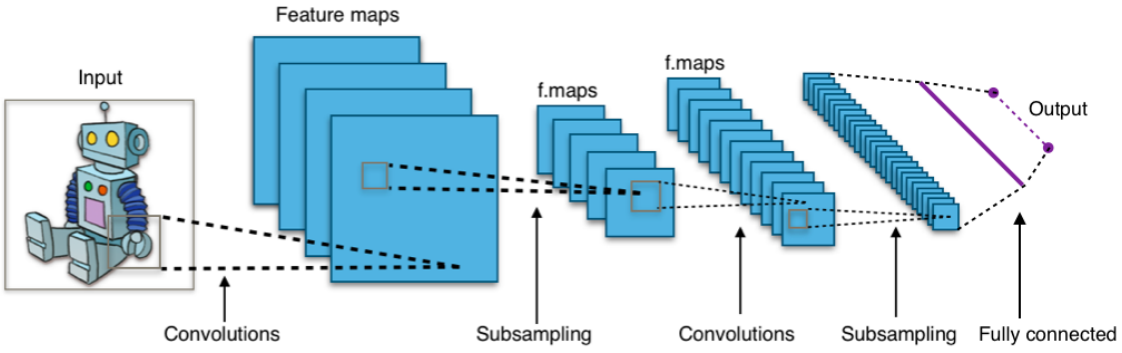


Figure 2.13: Typical CNN Architecture on Image Classification [57]

The success of CNNs in biomedical signal processing is well-documented, particularly for their ability to capture spatial and temporal patterns in high-dimensional and noisy data, such as

electrocardiogram (ECG) and EEG signals [58]. Their application in EEG signal classification has yielded significant advancements in BCI applications, sleep stage analysis [59]–[61], seizure detection [62]–[64], and emotion recognition [65]–[68]. In this study, CNNs are employed as one of the parallel branches in the EEG_SGNet model for classifying motor imagery and motor movement tasks, as detailed in Chapter 5.

2.9.2 Graph Convolutional Neural Network

Graph Neural Networks (GNNs) are a class of deep learning models specifically designed to perform convolutional operations on graph-structured data by leveraging the relationships between nodes. Unlike traditional deep learning methods, such as CNNs, which operate on fixed-dimensional data, GNNs can generalise to irregular domains, including social networks, molecular structures, and traffic networks. GNNs effectively learn meaningful representations of graph data, capturing both local and global graph properties. This makes them well-suited for EEG signal classification, as they can model spatial relationships between EEG signals in a graph structure, enabling multi-channel spatial analysis.

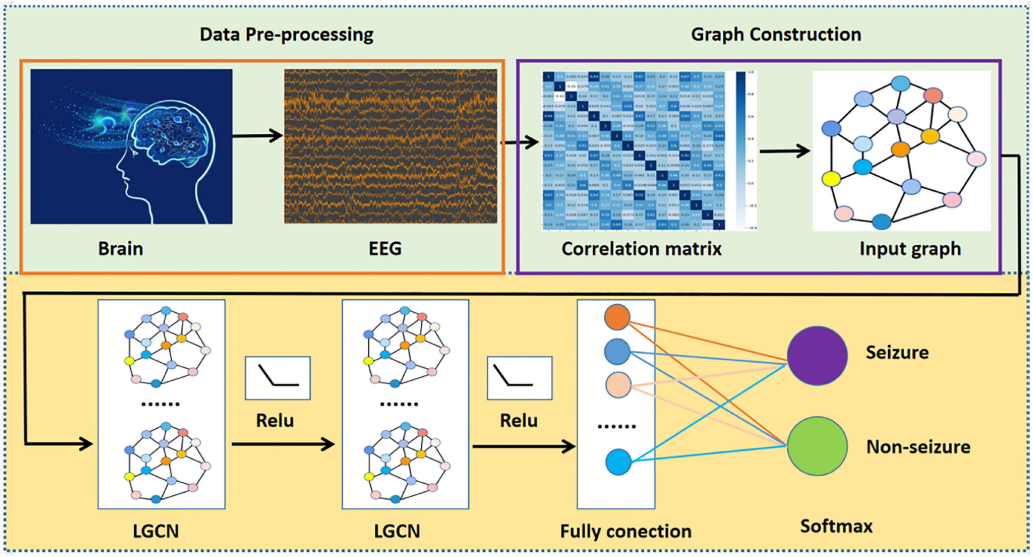


Figure 2.14: Typical GNN Architecture on EEG-based Seizure Detection [69]

In EEG applications, each channel can be represented as a node, while the relationships between channels are encoded as edges, which can be defined based on spatial proximity, functional connectivity, or physiological metrics such as phase-locking value (PLV) or Pearson correlation coefficient (PCC). GNNs have demonstrated success in numerous EEG signal classification tasks, including motor imagery, emotion recognition [70]–[72], seizure detection [73], [74], and neurological disorder diagnosis [75]. Despite their potential, selecting an optimal graph construction method remains a significant challenge and an active area of research. In Chapter 3, we propose a novel method for constructing an optimal graph, referred to as EEG_GLT.

2.9.3 Reinforcement Learning

Reinforcement Learning (RL) is a machine learning paradigm where an agent learns to make sequential decisions by interacting with its environment to maximise cumulative rewards. The agent takes actions, observes the effects through state transitions, and receives feedback in the form of rewards or penalties [76]. By balancing exploration (trying new actions) and exploitation (leveraging learned strategies), the agent iteratively refines its policy to achieve an optimal strategy. Unlike supervised learning, which relies on labeled datasets, RL enables autonomous learning through direct interaction with the environment, making it particularly suitable for complex and dynamic systems.

RL has been successfully applied in diverse domains, including healthcare, gaming, finance and trading, and robotics. In robotics, RL empowers robots to learn from environmental feedback and adapt to complex, varying conditions. One notable application is in rehabilitation robotics [77], where RL has had a significant impact by enabling robots to assist patients in regaining functionality and mobility after injuries, strokes, or surgeries. Rehabilitation robots can be personalised to address individual patient needs, abilities, and progress. For instance, a RL-driven robot can dynamically adjust the intensity of rehabilitation exercises based on a patient's feedback and performance during therapy sessions [78].

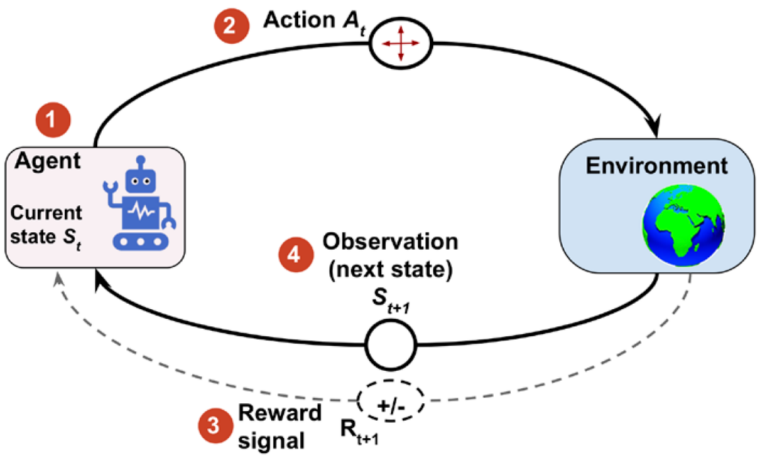


Figure 2.15: Interaction of RL Agent to its Environment [79]

The integration of RL and EEG signal classification further advances rehabilitation robotics through Brain-Robot Interfaces (BRIs). BRIs enable the decoding of motor intentions from a user's EEG signals, allowing direct brain control of robotic devices [80]. RL plays a critical role in BRIs by learning to interpret EEG signals in real time and making context-aware decisions. For example, an RL-based rehabilitation robot can optimise the therapeutic process by determining when to assist a patient with movements, when to pause for rest, and when to apply resistance for strength training [81]. By incorporating appropriate reward and penalty structures, RL allows robots to prioritise rehabilitation goals, such as minimising fatigue, maximising range of motion, and ensuring patient safety.

Despite these advancements, defining effective reward and penalty structures for RL remains a significant challenge, as they must balance therapeutic objectives with practical constraints. Furthermore, RL systems in rehabilitation robotics must address issues such as sample efficiency and safety during exploration, given the limited interaction opportunities with patients [82]. In this study, we propose EEG_RL-Net, a novel framework that leverages RL based on graph-based EEG features to decode EEG motor imagery (MI) signals in real time. The detailed methodologies, including graph-based EEG signal representation, reward design, and training strategies, are discussed in Chapter 6.

Chapter 3

EEG_GLT-Net: Optimising EEG Graphs for Real-time Motor Imagery Signals Classification

3.1 Introduction

Brain Computer Interfaces (BCIs) form an interdisciplinary bridge between engineering and neuroscience, enabling direct communication between the human brain and control devices. Originally designed to aid those with motor impairments [7], BCIs have expanded their applications to neurofeedback, gaming, and rehabilitation. Essentially, BCIs convert neural signals into actionable commands. The primary means of brain signal acquisition include electrocorticography (ECoG) and electroencephalography (EEG). Although ECoG boasts superior spatial resolution due to directly placing electrodes on the cortex, its invasive native limits its applications [83]. In contrast, EEG uses scalp placed electrodes to capture brain activity, making it more popular due to non-invasiveness and portability. This method captures various brain signals, from event-related to spontaneous and stimulus-evoked [84].

Motor Imagery (MI) pertains to the mental simulation of motor actions, such as moving one's hands or feet, without performing the actual movement [85], [86]. As highlighted by [87], action execution and its imagination share neural pathways. MI has prominent applications in rehabilitation and neuroscience. When paired with EEG, it captures neural signals generated from the intention to move. Integrating this with BCIs allows decoding EEG MI signals to control external devices such as a robotic exoskeleton. This technology is pivotal for those with motor impairments, especially stroke survivors, with the potential to restore quality of life and ability to perform daily activities. By accurately decoding EEG MI signals, BCIs can provide real-time feedback and communicate with assistive devices, to facilitate patient-intended movements [88].

Convolutional Neural Networks (CNNs) have consistently showcased superior results in computer vision tasks [89]–[91]. However, their effectiveness is largely constrained to regular Euclidean data, such as 2-dimensional grids and 1-dimensional sequences [91]. A drop in capability is experienced with non-Euclidean data, primarily because CNN cannot accurately capture the intrinsic structure and connectivity of this data. Graphs serve as powerful tools for representing relationships among entities, and are employed in diverse application areas including traffic systems, social networks, e-commerce platforms, biological structures, and trade networks. These graphs can highlight complex structures and be variable in nature such as being might be homogeneous or heterogeneous, having weight or not, and being signed or unsigned [92]. The Graph Convolutional Neural Network (GCN) is an adaptation of CNN operations that is, tailored for graphs. GCN excel in managing non-Euclidean data, incorporating topological relationships during convolution.

With the help of GCNs, the inherent connections among electrodes can be integrated through the adjacency matrix, a capability beyond the reach of traditional CNNs. Establishing relationships between nodes is essential before deploying the GCN method. Studies [93]–[95] have utilised Geodesic distances between electrodes to form the adjacency matrix, while others [96]–[100] have employed the Pearson coefficient correlation (PCC) to assess correlations between EEG channels. Additionally, [101] have utilised the phase lag index (PLI) in the adjacency matrix construction in their CSGNN model. Notably, [102] and [96] explored optimal adjacency matrices in EEG classification through a trainable matrix. [103] introduced a unified GNN sparsification technique (UGS), giving rise to a Graph Lottery Ticket (GLT) by pruning both the original adjacency matrix and GNN weights. This method decreases the Multiply Accumulate (MAC) inference, thus reducing computational overhead.

Existing methods for constructing adjacency matrices in EEG signal classification rely on prior knowledge of inter-channel relationships, which can be a limitation. This dependency is especially challenging for single time point classification of EEG motor imagery (MI) signals, where rapid signal interpretation at intervals as brief as $\frac{1}{160}$s is critical. To address this, our study proposes EEG_GLT which is a novel method for constructing adjacency matrices for GCNs specifically for EEG MI single time point classification, without requiring predefined inter-channel knowledge while enhancing both classification accuracy and computational efficiency.

3.2 Related Work

Traditional EEG MI classifiers typically rely on machine learning techniques that classify signals based on manually crafted features, such as wavelet transforms or analytic intrinsic mode functions [104], [105]. One widely used method is the filter bank common spatial pattern (FBCSP) [106], which applies common spatial patterns (CSP) across various frequency bands in EEG signals to extract discriminative features.

Deep neural networks (DNNs) have advanced EEG motor imagery (MI) classification by lever-

aging end-to-end architectures that combine feature extraction with classifier learning, eliminating the need for manual feature engineering. CNN-based models, such as those proposed by [107], [108], excel at extracting temporal features from 1D and 2D Euclidean data, achieving high accuracy. Further refinements, as seen in [109], [110], incorporate long short-term memory (LSTM) blocks to capture temporal dependencies effectively in EEG signals. The EEGProgress model [111] adopts a unique approach by applying CNN operations to individual brain regions for EEG MI signal classification, focusing on regional processing rather than all channels simultaneously. The ConTraNet model [112] combines Transformer and CNN blocks to capture both long- and short-term dependencies, fixed spatial patterns, and applies attention to non-stationary, time-varying inputs, resulting in improved performance for EEG-based emotion recognition. However, a common limitation of the methods discussed above is that they are applicable only to window-based EEG classification and not to single time-point classification.

GCNs have become increasingly popular in EEG signal classification due to their ability to encode non-Euclidean data, offering flexibility in analysing graph-structured information [92], [113]. GCNs can be applied across various graph analysis tasks:

- **Node-Level Tasks**: Predicting properties of individual nodes, used for both regression and classification.
- **Edge-Level Tasks**: Predicting edge properties, mainly for classification.
- **Graph-Level Tasks**: Classifying entire graphs based on their structure and properties.

Two main categories of GCNs are the spectral method [114]–[116] and the spatial method [117]–[120]. Studies [96] and [121] indicate challenges associated with the spatial method, particularly for matching local neighborhoods. GCNs have an important application in classifying EEG signals at the graph level, where EEG readings from individual electrodes are treated as node attributes. EEG feature extraction is broadly categorised into time and frequency domain features. Building on the work of [122], time-domain metrics such as Root Mean Square, skewness, minmax, variance, kurtosis, Hurst Exponent, Higuchi, and Petrosian fractal dimensions are derived within predefined time windows by [98]. Within the frequency domain, emphasis is placed on power spectral density (PSD) and power ratio (PR) across specific frequency bands: δ[0.5 - 4Hz], θ[4 - 8Hz], α[8 - 13Hz], β[13 - 30Hz], and γ[30 - 110Hz]. This is supplemented by other metrics such as total power, spectral entropy, and peak frequency, all captured within chosen time windows.

The DG-HAM [123] and EEG-ARNN [124] models classify EEG tasks using raw EEG signals within a specified window length, without extracting graph-based features, such as time-domain or frequency-domain features. In contrast, [100] introduced the state-of-the-art GCNs-Net for time-point classification, which treats each time point as an independent feature, enabling a more detailed time-resolved analysis of EEG MI signals. Although GCNs-Net performs well in classifying EEG MI single time points, it only considers the functional connectivity of EEG channels during GNN operations, which can limit its accuracy. Additionally, its fully dense

adjacency matrix requires high MAC operations, making it less efficient for single time-point EEG classification.

While dynamic graph models such as TodyNet [125] and SDGL [126] have demonstrated strong performance in multivariate time series tasks by learning evolving graph structures, they are primarily designed for window-based or sequential inputs. These methods are not suited for real-time EEG decoding, especially for single-time-point signals classification, where the model must operate without temporal context. This remains an open challenge in EEG research that calls for alternative approaches to graph construction and model design.

In this chapter, we propose the EEG_GLT method for adjacency matrix construction, integrated with a spectral GNN-based EEG_GLT-Net architecture, to classify EEG MI at the single-time-point level. Using the raw EEG MI single-time-point signals from the time-resolved PhysioNet dataset. The primary contributions of this study can be summarised as:

- **EEG Graph Lottery Ticket (EEG_GLT):** We present a novel method to construct an optimal adjacency matrix for EEG MI signal classification. Achieved through the iterative pruning of relationships among EEG channels, the EEG_GLT introduces a new direction in EEG adjacency matrix design.
- **Channel Relationship Optimisation:** Our approach reveals the most advantageous relationship between EEG channels. It is tailored for catering to individual subjects and the architecture of GCN models, eliminating the need for prior knowledge about the inter-relationships among EEG channels.
- **Computational Efficiency:** Recognising the computational intensity of classifying EEG at single time points, our strategy mitigates the high demand for computational resources, proving especially beneficial for real-time applications.
- **Performance Validation:** We benchmark the accuracy of our EEG_GLT method against two well-established techniques: the Geodesic method and the leading PCC method employed in the state-of-the-art GCNs-Net. This evaluation spans across six distinct spectral GCN models. Each model is distinguished by its unique specifications, including variations in GCN layer structures, polynomial degrees of filters, numbers of Fully Connected (FC) layers, and the amount of hidden nodes. Additionally, we compare the performance of our model with seven other state-of-the-art models to demonstrate its effectiveness.

3.3 Methodology

3.3.1 Overview

As shown in Figure 3.1, the project framework was as follows:

- EEG signals from 64 channels were captured at each time point $\frac{1}{160}s$ and used as input features for the EEG_GLT-Net.
- Additionally, the EEG_GLT-Net accepted the graph representation as another form of input. This representation included the graph Laplacian, derived using three different methods: PCC between EEG channels, Geodesic distance between EEG electrodes, and our newly proposed EEG Graph Lottery Ticket Adjacency Matrix Mask (m_{EEG_GLT}).
- The EEG_GLT-Net processed these inputs to decode the EEG MI time point signal, which was then categorised into one of the four MI types.

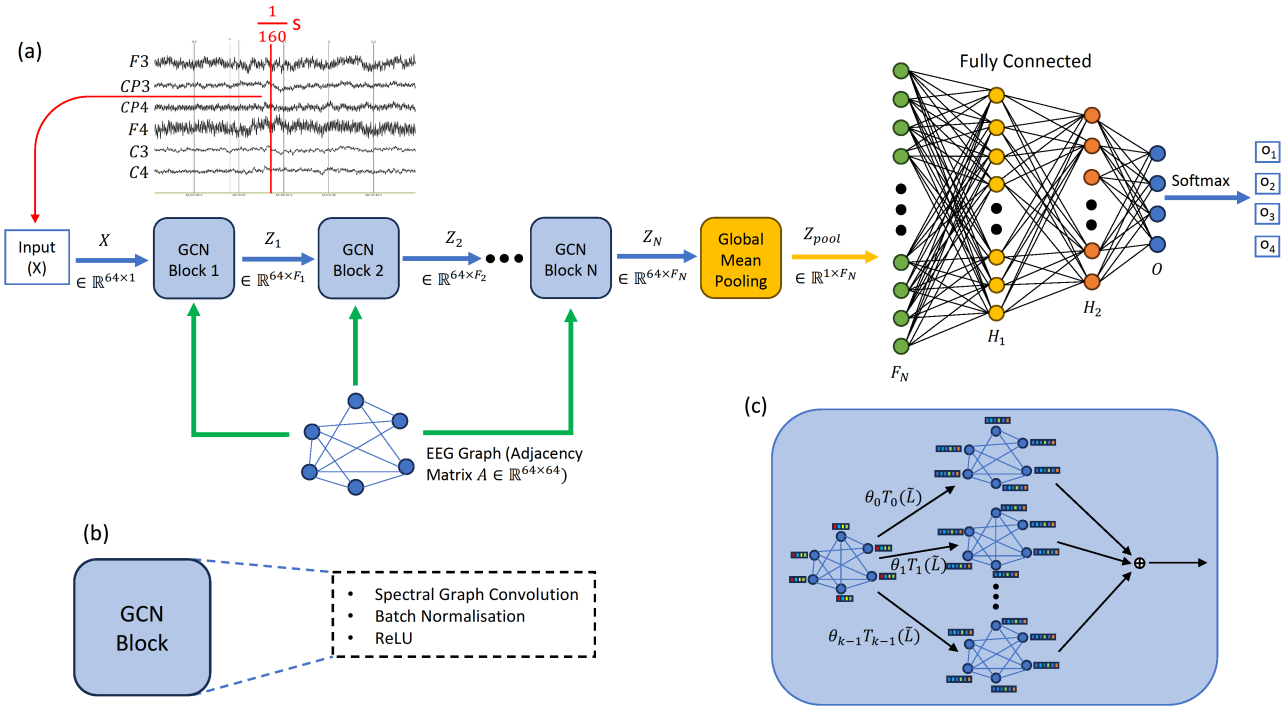


Figure 3.1: EEG_GLT-Net Model: (a) Overall architecture (classifying EEG MI of one time point $\frac{1}{160}s$ of signals from 64 EEG electrodes). Note that EEG Graph adjacency matrix can be $A^{Geodesic}$, A^{PCC} or A^{EEG_GLT}, (b) Components inside the spectral graph convolution block, (c) Chebyshev spectral graph convolution

3.3.2 Dataset Description

This chapter utilised the PhysioNet EEG MI dataset [127] encompassing over 1,500 EEG recordings sourced from 109 participants. The recordings were captured using 64 EEG electrodes, consistent with the international 10-10 system, with the exclusion of F9, Nz, F10, FT9, FT10,

A1, A2, TP9, TP10, P9, and P10 channels. Channel Iz is retained as the sole outermost sensor due to its minimal contamination from ocular and muscle artifacts and its proven utility in preserving relevant motor imagery information with high signal quality. Each participant executed 84 trials, broken down into 3 runs with, 7 trials per run, spanning 4 distinct tasks. The tasks included:

- Task 1: Imagining the act of opening and closing the left fist.
- Task 2: Imagining the act of opening and closing the right fist.
- Task 3: Imagining the act of opening and closing both fists simultaneously.
- Task 4: Imagining the act of opening and closing both feet.

Recordings in the dataset were originally sampled at 160 Hz and each recording had a duration of 4 seconds. Our study employed time point samples for classification, and our analysis was strictly conducted at the subject level. Although the original dataset comprised 109 participants, our study focused solely on 20 subjects, labelled S_1 to S_{20}.

3.3.3 Data Pre-processing and Feature Extraction

In the initial pre-processing phase, raw signals underwent only a notch filter at the 50 Hz power line frequency, foregoing typical filtering or denoising steps to maximise data integrity. Although each task lasted for a 4-second duration, only the time period from $t = 1s$ to $t = 3s$ was considered in our experiments. This is because subjects typically exhibited greater readiness post $t = 1s$. All 64 EEG channels were incorporated into our model. We utilised the signal values from each EEG channel at each time point as feature for each node. The construction methods of the adjacency matrix, which captures brain connectivity, are elaborated in Sections 3.3.4. The training data underwent normalisation, ensuring a mean $\mu = 0$ and a standard deviation $\sigma = 1$ for each channel. Following this, both the test and validation sets were adjusted in alignment with the normalisation parameters established from the training data.

3.3.4 Graph Preliminary

Graph Representation

Consider a directed weighted graph represented as $G = \{V, E\}$. Here, $|V| = N$ denoted the number of nodes and $|E|$ was the count of edges connecting the nodes. The node set was defined as $V = \{v_1, v_2, \ldots, v_n\}$. The node feature matrix of the entire graph was represented by $X \in \mathbb{R}^{N \times F}$, where F, the number of features, was equal to 1. The adjacency matrix, denoted as $A \in \mathbb{R}^{N \times N}$, captured the graph's overall topology. Specifically, if an edge existed between nodes v_i and v_j (i.e., $(v_i, v_j) \in E$), then $A[i, j] \neq 0$. Otherwise, $A[i, j] = 0$.

The adjacency matrix for the PCC method, $A^{PCC} \in \mathbb{R}^{N \times N}$ was defined in Equation 3.1, where

I was the identity matrix and $|P|$ was the absolute PCC matrix of dimension of $N \times N$. Each element P_{ij} of P was calculated using Equation 3.2. The absolute PCC matrix $|P|$ captured the linear correlations between EEG channel signals, with its elements satisfied $0 \leq |P_{ij}| \leq 1$:

$$A^{PCC} = |P| - I \tag{3.1}$$

where

$$P_{ij} = \frac{cov(x_i, x_j)}{\sigma_i \sigma_j}. \tag{3.2}$$

For the Geodesic-distance adjacency matrix method, the configuration of 64 electrodes into a unit sphere acted as a stand-in for spatial brain connectivity. This allowed the computation of geodesic distances between the electrodes placed on a sphere of radius r. If two electrodes have Cartesian coordinates $(x_i,\ y_i,\ z_i)$ and $(x_j,\ y_j,\ z_j)$, the geodesic distance for the adjacency matrix was calculated using Equation 3.3:

$$A_{ij}^{Geodesic} = arccos(\frac{(x_i\ x_j + y_i\ y_j + z_i\ z_j)}{r^2}) \tag{3.3}$$

These distances were standardised into the [0, 1] range. The degree matrix, D, was a diagonal representation of A, where the i^{th} diagonal element of D was computed as $D_{ii} = \sum_{j=1}^{N} A_{ij}$. The combinatorial Laplacian matrix, $L \in \mathbb{R}^{N \times N}$, was described as $L = D - A$. A normalised version of this combinatorial Laplacian can be obtained with $L = I_N - D^{-1/2} A D^{-1/2}$.

Spectral Graph Filtering

The eigenvectors of the graph Laplacian matrix can be expressed as graph Fourier modes, with $\{u_l\}_{l=0}^{N-1} \in \mathbb{R}^N$. The diagonal matrix of these Fourier frequencies, $\Lambda = diag[\lambda_0, ..., \lambda_{N-1}] \in \mathbb{R}^{N \times N}$. We defined the Fourier basis, $U = [u_0, ..., u_{N-1}] \in \mathbb{R}^{N \times N}$, which allows for the decomposition of the Laplacian matrix, L, into $L = U \Lambda U^T$. The signal x can be transformed by graph Fourier into $\hat{x} \in \mathbb{R}^N$ using $\hat{x} = U^T x$, while the inverse graph Fourier transform is given by $x = U\hat{x}$. The convolution operation on graph G is defined as:

$$x *_G g = U((U^T x) \odot (U^T g)) \tag{3.4}$$

where g represents the convolutional filter and $\odot$ denotes the Hadamard product. Given that $g_\theta(\Lambda) = diag(\theta)$, where $\theta \in \mathbb{R}^N$ represents the vector of Fourier coefficients, the Graph convolution operation can be implemented as follows:

$$x *_G g_\theta = g_\theta(U \Lambda U^T)x = U g_\theta(\Lambda) U^T x \tag{3.5}$$

where g_θ is a non-parametric filter, and polynomial approximation is employed to mitigate the excessive computational complexity. Chebyshev graph convolution, a specific instance of graph

convolution, utilises Chebyshev polynomials for filter approximation [115]. The approximation of $g_\theta(\Lambda)$ under the K^{th} order Chebyshev polynomial framework is given by:

$$g_\theta(\Lambda) = \sum_{k=0}^{K-1} \theta_k T_k(\hat{\Lambda}) \tag{3.6}$$

where

$$\hat{\Lambda} = \frac{2\Lambda}{\Lambda_{max}} - I_N \tag{3.7}$$

In the above Equation 3.6, $T_k(\hat{\Lambda}) \in \mathbb{R}^K$ is the k^{th} order Chebyshev polynomial evaluated using Equation 3.7. Then, the signal x is convolved with the defined filter g_θ using the following equation:

$$x *_G g_\theta = U \sum_{k=0}^{K-1} \theta_k T_k(\hat{\Lambda}) U^T x = \sum_{k=0}^{K-1} \theta_k T_k(\widetilde{L}) x \tag{3.8}$$

Normalising Λ can be achieved by using Equation 3.7, where Λ_{max} denotes the largest entry in the diagonal of Λ, and I_N represents the diagonal matrix of the scaled eigenvalues. In the above Equation 3.8, $T_k(\widetilde{L})$ is the Chebyshev polynomial of order k evaluated at the scaled Laplacian, $\widetilde{L}$ which is defined in the Equation 3.9 as:

$$\widetilde{L} = \frac{2L}{\lambda_{max}} - I_N \tag{3.9}$$

Let $\bar{x}_k = T_k(\widetilde{L})x \in \mathbb{R}^N$, where a recursive relation is used to compute $\bar{x}_k$ using Equation 3.10 with $\bar{x}_0 = x$ and $\bar{x}_1 = \widetilde{L}x$. One key advantage of using the Chebyshev polynomial to approximate convolutional filters is that it inherently avoids the need to compute the graph Fourier basis. The recursive relation is given by:

$$\bar{x}_k = 2\widetilde{L}\bar{x}_{k-1} - \bar{x}_{k-2} \tag{3.10}$$

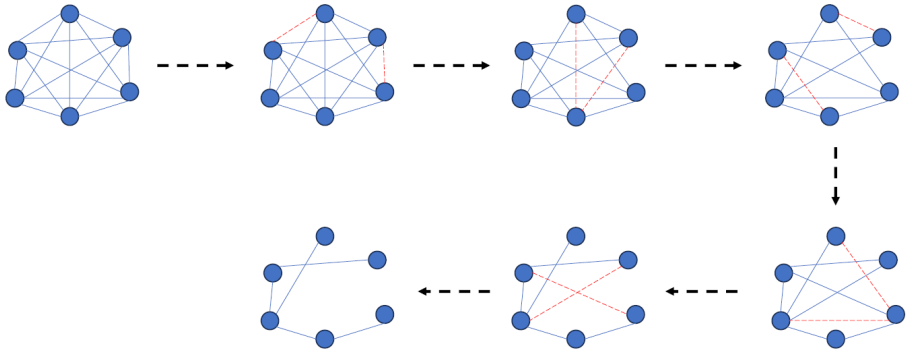


Figure 3.2: EEG graph (m_g) pruning using Algorithm 1: At each N_{ep} iteration, the bottom p_g% are pruned, reducing density from 100% until the lowest density s_g%. Solid lines indicate remaining edges, while red-dashed lines depict removed edges

3.3.5 EEG Graph Lottery Ticket (EEG_GLT)

In the process of executing a forward pass with the spectral GNN function, denoted as $f(., \Theta)$, and given a graph denoted as $G = \{A, X\}$, the method presented in [103] searches for an adjacency matrix mask $m_g \in \{0, 1\}$. The goal is to maximise sparsity while maintaining the highest prediction accuracy. In our model, the original matrix $A_{original\_ij} = \{0, if \quad i = j; 1, otherwise\}$ in the shape of $|V| \times |V|$ was not trainable. The adjacency matrix mask in our model $m_g \in \mathbb{R}^{|V| \times |V|}$ was trainable:

$$A = A_{original} \odot m_g \tag{3.11}$$

Once the model had undergone N epochs, the lowest p_g% ($p_g = 10$%) of the values in the trained m_g at highest accuracy of the validation dataset were pruned. These values were set to 0, while the remaining values were set to 1 as shown in Figure 3.2. Concurrently, the spectral filter weights, represented as Θ, were reset to their initial state, Θ_0. The trained m_g that yielded the highest accuracy of the validation set within the span of N epochs was designated as the GLT and duly noted. This process continued, and a GLT was recorded for each level of graph sparsity until the sparsity of m_g fell below the pre-determined final sparsity level, s_g. The EEG_GLT was ultimately identified as the GLT that achieves the highest accuracy alongside the highest level of graph sparsity. Moreover, it delineated the optimal adjacency matrix capable of producing the highest accuracy.

3.3.6 General Model Architecture

A GCN structure was designed to classify EEG MI signals. This architecture comprised three primary blocks: the GCN block, the Global Mean Pooling Block, and the Fully Connected Block. In the GCN Block, generalised graph features for each EEG electrode were extracted. Subsequently, the features from all 64 channels were consolidated using a mean in the Global Mean Pooling Block. The Fully Connected Block was employed for the final prediction. A

Algorithm 1 : EEG_GLT - Finding Optimal EEG Adjacency Matrix

Input: Graph $G = \{A, X\}$, GNN $f(G, \Theta)$, GNN initialisation Θ_0,
$A_{original\_ij} = \{0, if \;\; i = j; 1, otherwise\}$, initial Adjacency Matrix Mask
$m_g^0 = A_{original}$, learning rate $\eta = 0.01$, pruning rate $p_g = 10\%$, pre-defined lowest Graph Density Level $s_g = 13.39\%$.

Output: EEG Graph Lottery Ticket $(m_{g\_EEG\_GLT}) - m_g^{s,i}$ at the highest accuracy with the highest sparsity possible.

1: **while** $\frac{||m_g^s||_0}{||A_{original}||_0} \geq s_g$ **do**
2: **for** for iteration $i = 0, 1, 2, ..., N_{ep}$ **do**
3: Forward $f(., \Theta_i)$ with $G_s = \{m_g^{s,i} \odot A_{original}, X\}$ to compute Cross-Entropy Loss, L
4: Backpropagate and update, Θ_i and $m_g^{s,i}$ using Adam Optimiser
5: **end for**
6: Record $m_g^{s,i}$ with the highest accuracy in validation set during the N_{ep} iteration
7: Set $p_g = 10\%$ of the lowest absolute magnitude values in m_g^s to 0 and the others to 1, then obtain a new $m_g^{s+1,0}$
8: **end while**

detailed representation of this model architecture is provided in Figure 3.1 and Table 3.1.

Table 3.1: Generalised Architecture of EEG_GLT-Net Model

Layer	Type	Input Size	Polynomial Order	Weights	Bias	Output	Activation
Input	Input	$N \times 1$	-	-	-	-	-
Block A - GCN Block							
C1	Graph Convolution	$N \times 1$	K_1	$1 \times F_1 \times K_1$	$N \times F_1$	$N \times F_1$	-
BNC1	Batch Normalisation	$N \times F_1$	-	F_1	F_1	$N \times F_1$	ReLU
C2	Graph Convolution	$N \times F_1$	K_2	$F_1 \times F_2 \times K_2$	$N \times F_2$	$N \times F_2$	-
BNC2	Batch Normalisation	$N \times F_2$	-	F_2	F_2	$N \times F_2$	ReLU
C3	Graph Convolution	$N \times F_2$	K_3	$F_2 \times F_3 \times K_3$	$N \times F_3$	$N \times F_3$	-
BNC3	Batch Normalisation	$N \times F_3$	-	F_3	F_3	$N \times F_3$	ReLU
C4	Graph Convolution	$N \times F_3$	K_4	$F_3 \times F_4 \times K_4$	$N \times F_4$	$N \times F_4$	-
BNC4	Batch Normalisation	$N \times F_4$	-	F_4	F_4	$N \times F_4$	ReLU
C5	Graph Convolution	$N \times F_4$	K_5	$F_4 \times F_5 \times K_5$	$N \times F_5$	$N \times F_5$	-
BNC5	Batch Normalisation	$N \times F_5$	-	F_5	F_5	$N \times F_5$	ReLU
C6	Graph Convolution	$N \times F_5$	K_6	$F_5 \times F_6 \times K_6$	$N \times F_6$	$N \times F_6$	-
BNC6	Batch Normalisation	$N \times F_6$	-	F_6	F_6	$N \times F_6$	ReLU
Block B - Global Mean Pooling Block							
P	Global Mean Pool	$N \times F_6$	-	-	-	F_6	-
Block C - Fully Connected Block							
FC1	Fully Connected	F_6	-	$F_6 \times H_1$	H_1	H_1	-
BNFC1	Batch Normalisation	H_1	-	H_1	H_1	H_1	ReLU
FC2	Fully Connected	H_1	-	$H_1 \times H_2$	H_2	H_2	-
BNFC2	Batch Normalisation	H_2	-	H_2	H_2	H_2	ReLU
FC3	Fully Connected	$H_2 \times O$	-	$H_2 \times O$	O	O	-
S	Softmax Classification	O	-	-	-	O	-

N = Number of EEG Channels (i.e. 64); O = Number of EEG MI Classes (i.e. 4)

3.3.7 Model Setting

Let F_i represent the number of filters at each GCN level, given by $F_i \in [F_1, F_2, F_3, F_4, F_5, F_6]$. Similarly, K_i denotes the polynomial order of the filter for each i^{th} layer, and is defined as $K_i \in [K_1, K_2, K_3, K_4, K_5, K_6]$. O indicates the number of MI classes for prediction. Due to the large volume of instances in the training set, we employed a mini-batch size B of 1024. A batch normalisation (BN) layer was incorporated after both the spectral GCN and Fully Connected layers. This BN layer re-scales and re-centers normalised signals to match the original distribution within the mini-batch, addressing the internal covariate shift issue and helping to mitigate the gradient vanishing/exploding problem. Additionally, 50% dropout layers were integrated after the ReLU layers (Equation 3.12) within the Fully Connected Block

for regularisation. The details of the model settings can be found in Table 3.2, while the hyperparameter settings are provided in Table 3.3.

The ReLU activation function is defined as:

$$\text{ReLU}(x) = max(0, x) \tag{3.12}$$

The softmax function is given by

$$\text{Softmax}(\hat{y_i}) = \frac{e^{\hat{y_i}}}{\sum_{j=1}^{O} e^{\hat{y_j}}} \tag{3.13}$$

where $\hat{y}_i$ represent the predicted probability of an instance for each class, ranging over $\hat{y}_i \in [\hat{y_1}, ..., \hat{y_O}]$. O denotes the total number of classes. The loss function employed was the cross-entropy loss, defined as:

$$\text{Loss} = -\frac{1}{|B|} \sum_{b=1}^{B} \sum_{i=1}^{O} y_{b,i}. \log(\text{Softmax}(\hat{y}_{b,i})) \tag{3.14}$$

Table 3.2: Model Settings of EEG_GLT-Net

Model	Model Framework	Number of GCN Filters	GCN Filter Polynomial Order	Number of FC Hidden Nodes
A	$(C-BNC)\times 6-P-(FC-BNFC)\times 2-FC-S$	16, 32, 64, 128, 256, 512	5, 5, 5, 5, 5, 5	1024, 2048, 4
B	$(C-BNC)\times 6-P-(FC-BNFC)\times 2-FC-S$	16, 32, 64, 128, 256, 512	2, 2, 2, 2, 2, 2	1024, 2048, 4
C	$(C-BNC)\times 5-P-(FC-S)$	16, 32, 64, 128, 256	5, 5, 5, 5, 5	4
D	$(C-BNC)\times 5-P-(FC-S)$	16, 32, 64, 128, 256	2, 2, 2, 2, 2	4
E	$(C-BNC)\times 5-P-(FC-BNFC)\times 2-FC-S$	64, 128, 256, 512, 1024	5, 5, 5, 5, 5	512, 128, 4
F	$(C-BNC)\times 5-P-(FC-BNFC)\times 2-FC-S$	64, 128, 256, 512, 1024	2, 2, 2, 2, 2	512, 128, 4

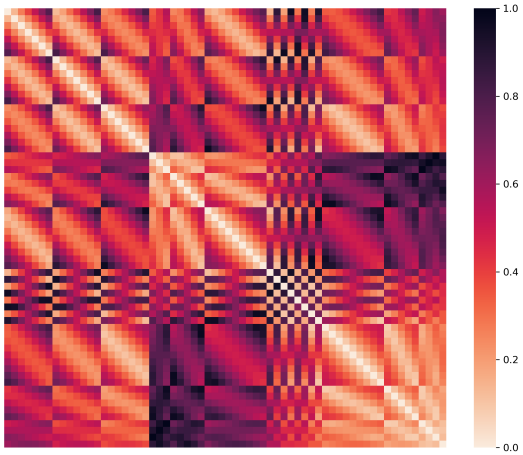


Figure 3.3: Geodesic Distance Adjacency Matrix ($A^{Geodesic}$)

Table 3.3: Hyperparameter Settings for Training EEG_GLT-Net

Hyperparameter	Value
Training Epochs (N_{ep})	1000
Batch Size (B)	1024
Dropout Rate	0.5
Optimizer	Adam
Initial Learning Rate (η)	0.01

Both accuracy and F1 score evaluation metrics were employed to assess the performance of models. The accuracy is defined as:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{3.15}$$

The recall is given by:

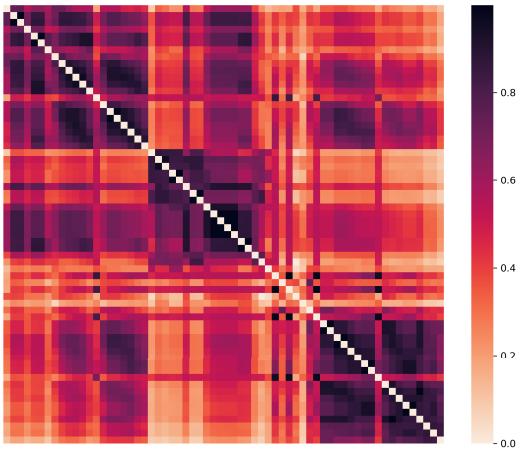
$$\text{Recall} = \frac{TP}{TP + FN} \tag{3.16}$$

The precision is defined as:

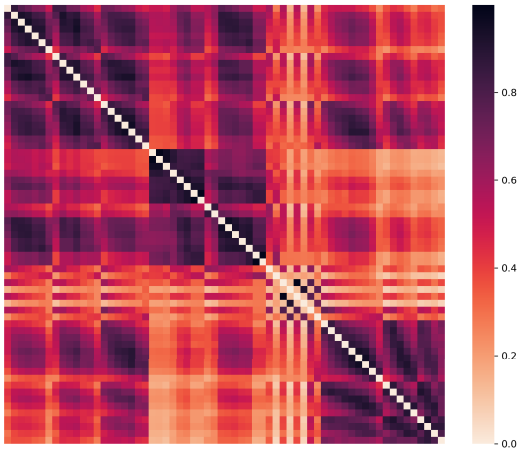
$$\text{Precision} = \frac{TP}{TP + FP} \tag{3.17}$$

The F1 score is given by:

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3.18}$$

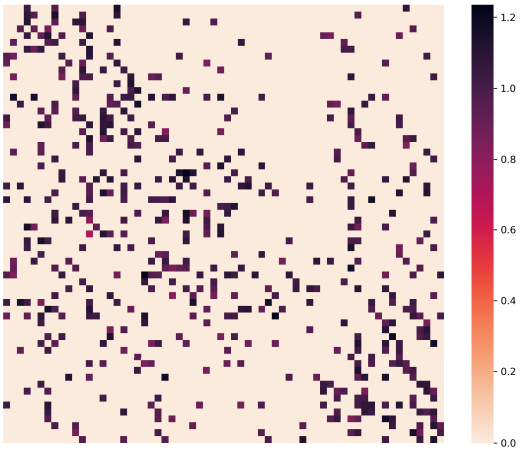


(a) Subject S_6

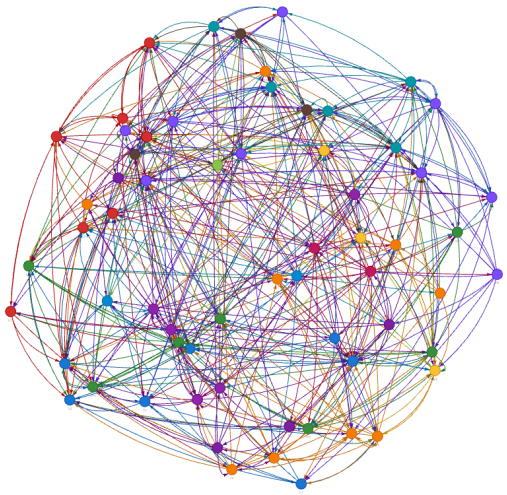


(b) Subject S_{14}

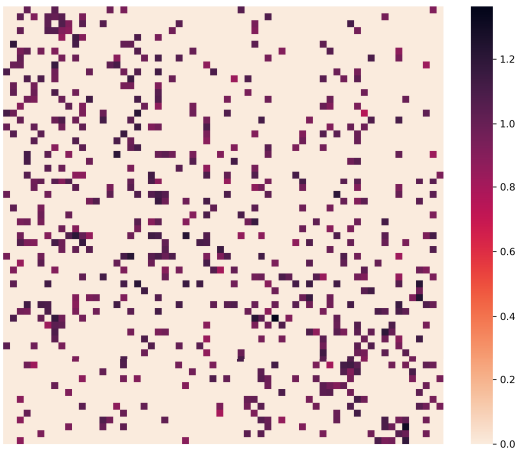
Figure 3.4: PCC Adjacency Matrix (A^{PCC}) of Subject S_6 and S_{14}



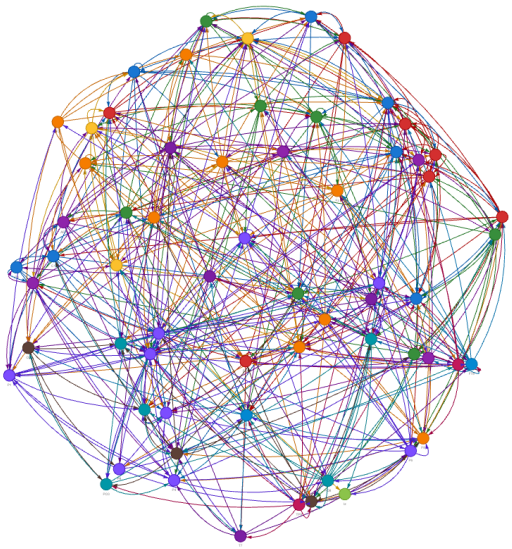
(a) Adjacency Matrix ($m_{g\_EEG\_GLT}$) - Subject S_6, 13.39% Density, Model A (Accuracy: 78.13%)



(b) Graph of $m_{g\_EEG\_GLT}$ - S_6, 13.39% Density, Model A



(c) Adjacency Matrix ($m_{g\_EEG\_GLT}$) - Subject S_6, 13.39% Density, Model E (Accuracy: 73.55%)



(d) Graph of $m_{g\_EEG\_GLT}$ - S_6, 13.39% Density, Model E

Table 3.4: Accuracy Comparison of Different Adjacency Matrix Construction Methods for Subjects S_1 to S_{10}

Subj	Accuracy (Mean±Std)			F1 Score (Mean±Std)		
	Geodesic	PCC	EEG_GLT (our method)	Geodesic	PCC	EEG_GLT (our method)
S_1	66.19% ± 4.17%	76.47% ± 9.94%	**98.51% ± 0.77%**	66.53% ± 4.36%	76.91% ± 9.78%	**98.53% ± 0.78%**
S_2	46.53% ± 1.33%	69.13% ± 7.05%	**76.18% ± 5.53%**	46.47% ± 1.46%	69.34% ± 7.37%	**76.19% ± 5.52%**
S_3	76.18% ± 4.98%	87.28% ± 9.19%	**99.17% ± 0.32%**	76.12% ± 5.00%	87.43% ± 8.97%	**99.19% ± 0.31%**
S_4	96.41% ± 1.97%	99.13% ± 1.01%	**99.97% ± 0.06%**	96.44% ± 1.98%	99.10% ± 1.12%	**99.97% ± 0.05%**
S_5	37.05% ± 1.04%	43.19% ± 3.03%	**50.95% ± 3.80%**	36.66% ± 0.97%	43.28% ± 2.73%	**50.86% ± 3.85%**
S_6	44.37% ± 1.59%	58.23% ± 5.19%	**69.60% ± 5.67%**	44.29% ± 1.65%	58.25% ± 5.49%	**69.50% ± 5.70%**
S_7	40.44% ± 1.19%	50.98% ± 3.80%	**59.45% ± 3.00%**	40.30% ± 1.23%	51.10% ± 3.49%	**59.34% ± 2.99%**
S_8	89.03% ± 7.04%	95.06% ± 5.96%	**99.95% ± 0.07%**	88.84% ± 6.88%	95.14% ± 5.81%	**99.96% ± 0.07%**
S_9	87.26% ± 14.26%	97.64% ± 3.33%	**99.95% ± 0.08%**	87.41% ± 14.49%	97.70% ± 3.78%	**99.95% ± 0.08%**
S_{10}	98.26% ± 0.31%	99.24% ± 0.19%	**99.99% ± 0.01%**	98.25% ± 0.32%	99.25% ± 0.20%	**99.99% ± 0.01%**

3.4 Results and Discussion

3.4.1 Geodesic vs PCC Adjacency Matrix Construction Method

The Table 3.6 presents the mean performance accuracy and F1 score across various models for different adjacency matrix construction methods, including Geodesic, PCC, and EEG_GLT, for each subject. Among the existing methods (PCC and Geodesic), the PCC adjacency method consistently outperformed the Geodesic method, enhancing the accuracy by 0.98% - 22.60% and the F1 score by 0.99% - 22.86%. Table V and Figure 6 detail the mean accuracies and F1 scores for 20 subjects (S_1 - S_{20}) across different matrix construction methods for each model setting. Notably, the PCC method outperformed the Geodesic method across all model settings, improving accuracy by 9.76% and the F1 score by 9.63%. The superiority of the PCC method in EEG MI adjacency matrix construction over the Geodesic method stems a major

Table 3.5: Accuracy Comparison of Different Adjacency Matrix Construction Methods for Subjects S_{11} to S_{20}

Subj	Accuracy (Mean±Std)			F1 Score (Mean±Std)		
	Geodesic	PCC	EEG_GLT (our method)	Geodesic	PCC	EEG_GLT (our method)
S_{11}	97.18% ± 1.12%	99.48% ± 0.70%	**99.99% ± 0.01%**	97.18% ± 1.13%	99.49% ± 0.74%	**99.99% ± 0.01%**
S_{12}	71.54% ± 3.44%	78.07% ± 8.95%	**99.69% ± 0.32%**	71.40% ± 3.37%	77.94% ± 8.76%	**99.70% ± 0.31%**
S_{13}	36.52% ± 0.32%	41.35% ± 1.23%	**44.50% ± 2.23%**	36.49% ± 0.45%	41.01% ± 1.34%	**44.47% ± 2.23%**
S_{14}	40.21% ± 1.80%	55.97% ± 6.47%	**72.39% ± 6.43%**	40.10% ± 1.88%	56.05% ± 6.57%	**72.71% ± 6.13%**
S_{15}	46.16% ± 1.28%	52.11% ± 3.96%	**67.55% ± 9.26%**	45.92% ± 1.93%	52.20% ± 3.66%	**67.52% ± 9.27%**
S_{16}	95.62% ± 3.87%	96.75% ± 5.00%	**99.98% ± 0.03%**	94.94% ± 5.25%	96.72% ± 5.07%	**99.98% ± 0.03%**
S_{17}	92.07% ± 8.10%	98.83% ± 2.33%	**99.98% ± 0.03%**	91.95% ± 8.31%	98.66% ± 2.76%	**99.98% ± 0.03%**
S_{18}	71.24% ± 5.96%	86.19% ± 10.95%	**99.92% ± 0.12%**	73.28% ± 3.28%	85.98% ± 11.10%	**99.93% ± 0.13%**
S_{19}	33.18% ± 0.40%	38.38% ± 2.27%	**41.41% ± 1.44%**	32.85% ± 0.32%	38.35% ± 2.32%	**41.27% ± 1.34%**
S_{20}	93.77% ± 2.08%	98.44% ± 0.68%	**99.94% ± 0.11%**	93.76% ± 2.06%	98.45% ± 0.72%	**99.95% ± 0.12%**

limitation in the latter: it considers only the geodesic distance between EEG electrodes, leading to identical adjacency matrices for all 20 subjects (Figure 3.3). In contrast, the PCC method produces unique matrices for each subject, offering tailored matrices that are better suited for subject-based EEG MI classification (Figure 3.4).

Our experiment revealed that using the relative physical distance between EEG electrodes was suboptimal due to limited accuracy. Since EEG electrodes do not have direct connections to brain tissue, electrical signals produced by large neuron groups that fire simultaneously or synchronously need to traverse multiple tissue layers such as the cerebral cortex, cerebrospinal fluid, skull, and scalp before detected by EEG electrodes. Given that the skull attenuates these signals, and causes a smearing effect [128], coupled with individual differences in skull thickness, scalp conductivity, and MI task approach, it was the most logical to use unique adjacency matrices for each individual.

In the $A^{Geodesic}$ adjacency matrix construction, we adopted a unit sphere assumption because the PhysioNet dataset lacks data on individual head shapes. Given natural variations in head structure, $A^{Geodesic}$ values could potentially differ for each subject.

Table 3.6: Accuracy Comparison Across Different Methods of Adjacency Matrix Construction for Each Model

Model	Adj Method	Avg. Accuracy	Avg. F1 Score
Model A	Geodesic	70.70%	70.14%
	PCC	79.82%	79.77%
	EEG_GLT	**85.90%**	**85.89%**
Model B	Geodesic	70.70%	70.65%
	PCC	78.69%	78.32%
	EEG_GLT	**83.84%**	**83.80%**
Model C	Geodesic	65.49%	65.43%
	PCC	74.13%	74.41%
	EEG_GLT	**83.27%**	**83.28%**
Model D	Geodesic	62.97%	63.08%
	PCC	68.13%	68.05%
	EEG_GLT	**81.52%**	**81.48%**
Model E	Geodesic	69.20%	69.16%
	PCC	78.90%	78.88%
	EEG_GLT	**85.91%**	**85.88%**
Model F	Geodesic	69.34%	69.28%
	PCC	76.89%	77.26%
	EEG_GLT	**83.26%**	**83.36%**

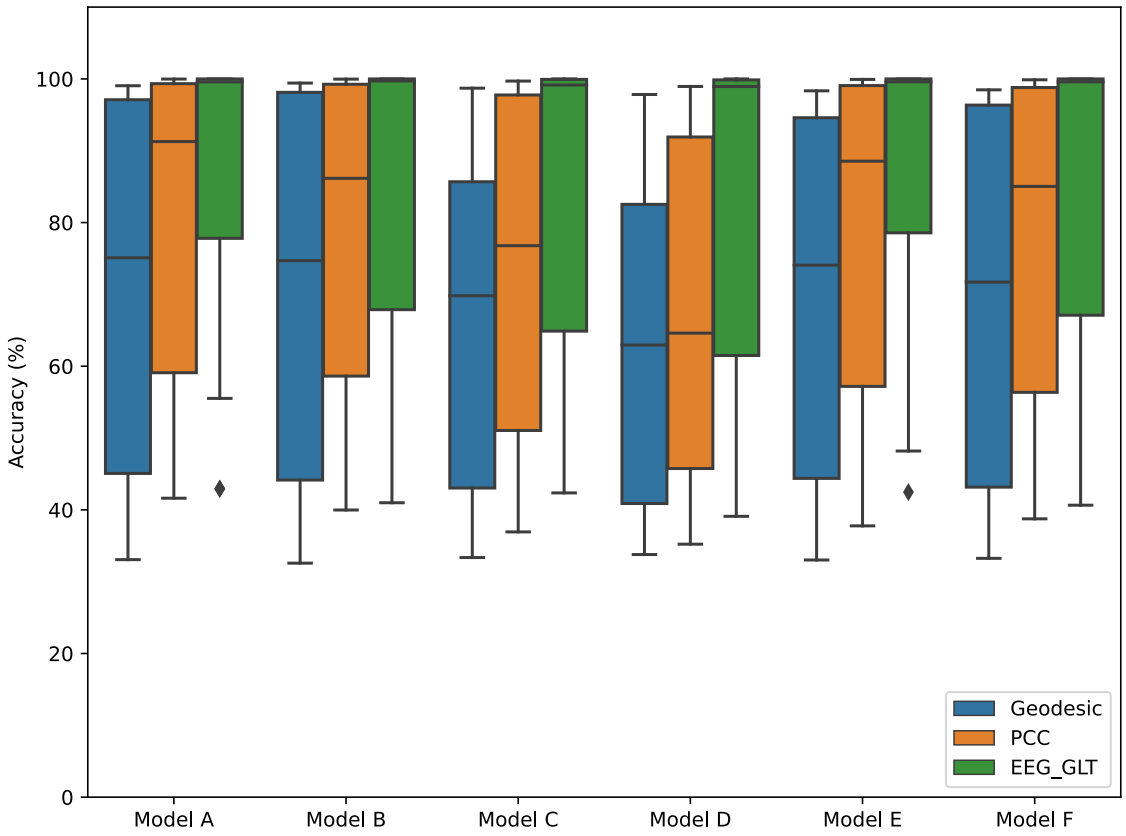


Figure 3.6: Comparison of Model Accuracy Across Different Adjacency Matrix Construction Methods

3.4.2 EEG_GLT Method vs PCC Method in Adjacency Matrix Construction

Our EEG_GLT method consistently surpassed the PCC method in both accuracy and F1 score. As shown in Table 3.4 and 3.5, EEG_GLT demonstrated substantial increase in accuracy and F1 score compared to the PCC method, by 0.52% - 22.04% and 0.50% - 21.76%, respectively. Unlike the PCC method, our EEG_GLT adjacency matrix is dynamic with the ability to adapt to both the individual subject and the model settings of GCNs (Table 3.2), as shown in Figure 3.5.

According to Table 3.6 and Figure 3.6, our EEG_GLT method improved the mean accuracies and F1 scores for 20 subjects by 13.39% and 13.43%, respectively compared to the PCC method. This underscores the necessity of model-specific adjustments, in addition to subject-based tailoring in the adjacency matrix construction, to attain the best possible outcomes. Distinctly, our EEG_GLT matrix is asymmetrical due to the iterative pruning process detailed in Algorithm 1, which refines the matrix until the optimal EEG Graph Lottery Ticket is identified.

Figures 3.7, 3.8, 3.9, 3.10, 3.11, 3.12 present the classification accuracy across various adjacency matrix densities for Subjects S_1, S_3, S_6, S_{12}, S_{14} and S_{15}. The data indicates an upward trend in classification accuracy with iterative pruning. Most importantly, the accuracy is notably lower at an adjacency matrix density of 100% in comparison to other densities. This observation suggests that some initial connections between EEG electrodes might be unnecessary, or even counterproductive, for achieving optimal classification. Removing these redundant links may boost the classification accuracy. Hence, a fully connected model between EEG channels may not be the most effective approach.

Table 3.6 displays the optimal EEG_GLT adjacency matrix ($m_{g\_EEG\_GLT}$) density for each subject. The transformation of the adjacency matrix mask m_g for the subjects S_6 and S_{14} at different densities is shown in Figure 3.13 and Figure 3.14 respectively. For subjects S_5, S_7, S_{13}, and S_{19}, their optimal $m_{g\_EEG\_GLT}$ were identified early at a 100% density. In contrast, other subjects attained their best results at densities below 22.53% for 2^{nd} order models. When considering 5^{th} order models, such as Model B, Model D, and Model F, the optimal EEG_GLTs emerged at densities of 59.00% or lower.

While our approach enhanced the accuracy for subjects S_5, S_7, S_{13}, and S_{19}, the results for both accuracy and F1 score lingered below 60.00%. A potential explanation is that relying on a single time point feature from EEG channels might not be adequate for MI tasks in these subjects, since there is inherent variability in the time required (or temporal dynamics) to execute the MI task among different individuals, as referenced in [129]. This variability might also explain why eliminating edges between EEG channels does not necessarily lead to improved performance accuracy for those subjects.

Table 3.7: Optimal EEG_GLT Adjacency Matrix ($m_{g\_EEG\_GLT}$) Density of Each Subject Across Models

Subject	Model A	Model B	Model C	Model D	Model E	Model F
S_1	18.43%	13.39%	31.30%	28.15%	18.43%	13.39%
S_2	16.57%	13.39%	13.39%	28.15%	18.43%	25.32%
S_3	18.43%	25.32%	34.80%	31.30%	25.32%	20.49%
S_4	13.39%	13.39%	14.91%	20.49%	14.91%	13.39%
S_5	100.00%	31.30%	100.00%	100.00%	100.00%	100.00%
S_6	13.39%	20.49%	100.00%	14.91%	14.91%	20.49%
S_7	100.00%	28.15%	100.00%	31.30%	100.00%	59.00%
S_8	20.49%	18.43%	13.39%	14.91%	31.30%	14.91%
S_9	13.39%	16.57%	16.57%	14.91%	13.39%	13.39%
S_{10}	13.39%	13.39%	22.77%	20.49%	13.39%	13.39%
S_{11}	13.39%	13.39%	16.57%	13.39%	13.39%	13.39%
S_{12}	14.91%	13.39%	34.80%	28.15%	16.57%	13.39%
S_{13}	80.98%	34.80%	100.00%	20.49%	100.00%	22.77%
S_{14}	13.39%	13.39%	18.43%	13.39%	13.39%	22.77%
S_{15}	14.91%	13.39%	28.15%	13.39%	22.77%	14.91%
S_{16}	14.91%	13.39%	20.49%	18.43%	13.39%	13.39%
S_{17}	14.91%	13.39%	20.49%	22.77%	13.39%	13.39%
S_{18}	14.91%	13.39%	28.15%	20.49%	22.77%	31.30%
S_{19}	100.00%	59.00%	100.00%	22.77%	100.00%	31.30%
S_{20}	25.32%	22.77%	34.80%	16.57%	20.49%	34.80%

3.4.3 Model Setting vs Adjacency Matrix Construction Methods

Based on Table 3.6, for the Geodesic method, 2^{nd} order GCN filters classify with higher average accuracy and F1 score than 5^{th} order filters. However, for the PCC and EEG GLT methods, 5^{th} order GCN filters perform better. As highlighted in Section 3.4.2, our EEG_GLT method consistently achieves better accuracy than both the PCC and Geodesic methods. This remains the case even when the EEG_GLT adjacency matrix is paired with Model D, characterised by its minimal complexity, encompassing just five spectral GCN layers with 2^{nd} order filters and a singular FC layer. These findings suggest that optimising the adjacency matrix is more importance than refining the GCN architecture when aiming for enhanced performance accuracy.

3.4.4 MACs Saving using EEG_GLT Method

The MACs inference for classifying a single-time-point EEG MI signal is influenced by several model settings, including the model framework, the number and polynomial order of GCN filters, and the specifications of FC layers as the number of layers and the node count. Among these, the count and polynomial orders of GCN filters at the GCN layers are the primary determinants of the MACs requirement. Both $A^{Geodesic}$ and A^{PCC} maintain 100% densities in their adjacency matrices. Consequently, the MACs inference for a single-time-point EEG MI signal, when using models A to F, are as follows: 81.89M, 42.26M, 22.64M, 11.32M, 291.62M, and 146.10M, respectively.

Our EEG_GLT method presents varied A^{EEG_GLT} densities due to the pruning employed by Algorithm 1. As elaborated in Section 3.4.2, the EEG_GLT approach enhances classification accuracy through pruning, which in turn decreases the MACs. Table 3.8 illustrates the percentage of MACs savings for each subject, comparing the top accuracy value from the PCC method to the EEG_GLT accuracies from models with adjacency matrix densities slightly exceeding PCC's best.

For performance equivalent to or surpassing PCC's optimal accuracy, only Models D and B with the sparsest adjacency matrix density (13.39%) are necessary. The PCC method requires between 42.26M to 291.62M for one-time-point inference across 20 subjects to reach peak accuracy. In contrast, our EEG_GLT approach needs only 8.76M to 80.67M to achieve equal or better accuracy, translating to savings in MACs of up to 97.00%.

3.4.5 Comparison with Current State-of-the-Art Models

In this paper, we compare our proposed method, EEG_GLT-Net, with seven other state-of-the-art (SOTA) models listed in Table 3.9, including FBCSP [106], EEGNet [108], CasCNN [109], DG-HAM [123], EEG-ARNN [124], SSDA [110], and GCNs-Net [100]. Our comparisons begin with the traditional FBCSP approach, which leverages CSP to extract features across multiple frequency bands and utilises SVM for classification. We then compare with EEGNet, a widely used model based solely on a CNN structure. Further, we assess CasCNN and SSDA, both of

Table 3.8: MACs Savings (%) for Each Subject: PCC's Best Model Accuracy vs. EEG_GLT Accuracy from Models with Adjacency Matrix Densities Just Surpassing PCC's Best Accuracy

Subj	PCC			EEG_GLT			MACs
	Model	Acc.	MACs	Model (Adj%)	Acc.	MACs	Saving
S_1	A	87.66%	81.89M	D (13.39%)	97.04%	**8.76M**	89.30%
S_2	B	75.43%	42.26M	B (13.39%)	78.09%	**36.97M**	12.52%
S_3	A	94.89%	81.89M	D (13.39%)	98.22%	**8.76M**	89.30%
S_4	A	99.88%	81.89M	B (13.39%)	99.98%	**36.97M**	54.85%
S_5	B	46.90%	42.26M	B (13.39%)	48.73%	**36.97M**	12.52%
S_6	E	62.92%	291.62M	B (13.39%)	70.17%	**36.97M**	87.32%
S_7	E	55.04%	291.62M	B (13.39%)	57.68%	**36.97M**	87.32%
S_8	B	98.71%	42.26M	D (13.39%)	99.78%	**8.76M**	79.27%
S_9	A	99.86%	81.89M	B (13.39%)	99.98%	**36.97M**	54.85%
S_{10}	E	99.44%	291.62M	D (13.39%)	99.97%	**8.76M**	97.00%
S_{11}	E	99.90%	291.62M	D (13.39%)	99.98%	**8.76M**	97.00%
S_{12}	A	86.76%	81.89M	D (13.39%)	99.05%	**8.76M**	89.30%
S_{13}	A	42.79%	81.89M	B (13.39%)	43.57%	**36.97M**	54.85%
S_{14}	B	63.58%	42.26M	D (13.39%)	66.25%	**8.76M**	79.29%
S_{15}	E	57.01%	291.62M	D (13.39%)	57.72%	**8.76M**	97.00%
S_{16}	B	99.80%	42.26M	D (13.39%)	99.85%	**8.76M**	79.27%
S_{17}	A	99.98%	81.89M	B (13.39%)	100.00%	**36.97M**	44.93%
S_{18}	A	96.05%	81.89M	D (16.57%)	99.58%	**8.76M**	76.14%
S_{19}	A	41.62%	81.89M	A (89.98%)	41.78%	**80.67M**	1.49%
S_{20}	B	99.17%	42.26M	D (13.39%)	99.68%	**8.76M**	79.27%

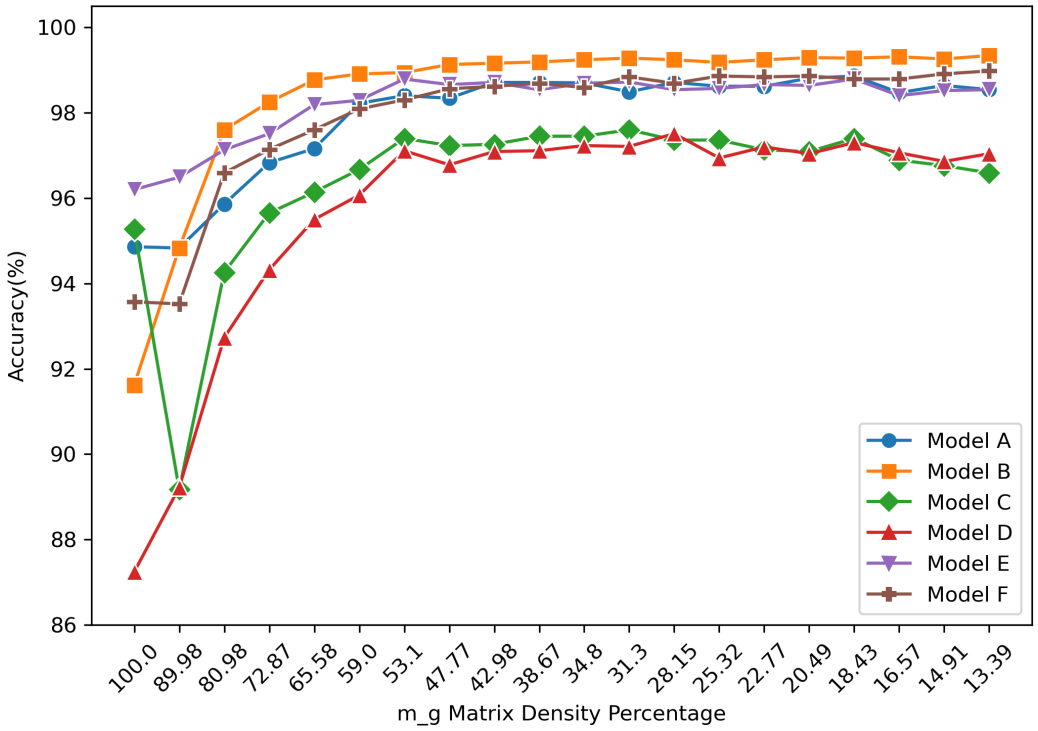


Figure 3.7: Performance Accuracy Across Different m_g Densities Using Different Models for Subject S_1

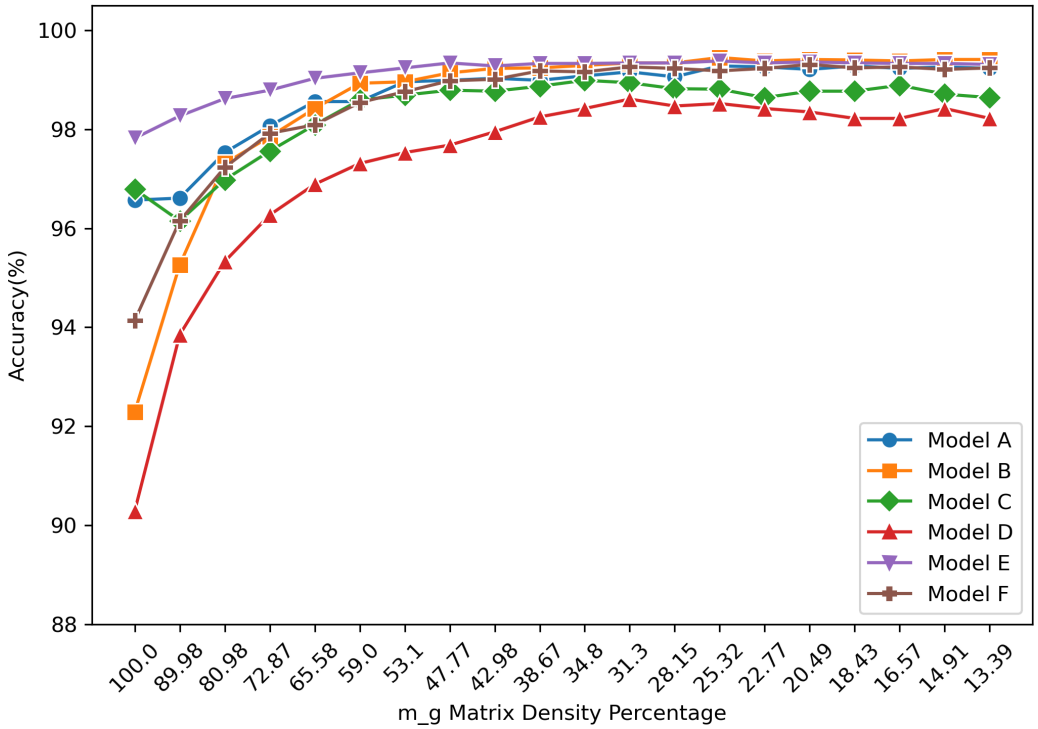


Figure 3.8: Performance Accuracy Across Different m_g Densities Using Different Models for Subject S_3

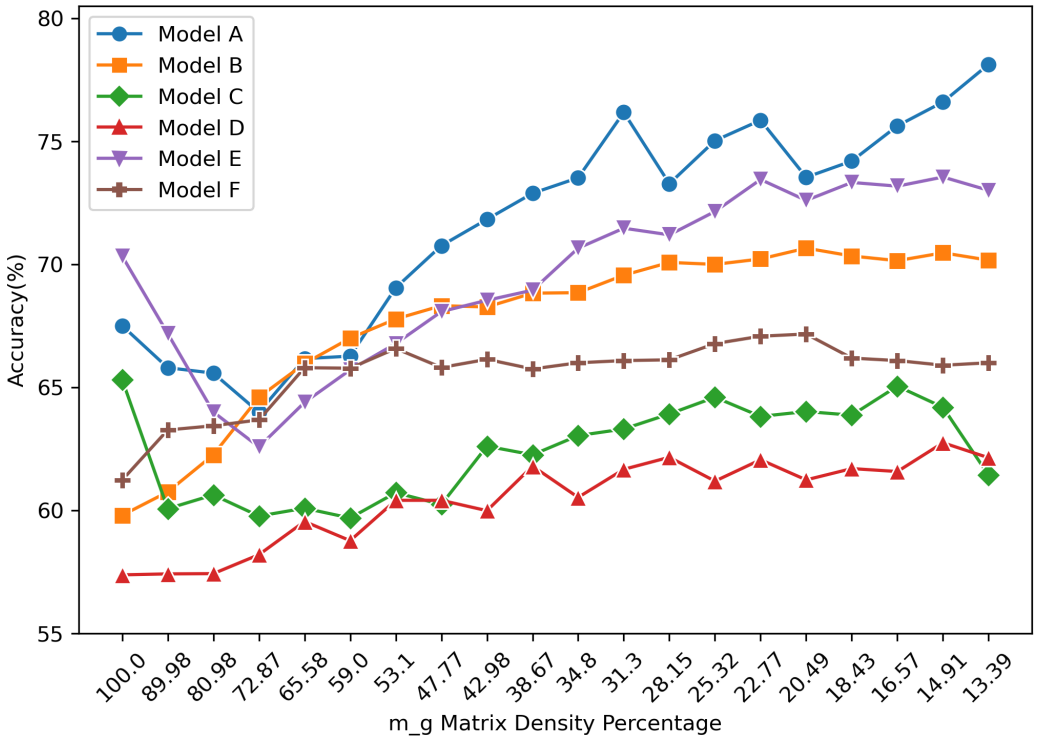


Figure 3.9: Performance Accuracy Across Different m_g Densities Using Different Models for Subject S_6

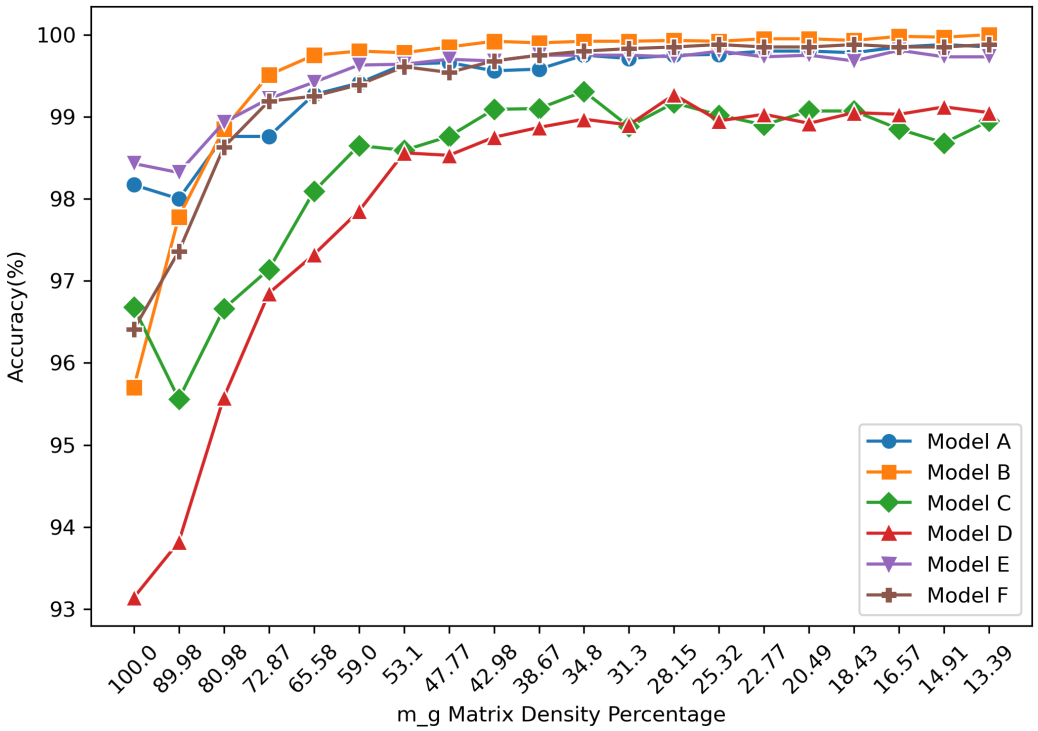


Figure 3.10: Performance Accuracy Across Different m_g Densities Using Different Models for Subject S_{12}

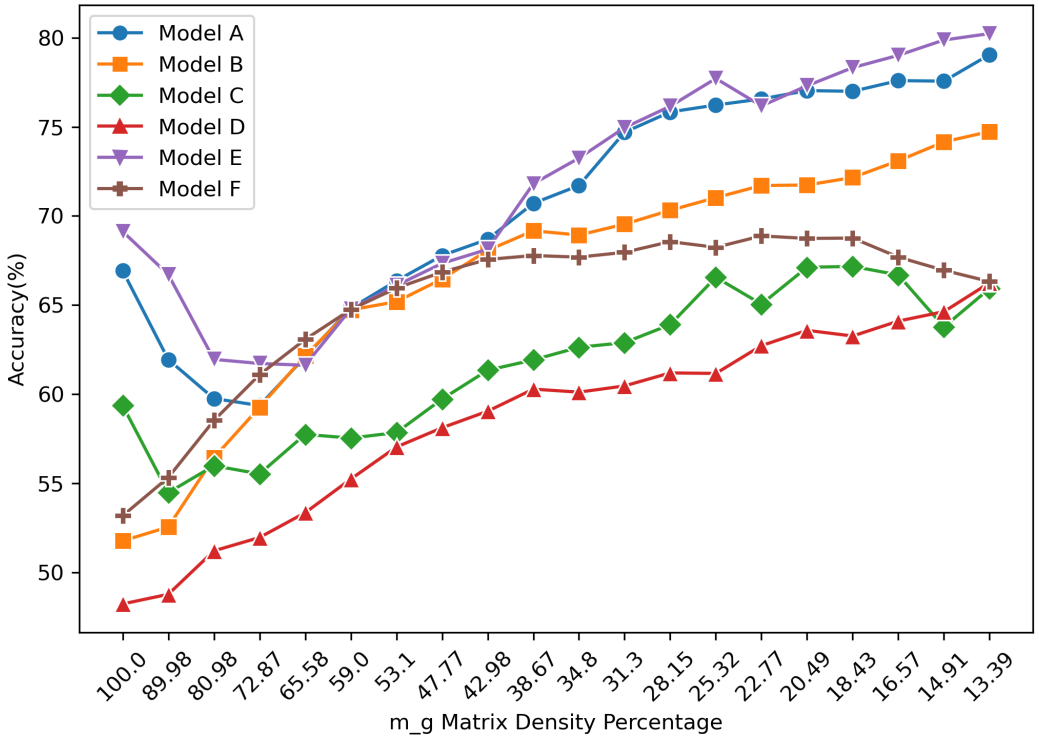


Figure 3.11: Performance Accuracy Across Different m_g Densities Using Different Models for Subject S_{14}

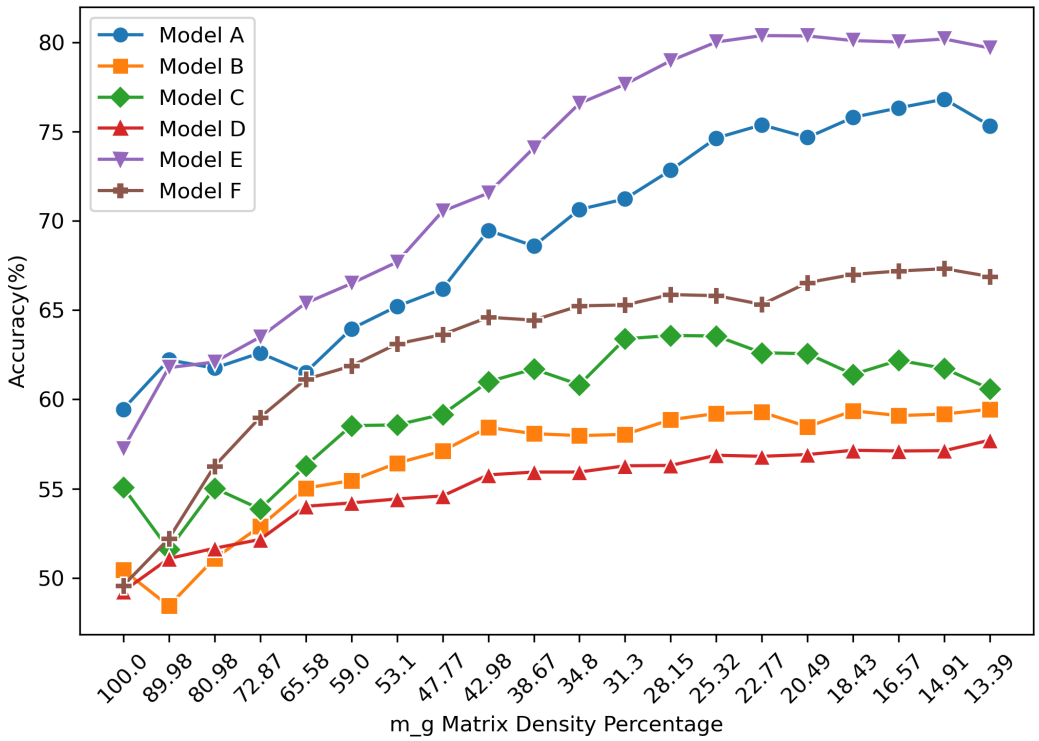


Figure 3.12: Performance Accuracy Across Different m_g Densities Using Different Models for Subject S_{15}

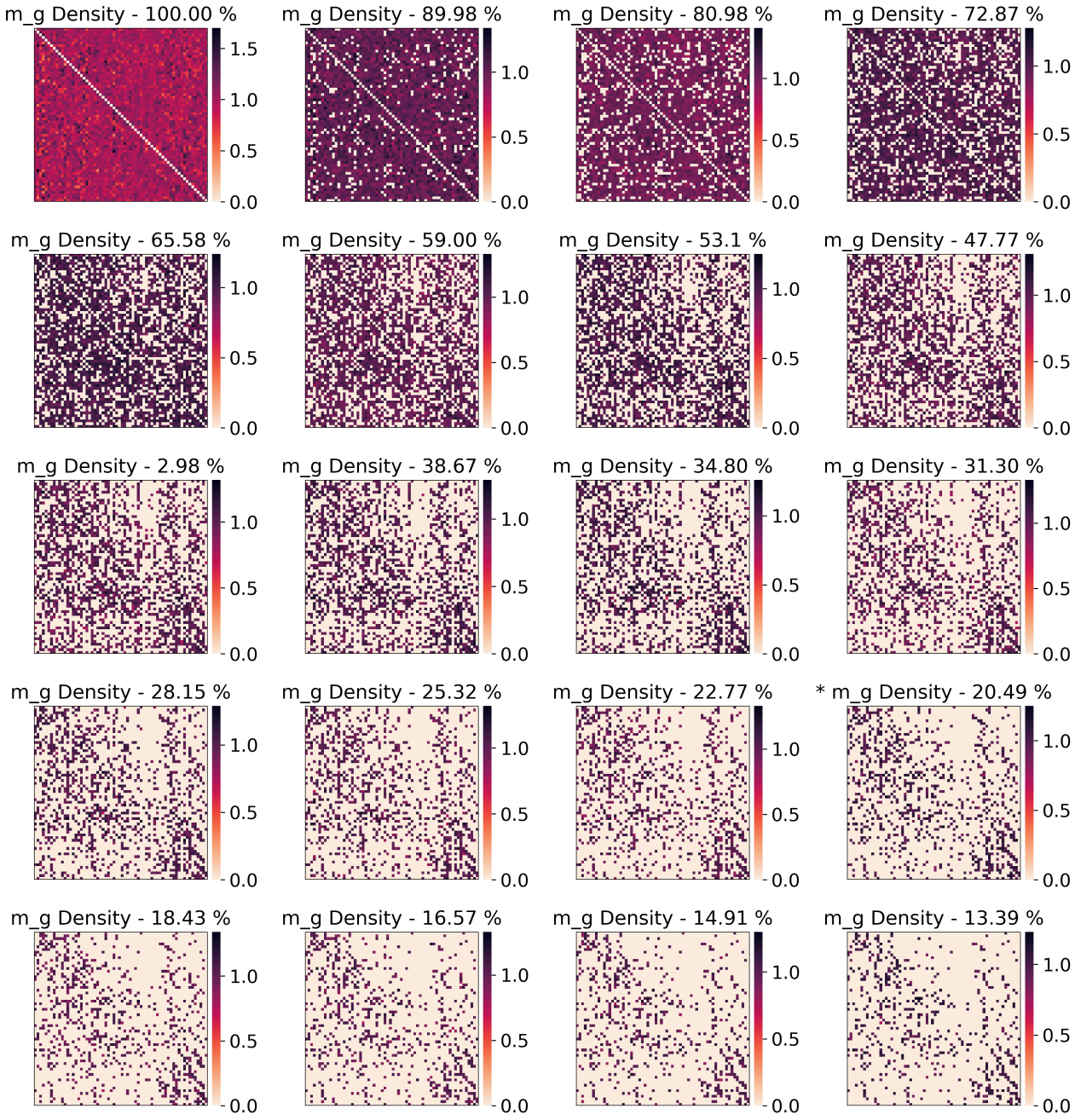


Figure 3.13: EEG_GLT Adjacency matrix mask (m_g) of Subject S_6 at different densities using Model A. The m_g density at 20.49% produces the highest accuracy of 78.13%

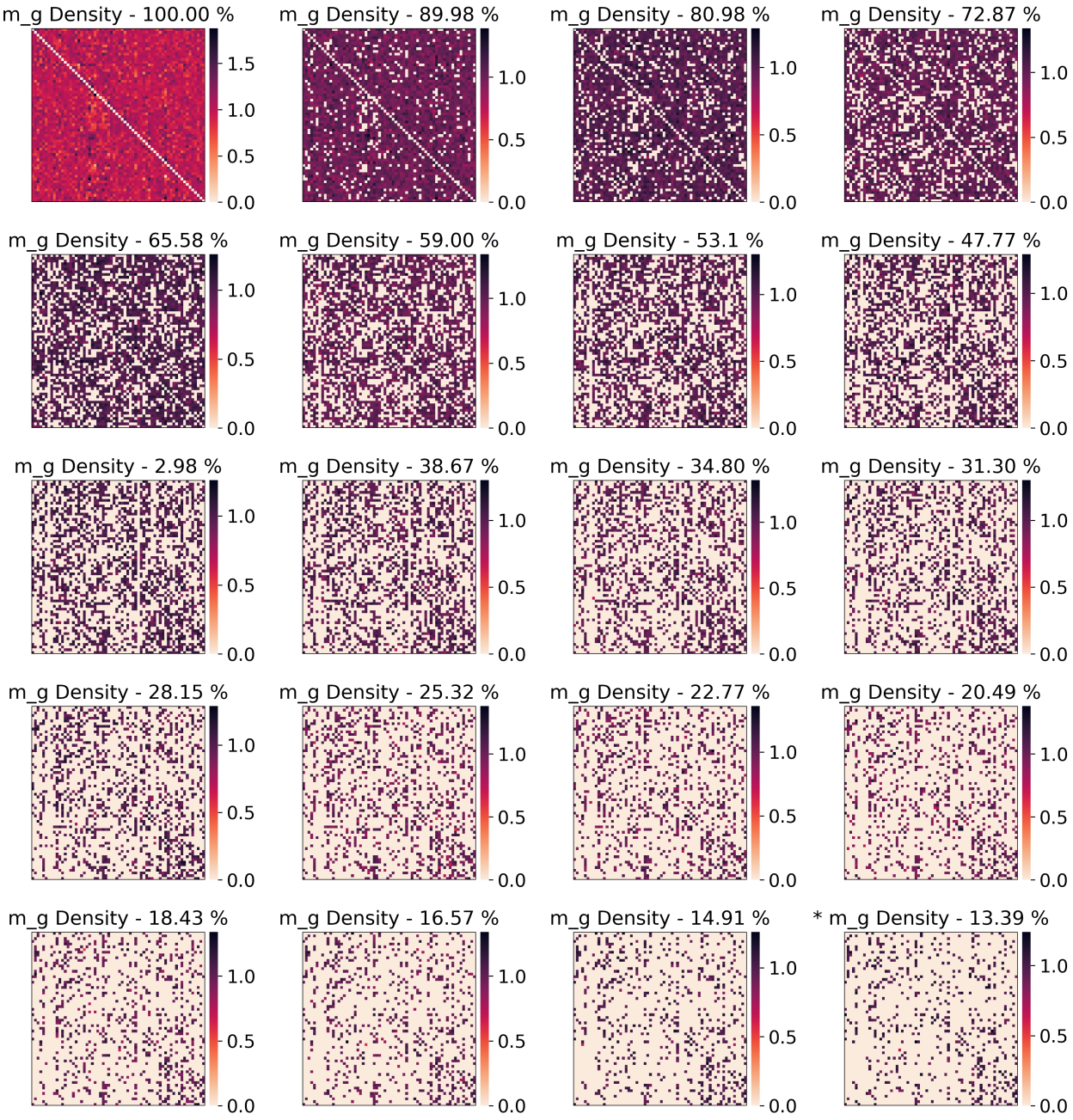


Figure 3.14: EEG_GLT Adjacency matrix mask (m_g) of Subject S_{14} at different densities using Model A. The m_g density at 13.39% produces the highest accuracy of 79.06%

which combine CNN and LSTM networks. Finally, we evaluate our method against DG-HAM, EEG-ARNN, and GCNs-Net, which are GNN-based networks.

Table 3.9: Performance Comparisons with State-of-the-Art Models

Method	Avg. Accuracy	Avg. F1 Score
FBCSP [106]	59.56%	60.04%
EEGNet [108]	72.20%	72.10%
CasCNN [109]	63.30%	63.18%
DG-HAM [123]	76.15%	76.08%
EEG-ARNN [124]	82.39%	82.17%
SSDA [110]	83.73%	83.24%
GCNs-Net [100]	80.16%	80.05%
Proposed EEG_GLT-Net	**86.43%**	**86.23%**

The traditional FBCSP method achieves 59.56%, the lowest accuracy among the SOTAs, likely due to its reliance on SVM as the classifier. The popular EEGNet achieves 72.20% accuracy, outperforming the CasCNN model, which achieves only 63.30%. Within the CNN-based SOTA models, SSDA reaches the highest accuracy at 83.73%. Among the GNN-based SOTA models, EEG-ARNN achieves the highest accuracy at 82.39%, followed by GCNs-Net and DG-HAM with accuracies of 80.16% and 76.15%, respectively.

From the perspective of adjacency matrix construction methods, the trainable adjacency matrix in EEG-ARNN outperforms the geodesic-based DG-HAM and PCC-based GCNs-Net. Our proposed EEG_GLT-Net, using single time-point classification at intervals of $(\frac{1}{160})$s, achieves the highest overall accuracy of 86.43% among all SOTAs. Notably, GCNs-Net is the only other model employing single time-point classification; however, while the GCNs-Net accuracy falls short of our EEG_GLT-Net using the EEG_GLT adjacency matrix, it surpasses our model when using a PCC-based adjacency matrix, reaching 79.82% which may be attributed to the application of pooling layer after every GNN layer within GCNs-Net.

3.5 Limitations and Future Works

We introduced a novel method for constructing an adjacency matrix in GNNs to classify single time-point EEG MI signals. The EEG_GLT-Net demonstrated high classification performance within a single-subject paradigm, where models are trained and tested on the same individual. This approach effectively captures subject-specific neural patterns, optimising classification accuracy by leveraging the unique signal characteristics of each participant. While this design enhances model precision for individualised EEG decoding, EEG signals naturally exhibit inter-subject variability due to anatomical, physiological, and cognitive differences. Exploring cross-subject settings in future work would provide further insights into the model's adaptability

across diverse users.

The EEG_GLT-Net architecture has shown strong performance in single time-point classification at ($\frac{1}{160}$s), particularly for MI tasks with distinct neural patterns, such as those in the PhysioNet dataset. However, for MI tasks with more subtle and overlapping patterns, as seen in datasets like BCIC_iv-2a, the model's effectiveness is reduced due to the inherent temporal dependencies in the data. In such cases, longer signal segments—typically exceeding one second—along with feature extraction techniques are essential to capture the meaningful temporal and spatial relationships required for accurate classification.

Nonetheless, the proposed EEG_GLT adjacency matrix construction method is not restricted to single time-point classification. A natural extension of this work is to apply EEG_GLT to additional EEG MI datasets, such as BCIC_iv-2a, and EEG movement datasets, including the High-Gamma dataset. Future adaptations will focus on integrating temporal embeddings to accommodate datasets that require sequential information while preserving the computational efficiency of the graph-based approach. Additionally, evaluating the method on large-scale multi-subject datasets will further establish its robustness, ensuring broader applicability across real-world neurorehabilitation and BCI systems.

3.6 Conclusion

Our EEG_GLT approach, developed for optimal adjacency matrix construction in EEG MI time-point signal classification, consistently outperforms both the Geodesic and PCC methods in accuracy and F1 score. It is important to note that the PCC method is currently employed in the state-of-the-art EEG time-point classification model, GCNs-Net. Specifically, our EEG_GLT method enhances accuracy and F1 score by margins ranging from 0.52% to 22.04% and 0.50% to 21.76%, respectively, compared to PCC. Furthermore, it improves the average accuracy across 20 subjects by 13.39%. With this method, optimal outcomes emerge when the adjacency matrix densities remain below 22.53%. Our study emphasises the pivotal role played by the configuration of the adjacency matrix in performance accuracy, overshadowing even model settings. In addition, our EEG_GLT approach has much higher computational efficiency, demanding between 8.76M and 80.67M MACs, which is significantly less than the 42.26M to 291.62M required by the PCC method for comparable or superior results.

While this research primarily focuses on identifying the optimal adjacency matrix, with pruning confined to the adjacency matrix, upcoming studies will explore pruning GNN and FC layers weights to further streamline computational costs. Additionally, we plan to expand the number of time points used for feature extraction, especially for subjects S_5, S_7, S_{13}, and S_{19}. In future work, we will refine Algorithm 1 to seamlessly integrate pooling layers within the GCN blocks under the EEG_GLT method, to further optimise computational efficiency. To achieve a more generalised understanding of the inter-relationships between EEG channels, it is essential to incorporate a broader range of tasks into models.

3.7 Significance and Novelty

This chapter introduces EEG_GLT-Net, a spectral Graph Neural Network (GNN) model for real-time classification of single time point EEG motor imagery (MI) signals. The significance of this work lies in its focus on addressing the computational and accuracy challenges inherent in real-time neurorehabilitation. By dynamically constructing adjacency matrices through the EEG Graph Lottery Ticket (EEG_GLT) method, the model identifies key inter-channel relationships without prior assumptions, enabling efficient and adaptive processing of EEG data. This approach demonstrates superior classification performance on benchmark datasets, making it a valuable contribution to personalised rehabilitation systems.

The novelty of EEG_GLT-Net lies in its integration of graph-based learning with a lottery ticket mechanism to enhance both interpretability and computational efficiency. Unlike traditional EEG models, which rely on static or predefined relationships, EEG_GLT-Net offers a flexible framework that adapts to individual datasets and subjects. This innovation sets a new benchmark for EEG analysis, paving the way for advanced neuroengineering applications in real-time contexts.

Chapter 4

Economic Graph Lottery Ticket: A GNN based Economic Forecasting Model

4.1 Introduction

The field of econometrics is constantly evolving due to on-going improvements in statistical methods, burgeoning data availability, and lately from rapid advances of machine learning techniques. The primary goal is to develop models that reveal the complex relationships among various economic variables, and enhance our understanding of the economic system.

Economic systems can be viewed as complex networks of interdependent relationships between variables or indicators such as inflation, wage price index, and equity returns. Simplifying these relationships between key economic variables can help create a logical network that facilitates a more intuitive understanding of the economic system.

The Wilkie Investment Model [130], a cascading stochastic economic model first developed in the actuary domain, has been extended to incorporate more economic variables, such as the SUPA model [131], [132] which is used for projecting accumulation and decumulation phases within the Australian superannuation system. However, the Wilkie model's assumptions of normally distributed asset returns and stationarity of economic variables may not always align with the real-world financial market dynamics [133]. Recently, there has been a shift towards using machine learning methods in economic modelling. For example, Scheidegger's framework combines Gaussian process regression with the active subspace method to address dynamic stochastic models [134].

The need to model non-linear relationships and to manage high-dimensional data has led to exploring the potential of Neural Networks in economic forecasting [135]. Despite criticism surrounding interpretability, these neural networks do have the potential to capture informa-

tion hidden in large datasets. However, standard implementation of Neural Networks does not provide the mechanism to account for inherent structure and connectivity in economic data and therefore cannot encapsulate complex interactions between these input economic variables. Addressing this limitation, Graph Neural Networks (GNNs) have been developed [136]. By incorporating data structure into their learning process, GNNs excel at uncovering complex patterns in structured data, such as networks of economic variables. Traditionally, these inter-relationships need to be delineated via an adjacency matrix before applying GNNs. However, establishing the adjacency matrix requires prior knowledge of the unknown relationships between the economic variables, the challenge is to utilise the delineation process to establish the complex relationship between the key economic variables without any preconceived knowledge about the relationship [113].

In this paper, we propose the Economic Graph Lottery Ticket (EGLT) algorithm, which is a novel GNN-based method that, without the prior knowledge of the inter-relationships between the economic variables, we can generate the optimal adjacency matrix as a by-product of the GNN training process from historical data only. The trained EGLT model is proven to provide more accurate economic forecasting. Compared with the SUPA model, the prediction RMSE reduced significantly for eight major Australian economic variables, highlighting the efficacy of EGLT in economic prediction.

4.2 Economic Graph Lottery Ticket Algorithm

In the proposed EGLT approach, the key difference is that we introduce a trainable adjacency matrix mask into the forward pass of the GNN. This would enable the adjacency matrix to be adaptively optimised. Such flexibility ensures that the model captures the most significant relationships without being constrained by preconceived notions. The EGLT algorithm also integrates the structure of economic data into the learning process, allowing for more accurate predictions.

The use of the adjacency matrix in GNNs is crucial for accurately modelling the relationships between nodes. By dynamically adjusting this matrix, the EGLT algorithm can better capture the complexities of economic interactions. This adaptability is particularly important in economic forecasting, where relationships between variables can change over time due to various factors such as policy changes, and market conditions or events.

More details of the EGLT approach are as follows. In the analysis, the forward pass of the Graph Neural Network (GNN) function $f(., \Theta)$ was performed on a given graph $G = \{A, X\}$, where A represents the overall adjacency matrix, as defined in Equation 4.1:

$$A = A_{original} \odot m_g \tag{4.1}$$

The original adjacency matrix $A_{original} \in \mathbb{1}^{|V| \times |V|}$ remained fixed as a unit matrix and was not

subject to training. To modify the graph structure, a trainable adjacency matrix mask, denoted as $m_g \in \mathbb{R}^{|V|\times|V|}$, was introduced. At the start of Algorithm 2, the initial adjacency matrix mask is defined as $m_g^0 = A_{original}$.

The GNN forward pass trained variables m_g, $\Theta^{(1)}$ and $\Theta^{(2)}$ across N_{ep} epochs for each adjacency matrix mask density during the training process. The lowest parts of m_g's absolute values at s^{th} density were pruned, leaving the rest set to 1, guided by the minimal MSE loss from the validation set. This cycle persisted until the m_g density fell below a pre-set minimum, κ_{min}, with all edges in the m_g at this lowest density deemed essential. The adjacency matrix density at the s^{th} pruned level, κ^s is given by

$$\kappa^s = \frac{||m_g^s||_0}{||A_{original}||_0} \tag{4.2}$$

where m_g^s represents the trainable adjacency matrix mask at the s^{th} pruned level. The state of m_g producing the smallest validation set MSE loss at the i^{th} iteration and lowest adjacency mask s^{th} density level during the GNN forward pass was defined as the Economic Graph Lottery Ticket (EGLT), with ties resolved by selecting the lower m_g density state.

The flowchart of a multiple layer GNN is shown in Figure 4.1. While the number of layers increases, the computational cost and complexity of the algorithm increases. Further details on the training EGLT algorithm are in Algorithm 2. All edges in the m_{g_EGLT} were necessary for the best predictions.

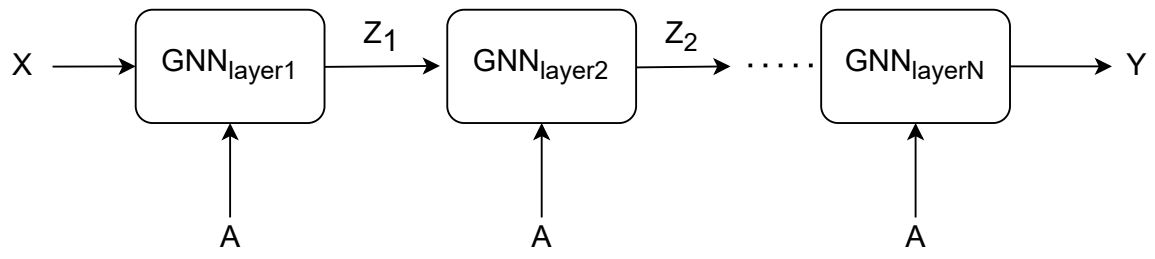


Figure 4.1: Flowchart of a multiple layer EGLT approach

Algorithm 2 Finding Economic Graph Lottery Ticket

Input: Graph $G = \{A, X\}$, GNN $f(G, \Theta)$, GNN initialisation $\Theta_0^{(1)}$ and $\Theta_0^{(2)}$, $A_{original} \in \mathbb{1}^{|V|\times|V|}$, initial Adjacency Matrix Mask $m_g^0 = A_{original}$, set learning rate η, set pruning rate p_g, pre-defined lowest Graph Density Level κ_{min}.

Output: Economic Graph Lottery Ticket $(m_{g\_EGLT}) - m_g^{s,i}$ at the lowest MSE loss with the highest sparsity possible.

1: **while** $\kappa^s \geq \kappa_{min}$ **do**
2: **for** for iteration $i = 0, 1, 2, ..., N_{ep}$ **do**
3: Forward $f(., \Theta_i)$ with $G_s = \{m_g^{s,i} \odot A_{original}, X\}$ to compute MSE Loss, L
4: Backpropagate to update $\Theta_{i+1} \leftarrow \Theta_i - \eta \nabla_{\Theta_i} L$
5: Update $m_g^{s,i+1} \leftarrow m_g^{s,i} - \eta \nabla_{m_g^{s,i}} L$
6: **end for**
7: Record $m_g^{s,i}$ with the lowest MSE loss in validation set during the N_{ep} iteration
8: Pre set p_g of the lowest absolute magnitude values in m_g^s to 0 and the others to 1, then obtain a new $m_g^{s+1,0}$
9: **end while**

4.3 Real Data Study: Implementation and Evaluation of the EGLT Approach

The EGLT algorithm provides a novel approach to economic forecasting by optimising the adjacency matrix with GNNs. In this section, we will detail the implementation of the EGLT model on real-world data, followed by an evaluation of its performance compared to the traditional SUPA model.

We implemented the EGLT approach to a comprehensive selection of eight Australian economic variables. These variables encompassed a wide range, including the Consumer Price Index (CPI), Wage Price Index (Wage), Short-Term Interest Rate (r_short), Long-Term Interest Rate (r_long), Domestic Bond Index (Dom_BI), International Bond Index (Int_BI), Unemployment Rate (Unemploy) and House Price Index (HPI).

Our analysis utilised monthly data, transforming daily variables into monthly averages and interpolating quarterly variables to match this frequency. Spanning 20 years (July 2000-June 2020), the dataset provided a substantial 240 months of data for exploration. The study used these variables' monthly values as input features. Since some variables are recorded quarterly, consecutive months weren't used as input features. For instance, data from months like January, April, and July 2008 were used to predict October 2008 values in our study. This approach helps capture seasonal trends.

Given their varying value scales, these data required normalisation via z-score normalisation to

ensure a mean of zero and a standard deviation of one. This step was crucial to prevent larger-scale features from dominating the learning process. We divided the dataset into a training set that includes data from July 2000 to December 2017, and a testing set that includes data from January 2018 to June 2020. Ten percent of the data in the training set was reserved for continuous performance monitoring as a validation set.

4.3.1 Model Initialisation and Hyper-parameter Settings

Our study employed a two-layer GNN model to optimise graph characteristics, as shown in Figure 4.2. The node features are represented by $F_{in} \in \mathbb{R}^3$ and the number of GNN hidden nodes, $M = 128$. We aimed to predict a single output month, thus $F_{out} \in \mathbb{R}^1$. Our research considered various scenarios, encompassing different input month and GNN weight combinations. The implementation details are listed in Table 4.1.

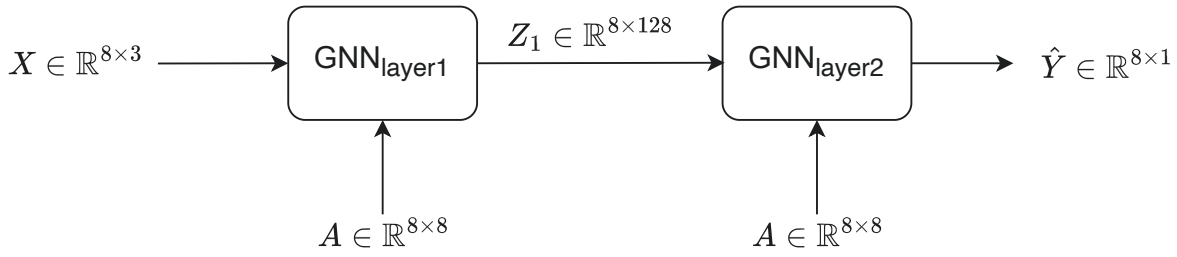


Figure 4.2: The implemented two-layer EGLT model for eight economic variables

Table 4.1: Implementation Details of the 2-layer EGLT Model

Layer	Type	Input Size	Weight	Output	Number of parameters
Input	Input graph features	$\|V\| \times F_{in}$	-	-	-
H1	First graph convolution layer	$\|V\| \times F_{in}$	$\Theta^{(1)} \in \mathbb{R}^{F_{in}\times M}$	$\|V\| \times M$	$F_{in} * M$
H2	Second graph convolution layer	$\|V\| \times M$	$\Theta^{(2)} \in \mathbb{R}^{M\times F_{out}}$	$\|V\| \times F_{out}$	$M * F_{out}$
$\widehat{Y}$	Predicted Output	$\|V\| \times F_{out}$	-	-	-

We used the PyTorch deep learning package to construct the two-layer GNN, with the training process running over 500 epochs (N_{ep}) and including a 0.5 dropout rate to reduce overfitting. The model was optimised using the Adam optimiser with a learning rate (η) of 0.01. The adjacency matrix mask m_g was set to prune at a rate (p_g) of 10% at each m_g density, down to a minimum density of 10.71%. The average MSE metric was used to assess the model's regression problem performance.

4.3.2 EGLT Model Implementation

Using the Algorithm 2, the RMSE losses corresponding to different adjacency matrix densities, κ^s, were obtained as shown in Figure 4.3. The optimal adjacency matrix, m_{g_EGLT}, was found at a density of 21.88%, as it yielded the lowest RMSE loss. Figure 4.4 (top) shows the heat map illustrating the interrelationships among the variables. Figure 4.4 (bottom) presents the nodes-edges graph, providing a visual understanding of how these economic factors are interconnected and their potential impact on the overall economic system. For example, the Unemployment Rate (Unemploy) directly affects the Wage Price Index (Wage) and House Price Index (HPI). This relationship aligns with economic observations where a high unemployment rate impacts the affordability of mortgages, hence influencing house prices.

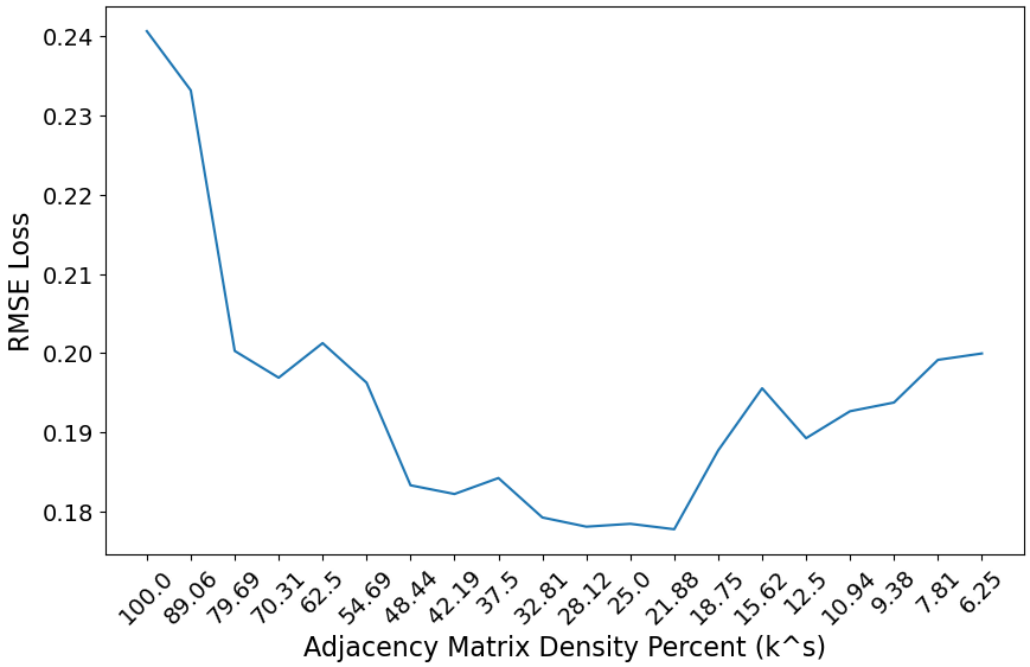
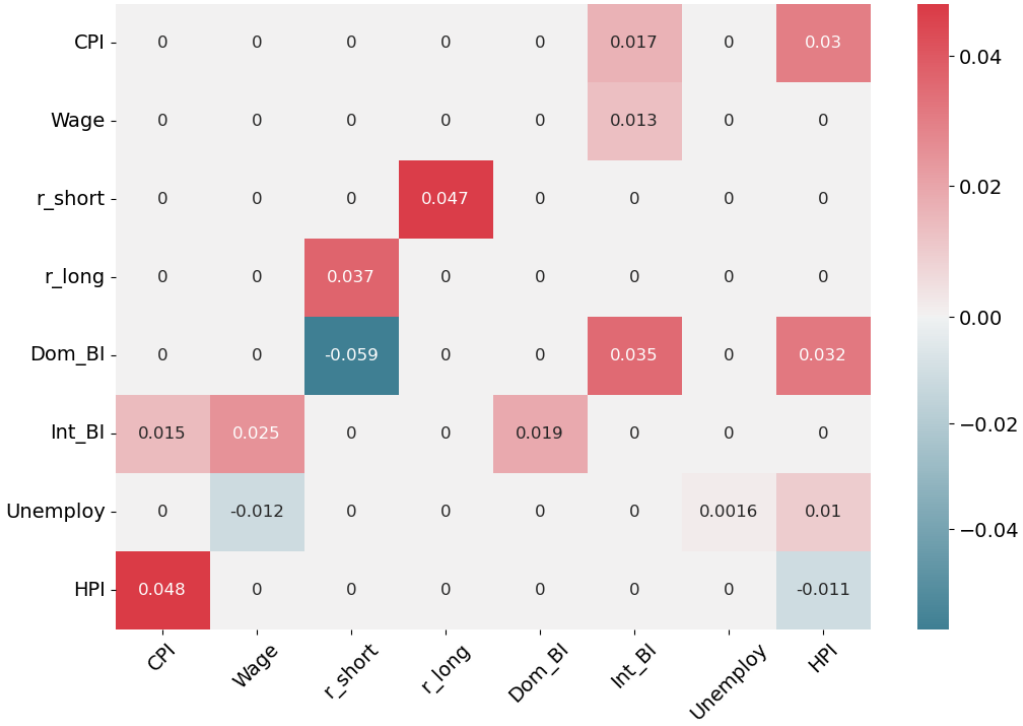


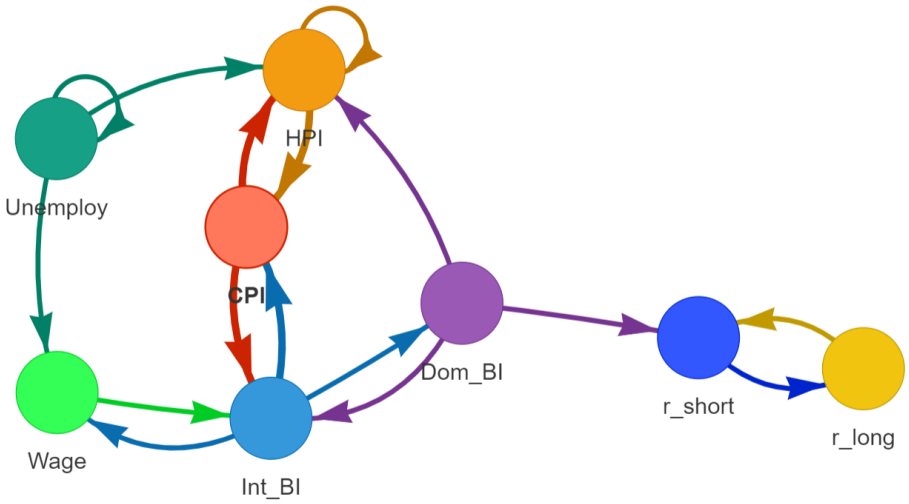
Figure 4.3: RMSE losses at different adjacency matrix densities, κ^s

In addition to the direct relationships, the heat map and nodes-edges graph reveal some indirect connections. For instance, the Domestic Bond Index (Dom_BI) shows interrelation with both short-term (r_short) and long-term interest rates (r_long). This reflects the real-world scenario where bond prices are sensitive to changes in interest rates. Similarly, the Consumer Price Index (CPI) and International Bond Index (Int_BI) have connections, indicating the influence of international economic conditions on domestic inflation.

The graphical representation of the adjacency matrix (Figure 4.4 Bottom) not only provides a clear depiction of the essential connections but also highlights the complexity of the economic network.



(a) Adjacency Matrix in Heatmap



(b) Directed Nodes-Edges Graph of Adjacency Matrix

Figure 4.4: The optimal adjacency matrix of the EGLT model, m_{g_EGLT}, is found at 21.88% density with Necessary Edges only. (a) Heatmap of the Adjacency Matrix. (b) Directed Nodes-Edges Graph representation of the Adjacency Matrix.

4.3.3 Prediction Analysis

The graphical representation of the optimised adjacency matrix shows the necessary connections between the provided economic variables. The EGLT algorithm identifies these critical relationships without prior knowledge. Such direct and indirect interactions can be crucial for accurate predictions. The performance of the trained EGLT model is validated against the current state-of-the-art SUPA model using data from January 2018 to June 2020, evaluated in terms of RMSE values. Table 4.2 lists the comparison across multiple variables. The average RMSE loss of the EGLT model is 0.1778, which is about 40% lower than the RMSE loss of the SUPA model.

Table 4.2: Comparison of the RMSE losses for the SUPA and the EGLT algorithm across various economic variables.

Variable	SUPA	EGLT
CPI	0.0204	0.0186
Wage	0.0186	0.0361
r_short	0.6170	0.0962
r_long	0.5303	0.3420
Dom_BI	0.1433	0.0795
Int_BI	0.5579	0.5536
Unemploy	0.3319	0.2108
HPI	0.0869	0.0852
Overall	**0.2883**	**0.1778**

We also show all variables' prediction results in more details as time series in Figures 4.5, 4.6, 4.7, 4.8, 4.9, 4.10, 4.11 and 4.12. In each figure, the green line represents the real value, the orange line the EGLT model prediction, and the blue line the SUPA model prediction.

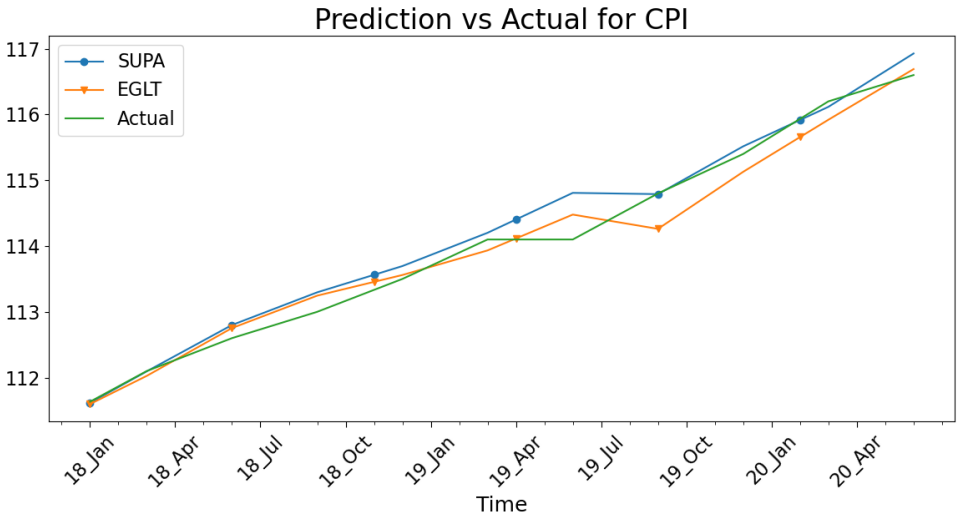


Figure 4.5: CPI Prediction Result

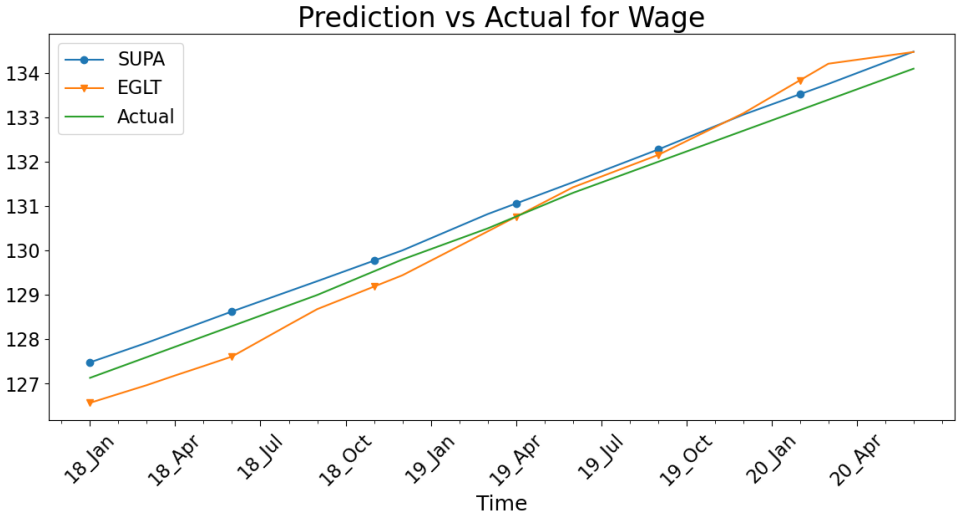


Figure 4.6: Wage Prediction Result

For the Consumer Price Index (CPI) and Wage Price Index (Figure 4.5 and 4.6), both datasets are linear and monotonically increasing. Here, the SUPA model predictions are better than the EGLT model results. The SUPA model closely follows the actual values, indicating its effectiveness in handling straightforward, linear trends.

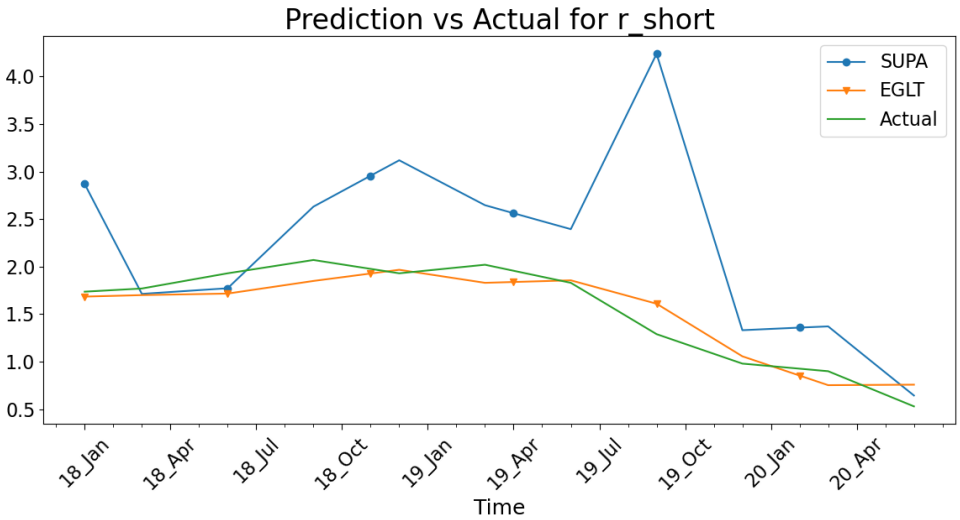


Figure 4.7: Short-term Interest Rate Prediction Result

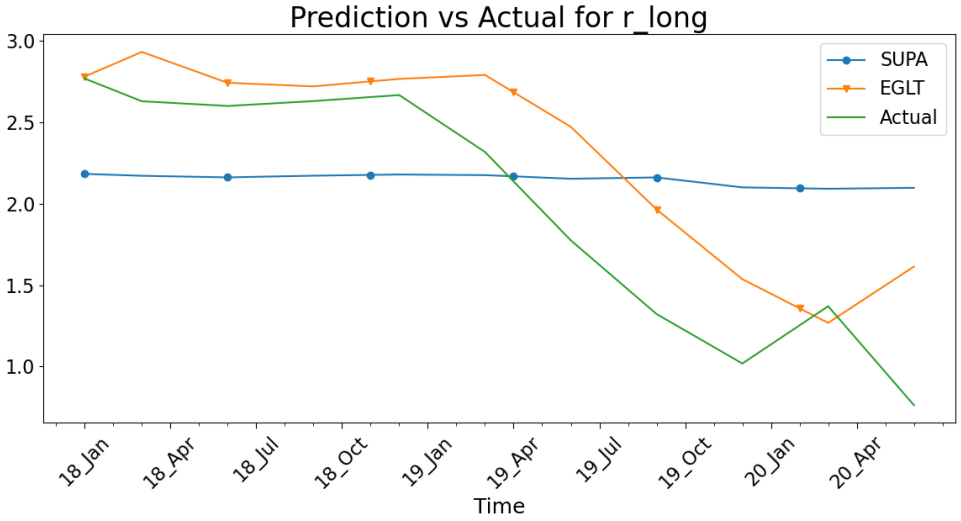


Figure 4.8: Long-term Interest Rate Prediction Result

In Figure 4.7 and 4.8, the Short-Term Interest Rate (r_short) and Long-Term Interest Rate (r_long) prediction results are depicted. The EGLT model outperforms the SUPA model, especially in capturing the rapid fluctuations in r_short. The SUPA model shows significant lag in adjusting to these changes, while EGLT adapts more quickly, highlighting its suitability for volatile datasets.

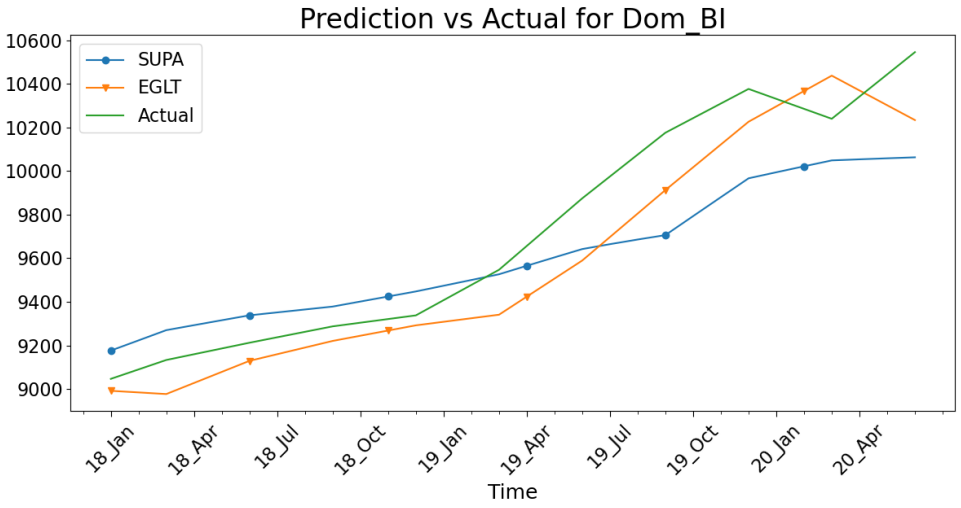


Figure 4.9: Domestic Bond Index Prediction Result

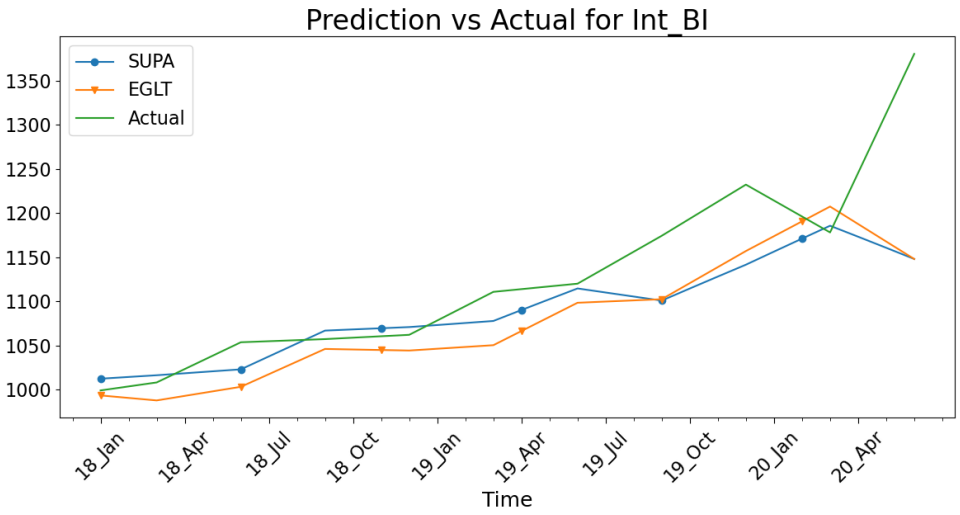


Figure 4.10: International Bond Index Prediction Result

Figure 4.9 and 4.10 shows the Domestic Bond Index (Dom_BI) and International Bond Index (Int_BI) prediction results. The EGLT model again provides better predictions, particularly for Dom_BI. The EGLT model captures both minor and major fluctuations more accurately than the SUPA model, demonstrating its robustness in dealing with the non-linear relationship.

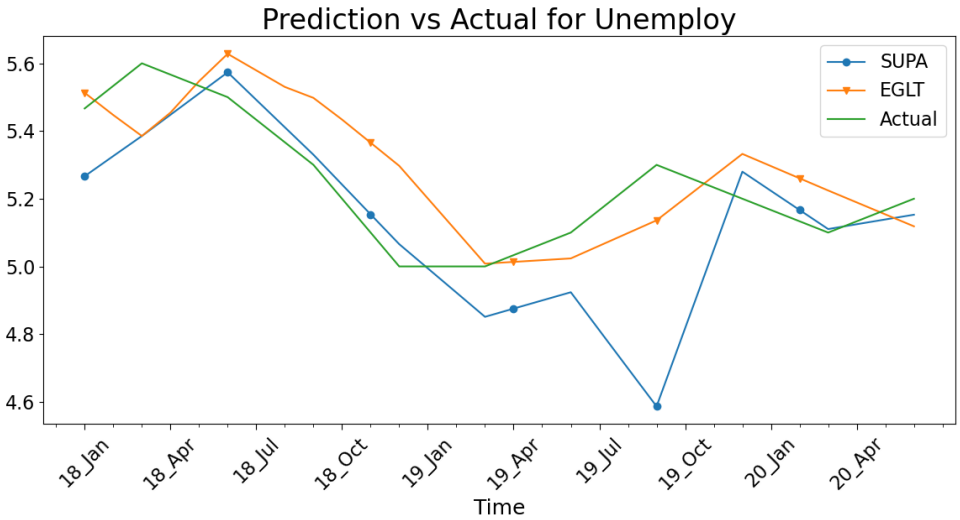


Figure 4.11: Unemployment Rate Prediction Result

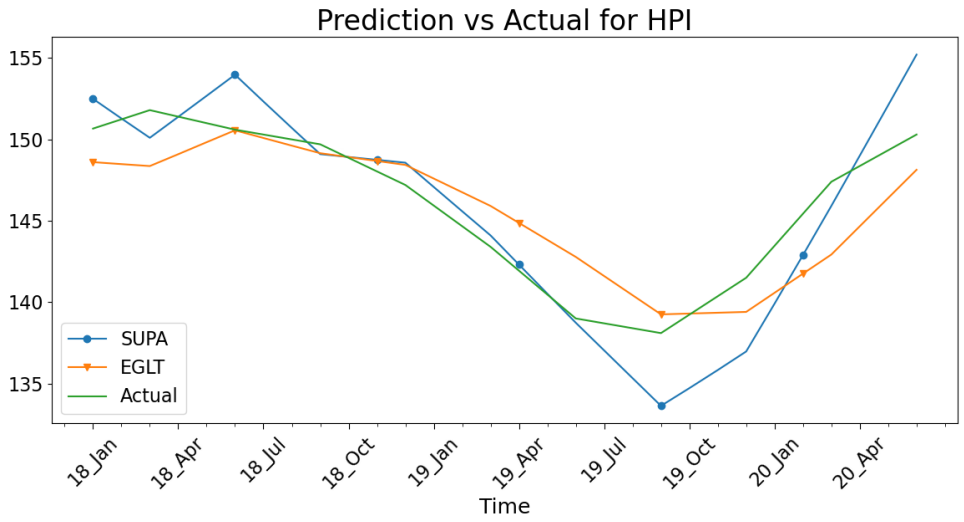


Figure 4.12: Housing Price Index Prediction Result

The Unemployment Rate (Unemploy) and House Price Index (HPI) prediction results are shown in Figure 4.11 and 4.12. The EGLT model significantly outperforms the SUPA model in predicting the Unemployment Rate. For the HPI, both models perform similarly, but EGLT has a slight edge in following the upward trends more closely.

Given the current small datasets, all the EGLT models are tested on single-step forecasting. Later on, when there are larger datasets covering longer time periods, the EGLT model can also be trained and evaluated for multiple-step long-term forecasting as well. The flexibility of the EGLT algorithm will allow it to adapt to different data conditions, making it a versatile tool for economic forecasting.

The EGLT algorithm can be readily applied to the task of identifying inter-dependency relationship structure of large number of time-series variables. We have used the EGLT algorithm to generate inter-dependency structure for eight major economic variables in Australia. For future work, an interesting test case would be for using the EGLT algorithm to quantify and reveal the complex inter-dependency relationship network among major commodity prices.

4.4 Conclusion

In this paper, within a GNN framework, we have presented the EGLT algorithm as more accurate method for economic forecasting. The EGLT algorithm automatically generates an optimal adjacency matrix without requiring prior knowledge of existing economic relationships. The EGLT algorithm can iteratively converge and reveal the most parsimonious inter-connecting relationships by minimising the Root Mean Square Error for the dataset and therefore improving the forecasting accuracy.

The forecasting accuracy of the EGLT model was compared directly with the state-of-the-art cascading SUPA model for eight major economic variables in Australia from January 2018 to

June 2020. The comparison shows that the EGLT model consistently outperforms the SUPA model across these economic variables, with a reduction in RMSE. Specifically, the EGLT model's average RMSE is 40% lower than that of the SUPA model.

In this paper, our study has focused on single-step forecasting. However, for applications in major decision making under uncertainty such as climate uncertainty, long-term multiple-step forecasting is required. In our subsequent study on EGLT algorithm, focus will be on extending the algorithm to the generation of uncertainty estimate for long-term forecasting.

The EGLT algorithm can be readily applied to the task of identifying inter-dependency relationship structure of large number of time-series variables. In this paper, we have used the EGLT algorithm to generate inter-dependency structure for eight major economic variables in Australia. For future work, an interesting test case would be for using the EGLT algorithm to quantify and reveal the complex inter-dependency relationship network among major commodity prices.

Although economic data is generally rich, the monthly resolution used in this study reduces the temporal density available for graph learning. In such settings, purely data-driven adjacency construction may overlook stable economic relationships. Incorporating domain knowledge could improve interpretability and model reliability in low-frequency forecasting tasks.

4.5 Significance and Novelty

This chapter extends the EEG_GLT framework to economic forecasting, introducing the Economic Graph Lottery Ticket (EGLT) model. The significance of this work lies in its ability to uncover complex relationships between economic variables, such as stock indices, commodity prices, and interest rates, using a graph-based representation. By improving prediction accuracy and providing actionable insights into variable interdependencies, EGLT offers a robust alternative to traditional econometric models, addressing key limitations in economic forecasting.

The novelty of this chapter is its application of GNNs to economic datasets, demonstrating the versatility of graph learning in non-biological domains. The dynamic construction of adjacency matrices through the lottery ticket framework ensures computational efficiency while maintaining high accuracy, even with large and complex datasets. This work highlights the transformative potential of graph-based methodologies in advancing economic modeling and decision-making tools.

Chapter 5

EEG_SGNet: A Parallel Architecture for Advanced EEG Signal Decoding

5.1 Introduction

Brain-Computer Interfaces (BCIs) bridge neuroscience and engineering, enabling direct communication between the human brain and external devices. Initially developed for individuals with motor impairments [7], BCIs now extend to diverse applications, including gaming [137], neurorehabilitation [138], smart homes [139], etc. The primary function of BCIs is to translate neural signals into actionable commands. Brain activity is typically measured using four modalities: magnetic, optical, metabolic, and electrical [140]. Popular methods include magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), functional near-infrared spectroscopy (fNIRS), electroencephalography (EEG), and electrocorticography (ECoG). The choice of modality depends on the application and associated costs.

Among these, ECoG and EEG are electrical-based methods for measuring brain activity. ECoG offers superior spatial resolution compared to EEG but requires invasive surgery to place electrodes directly on the cortex [83]. This invasiveness makes ECoG costly and less accessible. Conversely, EEG captures brain signals through electrodes placed on the scalp, making it non-invasive, portable, and cost-effective. These characteristics have contributed to EEG's widespread adoption. EEG can record various types of brain activity, including event-related, spontaneous, and stimulus-evoked signals [84].

Motor Imagery (MI) involves the mental simulation of body movements, such as imagining moving a hand or leg, without actual physical movement [85], [86]. When paired with EEG, MI captures neural signals generated by the intention to move. Integrating EEG-based MI with BCIs enables these signals to be decoded into commands for controlling external devices, such as robotic exoskeletons. This capability is particularly significant for stroke survivors, offering the potential to restore quality of life and support daily activities. Accurate decoding

of EEG MI signals is essential for facilitating patient-intended movements, necessitating robust machine learning models as a core component of BCIs.

Traditional EEG MI classifiers often rely on machine learning techniques that use manually engineered features, such as wavelet transforms or analytic intrinsic mode functions [104], [105]. A widely adopted method is the Filter Bank Common Spatial Pattern (FBCSP), which applies common spatial patterns (CSP) across various frequency bands in EEG signals to extract discriminative features.

Deep learning, a subset of machine learning, leverages multilayered neural networks to process diverse data types. Convolutional Neural Networks (CNNs), a prominent deep learning architecture, mimic the human visual system's neural image recognition processes [89]–[91]. A key element of CNNs is the convolutional layer, which uses filters to scan input data and extract meaningful features, enabling the network to learn complex patterns and representations. CNNs are particularly effective for processing Euclidean data, such as one-dimensional sequences and two-dimensional grids [91]. Beyond their widespread use in image recognition [90], CNNs have demonstrated success in speech recognition [141], natural language processing [142], drug to drug interaction [143], and physiological signal classification [107], [108], often eliminating the need for manual feature extraction.

In EEG signal classification, CNNs typically represent EEG data as two-dimensional inputs, with channels and time points forming the dimensions. This enables CNNs to effectively capture both spatial and temporal dependencies in brain signals, facilitating applications such as motor imagery classification, epileptic seizure detection, and motor movement classification. Several CNN-based models, such as ShallowNet [107] and EEGNet [108], excel at extracting temporal features from one-dimensional and two-dimensional Euclidean data, achieving high classification accuracy. Further advancements were made by incorporating Long Short-Term Memory (LSTM) blocks, as demonstrated in [109], [110]. Additionally, [32] proposed the ConTraNet model [112], which combines Transformer and CNN blocks, resulting in improved performance in EEG-based emotion recognition. The EEGProgress model [111] adopts a unique approach by applying CNN operations to individual brain regions for EEG MI signal classification, focusing on regional processing rather than all channels simultaneously.

Graph Convolutional Neural Networks (GNNs) represent another branch of deep learning, specifically designed for graph-structured data [92], [113]. Unlike CNNs, which are limited to Euclidean data, GNNs can process non-Euclidean data by incorporating topological relationships during convolutional operations. GNNs can operate on various graph types, including homogeneous or heterogeneous, weighted or unweighted, and signed or unsigned graphs, which are typically represented by adjacency matrices. GNNs are used in three main tasks: node-level, edge-level, and graph-level classifications. At the node level, GNNs predict properties of individual nodes, while edge-level tasks involve predicting edge attributes, and graph-level tasks classify entire graphs based on their structure and features. GNNs are broadly categorised into spatial [117]–[120] and spectral [114]–[116] methods, though spatial methods often

face challenges in matching local neighborhood structures [96] and [121].

In EEG applications, both time-domain and frequency-domain features are extracted as graph features for use in GNN operations. Time-domain features include root mean square (RMS), skewness, variance, min-max values, zero crossings, Petrosian fractal dimension, Higuchi fractal dimension, and Hurst exponent [98], [122]. Frequency-domain features, such as power spectral density (PSD) and power ratio (PR), are computed within specific frequency bands, including δ (0.5-4 Hz), θ (4-8 Hz), α (8-13 Hz), β (13-30 Hz), and γ (30-110 Hz).

Constructing the adjacency matrix is critical for GNN performance, particularly in EEG applications. In this context, EEG electrodes serve as nodes, and the relationships among them are represented in the adjacency matrix, enabling GNNs to capture connectivity patterns beyond the capabilities of traditional CNNs. Various methods have been proposed for constructing adjacency matrices in EEG-based applications. For example, Geodesic distances between electrodes have been used to define connections [93]–[95], while other studies, such as [101], have utilised the Phase Lag Index (PLI) in their CSGNN model. Pearson Correlation Coefficients (PCC) have also been employed to evaluate inter-channel relationships [96]–[100], and [102] introduced a trainable adjacency matrix for EEG classification tasks.

Aung et al. [144] proposed the EEG_GLT method for adjacency matrix construction in their EEG_GLT-Net model, which was specifically designed for single time point EEG MI classification. This method demonstrated superior performance accuracy and computational efficiency compared to Geodesic and PCC-based methods under various model configurations. However, EEG_GLT was applied in single time point signal classification, making it unsuitable for tasks requiring longer temporal features, such as EEG movement classification.

Other models, such as DG-HAM [123] and EEG-ARNN [124], classify EEG signals within predefined window lengths without extracting graph-based features like time-domain or frequency-domain metrics. In contrast, [100] introduced GCNs-Net for time-point classification, treating each time point as an independent feature, enabling detailed time-resolved EEG MI analysis. While GCNs-Net achieves strong performance in single-time-point classification, its reliance on functional connectivity alone during GNN operations limits its overall accuracy.

In this chapter, we propose EEG Synergistic Gated Network (EEG_SGNet), a novel architecture designed to address existing limitations in EEG signal classification by employing window-based analysis. The model integrates parallel branches of CNNs and GNNs, utilising the EEG_GLT method for adjacency matrix construction. This innovative approach captures both spatial and temporal features, enabling more comprehensive EEG signal classification and expanding the potential applications in EEG analysis. The primary contributions of this study are summarised as follows:

- **EEG_SGNet Architecture:** We introduce a novel method that optimally combines CNNs and GNNs with the EEG_GLT adjacency matrix for window-based EEG signal classification.

- **Optimal Parallel Branches of CNNs and GNNs:** We thoroughly investigated EEG_SGNet across nine model configurations, varying the number of GNN and CNN filters, to determine the optimal settings for each subject, ensuring tailored and effective performance.
- **Performance Validation:** The proposed EEG_SGNet model was rigorously evaluated using two EEG datasets: the BCIC_iv-2a EEG Motor Imagery (MI) dataset [145] and the High-Gamma (HGD) EEG movement dataset [58]. These evaluations validated the model's reliability and robustness. Furthermore, we compared EEG_SGNet against eight state-of-the-art models for each dataset, demonstrating its superior performance and effectiveness.

5.2 Methodology

5.2.1 Overview

The EEG_SGNet is composed of two parallel branches, GNN branch and CNN branch. The window size of 4 seconds which has 1000 time points. In the CNN path, the input signal of EEG signal is first pass through the CNN temporal filter block which extract temporal features using 25 filters at each EEG channel. Then, it goes through the CNN block 1 which extract the spatial relationship between EEG channels, which is followed by Max pooling layer. The signals have undertaken of CNN and max pooling until Block N. Then, signals CNN convoluted features are then flattened to become 1D array, $X_{CNN} \in R^{1400}$.

On the other branch in the GNN branch, the signal is pass through Feature Extractor Block which extract graph features of the signal manually in both time-domain and frequency-domain, wavelet transform, which are detailed in the Section 5.2.4, in total of 29 features for BCIC_iv-2a dataset and 31 features for HGD dataset. Then, the extracted features are passing through GNN blocks where graph convolution operation was undertaken using the defined EEG Graph which is represented with adjacency matrix. After undertaken of N^{th} GNN operation, it is then flatted to become 1D array, $X_{GNN} \in R^{1408}$.

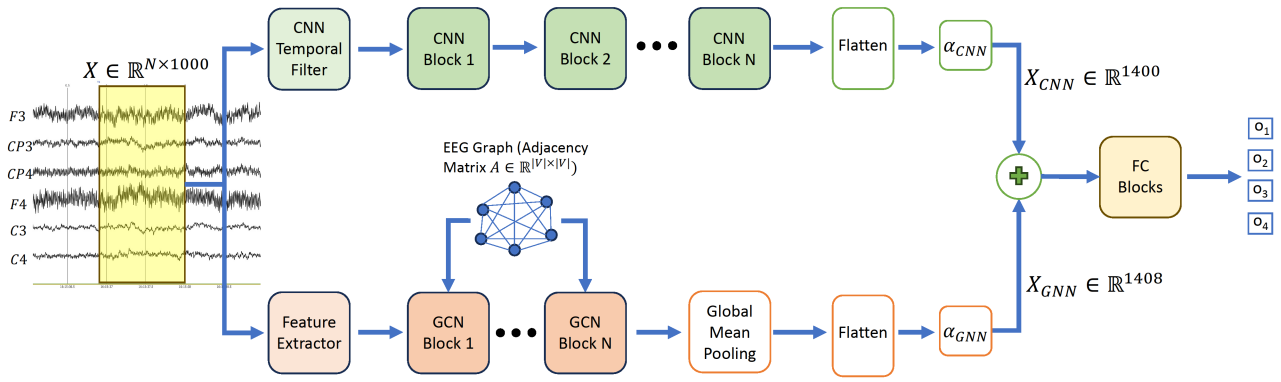


Figure 5.1: Overall Architecture of EEG_SGNet Model

5.2.2 Dataset Description

Two EEG datasets are used in the EEG_SGNet model. The first dataset is the BCIC_iv-2a dataset [145], which is widely used in BCI research, particularly for motor imagery (MI) studies. This dataset was organised by the Graz University of Technology. EEG signals were recorded from nine subjects, each of whom participated in two sessions on separate days. Each session included six runs, with a total of 288 trials of 4-second EEG MI per subject, resulting in 576 trials per subject across both sessions. EEG signals were captured using 22 EEG channels, alongside three monopolar EOG channels as shown in the Figure 5.2. The signals were sampled at 250 Hz and filtered with a bandpass filter of 0.5 Hz to 100 Hz, as well as a notch filter at the power line frequency of 50 Hz. The dataset contains four distinct motor imagery tasks:

- Class 1: Left hand MI
- Class 2: Right hand MI
- Class 3: Both feet MI
- Class 4: Tongue MI

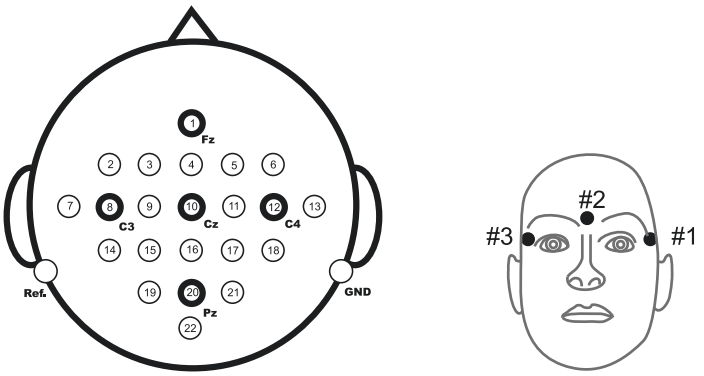


Figure 5.2: Electrode positions of BCIC_iv-2a dataset[145]. Left: Electrode montage corresponding to the international 10-20 system. Right: Electrode montage of the three monopolar EOG channels

The second dataset is the High-Gamma Dataset (HGD) [58], which consists of EEG recordings from 14 healthy subjects, with an average age of 27.2 years, including six female participants. Each subject performed 1000 trials of 4-second duration. Unlike the BCIC_iv-2a dataset, the High-Gamma EEG signals were recorded during motor execution tasks using 128 EEG channels, and were also sampled at 250 Hz. This dataset includes four motor movement tasks:

- Class 1: Left hand movement
- Class 2: Right hand movement
- Class 3: Both feet movement
- Class 4: Rest

5.2.3 Dataset Preprocessing

In the BCIC_iv-2a dataset, all 22 EEG channels are utilised in this project. For the HGD dataset, however, only 44 out of the original 128 channels are selected, specifically those covering the motor cortex, excluding the Cz channel. Given that the BCIC_iv-2a dataset pertains to motor imagery (MI), only EEG signals below 40 Hz are relevant. Although the original recordings underwent filtering between 0.5 Hz and 100 Hz, we further preprocess by first scaling the signals by 10^6 to convert the unit from V to μV. The data is then bandpass-filtered within the range of 4-38 Hz. For the HGD dataset, which involves motor execution tasks, a broader range is retained due to the importance of gamma band (γ) frequencies in movement-related EEG activity. Accordingly, we apply a bandpass filter from 4 Hz to 123 Hz.

Following bandpass filtering, we standardise the data within each trial using a channel-wise moving mean and variance, dynamically adjusting to local trends in each trial. Let x_t represent the original EEG signal at time t, and x'_t denote the standardised signal. The standardised signal x'_t is calculated as follows:

$$x'_t = \frac{x_t - \mu_t}{\sqrt{\sigma_t^2}} \tag{5.1}$$

where μ_t and σ_t^2 are the moving mean and variance at time t, respectively. To calculate these, we use a decay factor of 0.999 to balance immediate and historical data, ensuring that the standardisation remains responsive to new values while preserving stability across the trial. The moving mean μ_t and variance σ_t^2 are updated with the following recursive formulas:

$$\mu_t = 0.001 x_t + 0.999 \mu_{t-1} \tag{5.2}$$

$$\sigma_t^2 = 0.001 (x_t - \mu_t)^2 + 0.999 \sigma_{t-1}^2 \tag{5.3}$$

Here, the decay factor allows each new value to contribute 0.1% to the updated mean and variance, while the previous values retain 99.9% of the weight. This method provides a continuous standardisation that adapts to changes in the signal over time, improving the robustness of the feature extraction and classification stages.

5.2.4 Feature Extraction

In this project, we use a window-based approach to classify EEG signals, extracting a comprehensive set of features specifically for the GNN pathway. These features include time-domain features, frequency-domain features, wavelet transform coefficients, and band-specific power ratios.

For the GNN pathway, time-domain features capture fundamental statistical characteristics of the EEG signal within each window. Calculated features include the mean and standard deviation to provide an overall measure of the signal level and variability. Root mean square (RMS) represents signal power, which is especially relevant for analysing the energy content in EEG data. Additionally, skewness and kurtosis are computed to characterise the asymmetry and "tailedness" of the signal distribution, allowing us to identify any significant deviations or outliers. Zero-crossing rate (ZCR) measures the frequency content of the signal by counting the number of times it crosses the zero axis. Peak-to-peak amplitude captures the full range of signal fluctuations, and entropy serves as a measure of the signal's randomness or complexity, which can help differentiate structured brainwave activity from less predictable patterns.

The frequency-domain features are obtained by calculating the power spectral density (PSD) of the signal using Welch's method, providing a detailed breakdown of signal power across different frequency components. Band-specific power is calculated within key EEG frequency bands - θ (4-8 Hz), α (8-13 Hz), β (13-30 Hz), and γ (30-38 Hz) by integrating the PSD across each band's range, allowing us to capture cognitive-state-specific activities associated with these bands. Additional frequency-domain features include mean frequency, which indicates the central tendency of the frequency distribution by weighting each frequency by its power, and median frequency, which divides the power spectrum into two equal halves, indicating energy concentration. Spectral entropy measures the complexity of the frequency distribution by calculating the entropy of normalised PSD values. Hjorth mobility and Hjorth complexity are also computed, assessing the rate of change and complexity of the frequency components, respectively.

To capture transient features in the signal, wavelet transform coefficients are extracted using the Discrete Wavelet Transform (DWT), which decomposes the signal into various frequency bands. For each level of decomposition, the mean and standard deviation of the wavelet coefficients are calculated, offering a refined view of signal changes over time. This approach is beneficial for identifying non-stationary events in the EEG, which are common in brain activity.

Band-specific power ratios are also calculated to reveal the relative power contributions of different frequency bands, providing insights into the EEG's dynamic state. Ratios such as the beta-to-alpha ratio serve as indicators of cognitive load and alertness, while the gamma-to-theta ratio can reflect high-frequency engagement relative to slower theta activity. The alpha-to-theta ratio offers insights into relaxation versus alertness states.

Together, this feature extraction pipeline yields a total of 29 features for the BCIC_iv-2a dataset.

For the HGD dataset, additional gamma sub-bands are included - γ_{low} (30-60 Hz), γ_{mid} (60-90 Hz) and γ_{high} (90-123 Hz) resulting in a total of 31 features.

In contrast, the CNN pathway requires no manual feature extraction. Temporal features within each EEG window are automatically extracted using a temporal convolution layer (Temp_Conv), where 25 filters are applied to each EEG channel to capture temporal patterns. This is followed by spatial feature extraction across EEG channels using another set of 25 filters, capturing inter-channel relationships. This end-to-end CNN approach enables automatic learning of complex temporal and spatial patterns in the EEG data, which can significantly enhance classification accuracy.

5.2.5 Graph Preliminary

Graph Representation

Consider a directed weighted graph represented as $G = \{V, E\}$. Here, $|V| = N$ denoted the number of nodes and $|E|$ was the count of edges connecting the nodes. The node set was defined as $V = \{v_1, v_2, \ldots, v_n\}$. The node feature matrix of the entire graph was represented by $X \in \mathbb{R}^{N \times F}$, where F, the number of features, is equal to 1. The adjacency matrix, denoted as $A \in \mathbb{R}^{N \times N}$, captured the graph's overall topology. Specifically, if an edge existed between nodes v_i and v_j (i.e., $(v_i, v_j) \in E$), then $A[i, j] \neq 0$. Otherwise, $A[i, j] = 0$.

The adjacency matrix for the PCC method, $A^{PCC} \in \mathbb{R}^{N \times N}$ was defined in Equation 5.5, where I was the identity matrix and $|P|$ was the absolute PCC matrix of dimension of $N \times N$. Each element P_{ij} of P was calculated using Equation 5.4 as follows:

$$P_{ij} = \frac{cov(x_i, x_j)}{\sigma_i \sigma_j} \tag{5.4}$$

The absolute PCC matrix $|P|$ captured the linear correlations between EEG channel signals, with its elements satisfied $0 \leq |P_{ij}| \leq 1$. The adjacency matrix A^{PCC} is given by:

$$A^{PCC} = |P| - I \tag{5.5}$$

The degree matrix, D, was a diagonal representation of A, where the i^{th} diagonal element of D was computed as $D_{ii} = \sum_{j=1}^{N} A_{ij}$. The combinatorial Laplacian matrix, $L \in \mathbb{R}^{N \times N}$, was described as $L = D - A$. A normalised version of this combinatorial Laplacian can be obtained with $L = I_N - D^{-1/2} A D^{-1/2}$.

Spectral Graph Filtering

The eigenvectors of the graph Laplacian matrix can be expressed as graph Fourier modes, with $\{u_l\}_{l=0}^{N-1} \in \mathbb{R}^N$. The diagonal matrix of these Fourier frequencies, $\Lambda = diag[\lambda_0, ..., \lambda_{N-1}] \in$

$\mathbb{R}^{N\times N}$. We defined the Fourier basis, $U = [u_0, ..., u_{N-1}] \in \mathbb{R}^{N\times N}$, which allows for the decomposition of the Laplacian matrix, L, into $L = U\Lambda U^T$. The signal x can be transformed by graph Fourier into $\hat{x} \in \mathbb{R}^N$ using $\hat{x} = U^T x$, while the inverse graph Fourier transform is given by $x = U\hat{x}$. The convolution operation on graph G is defined as:

$$x *_G g = U((U^T x) \odot (U^T g)) \tag{5.6}$$

where g represents the convolutional filter and $\odot$ denotes the Hadamard product. Given that $g_\theta(\Lambda) = diag(\theta)$, where $\theta \in \mathbb{R}^N$ represents the vector of Fourier coefficients, the Graph convolution operation can be implemented as follows:

$$x *_G g_\theta = g_\theta(U\Lambda U^T)x = U g_\theta(\Lambda)U^T x \tag{5.7}$$

where g_θ is a non-parametric filter, and polynomial approximation is employed to mitigate the excessive computational complexity. Chebyshev graph convolution, a specific instance of graph convolution, utilises Chebyshev polynomials for filter approximation [115]. The approximation of $g_\theta(\Lambda)$ under the K^{th} order Chebyshev polynomial framework is given by:

$$g_\theta(\Lambda) = \sum_{k=0}^{K-1} \theta_k T_k(\hat{\Lambda}) \tag{5.8}$$

where

$$\hat{\Lambda} = \frac{2\Lambda}{\Lambda_{max}} - I_N \tag{5.9}$$

In the above Equation 5.8, $T_k(\hat{\Lambda}) \in \mathbb{R}^K$ is the k^{th} order Chebyshev polynomial evaluated using Equation 5.9. Then, the signal x is convolved with the defined filter g_θ using the following equation:

$$x *_G g_\theta = U \sum_{k=0}^{K-1} \theta_k T_k(\hat{\Lambda})U^T x = \sum_{k=0}^{K-1} \theta_k T_k(\widetilde{L})x \tag{5.10}$$

Normalising Λ can be achieved by using Equation 5.9, where Λ_{max} denotes the largest entry in the diagonal of Λ, and I_N represents the diagonal matrix of the scaled eigenvalues. In the above Equation 5.10, $T_k(\widetilde{L})$ is the Chebyshev polynomial of order k evaluated at the scaled Laplacian, $\widetilde{L}$ which is defined in the Equation 5.11 as:

$$\widetilde{L} = \frac{2L}{\lambda_{max}} - I_N \tag{5.11}$$

Let $\bar{x}_k = T_k(\widetilde{L})x \in \mathbb{R}^N$, where a recursive relation is used to compute $\bar{x}_k$ using Equation 5.12 with $\bar{x}_0 = x$ and $\bar{x}_1 = \widetilde{L}x$. One key advantage of using the Chebyshev polynomial to approximate convolutional filters is that it inherently avoids the need to compute the graph Fourier basis. The recursive relation is given by:

$$\bar{x}_k = 2\widetilde{L}\bar{x}_{k-1} - \bar{x}_{k-2} \tag{5.12}$$

5.2.6 Model Setting

The EEG signal $X \in \mathbb{R}^{N\times 1000}$ is processed through two distinct pathways: a CNN pathway and a GNN pathway. Here, N represents the number of EEG channels, with 22 channels in the BCIC_iv-2a dataset and 44 channels in the HGD dataset.

In the GNN pathway, X first passes through the Feature Extractor block, where node features for the GNN are extracted, as detailed in Section 5.2.4. The number of extracted features, denoted by F_0, is 29 for the BCIC_iv-2a dataset and 31 for the HGD dataset. Subsequently, the signal progresses through a series of GNN layers, reaching the GN layer, where it takes the shape $N \times F_N$ (with F_N equal to 64 for the BCIC_iv-2a dataset and 32 for the HGD dataset). Within the GNN block, an activation layer employing the Exponential Linear Unit (ELU) is applied, as shown in Equation 5.13, where β, a hyperparameter controlling saturation for negative inputs, is set to 0.1 in this project. Each GNN layer is followed by a batch normalisation (BN) layer, which rescale and recenter the signals to match their original distribution within a mini-batch (set to 64 in our project). This normalisation helps address internal covariate shift and mitigates potential gradient vanishing or exploding issues. Finally, after the GNN layers, the signal is flattened to 1408 dimensions.

The ELU activation function is defined as:

$$\mathrm{ELU}(x) = \begin{cases} x & \text{if } x > 0 \\ \beta(e^x - 1) & \text{if } x \leq 0 \end{cases} \tag{5.13}$$

In parallel, the EEG signal $X \in \mathbb{R}^{N\times 1000}$ is processed through the CNN pathway. This pathway begins with the Temp_Conv block, designed to extract temporal features from the EEG signal. Each EEG channel has M_0 filters, resulting in a total of $M_0 \times N$ filters in this block, producing an output with shape $M_0 \times N \times 991$. The signal then proceeds through the spatial CNN block, denoted C_1, containing M_1 2D filters that extract spatial features from the EEG signal, resulting in an output shape of $M_1 \times 991$. Each CNN operation is followed by a max pooling layer, along with an ELU activation layer and a BN layer. After passing through the final CNN block, the signal is flattened to 1400 dimensions.

The outputs from the GNN and CNN pathways are then concatenated, resulting in a combined dimension of 2808, which is fed into the Fully Connected (FC) blocks. Three FC layers are

applied, with dimensions H_1, H_2, and finally O, where O represents the number of classes (4 for both the BCIC_iv-2a and HGD datasets). Following the FC layers, a softmax layer is applied, as shown in Equation 5.14, to obtain the predicted probability for each class. The loss function used is cross-entropy loss, detailed in Equation 5.15. The softmax function is defined as:

$$\text{Softmax}(\hat{y}_i) = \frac{e^{\hat{y}_i}}{\sum_{j=1}^{O} e^{\hat{y}_j}} \tag{5.14}$$

The cross-entropy loss is given by:

$$\text{Loss} = -\frac{1}{|B|} \sum_{b=1}^{B} \sum_{i=1}^{O} y_{b,i}. \log(\text{Softmax}(\hat{y}_{b,i})) \tag{5.15}$$

A general overview of the model is provided in Figure 5.1 and Table 5.1. The model settings and hyperparameters are further detailed in Tables 5.2 and Table 5.3, respectively. Both accuracy and F1 score evaluation metrics were employed to assess the performance of models. The accuracy is defined as:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{5.16}$$

The recall is given by:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{5.17}$$

The precision is defined as:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{5.18}$$

The F1 score is given by:

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5.19}$$

Table 5.1: Generalised Architecture of EEG_SGNet Model

Layer	Type	Input Size	Weights	Output Size	Activation
Input	Input	$N \times 1000$	-	-	-
		GNN Pathway			
FE	Feature Extractor	$N \times 1000$	-	$N \times F_0$	-
G1	Graph Convolution	$N \times F_0$	$F_0 \times F_1$	$N \times F_1$	ELU
G2	Graph Convolution	$N \times F_1$	$F_1 \times F_2$	$N \times F_2$	ELU
GN	Graph Convolution	$N \times F_2$	$F_2 \times F_N$	$N \times F_N$	ELU
Flat	Flatten	$N \times F_N$	-	1408	-
		CNN Pathway			
Temp Conv	Conv1D, $(M_0 * N)$ filters, 10-kernel, stride 1	$N \times 1000$	$(M_0 * N) \times 10$	$M_0 \times N \times 991$	ELU
C1	Conv2D, M_1 filters, $(N \times 1)$ kernel, stride 1×1	$M_0 \times N \times 991$	$M_1 \times 1 \times 991$	$M_1 \times 991$	ELU
CP1	MaxPool1D, 3-kernel, stride 3	$M_1 \times 991$	-	$M_1 \times 330$	-
C2	Conv1D, M_2 filters, 10-kernel, stride 3	$M_1 \times 330$	$M_2 \times 3$	$M_2 \times 107$	ELU
CP2	MaxPool1D, 10-kernel, stride 4	$M_2 \times 321$	-	$M_2 \times 25$	-
Flat	Flatten	$M_2 \times 25$	-	1400	-
		Feature Combination			
Concat	Concatenation	1400 & 1408	-	2808	-
		Fully Connected Block			
FC1	Fully Connected	2808	$2808 \times H_1$	H_1	ELU
FC2	Fully Connected	H_1	$H_1 \times H_2$	H_2	ELU
FC3	Fully Connected	H_2	$H_2 \times O$	O	-
S	Softmax	O	-	O	-

Table 5.2: Model Settings of EEG_SGNet

Model	Model Framework	Number of GCN Filters	Number of CNN Filters	Number of FC Hidden Nodes
A	$[(G-BN)\times 2-P]//[TC-(C-BN-CP)\times 2]-(FC-BN)\times 2-S$	512, 64	25, 56	2808, 256, 4
B	$[(G-BN)\times 3-P]//[TC-(C-BN-CP)\times 2]-(FC-BN)\times 2-S$	512, 256, 64	25, 56	2808, 256, 4
C	$[(G-BN)\times 4-P]//[TC-(C-BN-CP)\times 2]-(FC-BN)\times 2-S$	512, 256, 128, 64	25, 56	2808, 256, 4
D	$[(G-BN)\times 2-P]//[TC-(C-BN-CP)\times 3]-(FC-BN)\times 2-S$	512, 64	25, 50, 50	2808, 256, 4
E	$[(G-BN)\times 3-P]//[TC-(C-BN-CP)\times 3]-(FC-BN)\times 2-S$	512, 256, 64	25, 50, 50	2808, 256, 4
F	$[(G-BN)\times 4-P]//[TC-(C-BN-CP)\times 3]-(FC-BN)\times 2-S$	512, 256, 128, 64	25, 50, 50	2808, 256, 4
G	$[(G-BN)\times 2-P]//[TC-(C-BN-CP)\times 4]-(FC-BN)\times 2-S$	512, 64	25, 50, 100, 200	2808, 256, 4
H	$[(G-BN)\times 3-P]//[TC-(C-BN-CP)\times 4]-(FC-BN)\times 2-S$	512, 256, 64	25, 50, 100, 200	2808, 256, 4
I	$[(G-BN)\times 4-P]//[TC-(C-BN-CP)\times 4]-(FC-BN)\times 2-S$	512, 256, 128, 64	25, 50, 100, 200	2808, 256, 4

Table 5.3: Hyperparameter Settings for Training EEG_SGNet

Hyperparameter	Value
Training Epochs (N_{ep})	500
Batch Size (B)	64
Optimizer	Adam
Initial Learning Rate (η)	0.001
L2 Regularisation Rate (λ)	0.0005

Table 5.4: Maximum Accuracy Achieved by Various Methods for Each Subject on the BCIC_iv-2a Dataset

Subj	PCC		PCC (alpha)		EEG_GLT		EEG_GLT (alpha)	
	Acc	F1	Acc	F1	Acc	F1	Acc	F1
S_1	80.34%	80.29%	80.35%	80.32%	86.13%	86.10%	**89.02%**	**88.92%**
S_2	63.05%	62.88%	**64.16%**	**64.35%**	58.96%	58.65%	59.54%	59.94%
S_3	84.39%	84.23%	85.55%	85.45%	90.76%	90.72%	**93.07%**	**93.06%**
S_4	80.35%	80.49%	**82.08%**	**82.21%**	79.77%	79.76%	80.35%	80.32%
S_5	61.27%	61.25%	63.01%	63.20%	64.16%	63.96%	**66.47%**	**66.32%**
S_6	71.10%	71.01%	73.41%	72.98%	74.94%	74.25%	**75.14%**	**75.22%**
S_7	91.36%	91.22%	92.49%	92.52%	94.22%	94.25%	**95.38%**	**95.39%**
S_8	85.17%	85.18%	86.49%	86.38%	88.44%	88.37%	**93.64%**	**93.65%**
S_9	82.02%	82.01%	83.02%	83.01%	84.97%	84.86%	**87.86%**	**87.81%**

5.3 Results and Discussion

5.3.1 PCC vs EEG_GLT Adjacency Construction Method

The EEG_SGNet model integrates parallel CNN and GNN pathways. Within the GNN pathway, two adjacency construction methods were employed: Pearson Correlation Coefficient (PCC) and the proposed EEG_GLT, introduced in Section 3.3.4 and Section 3.3.5 respectively. For the BCIC_iv-2a dataset, the average accuracy and F1 score achieved using EEG_GLT were consistently higher than those obtained with PCC across all model configurations (Models A to I). This demonstrates the superiority of EEG_GLT over PCC, regardless of the model architecture. On average, the EEG_GLT adjacency construction method improved accuracy by 3%.

When using the same model settings, the EEG_GLT method with the inclusion of an α parameter achieved the highest average accuracy across all configurations, except for Models B and C. Even in these exceptions, the highest average accuracy was achieved by EEG_GLT without the α parameter, as shown in Tables 5.6 and 5.7. Specifically, EEG_GLT with α achieved an average accuracy of 80.93% and an F1 score of 80.79%. Without α, EEG_GLT achieved an average accuracy of 79.42% and an F1 score of 79.24%. In comparison, PCC with α achieved an average accuracy and F1 score of 78.74% and 78.45%, respectively, while PCC without α achieved 78.74% and 78.69%.

In subject-specific performance, EEG_GLT, both with and without α values, consistently outperformed PCC in most subjects, except for S_2 and S_4 where PCC demonstrated better performance, as shown in Table 5.4. Notably, EEG_GLT improved accuracy by 8.68% for S_1 and S_3, and by 8.47% for S_8.

For the HGD dataset, the average accuracy across 14 subjects showed no significant differences

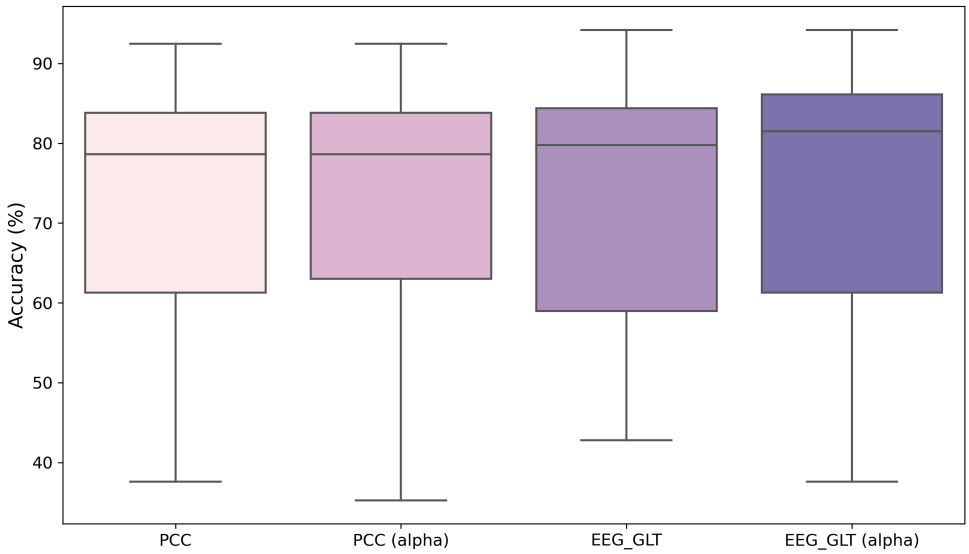


Figure 5.3: Boxplot of Average Accuracy Across Subjects for Different Adjacency Construction Methods Across Model Settings on the BCIC_iv-2a Dataset

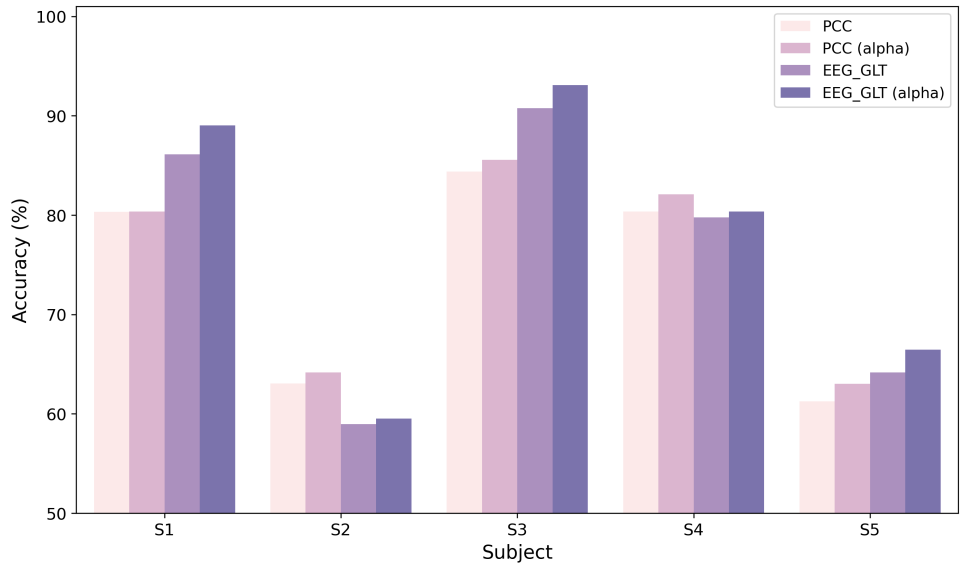


Figure 5.4: Maximum Accuracy Achieved with Each Adjacency Construction Method Across Different Model Settings for Subjects S_1 to S_5 on the BCIC_iv-2a Dataset

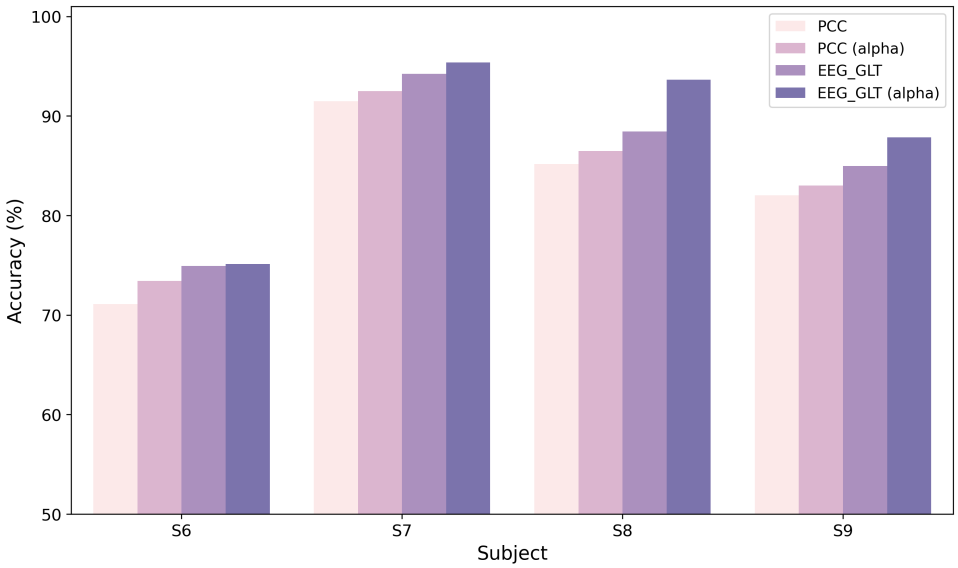


Figure 5.5: Maximum Accuracy Achieved with Each Adjacency Construction Method Across Different Model Settings for Subjects S_6 to S_9 on the BCIC_iv-2a Dataset

between adjacency construction methods for each model, as presented in Tables 5.8 and 5.9. This lack of differentiation may be due to the higher signal clarity and distinguishability in the HGD dataset compared to the BCIC_iv-2a dataset. Nevertheless, the EEG_GLT methods consistently outperformed PCC in terms of average accuracy and F1 scores. The overall model performance, depicted in Figure 5.6, indicates that EEG_GLT had narrower interquartile ranges (IQRs), reflecting more consistent accuracy compared to PCC and PCC with α values. Regarding outliers, all methods exhibited results below 90%; however, PCC showed more extreme outliers. In contrast, EEG_GLT had fewer and less extreme outliers, suggesting greater reliability under varying parameter settings and model architectures.

In terms of subject-specific performance, adjacency methods incorporating α values did not consistently outperform their counterparts, as shown in Table 5.5. However, the proposed EEG_GLT methods achieved higher accuracy than PCC for most subjects, except for S_3 and S_{13}. For some subjects (S_4, S_9, S_{14}), all methods achieved identical results, with perfect accuracy (100%).

5.3.2 Impact of CNNs and GNNs Layers

This project evaluates nine combinations of GNN and CNN pathways, as detailed in Table 5.2. Models A, B, and C use a fixed two-layer CNN while varying the number of GNN layers from two to four. Similarly, Models D, E, and F employ a fixed three-layer CNN with two to four GNN layers, and Models G, H, and I utilise a fixed four-layer CNN with the same variation in GNN layers. For the BCIC_iv-2a dataset, the lowest accuracy was observed in Models A, B, and C, which use two CNN layers, as shown in Tables 5.6 and 5.7.

The highest performance accuracy was achieved with three CNN layers. However, increasing

Table 5.5: Maximum Accuracy Achieved by Various Methods for Each Subject on the HGD Dataset

Subj	PCC		PCC (alpha)		EEG_GLT		EEG_GLT (alpha)	
	Acc	F1	Acc	F1	Acc	F1	Acc	F1
S_1	98.61%	98.60%	97.92%	97.93%	**100.00%**	**100.00%**	99.31%	99.31%
S_2	92.12%	92.09%	92.81%	92.80%	92.12%	92.10%	**93.49%**	**93.49%**
S_3	97.28%	97.28%	**97.44%**	**97.44%**	97.12%	97.11%	97.12%	97.12%
S_4	**100.00%**	**100.00%**	99.84%	99.84%	**100.00%**	**100.00%**	**100.00%**	**100.00%**
S_5	99.62%	99.62%	99.81%	99.81%	99.81%	99.81%	**100.00%**	**100.00%**
S_6	98.08%	98.08%	98.08%	99.08%	98.72%	98.72%	**99.36%**	**99.36%**
S_7	92.95%	92.97%	93.27%	93.30%	**93.91%**	**93.93%**	93.27%	93.27%
S_8	97.14%	97.14%	97.55%	97.55%	**98.37%**	**98.37%**	97.96%	97.96%
S_9	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**
S_{10}	**98.72%**	**98.72%**	98.40%	98.40%	**98.72%**	**98.72%**	**98.72%**	**98.72%**
S_{11}	99.68%	99.68%	99.68%	99.68%	**100.00%**	**100.00%**	**100.00%**	**100.00%**
S_{12}	99.68%	99.68%	99.68%	99.68%	**100.00%**	**100.00%**	**100.00%**	**100.00%**
S_{13}	**98.61%**	**98.61%**	97.57%	97.57%	98.26%	98.26%	97.92%	97.92%
S_{14}	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**

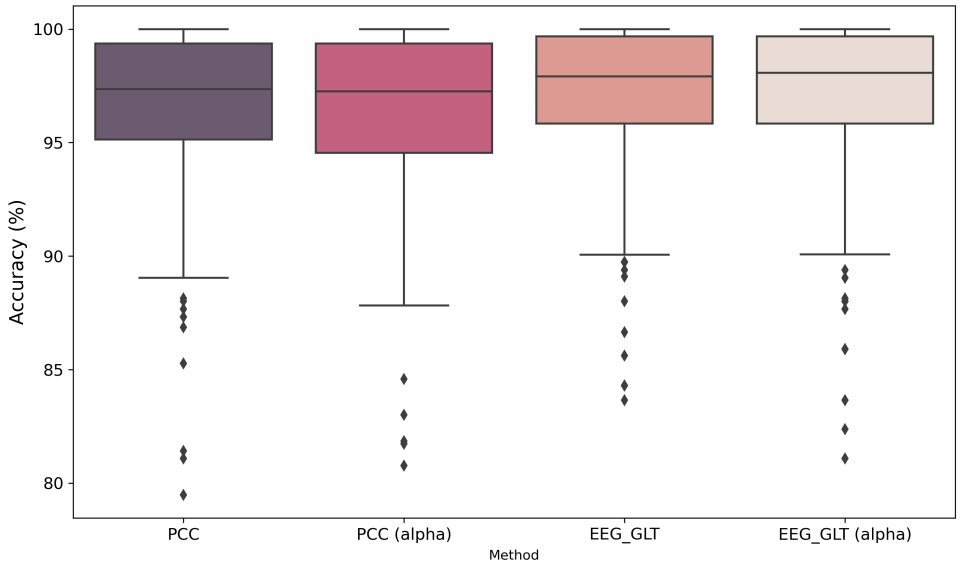


Figure 5.6: Boxplot of Average Accuracy Across Subjects for Different Adjacency Construction Methods Across Model Settings on the HGD Dataset

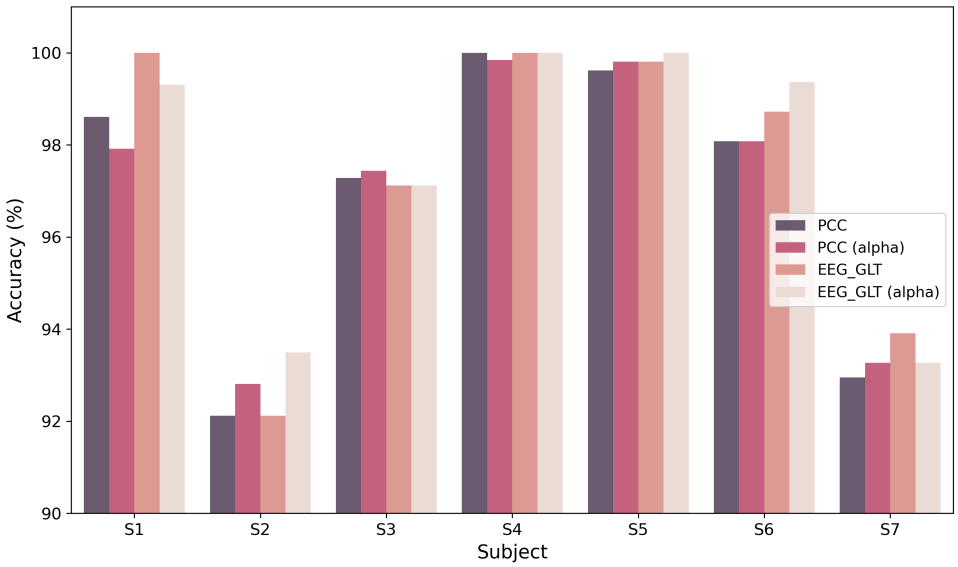


Figure 5.7: Maximum Accuracy Achieved with Each Adjacency Construction Method Across Different Model Settings for Subjects S_1 to S_7 on the HGD Dataset

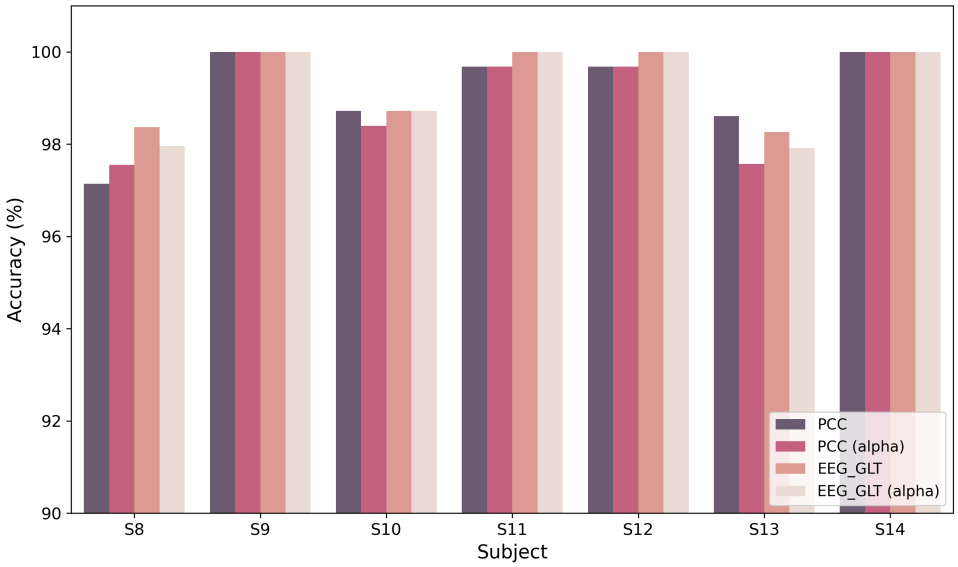


Figure 5.8: Maximum Accuracy Achieved with Each Adjacency Construction Method Across Different Model Settings for Subjects S_8 to S_{14} on the HGD Dataset

Table 5.6: Average Accuracy and F1 Score Across Subjects for Models A to E with Different Adjacency Matrix Construction Methods on the BCIC_iv-2a Dataset

Model	Method	Accuracy (Mean ± Std)	F1 Score (Mean ± Std)
Model A	PCC	68.46% ± 16.34%	68.36% ± 16.34%
	PCC (alpha)	69.04% ± 15.28%	68.98% ± 15.50%
	EEG_GLT	70.33% ± 13.85%	70.04% ± 14.03%
	EEG_GLT (alpha)	**72.77% ± 13.55%**	**72.79% ± 13.41%**
Model B	PCC	68.40% ± 13.90%	68.31% ± 13.76%
	PCC (alpha)	68.85% ± 15.69%	68.78% ± 15.62%
	EEG_GLT	**70.71% ± 13.70%**	**70.70% ± 13.67%**
	EEG_GLT (alpha)	69.62% ± 17.47%	69.51% ± 17.59%
Model C	PCC	68.59% ± 15.76%	68.36% ± 15.81%
	PCC (alpha)	67.12% ± 16.77%	66.99% ± 16.77%
	EEG_GLT	**70.01% ± 15.04%**	**69.89% ± 15.12%**
	EEG_GLT (alpha)	69.81% ± 16.18%	69.59% ± 16.37%
Model D	PCC	77.58% ± 12.41%	77.47% ± 12.48%
	PCC (alpha)	78.45% ± 13.26%	78.49% ± 13.14%
	EEG_GLT	79.42% ± 11.24%	79.24% ± 11.52%
	EEG_GLT (alpha)	**80.92% ± 14.09%**	**80.94% ± 14.01%**
Model E	PCC	77.83% ± 13.26%	77.69% ± 13.14%
	PCC (alpha)	78.74% ± 10.65%	78.69% ± 10.77 %
	EEG_GLT	79.00% ± 12.26%	78.88% ± 12.35%
	EEG_GLT (alpha)	**80.93% ± 12.44%**	**80.79% ± 12.54%**

Table 5.7: Average Accuracy and F1 Score Across Subjects for Models F to I with Different Adjacency Matrix Construction Methods on the BCIC_iv-2a Dataset

Model	Method	Accuracy (Mean ± Std)	F1 Score (Mean ± Std)
Model F	PCC	77.93% ± 12.29%	77.83% ± 12.32%
	PCC (alpha)	78.45% ± 11.65%	78.49% ± 11.59%
	EEG_GLT	79.10% ± 13.76%	70.03% ± 13.88%
	EEG_GLT (alpha)	**80.15% ± 12.81%**	**80.11% ± 12.73%**
Model G	PCC	68.59% ± 16.95%	68.42% ± 17.10%
	PCC (alpha)	68.21% ± 17.78%	68.03% ± 17.54%
	EEG_GLT	69.62% ± 16.98%	69.47% ± 17.00%
	EEG_GLT (alpha)	**71.48% ± 18.94%**	**71.27% ± 19.13%**
Model H	PCC	69.78% ± 16.28%	69.72% ± 16.20%
	PCC (alpha)	68.66% ± 18.21%	68.43% ± 18.09%
	EEG_GLT	70.52% ± 16.67%	70.44% ± 16.53%
	EEG_GLT (alpha)	**71.80% ± 18.28%**	**71.63% ± 18.41%**
Model I	PCC	68.68% ± 18.70%	68.40% ± 18.98%
	PCC (alpha)	68.14% ± 16.66%	67.98% ± 16.54%
	EEG_GLT	70.91% ± 16.30%	70.80% ± 16.33%
	EEG_GLT (alpha)	**71.93% ± 18.21%**	**71.82% ± 18.11%**

Table 5.8: Average Accuracy and F1 Score Across Subjects for Models A to E with Different Adjacency Matrix Construction Methods on the HGD Dataset

Model	Method	Accuracy (Mean ± Std)	F1 Score (Mean ± Std)
Model A	PCC	95.23% ± 5.07%	95.23% ± 5.05%
	PCC (alpha)	94.75% ± 4.63%	94.36% ± 5.29%
	EEG_GLT	95.57% ± 4.57%	95.57% ± 4.56%
	EEG_GLT (alpha)	**95.84% ± 4.62%**	**95.85% ± 4.59%**
Model B	PCC	94.62% ± 5.70%	94.62% ± 5.70%
	PCC (alpha)	94.19% ± 5.90%	94.17% ± 5.90%
	EEG_GLT	**95.25% ± 4.86%**	**95.25% ± 4.86%**
	EEG_GLT (alpha)	95.23% ± 5.18%	95.25% ± 5.12%
Model C	PCC	94.98% ± 4.97%	94.98% ± 4.96%
	PCC (alpha)	95.19% ± 5.12%	95.17% ± 5.16%
	EEG_GLT	95.39% ± 4.73%	95.38% ± 4.75%
	EEG_GLT (alpha)	**95.51% ± 4.89%**	**95.48% ± 4.98%**
Model D	PCC	97.47% ± 3.04%	97.46% ± 3.05%
	PCC (alpha)	97.44% ± 2.74%	97.44% ± 2.74%
	EEG_GLT	**97.98% ± 2.49%**	**97.98% ± 2.49%**
	EEG_GLT (alpha)	97.71% ± 2.76%	97.70% ± 2.77%
Model E	PCC	97.36% ± 2.57%	97.37% ± 2.57%
	PCC (alpha)	97.33% ± 2.97%	97.33% ± 2.96 %
	EEG_GLT	97.90% ± 2.64%	98.01% ± 2.61%
	EEG_GLT (alpha)	**98.18% ± 2.24%**	**98.18% ± 2.25%**

the number of CNN layers to four resulted in a slight drop in accuracy, though it remained higher than the accuracy achieved with two-layer CNNs. Thus, the optimal number of CNN layers for the BCIC_iv-2a dataset is three. A similar trend was observed for the HGD dataset, where the best performance was also achieved with three CNN layers, as shown in Tables 5.8 and 5.9.

In contrast, the number of GNN layers did not have a significant impact on performance accuracy. However, this does not imply that GNNs are unimportant in the architecture. The results demonstrate that the choice of adjacency matrix construction method plays a crucial role in model performance. This is likely because the BCIC_iv-2a dataset contains only 22 nodes, and the HGD dataset contains 44 nodes. With just two hops of graph embedding, the features of all associated nodes are already embedded within the graph, reducing the marginal benefit of additional GNN layers.

Table 5.9: Average Accuracy and F1 Score Across Subjects for Models F to I with Different Adjacency Matrix Construction Methods on the HGD Dataset

Model	Method	Accuracy (Mean ± Std)	F1 Score (Mean ± Std)
Model F	PCC	97.58% ± 2.70%	97.58% ± 2.70%
	PCC (alpha)	97.29% ± 2.86%	97.29% ± 2.85%
	EEG_GLT	97.86% ± 2.62%	97.86% ± 2.61%
	EEG_GLT (alpha)	**98.07% ± 2.40%**	**98.07% ± 2.39%**
Model G	PCC	96.40% ± 4.27%	96.40% ± 4.27%
	PCC (alpha)	96.48% ± 3.86%	96.47% ± 3.88%
	EEG_GLT	96.95% ± 3.66%	96.96% ± 3.64%
	EEG_GLT (alpha)	**96.78% ± 4.14%**	**96.78% ± 4.15%**
Model H	PCC	96.46% ± 4.05%	96.47% ± 4.04%
	PCC (alpha)	96.19% ± 3.96%	96.19% ± 3.96%
	EEG_GLT	**97.03% ± 3.64%**	**97.03% ± 3.65%**
	EEG_GLT (alpha)	96.68% ± 4.22%	96.67% ± 4.24%
Model I	PCC	96.53% ± 4.13%	96.52% ± 4.14%
	PCC (alpha)	95.83% ± 4.04%	95.84% ± 4.03%
	EEG_GLT	**97.14% ± 3.47%**	**97.12% ± 3.50%**
	EEG_GLT (alpha)	96.79% ± 4.29%	96.79% ± 4.31%

Table 5.10: Performance Comparison of EEG_SGNet Model with State-of-the-Art Models on BCIC_iv-2a Dataset

Method	Avg. Accuracy	Avg. F1 Score
ShallowNet [58]	72.05%	71.93%
FBCSP+CNN [146]	74.46%	74.38%
WaSFNet [147]	68.18%	68.04%
EEGNet [108]	72.40%	72.15%
EEG-Conformer [150]	74.56%	74.38%
EEG-TCNet [148]	77.35%	77.18%
4DDFM+3M3DCNN [149]	80.04%	79.95%
SSDA [110]	61.56%	59.84%
DG-HAM [123]	76.36%	76.56%
GSAN [151]	77.15%	77.08%
Proposed EEG_SGNet	**82.98%**	**82.88%**

5.3.3 Comparison with Current State-of-the-Art Models

For the BCIC_iv-2a dataset, we compare our proposed EEG_SGNet with ten state-of-the-art (SOTA) models, as summarised in Table 5.10. These models include ShallowNet [58], FBCSP+CNN [146], WaSFNet [147], EEGNet [108], EEG-TCNet [148], 4DDFM+3M3DCNN [149], SSDA [110], DG-HAM [123], EEG-Conformer [150], and GSAN [151]. Among these, SSDA, which employs a columnar spatio-temporal auto-encoder with CNN and LSTM layers in both the encoder and decoder, achieves the lowest average accuracy at 61.56%. This performance is lower than that of ShallowNet, a simple two-layer CNN model, which achieves 72.05%, and FBCSP+CNN, which combines Filter Bank Common Spatial Pattern (FBCSP) with CNNs to achieve a slightly higher accuracy of 74.46%.

WaSFNet, incorporating a time-frequency convolution layer and a spatial convolution layer, achieves an average accuracy of 68.18%. EEGNet, a compact 2D CNN architecture, achieves a slightly better accuracy of 72.4%. EEG-Conformer, which combines CNNs and Transformers, outperforms EEGNet with an accuracy of 74.56%. EEG-TCNet, which introduces a residual block on EEGNet, further improves the accuracy to 77.35%. Similarly, GSAN, a subdomain adversarial network utilising EEG graph data, achieves a high accuracy of 77.15%. The highest accuracy among the existing SOTA models is achieved by 4DDFM+3M3DCNN, which combines a 4D Dipole Feature Matrix with 3D CNNs, reaching 80.04%. In comparison, our EEG_SGNet, a parallel architecture of CNNs and GNNs, surpasses all these models with an average accuracy of 82.98%.

For the HGD dataset, we compare EEG_SGNet with eight SOTA models, as detailed in Table 5.11. These include ShallowNet [58], EEGNet [108], EEG-ITNet [152], CRGNet [153], EEG-Conformer [150], LMDA-Net [154], IFNet [155], and GCNs-Net [100]. The HGD dataset,

Table 5.11: Performance Comparison of EEG_SGNet Model with State-of-the-Art Models on HGD Dataset

Method	Avg. Accuracy	Avg. F1 Score
ShallowNet [58]	93.04%	92.87%
EEGNet [108]	87.99%	87.85%
EEG-ITNet [152]	84.87%	84.74%
CRGNet [153]	94.33%	94.28%
EEG-Conformer [150]	93.48%	93.32%
LMDA-Net [154]	87.68%	87.49%
IFNet [155]	93.62%	93.28%
GCNs-Net [100]	96.24%	96.18%
Proposed EEG_SGNet	**98.56%**	**98.56%**

which comprises EEG signals of motor movements, has more distinguishable class signals compared to motor imagery datasets like BCIC_iv-2a. As a result, most SOTA models achieve over 90% accuracy, except for EEG-ITNet (84.87%), EEGNet (87.99%), and LMDA-Net (87.68%). EEG-ITNet, a CNN-based model that includes inception and temporal convolution (TC) blocks, underperforms compared to the simpler two-layer CNN model ShallowNet, which achieves 93.04%.

EEG-Conformer performs slightly better than ShallowNet, while IFNet, a model employing 1D CNN operations for low- and high-frequency EEG bands with cross-frequency integration, achieves an accuracy of 93.62%. CRGNet, which incorporates Riemannian manifold embedding after CNN operations, achieves the second-highest accuracy among SOTA models at 94.33%. The highest accuracy among the existing SOTA models is achieved by GCNs-Net, a spectral GNN approach using PCC for adjacency matrix construction, with 96.24%. In comparison, our proposed EEG_SGNet outperforms all existing SOTA models, achieving the highest accuracy of 98.56%.

5.4 Limitations and Future Works

EEG_SGNet has demonstrated strong performance in EEG-based MI and movement classification by integrating CNN and GNN with EEG_GLT adjacency matrix construction method. The single-subject paradigm used in this study allows for a highly personalised model that captures individual-specific neural patterns effectively. This approach optimises classification accuracy by leveraging subject-specific signal characteristics. However, EEG signals naturally vary across individuals due to anatomical, physiological, and cognitive differences, which are important factors to consider when extending the model to broader user populations. While EEG_SGNet is designed to extract meaningful spatial-temporal features, further evaluations in cross-subject settings would provide valuable insights into its adaptability.

The computational efficiency of EEG_SGNet has been significantly improved through the incorporation of the EEG_GLT method for adjacency matrix construction, which reduces complexity in the GNN pathway. However, for real-time rehabilitation applications, further optimisations like lightweight model compression techniques. Future research could also focus on improving model adaptability for multiple users through adaptive calibration or transfer learning, ensuring that EEG_SGNet remains robust and scalable for practical neurorehabilitation settings.

5.5 Conclusion

The EEG_SGNet architecture represents a significant advancement in EEG signal classification, particularly for datasets with inherent temporal information. By utilising a parallel architecture of CNN and GNN pathways, our model effectively extracts valuable features from EEG signals. The CNN pathway captures both spatial and temporal features through the application of optimised filters, while the GNN pathway extracts a diverse set of features, including time-domain and frequency-domain characteristics, wavelet transform coefficients, and band-specific power ratios within each window.

To further enhance the model's performance, the EEG_GLT method was employed for constructing the adjacency matrix in the GNN pathway. This approach improved both performance accuracy and computational efficiency compared to traditional PCC-based methods. Consequently, EEG_SGNet outperformed state-of-the-art models, achieving an average accuracy improvement of 2.95% on the BCIC_iv-2a dataset (EEG motor imagery classification) and 2.32% on the HGD dataset (EEG movement classification).

While this study focuses on window-based classification, it underscores EEG_SGNet's ability to leverage temporal information for robust and reliable performance. Although a four-second window was used in this study, future research will aim to optimise EEG_SGNet for shorter window sizes, such as 500 ms, making it better suited for real-time applications and expanding its applicability to a broader range of EEG analysis tasks.

5.6 Significance and Novelty

This chapter introduces EEG_SGNet, a parallel deep learning architecture that combines Convolutional Neural Networks (CNNs) and Graph Neural Networks (GNNs) for EEG signal decoding. The significance of EEG_SGNet lies in its ability to address both temporal and spatial complexities of EEG data. By leveraging the CNN pathway for temporal feature extraction and the GNN pathway for spatial dependencies, the model achieves state-of-the-art performance on motor imagery and movement-related datasets, making it highly relevant for applications in rehabilitation robotics and brain-computer interfaces (BCIs).

The novelty of this work lies in its unified parallel architecture, which integrates spatial and temporal analyses in a single framework. Unlike existing approaches that prioritise one type of

feature extraction over the other, EEG_SGNet balances both, resulting in superior classification accuracy and adaptability to diverse EEG datasets. This innovation sets a precedent for more holistic and effective EEG analysis techniques.

Chapter 6

A Real-time Framework for EEG Signal Decoding with Graph Neural Networks and Reinforcement Learning

6.1 Introduction

Brain-Computer Interfaces (BCIs) establish a connection between the brain and external control devices. Originally developed to assist individuals with motor impairments [7], BCIs translate brain signals acquired through measurements such as electrocorticography (ECoG) and electroencephalogram (EEG) into actionable commands for electronic control devices including wheelchairs and exoskeleton robots. Although ECoG offers superior signal quality over EEG, its application in BCIs is limited due to invasive route of acquisition, requiring the placement of electrodes directly on the cerebral cortex [83]. Meanwhile, EEG is a much more accessible and hence popular signal acquisition method as it involves non-invasive placement of electrodes on the scalp. EEG is widely used to record various types of brain signals, from spontaneous and stimulus-evoked signals to event-related potentials [84]. Its clinically relevant applications extend to dementia classification [156], depression state assessment [157], seizure detection [158], and the classification of cognitive and motor tasks [159], including motor imagery (MI) tasks [100], [157], [160].

MI involves the mental simulation of motor actions, such as movements of the hands, feet, or tongue, without performing the physical movements [85], [86]. This technique is crucial in neuroscience and rehabilitation, with real-world relevance especially for individuals with motor impairments, such as stroke survivors. Through integration with an external control device, MI enables the physically impaired to perform daily activities that are not otherwise possible, leading to potentially life-changing benefits by improving quality of life and reducing the level of chronic care. By integrating MI and BCIs, EEG based MI signals can be decoded and used to control external devices, enabling real-time feedback and facilitating patient-intended

movements through accurate signal interpretation [88].

Deep learning, a subset of machine learning, utilises multiple layers of neural networks to process a variety of data forms. Convolutional Neural Networks (CNNs), which mimic natural image recognition in the human visual system, are part of the deep learning family and excel in computer vision tasks [89]–[91]. However, their application is restricted to Euclidean data, such as 1-dimensional sequences and 2-dimensional grids [91]. CNNs struggle with non-Euclidean data, failing to accurately capture the intrinsic structure and connectivity of the data.

Graph Convolutional Networks (GCNs) have been developed to perform convolutional operations on graphs, which can handle non-Euclidean data due to incorporating topological relationships during convolution. GCNs can represent complex structures and variations in these structures, which may be heterogeneous or homogeneous, weighted or unweighted, signed or unsigned [92]. They support various types of graph analyses, including node-level, edge-level, and graph-level tasks [92], [113]. GCNs are particularly effective at classifying EEG signals as a graph-level task [100], [144]. For this application, EEG signal readings from each channel are treated as node attributes, and the relationships between EEG electrodes are represented by an adjacency matrix, hence surpassing the capabilities of traditional CNNs.

There are two primary categories of GCNs: spatial [117]–[120] and spectral methods [114]–[116]. Some challenges are encountered with the spatial method [96], [121] especially in matching local neighbourhoods. Both time domain and frequency domain features can be extracted from EEG signals to perform GCN operations [98], [122], [156], [161]. Frequency domain features include Power Spectral Density (PSD) and Power Ratio (PR) for various bands, such as δ (0.5-4Hz), θ (4-8Hz), α (8-13Hz), β (13-30Hz), and γ (30-110Hz) within specified time windows. Time domain features, such as Root Mean Square (RMS), skewness, minmax, variance, number of zero crosses, Hurst Exponent, Petrosian fractal, and Higuchi, are also extracted for GCN operations during specific time windows. These features are integral to window-based GCN methods.

In the GCNs-Net [100], individual time point signals at each channel are treated as distinct features. This method is designed for real-time EEG MI signal classification, focusing on $\frac{1}{160}s$ time point signals. The constructing of an effective adjacency matrix is crucial for GCN operations, and different methods have been explored in various studies, including: Geodesic method, which relies on geodesic distances between EEG channels [93]–[95], [160]; using Pearson coefficient correlation (PCC) to evaluate interchannel correlations [96]–[100] ; and experimenting with a trainable matrix approach [96], [102].

In the EEG_GLT-Net [144], a sophisticated algorithm known as the EEG Graph Lottery Ticket (EEG_GLT) is used to optimise the adjacency matrix by exploring various density levels, inspired from the unified GNN sparsification technique (UGS) [103]. This method represents the current state-of-the-art in adjacency matrix construction, significantly enhancing accuracy, F1 score, and computational efficiency on the EEG MI PhysioNet dataset [127] compared to the

PCC and Geodesic methods. However, despite the overall superiority of this method, it remains challenging to classify the EEG MI time points remains challenging for some subjects due to signal ambiguity among different MI tasks at specific time points. Consequently, supervised learning on these subjects involves training that forces classification of all the time points.

Reinforcement Learning (RL), another subset of machine learning, enables an RL agent to learn sequential decision-making in dynamic environments to maximise cumulative rewards [76]. RL has been primarily applied in robotics and autonomous systems, which require complex sequential decision-making. Deep RL principles have been applied to optimise feature selection for the Classification with Costly Features (CwCF) problem [162], across various public datasets including miniboone [163], forest [164], cifar [165], wine [166], and mnist [90]. Others [167] have trained an RL agent to minimise feature extraction costs in classifying electromyography (EMG) signals [168], although this reduction in features compromised accuracy.

In this paper, we introduce EEG_RL-Net as a new algorithm, with more advanced capability than existing methods for classifying EEG MI time point signals by combining GNNs and RL. Initially, optimal graph features of EEG MI time point signals are extracted using the best weights and adjacency matrix from an EEG_GCN block, refined to 13.39% density using the EEG_GLT algorithm. Subsequently, the RL agent makes sequential decisions within an episode of pre-defined horizon length to accurately classify the EEG MI signals. The main contributions of this study are:

- **EEG_RL-Net:** A new approach for classifying EEG MI time point signals, using a trained RL agent that determines whether to classify or skip each time point based on GNN features. This method greatly enhances performance accuracy by achieving classification as swiftly as possible within predefined episode lengths.
- **Optimal Reward and Max Episode Length Setting:** We evaluated the accuracy and classification speed under various reward settings and maximum episode lengths for each subject, identifying the optimal combinations for simultaneously achieving high accuracy and efficiency.
- **Performance Validation:** We evaluated the performance of each subject under optimal settings against the state-of-the-art EEG_GLT-Net with m_{g_GLT} matrix and PCC adjacency matrix. Our results showed significant enhancement of accuracy and efficiency on the PhysioNet dataset.

6.2 Methodology

6.2.1 Overview

The EEG_RL-Net is comprised of two distinct parts. The overview of the EEG_RL-Net can be seen in the Figure. The first part is training the EEG_GLT-Net to obtain the trained EEG_GCN

block across different adjacency matrix density levels (100% - 13.39%) using Algorithm 1. The main purpose of the EEG_GCN block was to extract EEG MI time points graph signal features. This phase of training spanned from $t = 1s$ to $t = 3s$. Then, the optimal adjacency matrix and GNN weights at the 13.39% of the adjacency matrix density was selected for the purpose of the extracting graph features to save computation efficiency.

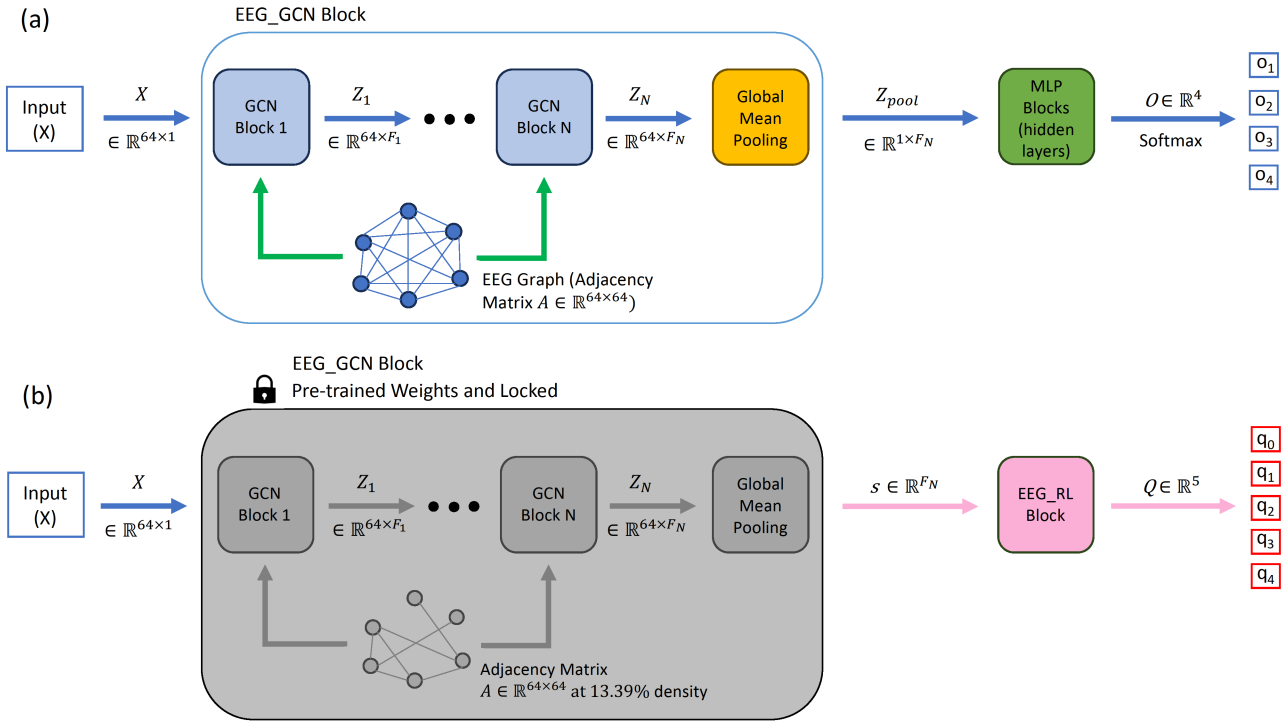


Figure 6.1: (a) EEG_GLT-Net model [144]: Overall architecture (classifying EEG MI of one time point $\frac{1}{160}s$ of signals from 64 EEG electrodes), (b) EEG_RL-Net model (our proposed): Incorporation of an optimal pre-trained EEG_GCN Block at a 13.39% m_g density from the EEG_GLT-Net [144], coupled with an EEG_RL Block

The Multilayer Perceptron (MLP) part within the EEG_GLT-Net is replaced with RL block resulting in the EEG_RL-Net as shown in the Figure 6.1. In this part, all time points from $t = 0s$ to $t = 4s$ were utilised, with these points organised into groups spanning a horizon of 20 states, where each point represented a single state. The RL agent performed action at every state within each episode's horizon based on the graph features generated by the EEG_GCN segment. The agent has five option of actions involving classifying the state such as classifying actions (Task 1 through to Task 4) or skipping action (Task 0) if the agent determined that the current state is not suitable to classify based on the graph features generated by the EEG_GCN segment.

6.2.2 Dataset Description and Pre-processing

Following the approach of papers [100] and [144], this study employed the PhysioNet EEG MI dataset [127], which comprises EEG recordings from 109 subjects acquired using the interna-

tional 10-10 system with 64 EEG channels. The dataset is structured around four distinct EEG MI tasks, which involve the subject imagining the actions of:

- Task 1: Opening and closing the left fist.
- Task 2: Opening and closing the right fist.
- Task 3: Opening and closing both fists simultaneously.
- Task 4: Opening and closing both feet.

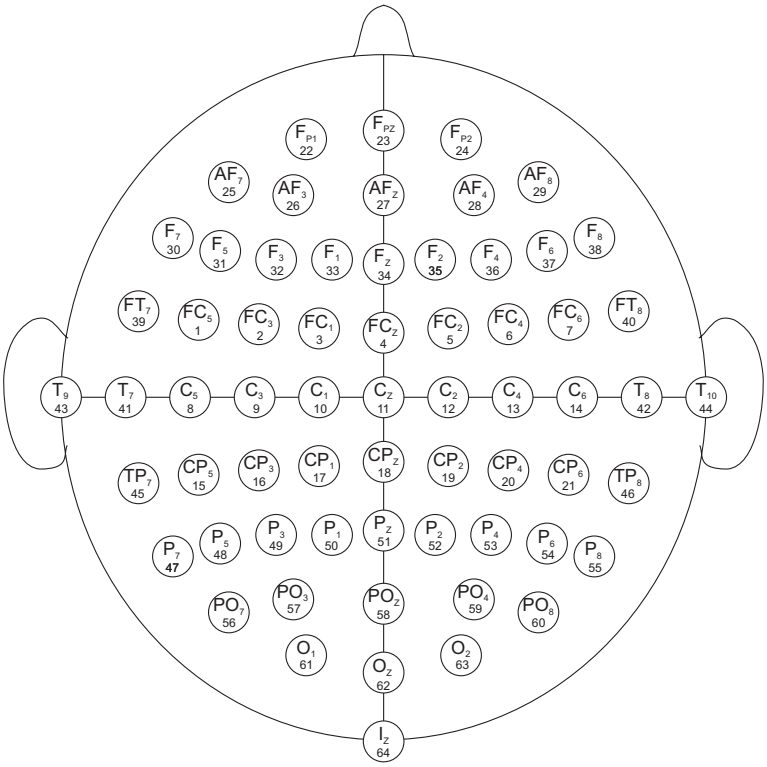


Figure 6.2: Electrode positions of PhysioNet dataset[127]

Each participant completed 84 trials, divided into 3 runs with 7 trials per run for each task type. The duration of each trial's recording was $4s$, sampled at 160 Hz. In our study analyses were specifically conducted on a subset of 20 subjects, labelled S_1 to S_{20}. Initially, the raw signals were processed solely through a notch filter at the 50Hz power line frequency to eliminate electrical interference, deliberately avoiding other common filtering or denoising techniques to preserve data integrity. Signals from all 64 channels were utilised, with each channel treated as a node and the signal at each time point considered as the node's feature. Additionally, the signals at each channel were normalised to achieve a mean (μ) of 0 and a standard deviation (σ) of 1.

The second dataset used in this study is the High-Gamma Dataset (HGD) [58], which contains EEG recordings from 14 healthy subjects with an average age of 27.2 years, including six females. Each subject participated in 1000 trials, each lasting $4s$. The HGD signals were

recorded during motor execution tasks using 128 EEG channels and sampled at 250 Hz. This dataset encompasses four motor movement tasks:

- Task 1: Left hand movement
- Task 2: Right hand movement
- Task 3: Both feet movement
- Task 4: Rest

Only 44 channels covering the motor cortex were selected from the original 128 channels, excluding the Cz channel. Similar to the PhysioNet dataset, the raw EEG signals were processed with a notch filter at the 50 Hz power line frequency. A bandpass filter was then applied to retain a broad frequency range of 4-123 Hz, which is suitable for motor execution tasks. Following the filtering, the signals for each trial were standardised using a channel-wise moving mean and variance to dynamically adjust for local trends within each trial.

6.2.3 Graph Feature Extraction

Graph Representation

In a directed graph, $G = \{V, E\}$ where $V = \{v_1, v_2, ..., v_N\}$ represents the set of nodes and $|E|$ signifies the total number of edges connecting these nodes. The structure of the graph can be illustrated using an adjacency matrix $A \in \mathbb{R}^{N \times N}$. Every node within the graph is associated with F_N features, and the matrix encapsulating these node features is expressed as $X \in \mathbb{R}^{N \times F_N}$. A normalised combinatorial Laplacian matrix is $L = I_N - D^{-1/2}AD^{-1/2}$. This involves the use of the degree matrix of A, symbolised as D, which is calculated using $D_{ii} = \sum_{j=1}^{N} A_{ij}$.

Spectral Graph Filtering

The eigenvectors of the graph Laplacian matrix can be expressed as graph Fourier modes, with $\{u_l\}_{l=0}^{N-1} \in \mathbb{R}^N$. The diagonal matrix of these Fourier frequencies, $\Lambda = diag[\lambda_0, ..., \lambda_{N-1}] \in \mathbb{R}^{N \times N}$. We defined the Fourier basis, $U = [u_0, ..., u_{N-1}] \in \mathbb{R}^{N \times N}$, which allows for the decomposition of the Laplacian matrix, L, into $L = U\Lambda U^T$. The signal x can be transformed by graph Fourier into $\hat{x} \in \mathbb{R}^N$ using $\hat{x} = U^T x$, while the inverse graph Fourier transform is given by $x = U\hat{x}$. The convolution operation on graph G is defined as:

$$x *_G g = U((U^T x) \odot (U^T g)) \tag{6.1}$$

where g represents the convolutional filter and $\odot$ denotes the Hadamard product. Given that $g_\theta(\Lambda) = diag(\theta)$, where $\theta \in \mathbb{R}^N$ represents the vector of Fourier coefficients, the Graph convolution operation can be implemented as follows:

$$x *_G g_\theta = g_\theta(U\Lambda U^T)x = U g_\theta(\Lambda)U^T x \tag{6.2}$$

where g_θ is a non-parametric filter, and polynomial approximation is employed to mitigate the excessive computational complexity. Chebyshev graph convolution, a specific instance of graph convolution, utilises Chebyshev polynomials for filter approximation [115]. The approximation of $g_\theta(\Lambda)$ under the K^{th} order Chebyshev polynomial framework is given by:

$$g_\theta(\Lambda) = \sum_{k=0}^{K-1} \theta_k T_k(\hat{\Lambda}) \tag{6.3}$$

where

$$\hat{\Lambda} = \frac{2\Lambda}{\Lambda_{max}} - I_N \tag{6.4}$$

In the above Equation 6.3, $T_k(\hat{\Lambda}) \in \mathbb{R}^K$ is the k^{th} order Chebyshev polynomial evaluated using Equation 6.4. Then, the signal x is convolved with the defined filter g_θ using the following equation:

$$x *_G g_\theta = U \sum_{k=0}^{K-1} \theta_k T_k(\hat{\Lambda}) U^T x = \sum_{k=0}^{K-1} \theta_k T_k(\widetilde{L}) x \tag{6.5}$$

Normalising Λ can be achieved by using Equation 6.4, where Λ_{max} denotes the largest entry in the diagonal of Λ, and I_N represents the diagonal matrix of the scaled eigenvalues. In the above Equation 6.5, $T_k(\widetilde{L})$ is the Chebyshev polynomial of order k evaluated at the scaled Laplacian, $\widetilde{L}$ which is defined in the Equation 6.6 as:

$$\widetilde{L} = \frac{2L}{\lambda_{max}} - I_N \tag{6.6}$$

Let $\bar{x}_k = T_k(\widetilde{L})x \in \mathbb{R}^N$, where a recursive relation is used to compute $\bar{x}_k$ using Equation 6.7 with $\bar{x}_0 = x$ and $\bar{x}_1 = \widetilde{L}x$. One key advantage of using the Chebyshev polynomial to approximate convolutional filters is that it inherently avoids the need to compute the graph Fourier basis. The recursive relation is given by:

$$\bar{x}_k = 2\widetilde{L}\bar{x}_{k-1} - \bar{x}_{k-2} \tag{6.7}$$

Training EEG_GLT-Net

In the EEG_GLT-Net study [144], the classification of EEG MI signals, X is facilitated through a forward pass using the Spectral GNN function, denoted as $f(., \Theta)$, with a given graph $G = \{A, X\}$. The adjacency matrix, A, integrates $A_{original}$ and m_g as outlined in Equation 6.8. The matrix $A_{original}$, defined as $A_{original\_ij} = \{0, \text{ if } i = j;\ 1, \text{ otherwise}\}$, is fixed and not subject to training, structured in the dimension of $\mathbb{R}^{64\times64}$. Meanwhile, the adjacency matrix mask $m_g \in \mathbb{R}^{64\times64}$ is designated as trainable. The combined adjacency matrix is given by:

$$A = A_{original} \odot m_g \tag{6.8}$$

Table 6.1: Details of EEG_GLT-Net [144] Model

Layer	Type	Input Size	Polynomial Order	Weights	Output
Input	Input	64×1	-	-	-
		GCN Blocks			
GC1	Graph Convolution	64×1	5	$1 \times 16 \times 5$	64×16
BNC1	Batch Normalisation	64×16	-	16	64×16
GC2	Graph Convolution	64×16	5	$16 \times 32 \times 5$	64×32
BNC2	Batch Normalisation	64×32	-	32	64×32
GC3	Graph Convolution	64×32	5	$32 \times 64 \times 5$	64×64
BNC3	Batch Normalisation	64×64	-	64	64×64
GC4	Graph Convolution	64×64	5	$64 \times 128 \times 5$	64×128
BNC4	Batch Normalisation	64×128	-	128	64×128
GC5	Graph Convolution	64×128	5	$128 \times 256 \times 5$	64×256
BNC5	Batch Normalisation	64×256	-	256	64×256
GC6	Graph Convolution	64×256	5	$256 \times 512 \times 5$	64×512
BNC6	Batch Normalisation	64×512	-	512	64×512
		Global Mean Pooling Block			
P	Global Mean Pool	64×512	-	-	512
		Fully Connected Blocks			
FC1	Fully Connected	512	-	512×1024	1024
BNFC1	Batch Normalisation	1024	-	1024	1024
FC2	Fully Connected	1024	-	1024×2048	2048
BNFC2	Batch Normalisation	2048	-	2048	2048
FC3	Fully Connected	2048×4	-	2048×4	4
S	Softmax Classification	4	-	-	4

EEG MI signals from individual subjects, recorded between $t = 1s$ and $t = 3s$, are used for training with Algorithm 1 on PhysioNet dataset. This time window was chosen because subjects

Table 6.2: Hyperparameter Settings for Training EEG_GLT-Net [144]

Hyperparameter	Value
Training Epochs (N_{ep})	1000
Batch Size (B)	1024
Dropout Rate	0.5
Optimiser	Adam
Initial Learning Rate (η)	0.01

typically demonstrated greater readiness after $t = 1s$, as highlighted in the EEG_GLT-Net study [144]. The detailed structure of the EEG_GLT-Net is depicted in Figure 6.1 and Table 6.1, with the specific hyperparameter configurations for the training outlined in Table 6.2. The optimally trained GNN weights (Θ) and the trained adjacency matrix mask (m_g) are recorded across various adjacency matrix density levels, ranging from 100% to 13.39%.

EEG MI Time Points GNN Features

The pre-trained GNN weights and optimal adjacency matrices were evaluated across various m_g densities, ranging from 100% to 13.39%, using the PhysioNet dataset. Among these, the set with a density of 13.39% was selected for extracting graph features from EEG MI signals. This choice was based on findings from the EEG_GLT-Net study [144], which reported optimal performance at m_g densities were found below 31.30% for most subjects. Specifically, 11 out of 20 subjects achieved their highest accuracy at a density of 13.39%. Additionally, this density significantly reduces computational requirements, a crucial consideration given the increased computational demands anticipated in the subsequent EEG_RL-Net stage.

GNN features were then extracted for all EEG MI time points, spanning from $t = 0s$ to $t = 4s$ for all 84 trials of each subject, was conducted. The GNN feature corresponding to each time point had a dimensionality of $\mathbb{R}^{512}$.

6.2.4 Problem Redefinition

The EEG_GLT-Net underwent training for the classification of EEG MI time-point signals. Integration the GNN and an optimally trained adjacency matrix significantly enhanced the classification accuracy compared to traditional PCC adjacency matrix method. Nonetheless, ambiguities in signal clarity between different classes at certain time points could adversely affect the model accuracy. Leveraging the high efficacy of the EEG_GLT-Net model, the pre-trained weights from the GNN and adjacency matrix components were integrated with an RL block, resulting in the formation of the EEG_RL-Net, as depicted in Figure 6.1.

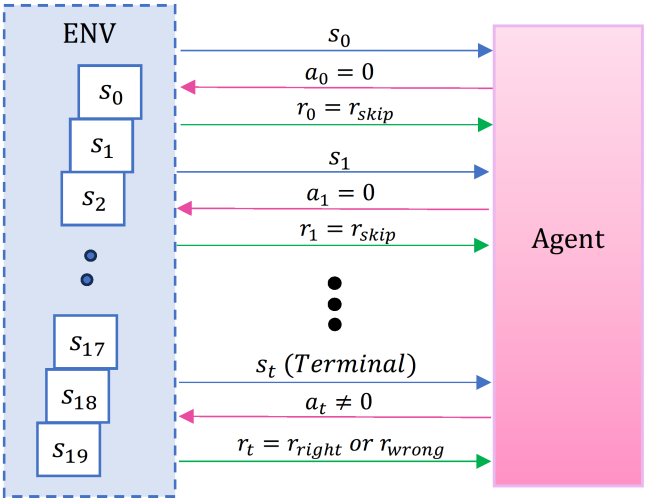


Figure 6.3: Agent interaction with the EEG_RL Environment. The agent transitions to the next state within the current episode after performing the skipping action ($a = 0$), receiving a reward of r_{skip}. The agent transitions to the terminal state after performing a classifying action ($a \neq 0$), receiving a reward of r_{right} for a correct classification or r_{wrong} for an incorrect classification

Algorithm 3 EEG_RL Environment

function STEP(s_t, a_t, y_t, s'_t)
  if $a_t = 0$ **then**
    $r_t = -0.1$
    Return(s'_t, r_t)
  else
    $r_t = \begin{cases} r_{right}, \text{ eg. } +10 & \text{if } a_t = y_t \\ r_{wrong}, \text{ eg. } -10 & \text{if } a_t \neq y_t \end{cases}$
    Return $(s'_t = Terminal, r_t)$
  end if
end function

A reinforcement learning approach is used to train an RL agent for classifying EEG MI time-point signals. Beyond the four initial classes, the RL agent has the capability to defer classification of a current time point if it determines that it is not ready. In each state s_t, the RL agent can perform one of five discrete actions $a_t \in \{0, 1, 2, 3, 4\}$ within the EEG_RL environment, guided by the GNN features extracted from s_t. The actions a_t are described as follows:

- $a_t = 0$: Skip the current state s_t
- $a_t = 1$: Classify the signal as Class 1
- $a_t = 2$: Classify the signal as Class 2
- $a_t = 3$: Classify the signal as Class 3

- $a_t = 4$: Classify the signal as Class 4

Following action a_t, the RL agent is rewarded with r_t and transitions to the next state s'_t, as illustrated in Figure 6.3. Choosing $a_t = 0$ indicates the agent's hesitance to classify due to uncertainty, leading to a decision to skip the current state with a minimal penalty until it is deemed ready to classify or the episode ends. Upon selecting an action $a_t > 0$, s'_t is marked as $Terminal$, which concludes the episode and the agent receives r_t, a positive reward (r_{right}) for correct classification or a negative reward (r_{wrong}) for incorrect classification. The dynamics of the EEG_RL environment are elaborated in Algorithm 3. The ultimate goal is for the RL agent to accurately classify EEG MI signals within the designated horizon (H) of 20 (120 ms) as swiftly as possible.

The Dueling Deep Q Network (DQN) was selected for this study due to its Advantage function, which separates the state-value and advantage functions. This design is particularly effective for EEG time-point signal classification, where consecutive states have minimal changes, allowing the model to focus on the relative importance of actions. Unlike the standard DQN, Dueling DQN provides better stability and performance. Additionally, it is computationally more efficient and less sensitive to sparse rewards compared to policy-gradient methods, making it the most suitable choice for this task.

6.2.5 Data Preprocessing and Data Splitting

In the EEG_RL-Net, the entire duration of the EEG MI signals, spanning four seconds at a sampling rate of 160 Hz, was included. As outlined in Section 6.2.3, GNN features of EEG MI time points were extracted using pre-trained weights with an adjacency matrix density, m_g, of 13.39%. The GNN features of each time point were treated as states, $s \in \mathbb{R}^{512}$. For all 82 trials, from $t = 0s$ to $t = 4s$, groups of consecutive $H = 20$ states were organised into episodes without time point overlap between subsequent episodes, forming an episode set, $E = e_0, e_1, ..., e_n$, as illustrated in Figure 6.4.

The episodes, E, were stratified split into training (E_{train}, 80%), validation (E_{val}, 10%), and test (E_{test}, 10%) sets. The stratified split ensures class balance across all subsets. To assess the stability and robustness of the results, the entire experiment was repeated five times using different random seeds.

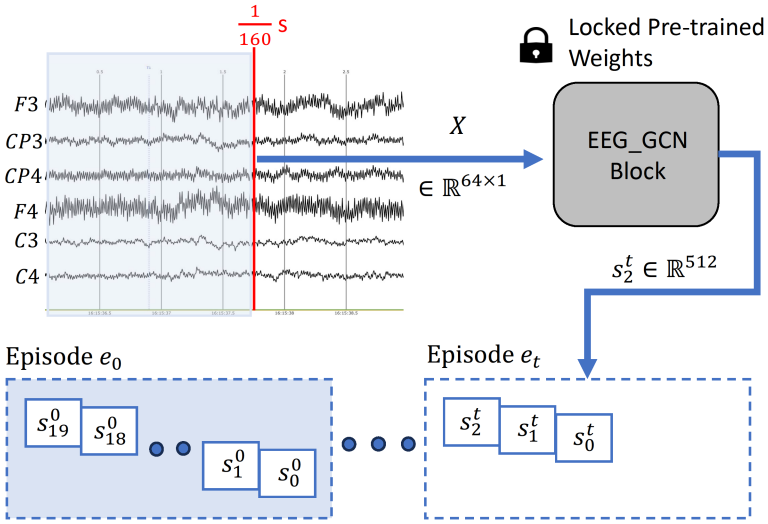


Figure 6.4: Conversion of EEG MI time points into states using the pre-trained EEG_GCN Block, grouped into episodes comprising 20 states each.

6.2.6 Dueling Deep Q-Learning

The DQN method, a value-based RL approach, was employed in this study to learn an optimal policy to enable more accurate classification of EEG MI signals. A state-action value, $Q(s, a)$, represents the expected discounted reward when the agent is in state s, and takes action a according to policy π. With the optimal policy (π^*), the agent aims to achieve the maximum expected discounted reward $Q^*(s, a)$, fulfilling the Bellman equation:

$$Q^*(s, a) = \mathbb{E}_{\pi^*}\left[r + \gamma . \max_{a'} Q^*(s', a') \,\middle|\, s, a\right] \tag{6.9}$$

here r is the immediate reward, and γ is the discount factor. The state-action value, $\hat{Q}(s, a)$, for state s and action a can be approximated using a deep neural network parameterised by θ. The loss function is defined as:

$$\text{Loss}(\theta) = \left(\hat{y}^{DQN} - \hat{Q}(s, a; \theta)\right)^2 \tag{6.10}$$

where $\hat{y}^{DQN}$ is the target value, calculated as follows:

$$\hat{y}^{DQN} = \begin{cases} r_t, & \text{if } s'_t \text{ is Terminal} \\ r_t + \gamma .\ \max_{a'_t} \hat{Q}(s'_t, a'_t; \theta_{\text{target}}), & \text{otherwise} \end{cases} \tag{6.11}$$

The θ_{target} denotes the parameters of the target network, which are kept constant. The approximation $\hat{Q}(s, a; \theta)$ shares the architecture with the target network. Our study utilises Dueling DQN, a variant of DQN that enhances training stability and efficiency by separating the estimation of $\hat{Q}(s, a; \theta)$ into state values $V(s)$ and action advantages $A(s, a)$, as follows:

$$\hat{Q}(s, a; \theta) = \hat{V}(s; \alpha) + \hat{A}(s, a; \beta) \tag{6.12}$$

The network separately estimates the state values and action advantages, which then converge into a single output. The parameters θ represent the overall network parameters, with α and β specifically used for estimating state values and action advantages, respectively. To enhance stability, the equation subtracts the average advantage values from $\hat{Q}(s, a; \theta)$:

$$\hat{Q}(s, a; \theta) = \hat{V}(s; \alpha) + \left[\hat{A}(s, a; \beta) - \frac{1}{|A|} \sum_{a} \hat{A}(s, a; \beta)\right] \tag{6.13}$$

6.2.7 EEG_RL Algorithm

To generate training data for the EEG_RL Block, all possible actions $a_t = \{0, 1, 2, 3, 4\}$ are executed at each state s_t within an episode e_i in E_{train}, interacting with the EEG_RL environment to determine the reward r_t and the subsequent state s'_t. Each transition records a tuple (s, a, r, s'). This study employs the Dueling DQN method for the RL block, as illustrated in Figure 6.5. The Dueling DQN agent undergoes training according to the procedure outlined in Algorithm 4, utilising the Adam optimiser until convergence is achieved. The configuration of the entire EEG_RL-Net model is outlined in Table 6.3. The parameters of the fixed target network, θ_{target}, for the Dueling DQN network, are refreshed after every 50 batch updates of θ.

Algorithm 4 Training EEG_RL-Net's Dueling DQN Agent

1: Initialise randomly Dueling DQN network parameter (θ) and target network parameter (θ_{target}).
2: Set of train episodes $E_{TRAIN} \in \{e_0, e_1, ..., e_N\}$ where each e_i has set of states, $S = \{s_0, s_1, ..., s_{H-1}\}$. Each state, $s_t \in \mathbb{R}^{512}$.
3: At each state s_t, simulate one step with all possible actions from action set, $A \in \{0, 1, 2, 3, 4\}$ to observe next state, s'_t and reward, r_t. Record all the (s_t, a_t, r_t, s'_t) tuples to the Buffer, B.
4: Shuffle the state transitions in the B using random seed, and group into mini-batches in size of 64 transitions.
5: **for** $epoch = 0$ to $EPOCHS$ **do**
6: Compute $\hat{y}^{DQN}$ for each mini-batch:
7: $$\hat{y}^{DQN} = \begin{cases} r_t, & \text{if } s'_t \text{ is Terminal} \\ r_t + \gamma \max_{a'_t} \hat{Q}(s'_t, a'_t; \theta_{\text{target}}) & \text{otherwise} \end{cases}$$
8: $Loss(\theta) = \left(\hat{y}^{DQN} - \hat{Q}(s_t, a_t; \theta)\right)^2$
9: Backpropagate to update θ using $Adam$ optimiser
10: Update $\theta_{target} = \theta$ at every 50 updates of θ
11: **end for**

Performance evaluation of the RL agent on E_{val} and E_{test} is described in Algorithm 5. At every time step, the agent selects an action based on the q-values predicted by the EEG_RL-Net. In this study, a correct classification by the agent yields a reward r_{right}, while an incorrect classification results in r_{wrong}. The agent's objective in each episode e_i is to maximise the cumulative reward r_{sum} within the predefined horizon $H = 20$. This requires the agent to make classifications as quickly as possible, since it incurs a penalty of $r = -0.1$ for each skipped step. However, at time $t = H - 1$, skipping is no longer an option, and the agent must make a classification action.

Algorithm 5 Evaluation of DQN Agent for a Validation or Test Episode

Episode, e_i has horizon of $H = 20$
At e_i, the set of states $S = \{s_0, s_1, ..., s_{H-1}\}$, where each $s_t \in \mathbb{R}^{512}$
At e_i, the set of labels $Y = \{y_0, y_1, ..., y_{H-1}\}$, where each $y_t \in \{1, 2, 3, 4\}$
Action $a' \in \{0, 1, 2, 3, 4\}$, and $a'' \in \{1, 2, 3, 4\}$
Initialise $t = 0$, $r_{sum} = 0$
while $t < H$ **do**

$$a_t = \begin{cases} argmax_{a'} \hat{q}_{DQN}(s_t, a'), & \text{if } t < H - 1 \\ argmax_{a''} \hat{q}_{DQN}(s_t, a''), & \text{otherwise} \end{cases}$$

$s'_t, r_t = STEP(s_t, a_t, y_t, s'_t)$
$r_{sum} \leftarrow r_{sum} + r_t$
if $r_t = Terminate$ **then**
Terminate the Episode, e_i
else
$t \leftarrow t + 1$
end if
end while

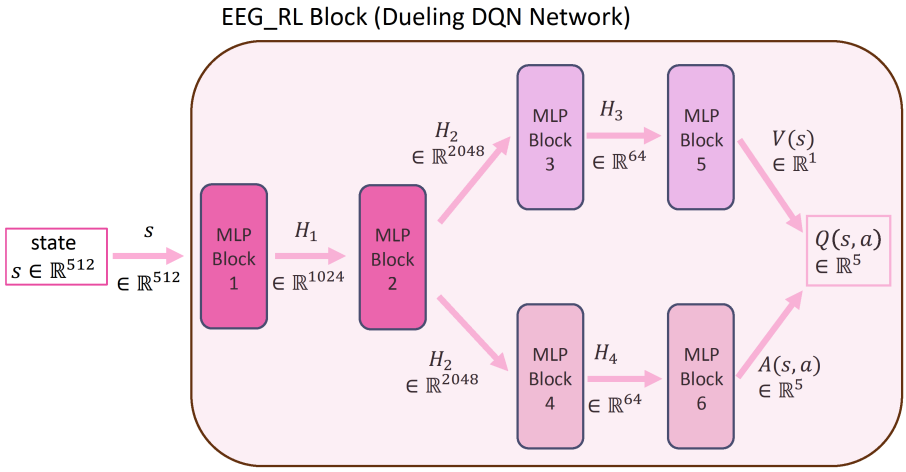


Figure 6.5: EEG_RL Block: Featuring the Dueling DQN, this component predicts the q-values linked to various actions

Table 6.3: Details of EEG_RL-Net Model

Layer	Type	Input Size	Weights	Output
Input	Input	64×1	-	-
EEG_GCN Block				
EEG_GCN	Graph Convolution and Global Pooling	64×1	-	512
RL Block (Dueling DQN Network)				
MLP1	Fully Connected	512	512×1024	1024
MLP2	Fully Connected	1024	1024×2048	2048
MLP3	Fully Connected	2048	2048×64	64
MLP4	Fully Connected	64	64×1	1
MLP5	Fully Connected	2048	2048×64	64
MLP6	Fully Connected	64	64×5	5
Q	Dueling DQN	64×1 & 64×5	-	5

6.2.8 Model Setting and Evaluation Metrics

The structure of EEG_RL-Net is defined by two principal components: the spectral EEG_GCN block, which extracts graph features from EEG MI time point signals using pre-trained weights, and the RL block, embodied by the Dueling DQN network. The specifics of the EEG_RL-Net's design are provided in Table 6.3. The RL block comprises six MLP (Multi-Layer Perceptron) layers, or Fully Connected Layers, each followed by a Rectified Linear Unit (ReLU) layer, as described in Equation 3.12. Information on the training hyperparameters is presented in Table 6.4. The performance of the different methods was evaluated using both accuracy and F1 score metrics.

Table 6.4: Hyperparameter Settings for Training EEG_RL-Net

Hyperparameter	Value
Reward Right (r_{right})	+10
Reward Wrong (r_{wrong})	−10
Reward Skip (r_{skip})	−0.1
Discount Factor (γ)	0.99
Training Epoch ($EPOCHS$)	150
Batch Size	63
Target Network Update Frequency	50
Initial Learning Rate (η)	0.0001
L2 Regularisation Rate (λ)	0.001
Optimiser	Adam

The ReLU activation function is defined as:

$$\text{ReLU}(x) = max(0, x) \tag{6.14}$$

The accuracy is defined as:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{6.15}$$

The recall is given by:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{6.16}$$

The precision is defined as:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{6.17}$$

The F1 score is given by:

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{6.18}$$

6.3 Results and Discussion

6.3.1 EEG_RL-Net vs EEG_GLT-Net

Table 6.5 shows comparative analysis of mean accuracy between the EEG_GLT-Net and the EEG_RL-Net. The EEG_GLT-Net incorporates two adjacency matrix types: the PCC and the m_{g_GLT}. The latter is identified as the most optimal adjacency matrix, after searching through 100% to 13.39% of adjacency matrix density using the EEG_GLT algorithm. According to paper [144], employing the m_{g_GLT} adjacency matrix yields an accuracy improvement ranging between 0.51% and 22.04% over the PCC adjacency matrix, with significant enhancements noted for subjects S_1 and S_{12}, at 22.04% and 21.62% respectively. Despite seeing notable improvements in accuracy and F1 score with the m_{g_GLT} matrix, certain subjects, specifically S_5, S_6, S_7, S_{13}, S_{15}, and S_{19}, exhibited classification accuracies below 70%.

Using baseline parameters ($r_{right} = +10$, $r_{wrong} = -10$, $r_{skip} = -0.1$ and $H = 20$), the EEG_RL-Net framework advances the accuracy beyond the current state-of-the-art EEG_GLT-Net employing the m_{g_GLT} adjacency matrix, with improvements spanning 0.01% to 44.95%. A total of 12 out of 20 subjects, namely S_1, S_3, S_4, S_8, S_9, S_{10}, S_{11}, S_{12}, S_{16}, S_{17}, S_{18}, and S_{20}, achieved perfect classification. The EEG_RL-Net also significantly elevated the accuracies for S_{13} and S_{19} to 89.45% and 79.65%, respectively. Even for subjects S_{13} and S_{19}, who initially demonstrated low accuracies, modest improvement in accuracy at 44.50% and 41.41%, respectively was achieved using the EEG_GLT-Net with the m_{g_GLT} matrix.

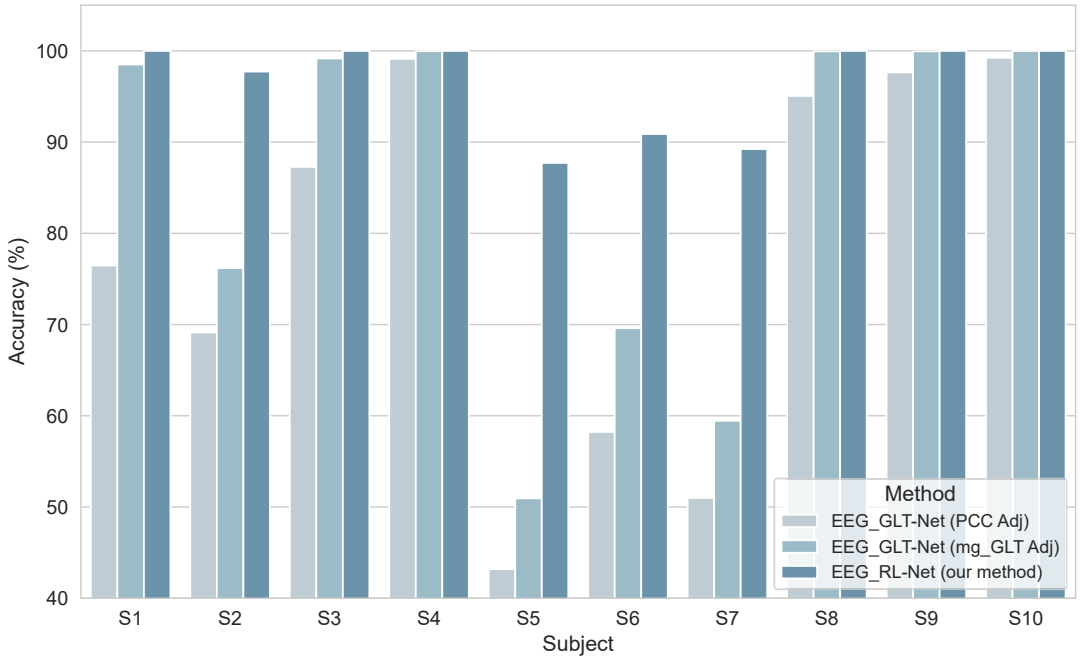


Figure 6.6: Performance Mean Accuracy Comparison between EEG_GLT-Net (PCC Adj), EEG_GLT-Net (m_{g_GLT} Adj) and EEG_RL-Net Methods for Subjects S_1 - S_{10}

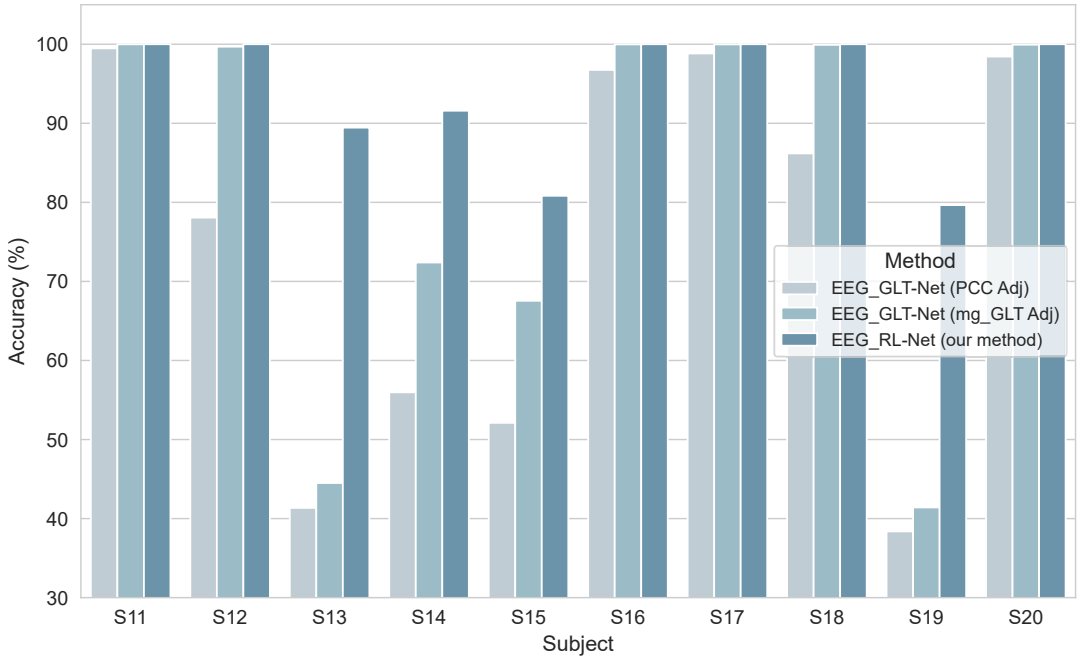


Figure 6.7: Performance Mean Accuracy Comparison between EEG_GLT-Net (PCC Adj), EEG_GLT-Net (m_{g_GLT} Adj) and EEG_RL-Net Methods for Subjects S_{11} - S_{20}

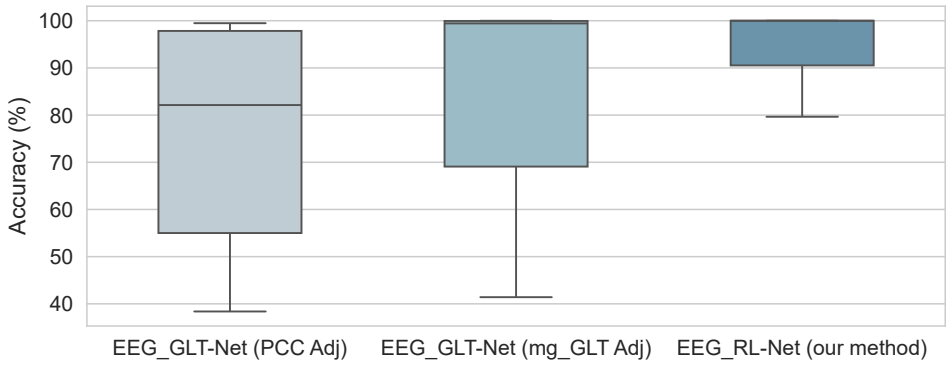


Figure 6.8: Overall Performance Mean Accuracy Comparison between EEG_GLT-Net (PCC Adj), EEG_GLT-Net (m_{g_GLT} Adj) and EEG_RL-Net Methods

The EEG_GLT-Net with the m_{g_GLT} matrix boosted accuracy across the 20 subjects, increasing the average accuracy by 7.85% (from 76.10% to 83.95%). Given the inherent noise in EEG MI time-point signals and the challenge of classifying signals representing $\frac{1}{160}s$, the EEG_GLT-Net showed a decline in performance accuracy due to its attempt to classify all time points. Comparatively, the EEG_RL-Net achieved remarkable increase in average accuracy across the 20 subjects to 95.35%. This substantial improvement is the result of the RL agent's capacity to discern the appropriateness of the current signal for classification. The agent has been optimised to classify signals as swiftly as possible within a 20 time-point window, averaging a classification time of 2.91 time points in the EEG_RL-Net setup.

Table 6.5: Accuracy Assessment: EEG_RL-Net versus EEG_GLT-Net

Subj	Accuracy (Mean±Std)		
	EEG_GLT-Net (PCC Adj)	EEG_GLT-Net (m_{g_GLT} Adj)	EEG_RL-Net* (our method)
S_1	76.47% ± 9.94%	98.51% ± 0.77%	**100.00% ± 0.00%**
S_2	69.13% ± 7.05%	76.18% ± 5.53%	**97.73% ± 0.20%**
S_3	87.28% ± 9.19%	99.17% ± 0.32%	**100.00% ± 0.00%**
S_4	99.13% ± 1.01%	99.97% ± 0.06%	**100.00% ± 0.00%**
S_5	43.19% ± 3.03%	50.95% ± 3.80%	**87.72% ± 0.70%**
S_6	58.23% ± 5.19%	69.60% ± 5.67%	**90.89% ± 1.50%**
S_7	50.98% ± 3.80%	59.45% ± 3.00%	**89.24% ± 2.10%**
S_8	95.06% ± 5.96%	99.95% ± 0.07%	**100.00% ± 0.00%**
S_9	97.64% ± 3.33%	99.95% ± 0.08%	**100.00% ± 0.00%**
S_{10}	99.24% ± 0.19%	99.99% ± 0.01%	**100.00% ± 0.00%**
S_{11}	99.48% ± 0.70%	99.99% ± 0.01%	**100.00% ± 0.00%**
S_{12}	78.07% ± 8.95%	99.69% ± 0.32%	**100.00% ± 0.00%**
S_{13}	41.35% ± 1.23%	44.50% ± 2.23%	**89.45% ± 0.90%**
S_{14}	55.97% ± 6.47%	72.39% ± 6.43%	**91.59% ± 2.10%**
S_{15}	52.11% ± 3.96%	67.55% ± 9.26%	**80.83% ± 1.50%**
S_{16}	96.75% ± 5.00%	99.98% ± 0.03%	**100.00% ± 0.00%**
S_{17}	98.83% ± 2.33%	99.98% ± 0.03%	**100.00% ± 0.00%**
S_{18}	86.19% ± 9.95%	99.92% ± 0.12%	**100.00% ± 0.00%**
S_{19}	38.38% ± 2.27%	41.41% ± 1.44%	**79.65% ± 1.40%**
S_{20}	98.44% ± 0.68%	99.94% ± 0.11%	**100.00% ± 0.00%**
Overall	76.10% ± 22.71%	83.95% ± 21.43%	**95.36% ± 6.83%**

* $r_{right} = +10$, $r_{wrong} = -10$, $r_{skip} = -0.1$, $H = 20$

6.3.2 Study of Changing r_{right} Values

Table 6.6 demonstrates the effect of varying the r_{right} value $(+5, +10, +15)$, on average accuracy while keeping $r_{wrong} = -10$ constant. The results show average accuracies of 95.57%, 95.36%, and 94.94% for $r_{right} = +5, +10$, and $+20$, respectively. Notably, the accuracy tends to improve when r_{right} is less than r_{wrong}, but declines when r_{right} exceeds r_{wrong}, although the level of variance is minimal at just 0.63%.

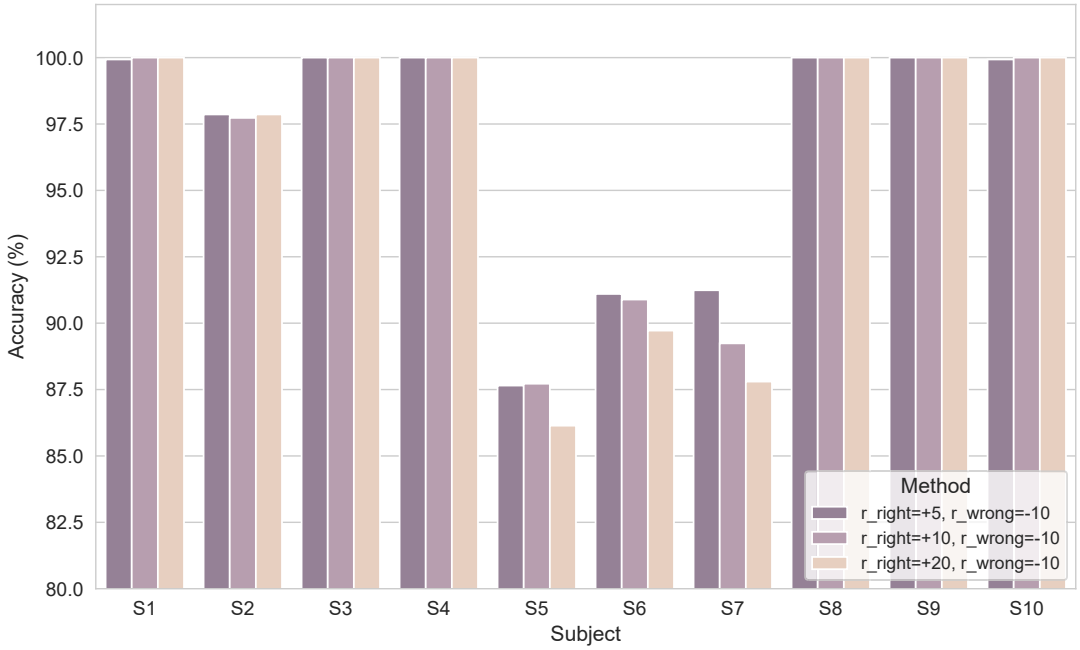


Figure 6.9: Performance Mean Accuracy Comparison Varying r_{right} Values for Subjects S_1 - S_{10}

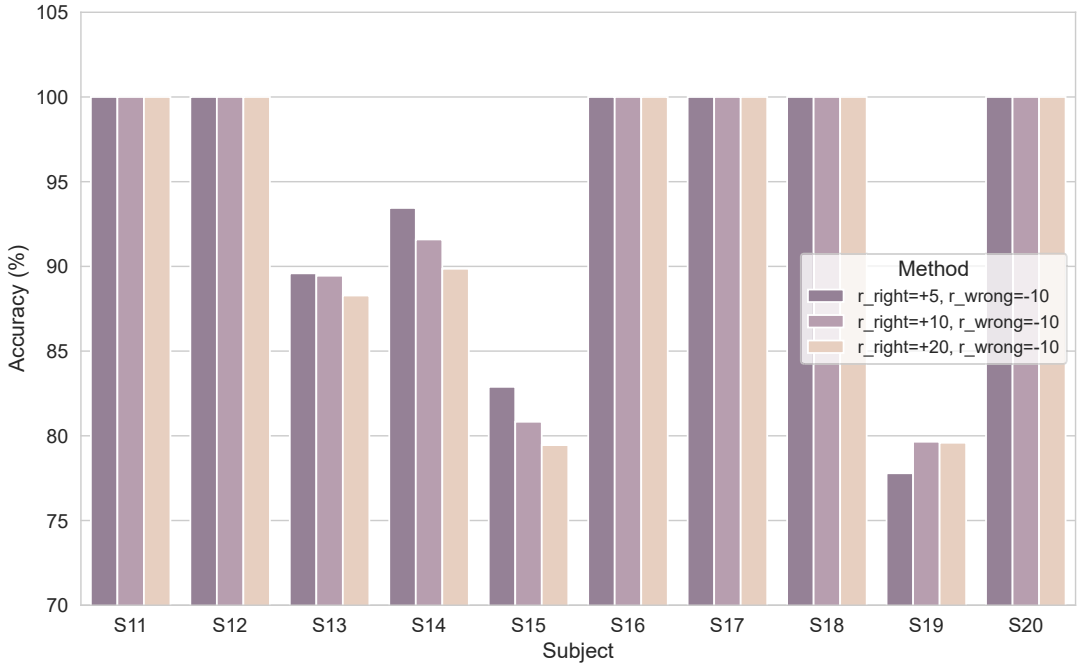


Figure 6.10: Performance Mean Accuracy Comparison Varying r_{right} Values for Subjects S_{11} - S_{20}

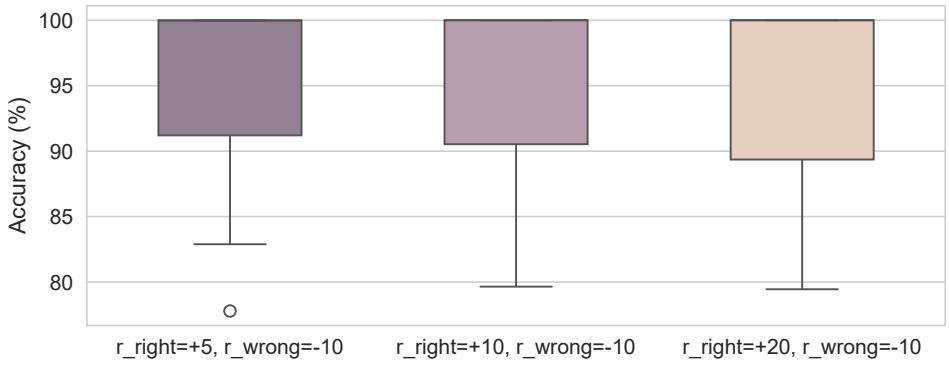


Figure 6.11: Overall Performance Mean Accuracy Comparison Varying r_{right} Values

Table 6.6: Impact of Varying r_{right} Values on Accuracy and Classification Time

Subj	Mean Accuracy (Mean Classification Time)		
	$r_{right} = +5$ $r_{wrong} = -10$	$r_{right} = +10$ $r_{wrong} = -10$	$r_{right} = +20$ $r_{wrong} = -10$
S_1	99.93% (1.70)	100.00% (1.80)	100.00% (1.50)
S_2	97.86% (1.51)	97.73% (1.87)	97.86% (1.65)
S_3	100.00% (2.10)	100.00% (2.20)	100.00% (1.90)
S_4	100.00% (2.80)	100.00% (2.50)	100.00% (1.80)
S_5	87.65% (5.71)	87.72% (4.55)	86.14% (3.45)
S_6	91.10% (3.37)	90.89% (2.80)	89.72% (2.07)
S_7	91.24% (3.64)	89.24% (3.23)	87.80% (3.66)
S_8	100.00% (2.90)	100.00% (2.80)	100.00% (2.10)
S_9	100.00% (3.70)	100.00% (3.00)	100.00% (2.20)
S_{10}	99.93% (2.40)	100.00% (2.20)	100.00% (1.80)
S_{11}	100.00% (2.00)	100.00% (1.50)	100.00% (2.30)
S_{12}	100.00% (2.80)	100.00% (2.40)	100.00% (2.60)
S_{13}	89.59% (4.90)	89.45% (3.97)	88.28% (3.58)
S_{14}	93.45% (3.75)	91.59% (3.14)	89.86% (2.81)
S_{15}	82.89% (5.23)	80.83% (4.80)	79.45% (4.51)
S_{16}	100.00% (2.40)	100.00% (2.70)	100.00% (2.20)
S_{17}	100.00% (3.00)	100.00% (2.00)	100.00% (1.90)
S_{18}	100.00% (3.30)	100.00% (2.40)	100.00% (1.50)
S_{19}	77.79% (6.89)	79.65% (5.64)	79.59% (4.80)
S_{20}	100.00% (2.20)	100.00% (2.70)	100.00% (2.10)
Mean	95.57% (3.32)	95.36% (2.91)	94.94% (2.51)
Std	± 6.72%	± 6.83%	± 7.32%

$r_{skip} = -0.1$ and $H = 20$

On an individual basis, $r_{right} = +5$ yielded higher accuracies for most subjects, except for S_1, S_5, and S_{19}, where $r_{right} = +10$ performed marginally better. No subjects showed improved accuracy when r_{right} was greater than r_{wrong}. Therefore, for optimal performance, the magnitude of r_{right} should not exceed r_{wrong}. It appears that accuracy is enhanced by a higher penalty for incorrect classifications (r_{wrong}) rather than a higher reward for correct ones (r_{right}) enhances accuracy, likely motivating the agent to avoid misclassifications more stringently.

Regarding the time points required to classify EEG MI signals, the configuration with $r_{right} = +10$ and $r_{wrong} = -10$ averages at 2.91 time points. Increasing r_{right} to +20 (while r_{wrong} remains at -10) reduces the classification time to 2.51 time points. Conversely, lowering r_{right} to +5 increases the average classification time to 3.32 time points, indicating a more cautious approach by the agent, likely due to prioritising accuracy over speed by utilising the option to skip uncertain classifications.

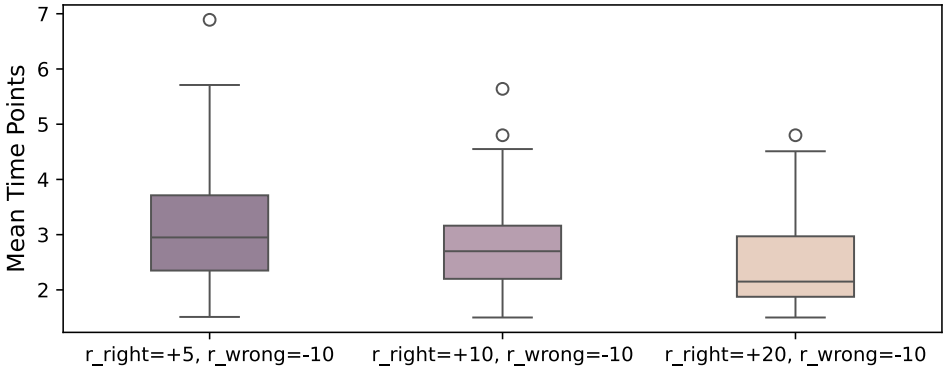


Figure 6.12: Overall Mean Classification Time (Time Points) Comparing Varying r_{right} Values

6.3.3 Study of Changing r_{wrong} Values

In this study, we examined the impact of altering the r_{wrong} values while keeping the r_{right} constant at +10, as shown in Table 6.7. We observed the r_{wrong} values at -10, -20, -30, and -40, correlating with an average performance accuracy of 95.35%, 95.18%, 95.11%, and 94.88%, respectively. This indicates that simply increasing the negative magnitude of r_{wrong} beyond that of r_{right} does not invariably lead to enhanced performance accuracy. Additionally, we found that the time required for signal classification was directly related to the difference in rewards.

Despite the reward configuration of $\{r_{right} = +10, r_{wrong} = -10\}$ achieving the highest average performance accuracy among the four settings, it does not universally outperform across all test subjects. Specifically, this configuration was only superior for subjects S_5 and S_6. Conversely, the configuration of $\{r_{right} = +10, r_{wrong} = -20\}$ exhibited higher performance accuracy in subjects S_2, S_7, S_{13}, and S_{15}. For subject S_{14}, the $\{r_{right} = +10, r_{wrong} = -30\}$ setting was more advantageous.

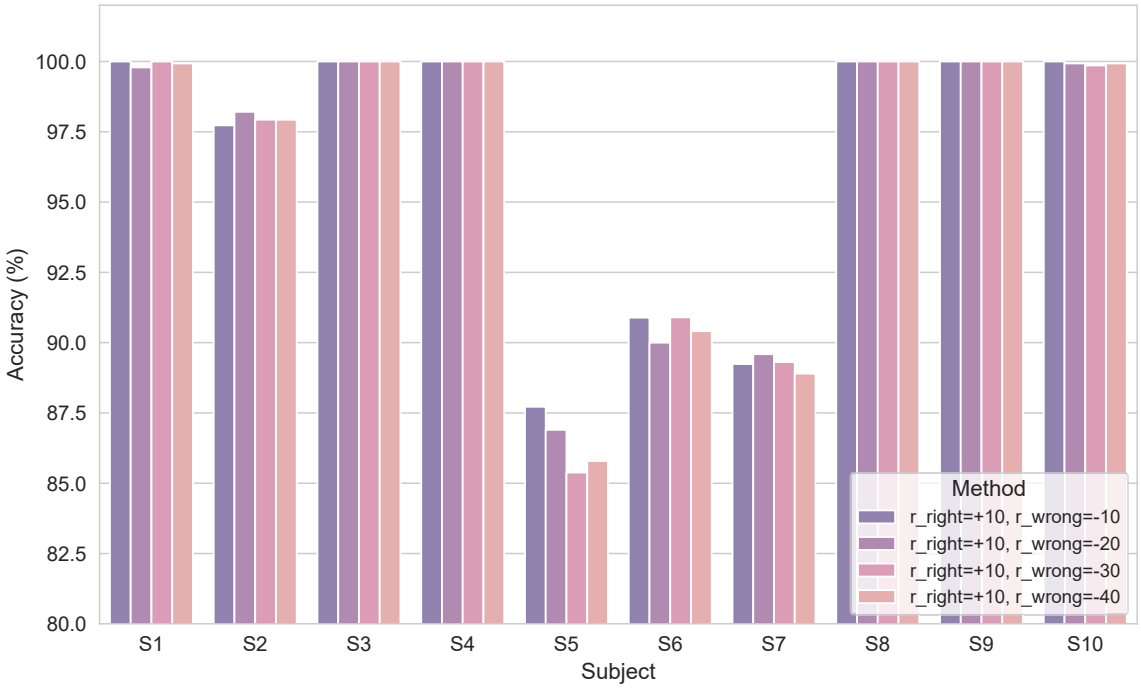


Figure 6.13: Performance Mean Accuracy Comparison Varying r_{wrong} Values for Subjects S_1 - S_{10}

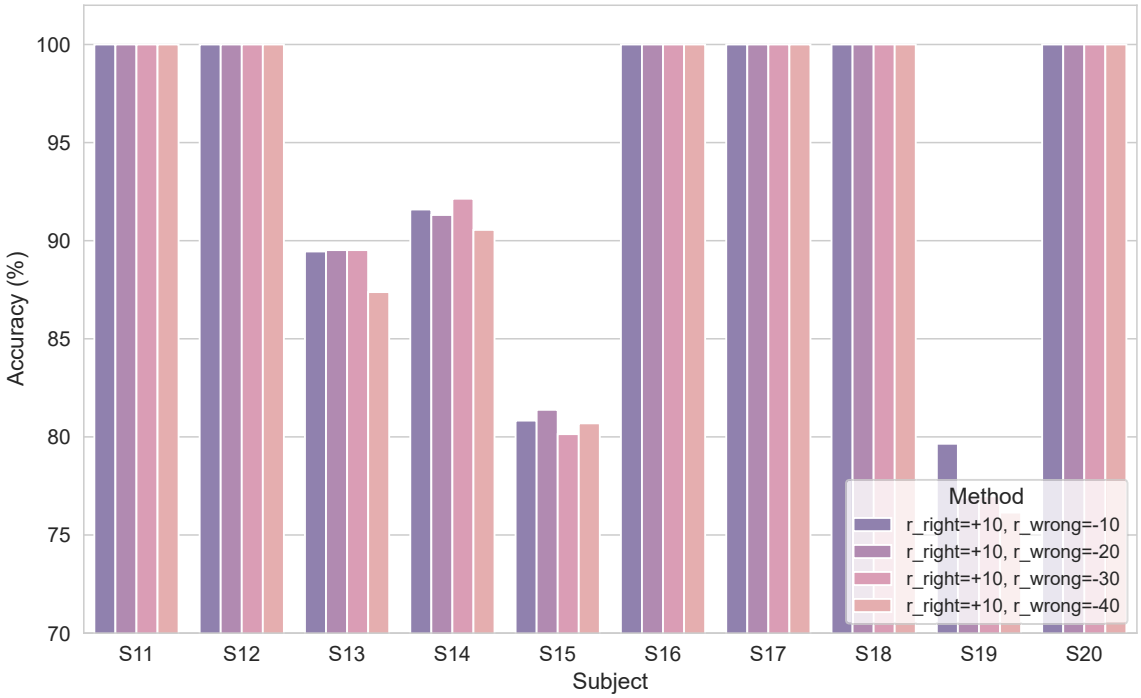


Figure 6.14: Performance Mean Accuracy Comparison Varying r_{wrong} Values for Subjects S_{11} - S_{20}

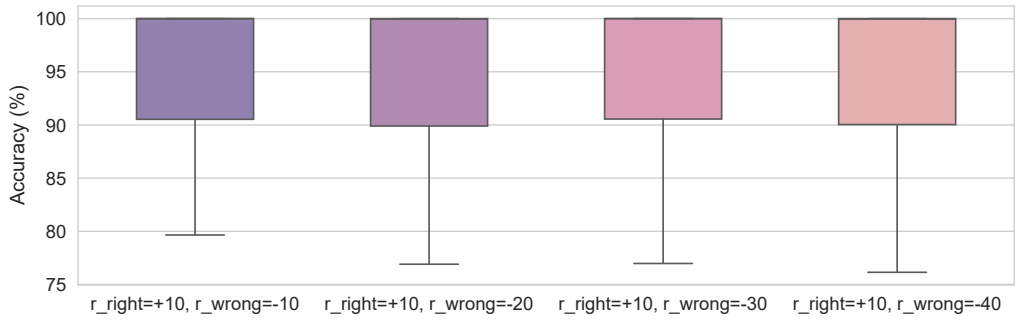


Figure 6.15: Overall Performance Mean Accuracy Comparison Varying r_{wrong} Values

Table 6.7: Impact of Varying r_{wrong} Values on Accuracy and Classification Time

Subj	Mean Accuracy (Mean Classification Time)			
	$r_{right} = +10$ $r_{wrong} = -10$	$r_{right} = +10$ $r_{wrong} = -20$	$r_{right} = +10$ $r_{wrong} = -30$	$r_{right} = +10$ $r_{wrong} = -40$
S_1	100.00% (1.80)	99.79% (1.60)	100.00% (2.00)	99.93% (2.00)
S_2	97.73% (1.87)	98.21% (1.97)	97.93% (2.49)	97.93% (2.88)
S_3	100.00% (2.20)	100.00% (1.70)	100.00% (2.70)	100.00% (2.20)
S_4	100.00% (2.50)	100.00% (3.30)	100.00% (4.30)	100.00% (5.70)
S_5	87.72% (4.55)	86.90% (6.26)	85.38% (6.28)	85.79% (7.31)
S_6	90.89% (2.80)	90.00% (2.63)	90.90% (4.94)	90.41% (4.25)
S_7	89.24% (3.23)	89.59% (5.09)	89.31% (5.20)	88.90% (6.70)
S_8	100.00% (2.80)	100.00% (2.60)	100.00% (3.70)	100.00% (4.30)
S_9	100.00% (3.00)	100.00% (3.30)	100.00% (5.00)	100.00% (5.30)
S_{10}	100.00% (2.20)	99.93% (1.90)	99.86% (2.00)	99.93% (2.52)
S_{11}	100.00% (1.50)	100.00% (2.20)	100.00% (3.00)	100.00% (3.30)
S_{12}	100.00% (2.40)	100.00% (2.70)	100.00% (4.10)	100.00% (5.10)
S_{13}	89.45% (3.97)	89.52% (5.33)	89.52% (6.24)	87.38% (6.07)
S_{14}	91.59% (3.14)	91.31% (3.00)	92.14% (3.00)	90.55% (4.37)
S_{15}	80.83% (4.80)	81.38% (4.53)	80.14% (3.48)	80.69% (4.13)
S_{16}	100.00% (2.70)	100.00% (3.10)	100.00% (5.00)	100.00% (4.30)
S_{17}	100.00% (2.00)	100.00% (3.70)	100.00% (3.20)	100.00% (4.60)
S_{18}	100.00% (2.40)	100.00% (3.20)	100.00% (2.40)	100.00% (4.30)
S_{19}	79.65% (5.64)	76.90% (7.21)	76.97% (8.22)	76.14% (9.20)
S_{20}	100.00% (2.70)	100.00% (2.80)	100.00% (3.40)	100.00% (4.80)
Mean	95.36% (2.91)	95.18% (3.41)	95.11% (4.03)	94.88% (4.66)
Std	± 6.83%	± 7.19%	± 7.37%	± 7.57%

$r_{skip} = -0.1$ and $H = 20$

Although a smaller magnitude of r_{wrong} relative to r_{right} appears beneficial, a higher r_{wrong}

to r_{right} ratio does not necessarily equate to improved accuracy. As demonstrated in Table 6.7, performance accuracy diminishes with an increasing ratio, identifying the optimal ratio as twice the magnitude of r_{wrong} to r_{right}. Furthermore, comparing different of reward configurations with equivalent magnitude ratios, such as $\{r_{right} = +5, r_{wrong} = -10\}$ and $\{r_{right} = +10, r_{wrong} = -20\}$, reveal subtle differences are noted in average performance accuracy and classification time. The former configuration outperforms in both average accuracy and time efficiency for classification.

According to Table 6.7, the classification time escalates with the increases in r_{wrong} magnitude, where average times of 2.91, 3.41, 4.03, and 4.66 seconds were recorded for r_{wrong} values of -10, -20, -30, and -40, respectively. This trend suggests that as the penalty for incorrect classification outweighs the reward for correct answers, the agents proceed with increased caution, hence extending the classification time.

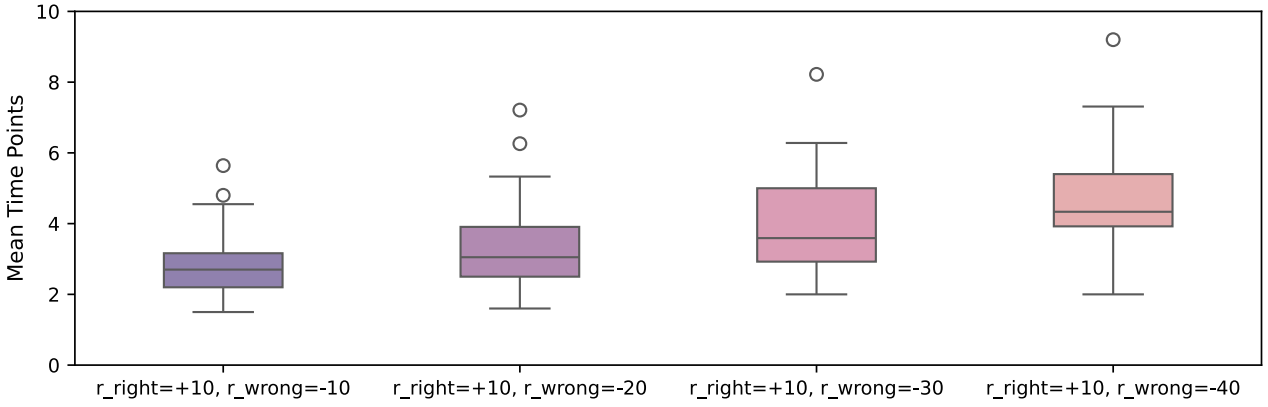


Figure 6.16: Overall Mean Classification Time (Time Points) Comparing Varying r_{wrong} Values

6.3.4 Effects of Episode Length Variation and Optimisation on Classification Performance

In this study, we examined the mean accuracy, F1 score, and mean classification time across various episode lengths (H), including 10, 20, 30, and 40, as presented in Table 6.8. We observed that both accuracy and F1 scores increased with extension of the episode horizon extends. Conversely, classification time per point increased with longer episode lengths. These finding suggests that larger episode lengths contribute to improvements in accuracy and F1 scores.

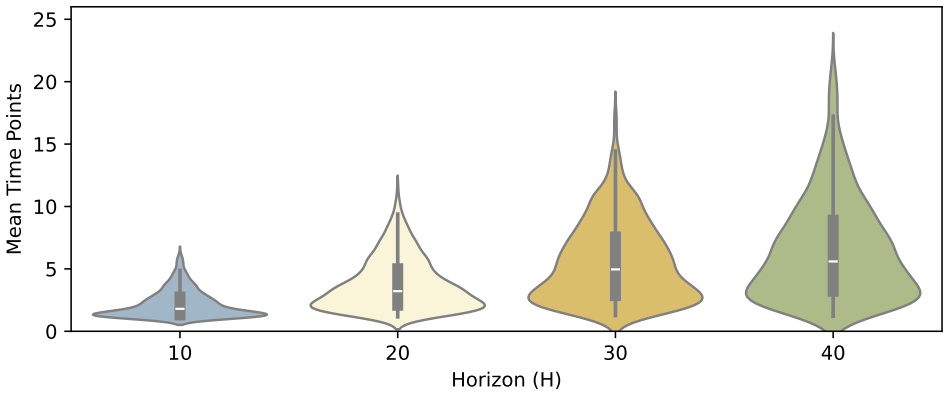


Figure 6.17: Overall Mean Classification Time (Time Points) Comparing Varying Horizon (H) Values

Table 6.8: Impact of Varying Episode Lengths (H) Values on Accuracy, F1 Score and Classification Time

Horizon (H)	Accuracy (Mean ± Std)	F1 Score (Mean ± Std)	Mean Classification Time
10	94.46% ± 8.10%	94.42% ± 8.15%	2.18
20	95.14% ± 7.14%	95.10% ± 7.18%	3.76
30	95.56% ± 6.54%	95.53% ± 5.53%	5.53
40	95.82% ± 6.16%	95.79% ± 6.54%	6.54

Table 6.9 delineates the optimal configuration of reward for correct (r_{right}) and incorrect (r_{wrong}) decisions, and episode horizon (H) that achieves the highest accuracy and F1 score in the shortest classification time possible. In this optimal setting, the RL agent demonstrates superior performance, achieving an average accuracy of 96.40% and an average classification time of less than 25 milliseconds across all 20 subjects.

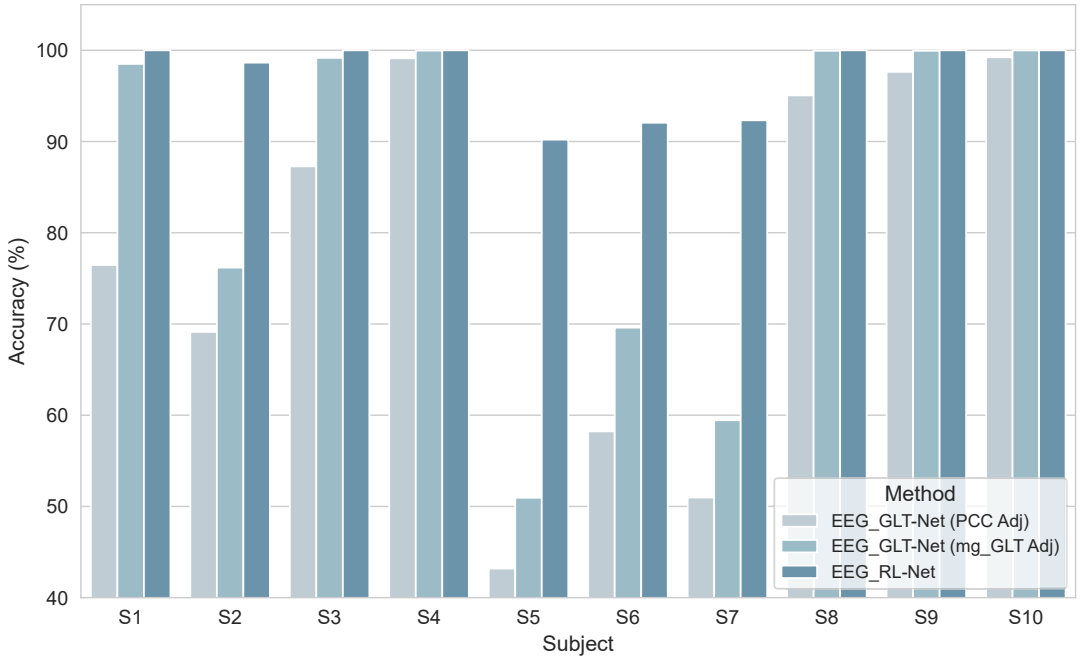


Figure 6.18: Performance Mean Accuracy Comparison between EEG_GLT-Net (PCC Adj), EEG_GLT-Net (m_{g_GLT} Adj) and EEG_RL-Net Methods at Optimal Reward Settings for Subjects S_1 - S_{10}

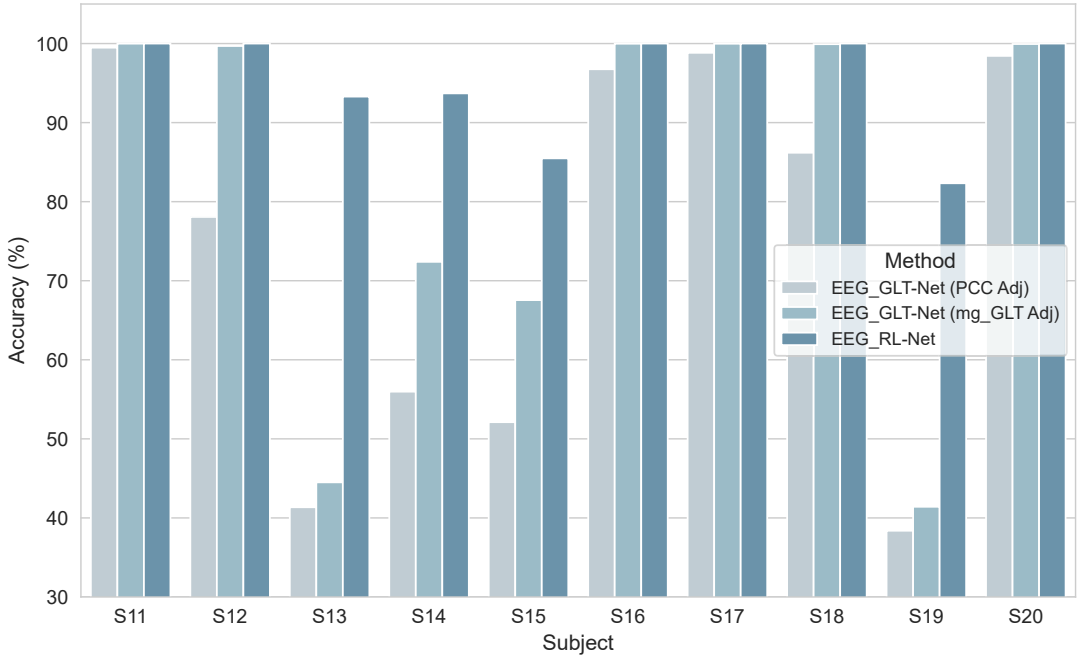


Figure 6.19: Performance Mean Accuracy Comparison between EEG_GLT-Net (PCC Adj), EEG_GLT-Net (m_{g_GLT} Adj) and EEG_RL-Net Methods at Optimal Reward Settings for Subjects S_{11} - S_{20}

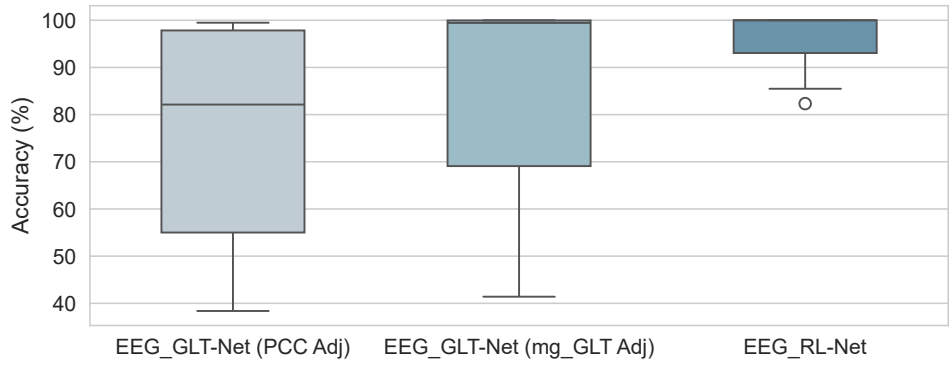


Figure 6.20: Overall Performance Mean Accuracy Comparison between EEG_GLT-Net (PCC Adj), EEG_GLT-Net (m_{g_GLT} Adj) and EEG_RL-Net Methods at Optimal Reward Settings

Table 6.9: Subject-wise Classification Accuracy and Time with Optimal Reward Settings and Episode Lengths

Subj	Mean Accuracy	Mean F1 Score	Mean Classification Time	(r_{right}, r_{wrong})	Episode Horizon
S_1	100.00%	100.00%	1.45	(20, -10)	10
S_2	98.65%	98.62%	2.93	(20, -30)	30
S_3	100.00%	100.00%	1.13	(20, -10)	10
S_4	100.00%	100.00%	1.32	(20, -30)	10
S_5	90.21%	90.05%	4.85	(10, -10)	30
S_6	92.06%	92.06%	4.25	(5, -10)	40
S_7	92.33%	92.29%	9.94	(10, -30)	40
S_8	100.00%	100.00%	1.23	(10, -10)	10
S_9	100.00%	100.00%	1.27	(20, -10)	10
S_{10}	100.00%	100.00%	1.17	(20, -10)	10
S_{11}	100.00%	100.00%	1.11	(20, -20)	10
S_{12}	100.00%	100.00%	1.19	(20, -10)	10
S_{13}	93.29%	93.27%	5.95	(10, -10)	40
S_{14}	93.70%	93.69%	4.17	(5, -10)	40
S_{15}	85.48%	85.43%	7.51	(10, -20)	40
S_{16}	100.00%	100.00%	1.25	(20, -10)	10
S_{17}	100.00%	100.00%	1.28	(20, -10)	10
S_{18}	100.00%	100.00%	1.27	(10, -10)	10
S_{19}	82.33%	82.20%	9.69	(20, -30)	40
S_{20}	100.00%	100.00%	1.15	(20, -10)	10
Mean	96.40%	96.38%	3.21	-	-
Std	± 5.47	± 5.50	-	-	-

Our analysis, as indicated in Table 6.9 shows that the RL agent achieved accuracy exceeding 90.00% for each subject, with the exceptions of S_{15} and S_{19} whose accuracies were 85.48% and 82.33%, respectively. Subjects such as S_1, S_3, S_4, S_8, S_9, S_{10}, S_{11}, S_{12}, S_{16}, S_{17}, S_{18}, and S_{20}, where the RL agent achieved perfect classification, had notably clearer EEG MI signals. For these subjects, the agent performed consistently well across most reward and episode horizon configurations. Classifications were achieved within an average of 2 time points, where an optimal episode horizon of 10 and a reward configuration where r_{right} significantly exceeded r_{wrong} were conducive to faster classification decisions.

Particularly noteworthy was the performance of EEG_RL-Net on subject S_{13}, where the RL agent achieved a classification accuracy of 93.29%. This represented an exceptional improvement by 48.79% over EEG_GLT-Net with m_{g_GLT}, the current state-of-the-art EEG MI time point classification method. The classification for S_{13} took 6 time points on average, possibly reflecting the only subtle distinctions between EEG MI tasks for this subject.

6.3.5 Comparison with Current State-of-the-Art Models

In this paper, we compare our proposed method, EEG_GLT-Net, with eight other state-of-the-art (SOTA) models on PhysioNet dataset listed in Table 6.10, including FBCSP [106], EEGNet [108], CasCNN [109], DG-HAM [123], EEG-ARNN [124], SSDA [110], GCNs-Net [100], and EEG_GLT-Net [144]. Our comparisons begin with the traditional FBCSP approach, which leverages CSP to extract features across multiple frequency bands and utilises SVM for classification. We then compare with EEGNet, a widely used model based solely on a CNN structure. Further, we assess CasCNN and SSDA, both of which combine CNN and LSTM networks. Finally, we evaluate our method against DG-HAM, EEG-ARNN, and GCNs-Net, EEG_GLT-Net, which are GNN-based networks.

Table 6.10: Performance Comparison of EEG_RL-Net Model with State-of-the-Art Models on PhysioNet Dataset

Method	Avg. Accuracy	Avg. F1 Score
FBCSP [106]	59.56%	60.04%
EEGNet [108]	72.20%	72.10%
CasCNN [109]	63.30%	63.18%
DG-HAM [123]	76.15%	76.08%
EEG-ARNN [124]	82.39%	82.17%
SSDA [110]	83.73%	83.24%
GCNs-Net [100]	80.16%	80.05%
EEG_GLT-Net	86.43%	86.23%
Proposed EEG_RL-Net	**96.40%**	**96.38%**

The traditional FBCSP method achieves 59.56%, the lowest accuracy among the SOTAs, likely due to its reliance on SVM as the classifier. The popular EEGNet achieves 72.20% accuracy, outperforming the CasCNN model, which achieves only 63.30%. Within the CNN-based SOTA models, SSDA reaches the highest accuracy at 83.73%. Among the GNN-based SOTA models, EEG_GLT-Net achieves the highest accuracy at 86.43% followed by EEG-ARNN, GCNs-Net, DG-HAM with accuracies of 82.39%, 80.16% and 76.15%, respectively. Our proposed EEG_RL-Net, achieves the highest overall accuracy of 96.40% among all SOTAs.

For the HGD dataset, we compare EEG_RL-Net with eight SOTA models, as detailed in Table 6.11. These include ShallowNet [58], EEGNet [108], EEG-ITNet [155], CRGNet [169], EEG-Conformer [170], LMDA-Net [171], IFNet [172], and GCNs-Net [100]. The HGD dataset, which comprises EEG signals of motor movements, has more distinguishable class signals compared to motor imagery datasets like PhysioNet dataset. As a result, most SOTA models achieve over 90% accuracy, except for EEG-ITNet (84.87%), EEGNet (87.99%), and LMDA-Net (87.68%). EEG-ITNet, a CNN-based model that includes inception and temporal convolution (TC) blocks, underperforms compared to the simpler two-layer CNN model ShallowNet, which achieves 93.04%.

Table 6.11: Performance Comparison of EEG_RL-Net Model with State-of-the-Art Models on HGD Dataset

Method	Avg. Accuracy	Avg. F1 Score
ShallowNet [58]	93.04%	92.87%
EEGNet [108]	87.99%	87.85%
EEG-ITNet [173]	84.87%	84.74%
CRGNet [169]	94.33%	94.28%
EEG-Conformer [170]	93.48%	93.32%
LMDA-Net [171]	87.68%	87.49%
IFNet [172]	93.62%	93.28%
GCNs-Net [100]	96.24%	96.18%
EEG_SGNet	98.56%	98.56%
Proposed EEG_RL-Net	**99.24%**	**99.24%**

EEG-Conformer performs slightly better than ShallowNet, while IFNet, a model employing 1D CNN operations for low- and high-frequency EEG bands with cross-frequency integration, achieves an accuracy of 93.62%. CRGNet, which incorporates Riemannian manifold embedding after CNN operations, achieves the second-highest accuracy among SOTA models at 94.33%. The highest accuracy among the existing SOTA models is achieved by GCNs-Net, a spectral GNN approach using PCC for adjacency matrix construction, with 96.24%. In comparison, our proposed EEG_RL-Net outperforms all existing SOTAs models, achieving the highest accuracy of 99.24%.

6.4 Limitations and Future Works

EEG_RL-Net introduces reinforcement learning into EEG-based classification, providing an adaptive decision-making framework for stroke rehabilitation applications. The single-subject paradigm used in this study ensures that the model effectively learns individualised neural patterns, which is particularly beneficial for fine-tuned classification. While the current training approach is focused on optimising performance within individual subjects, evaluating its adaptability across a broader population would further validate its robustness. Given that EEG signals exhibit natural variations across users, further assessments in diverse settings could provide deeper insights into how EEG_RL-Net generalises to new individuals.

Despite these considerations, EEG_RL-Net presents a flexible and promising approach for real-time EEG classification. Its reinforcement learning framework is well-suited to incremental learning and adaptive policy refinement, which can support generalisation to new users with minimal calibration. Future work could explore subject-independent training strategies or transfer learning techniques to further improve cross-subject robustness. Additionally, optimising reinforcement learning models for computational efficiency will be essential for their real-time deployment in portable or resource-limited rehabilitation devices. These advancements would enhance EEG_RL-Net's scalability, making it a more practical solution for EEG-based rehabilitation technologies.

6.5 Conclusion

Our study introduces EEG_RL-Net, an innovative approach for the real-time classification of EEG MI signals utilising RL techniques. Building on the foundation of EEG_GLT-Net's EEG_GCN block and optimising computational efficiency with an adjacency matrix density of just 13.39%, EEG_RL-Net not only achieves accurate classification of EEG MI signals but also identifies signals that are unsuitable for classification. Remarkably, it achieved 100.00% classification accuracy for 12 out of 20 subjects within less than 12.5 ms. For challenging subjects (S_{13} and S_{19} in this study), where previous state-of-the-art methods such as EEG_GLT-Net could classify with accuracies of only 44.50% and 41.41% respectively, EEG_RL-Net achieved unprecedented improvement in performance, reaching classification accuracies of 93.29% and 82.33% in less than 62.5 ms. These results underscore the robustness and efficacy of EEG_RL-Net in enhancing classification rates, filling a gap for subjects previously deemed difficult by existing classification methods. In future work, we will further explore the integration of the optimal adjacency matrix m_{g_GLT} for advanced graph feature extraction in the EEG_GCN block, aiming to unlock even greater improvements in the classification capabilities of our EEG_RL-Net system.

6.6 Significance and Novelty

This chapter presents EEG_RL-Net, a real-time framework combining Graph Neural Networks (GNNs) with reinforcement learning (RL) to optimise EEG signal decoding. The significance of this work lies in its innovative use of a dueling deep Q-network (DQN) that allows the model to dynamically skip uncertain time points, focusing only on high-value segments of the signal. This approach enhances classification accuracy while reducing computational costs, making it well-suited for time-critical applications like neurorehabilitation and assistive technologies.

The novelty of EEG_RL-Net is its seamless integration of GNN-based feature extraction and RL for decision-making. By dynamically adapting to signal variability and selectively prioritising key data segments, the framework offers a scalable and efficient solution for real-time EEG classification. This chapter advances the field by providing a novel approach to balancing speed and precision in EEG applications.

Chapter 7

Conclusions and Future Work

7.1 Conclusion

In Chapter 3, the EEG_GLT approach for constructing adjacency matrices was introduced, demonstrating its superiority in both performance accuracy and computational efficiency. Compared to existing methods like Geodesic and PCC, EEG_GLT consistently outperformed in six different EEG_GLT-Net model configurations using time-point EEG MI signals from the public PhysioNet dataset. This classification method, applicable in rehabilitation robotics, utilised data from 20 of 109 subjects. Notably, EEG_GLT enhanced accuracy and F1 scores by 0.52% to 22.04% and 0.50% to 21.76%, respectively, surpassing the PCC method employed in the state-of-the-art GCNs-Net model. The method also achieved a 13.39% improvement in average accuracy across subjects while maintaining adjacency matrix densities below 22.53%. Additionally, EEG_GLT required only 8.76M to 80.67M MACs, significantly lower than PCC and Geodesic methods (42.26M to 291.62M MACs). Furthermore, EEG_GLT outperformed the state-of-the-art SSDA model, achieving 86.43% accuracy compared to SSDA's 83.73%.

Chapter 4 extended the application of EEG_GLT to economic variable forecasting. Using 20 years of Australian economic data (July 2000-June 2020) across eight variables-CPI, Wage Index, short- and long-term interest rates, domestic and international bond indices, housing price index, and unemployment rate—the method demonstrated its versatility. In a 2-layer GNN forecasting model, EEG_GLT utilised past three-quarter values to predict the next quarter. The optimal network density was 21.88%. Compared to the SUPA model, the EEG_GLT-based model achieved a 40% lower average RMSE, underscoring its predictive accuracy.

Finally, in Chapter 5, EEG_SGNet was introduced to classify window-based EEG MI and movement signals, with potential applications in rehabilitation robotics. This parallel CNN-GNN architecture utilised BCIC_iv-2a and HGD public datasets. Nine model configurations were explored, comparing PCC and EEG_GLT methods in the GNN pathway. The EEG_GLT-based EEG_SGNet achieved up to 4.31% and 4.43% improvements in accuracy and F1 scores on the

BCIC_iv-2a dataset, and up to 1.31% and 1.28% on the HGD dataset. Overall, EEG_SGNet achieved 82.98% accuracy and 82.88% F1 score on BCIC_iv-2a, outperforming the state-of-the-art 4DDFM+3M3DCNN model (80.04% accuracy, 79.95% F1 score). On the HGD dataset, EEG_SGNet achieved 98.56% accuracy and F1, surpassing the GCNs-Net model's 96.24% accuracy and 96.18% F1 score.

In Chapter 6, the performance of EEG time-point classification was significantly enhanced using the EEG RL-Net model, which integrates the EEG GCN block for feature extraction with a Dueling DQN RL agent for classification. Unlike traditional supervised learning models, the RL agent's ability to skip noisy time points improved robustness and accuracy. On the PhysioNet dataset, EEG_RL-Net achieved 100% accuracy for 12 out of 20 subjects within 12.5 milliseconds. For challenging cases such as S_{13} and S_{19}, where EEG_GLT-Net achieved only 44.50% and 41.41% accuracy, EEG_RL-Net demonstrated remarkable improvements, reaching 93.29% and 82.33%, respectively, in under 62.5 milliseconds. Across all subjects, EEG_RL-Net increased the average accuracy to 96.40% and the F1 score to 96.38%, outperforming EEG_GLT-Net by approximately 10% on the PhysioNet dataset. Furthermore, on the HGD dataset, EEG_RL-Net achieved nearly perfect accuracy of 99.24%, surpassing the existing state-of-the-art model, GCNs-Net, which achieved 96.24%.

Finally, this research makes significant contributions by addressing critical challenges in signal processing and graph-based learning. It introduces a domain-agnostic adjacency matrix construction method, enabling scalable graph-based models applicable across diverse fields. Additionally, it demonstrates real-time, time-point-level signal classification, overcoming limitations in time-sensitive decision-making and paving the way for more efficient and adaptive systems. These advancements enhance quality of life by improving neurorehabilitation technologies and enabling accessible health monitoring tools. By optimising computational efficiency and resource utilisation, the research also reduces costs, making advanced technologies more practical and widely adoptable. Moreover, it inspires innovation in interdisciplinary domains, driving progress in areas such as healthcare, engineering, environmental systems, and beyond.

7.2 Future Work

This research has made significant progress in EEG signal classification, particularly for rehabilitation applications. The proposed models such as EEG_GLT-Net, EEG_SGNet, and EEG_RL-Net have advanced the state-of-the-art in decoding user intentions from EEG data, paving the way for more effective and efficient rehabilitation systems. Building on this foundation, several promising directions for future exploration have emerged, both within rehabilitation and in broader domains.

7.2.1 Rehabilitation Applications

- **Integration with Robotic Exoskeletons**
  The integration of the proposed models with robotic exoskeletons presents an exciting opportunity to advance real-world rehabilitation technologies. Future work should focus on testing these models in real-time scenarios to validate their effectiveness in controlling exoskeletons for motor recovery tasks. This step is critical for translating the research into practical rehabilitation settings and demonstrating its potential to enhance patient outcomes.

- **Patient-Specific Model Training**
  While this study benchmarked the models using public datasets from healthy individuals, the next phase involves collecting EEG data from stroke survivors. Training the models on patient-specific data will adapt them to the unique needs of individuals undergoing rehabilitation. This step is crucial for ensuring the models are clinically applicable and capable of delivering personalized rehabilitation solutions.

- **Multimodal Signal Integration**
  Incorporating additional physiological signals, such as Electromyography (EMG) or functional Near-Infrared Spectroscopy (fNIRS), can further improve the accuracy and robustness of the models. By combining these multimodal signals with EEG data, future research can achieve a deeper understanding of user intentions, enhancing the reliability and adaptability of the systems in diverse scenarios.

- **Deployment on Portable and Wearable Devices**
  Deploying these systems on portable and wearable devices, such as compact EEG headsets, is essential for making them accessible for home-based rehabilitation. While the models are computationally efficient, developing lightweight versions optimized for mobile and embedded platforms remains a key challenge. Addressing this will expand the reach and usability of these systems, enabling broader adoption in non-clinical environments.

- **Advancing Reinforcement Learning Strategies**
  The reinforcement learning strategies employed in EEG_RL-Net can be further refined to handle more complex and dynamic scenarios. Exploring advanced reward mechanisms, multi-agent frameworks, or hierarchical reinforcement learning approaches could enhance adaptability and versatility. These improvements would enable the systems to perform effectively in a wider range of real-time applications, creating smarter and more adaptable rehabilitation technologies that improve patients' quality of life.

7.2.2 Broader Applications Beyond Rehabilitation

The methodologies developed in this thesis demonstrate potential far beyond EEG classification. Future work could explore applications such as:

- **Healthcare Beyond Neurorehabilitation**
  Extending real-time classification to other physiological signals, such as EMG for muscle analysis, ECG for cardiac monitoring, and wearable sensors for personalised health tracking.
- **Human-Computer Interaction (HCI)**
  Adapting these methods for brain-computer interfaces in gaming, augmented reality, or assistive technologies for individuals with disabilities.
- **Autonomous Systems**
  Enhancing decision-making in robotics or autonomous vehicles by integrating real-time adaptive classification of sensor data.
- **Environmental Monitoring and Financial Systems**
  Applying the adjacency matrix construction and classification techniques to model complex relationships in climate change forecasting and economic prediction.

References

[1] R. L. Sacco, S. E. Kasner, J. P. Broderick, *et al.*, "An updated definition of stroke for the 21st century: A statement for healthcare professionals from the american heart association/american stroke association," *Stroke*, vol. 44, no. 7, pp. 2064–2089, 2013.

[2] J. L. Saver, "Time is brain—quantified," *Stroke*, vol. 37, no. 1, pp. 263–266, 2006.

[3] Deloitte, *The economic impact of stroke in australia, 2020*, https://strokefoundation.org.au/, Viewed 16th October 2022, Nov. 2020.

[4] E. Akdoğan and M. A. Adli, "The design and control of a therapeutic exercise robot for lower limb rehabilitation: Physiotherabot," *Mechatronics*, vol. 21, no. 3, pp. 509–522, 2011, ISSN: 0957-4158. DOI: https://doi.org/10.1016/j.mechatronics.2011.01.005. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957415811000195.

[5] I. Díaz, J. J. Gil, and E. Sánchez, "Lower-limb robotic rehabilitation: Literature review and challenges," *Journal of Robotics*, vol. 2011, no. 1, p. 759 764, 2011. DOI: https://doi.org/10.1155/2011/759764. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1155/2011/759764. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1155/2011/759764.

[6] M. S. AL-Quraishi, I. Elamvazuthi, S. A. Daud, S. Parasuraman, and A. Borboni, "Eeg-based control for upper and lower limb exoskeletons and prostheses: A systematic review," *Sensors*, vol. 18, no. 10, 2018, ISSN: 1424-8220. DOI: 10.3390/s18103342. [Online]. Available: https://www.mdpi.com/1424-8220/18/10/3342.

[7] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain–computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002, ISSN: 1388-2457. DOI: https://doi.org/10.1016/S1388-2457(02)00057-3. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1388245702000573.

[8] IBM, *What is machine learning (ml)?* Accessed on Friday 22nd November 2024, Aug. 2024. [Online]. Available: https://www.ibm.com/topics/machine-learning.

[9] Expert.ai Team, *What is the definition of machine learning?* Accessed on Friday 22nd November 2024, Mar. 2022. [Online]. Available: https://www.expert.ai/blog/machine-learning-definition/.

[10] W. Huo, S. Mohammed, J. C. Moreno, and Y. Amirat, "Lower limb wearable robots for assistance and rehabilitation: A state of the art," *IEEE Systems Journal*, vol. 10, no. 3, pp. 1068–1081, 2016. DOI: 10.1109/JSYST.2014.2351491.

[11] T. D. Lalitharatne, K. Teramoto, Y. Hayashi, K. Tamura, and K. Kiguchi, "Eeg-based evaluation for perception-assist in upper-limb power-assist exoskeletons," in *2014 World Automation Congress (WAC)*, 2014, pp. 307–312. DOI: 10.1109/WAC.2014.6935909.

[12] C. Cordonnier, A. Demchuk, W. Ziai, and C. S. Anderson, "Intracerebral haemorrhage: Current approaches to acute management," *The Lancet*, vol. 392, no. 10154, pp. 1257–1268, 2018.

[13] D. Purves, G. J. Augustine, D. Fitzpatrick, *et al.*, "Circuits within the basal ganglia system," in *Neuroscience. 2nd edition*, Sinauer Associates, 2001.

[14] S. Foundation, *Stroke attacks the brain - the human control centre - changing lives in an instant.* en, publisher: Stroke Foundation. [Online]. Available: https://strokefoundation.org.au/about-stroke/learn/what-is-a-stroke (visited on 09/27/2024).

[15] C. O. Johnson, M. Nguyen, G. A. Roth, *et al.*, "Global, regional, and national burden of stroke, 1990–2016: A systematic analysis for the global burden of disease study 2016," *The Lancet Neurology*, vol. 18, no. 5, pp. 439–458, 2019.

[16] L.-F. Zhang, J. Yang, Z. Hong, *et al.*, "Proportion of different subtypes of stroke in china," *Stroke*, vol. 34, no. 9, pp. 2091–2096, 2003.

[17] B. C. V. Campbell and P. Khatri, "Stroke," *The Lancet (British edition)*, vol. 396, no. 10244, pp. 129–, 2020.

[18] Australian Bureau of Statistics, *National health survey: First results*, https://www.abs.gov.au/statistics/health/health-conditions-and-risks/national-health-survey/2017-18, Viewed 20 October 2022, 2018.

[19] V. L. Feigin, M. Brainin, B. Norrving, *et al.*, "World stroke organization (wso): Global stroke fact sheet 2022," *International Journal of Stroke*, vol. 17, no. 1, pp. 18–29, 2022, PMID: 34986727. DOI: 10.1177/17474930211065917. eprint: https://doi.org/10.1177/17474930211065917. [Online]. Available: https://doi.org/10.1177/17474930211065917.

[20] Australian Institute of Health and Welfare, *Australian Burden of Disease Study: Impact and Causes of Illness and Death in Australia 2015*. AIHW, Canberra, 2019. DOI: 10.25816/5ebca2a4fa7dc.

[21] Australian Institute of Health and Welfare, *Australia's Health 2020: Data Insights*. AIHW, Canberra, 2020. DOI: 10.25816/5f05371c539f3.

[22] Y. Cheng, Y. Lin, H. Shi, *et al.*, "Projections of the stroke burden at the global, regional, and national levels up to 2050 based on the global burden of disease study 2021," *Journal of the American Heart Association*, vol. 13, no. 23, e036142, 2024. DOI: 10.1161/JAHA.124.036142. eprint: https://www.ahajournals.org/doi/pdf/10.1161/JAHA.124.036142. [Online]. Available: https://www.ahajournals.org/doi/abs/10.1161/JAHA.124.036142.

[23] S. Foundation and D. A. Economics, *No postcode untouched: Stroke in australia 2020*, Prepared for the Stroke Foundation., Sydney, Australia, Nov. 2020. [Online]. Available: https://strokefoundation.org.au/What-we-do/Research/Economic-impact-of-stroke-in-Australia.

[24] Australian Bureau of Statistics, *Disability, ageing and carers, australia: Summary of findings*, https://www.abs.gov.au/statistics/health/disability/disability-ageing-and-carers-australia-summary-findings/2018, Viewed 20 October 2022, 2018.

[25] G. Kwakkel, B. Kollen, and E. Lindeman, "Understanding the pattern of functional recovery after stroke: Facts and theories," *Restorative neurology and neuroscience*, vol. 22, no. 3-5, pp. 281–299, 2004.

[26] *Stroke rehabilitation*, Accessed 21 November 2024, 2023. [Online]. Available: https://www.healthdirect.gov.au/stroke-rehabilitation.

[27] LG Electronics, *Lg to usher in new era of ai robotics with wearable robot*, https://www.lg.com/ae/about-lg/press-media/lg-to-usher-in-new-era-of-ai-robotics-with-wearable-robot, Accessed: 2024-11-28, 2018.

[28] NASA Technology Transfer Program, *X1 robotic exoskeleton*, https://technology.nasa.gov/patent/msc-tops-85, Accessed: 2024-11-28, 2024.

[29] M. Tariq, P. M. Trivailo, and M. Simic, "Eeg-based bci control schemes for lower-limb assistive-robots," *Frontiers in human neuroscience*, vol. 12, p. 312, 2018.

[30] J. S. Kumar and P. Bhuvaneswari, "Analysis of electroencephalography (eeg) signals and its categorization–a study," *Procedia Engineering*, vol. 38, pp. 2525–2536, 2012, INTERNATIONAL CONFERENCE ON MODELLING OPTIMIZATION AND COMPUTING, ISSN: 1877-7058. DOI: https://doi.org/10.1016/j.proeng.2012.06.298. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877705812022114.

[31] R. Palaniappan and P. Raveendran, "Cognitive task prediction using parametric spectral analysis of eeg signals," *Malaysian Journal of Computer Science*, vol. 14, no. 1, pp. 58–67, Jun. 2001. [Online]. Available: https://mjes.um.edu.my/index.php/MJCS/article/view/5852.

[32] C. Henley, *External Brain Anatomy*, en, Jan. 2021. [Online]. Available: https://openbooks.lib.msu.edu/neuroscience/chapter/external-brain-anatomy/ (visited on 09/27/2024).

[33] C. Chiron, I. Jambaque, R. Nabbout, R. Lounes, A. Syrota, and O. Dulac, "The right brain hemisphere is dominant in human infants.," *Brain: a journal of neurology*, vol. 120, no. 6, pp. 1057–1065, 1997.

[34] Y. Gil, S. Seo, and J. Lee, "Eeg analysis of frontal lobe activities by decision stimuli," in *2008 Second International Conference on Future Generation Communication and Networking*, vol. 3, 2008, pp. 30–34. DOI: 10.1109/FGCN.2008.115.

[35] M. B. Westover, D. M. Greer, E. Choi, and K. M. Awad, *Pocket neurology*. Lippincott Williams & Wilkins, 2010.

[36] T. Yamada and E. Meng, *Practical guide for clinical neurophysiologic testing: EEG*. Lippincott Williams & Wilkins, 2012.

[37] *Brain Facts: The Four Lobes*. [Online]. Available: https://www.brainframe-kids.com/brain/facts-lobes.htm (visited on 09/27/2024).

[38] E. Goldberg, *The executive brain: Frontal lobes and the civilized mind*. Oxford University Press, USA, 2001.

[39] L. J. Hirsch and R. P. Brenner, *Atlas of EEG in critical care*. John Wiley & Sons, 2011.

[40] A. S. AlMejrad, "Human emotions detection using brain wave signals: A challenging," *European Journal of Scientific Research*, vol. 44, no. 4, pp. 640–659, 2010.

[41] "Treatment of status epilepticus: A survey of neurologists," *Journal of the Neurological Sciences*, vol. 211, no. 1, pp. 37–41, 2003, ISSN: 0022-510X. DOI: https://doi.org/10.1016/S0022-510X(03)00036-4. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0022510X03000364.

[42] D. L. Trudeau, "Applicability of brain wave biofeedback to substance use disorder in adolescents," *Child and Adolescent Psychiatric Clinics*, vol. 14, no. 1, pp. 125–136, 2005.

[43] I. Gutberlet, S. Debener, T. Jung, and S. Makeig, "Techniques of eeg recording and preprocessing," *Quantative EEG Analysis Methods and Clinical Applications*, pp. 23–49, 2009.

[44] TimesPro, *Processing brain signals in python*, https://timespro.com/blog/processing-brain-signals-in-python, Accessed: 2024-11-28, 2023.

[45] P. Tallgren, S. Vanhatalo, K. Kaila, and J. Voipio, "Evaluation of commercially available electrodes and gels for recording of slow eeg potentials," *Clinical Neurophysiology*, vol. 116, no. 4, pp. 799–806, 2005, ISSN: 1388-2457. DOI: https://doi.org/10.1016/j.clinph.2004.10.001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1388245704003906.

[46] B. A. Taheri, R. T. Knight, and R. L. Smith, "A dry electrode for eeg recording," *Electroencephalography and Clinical Neurophysiology*, vol. 90, no. 5, pp. 376–383, 1994, ISSN: 0013-4694. DOI: https://doi.org/10.1016/0013-4694(94)90053-1. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0013469494900531.

[47] J. M. Stern, *Atlas of EEG patterns*. Lippincott Williams & Wilkins, 2005.

[48] H. H. Jasper, "Ten-twenty electrode system of the international federation," *Electroencephalogr Clin Neurophysiol*, vol. 10, pp. 371–375, 1958.

[49] TMSi, *The 10-20 System for EEG*, en. [Online]. Available: https://info.tmsi.com/blog/the-10-20-system-for-eeg (visited on 09/27/2024).

[50] G. E. Chatrian, E. Lettich, and P. L. Nelson, "Ten percent electrode system for topographic studies of spontaneous and evoked eeg activities," *American Journal of EEG technology*, vol. 25, no. 2, pp. 83–92, 1985.

[51] R. Oostenveld and P. Praamstra, "The five percent electrode system for high-resolution eeg and erp measurements," *Clinical Neurophysiology*, vol. 112, no. 4, pp. 713–719, 2001, ISSN: 1388-2457. DOI: https://doi.org/10.1016/S1388-2457(00)00527-7. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1388245700005277.

[52] J. N. Acharya, A. J. Hani, J. Cheek, P. Thirumala, and T. N. T. and, "American clinical neurophysiology society guideline 2: Guidelines for standard electrode position nomenclature," *The Neurodiagnostic Journal*, vol. 56, no. 4, pp. 245–252, 2016, PMID: 28436791. DOI: 10.1080/21646821.2016.1245558. eprint: https://doi.org/10.1080/21646821.2016.1245558. [Online]. Available: https://doi.org/10.1080/21646821.2016.1245558.

[53] C. Noronha, *A brief introduction to EEG and the types of electrodes*, en, Mar. 2024. [Online]. Available: https://brainlatam.com/blog/a-brief-introduction-to-eeg-andthe-types-of-electrodes-75 (visited on 09/27/2024).

[54] B. Farnsworth, *Eeg headset prices – an overview of 15+ eeg devices*, en-US, Jul. 2019. [Online]. Available: https://imotions.com/blog/learning/product-guides/eeg-headset-prices/ (visited on 09/27/2024).

[55] Y. LeCun, B. Boser, J. Denker, *et al.*, "Handwritten digit recognition with a back-propagation network," *Advances in neural information processing systems*, vol. 2, 1989.

[56] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 1, 2015. DOI: 10.1038/nature14539. [Online]. Available: https://doi.org/10.1038/nature14539.

[57] L. L. Iglesias, P. S. Bellón, A. P. del Barrio, *et al.*, "A primer on deep learning and convolutional neural networks for clinicians," *Insights into Imaging*, vol. 12, no. 1, p. 117, Aug. 12, 2021. DOI: 10.1186/s13244-021-01052-z. [Online]. Available: https://doi.org/10.1186/s13244-021-01052-z.

[58] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, *et al.*, "Deep learning with convolutional neural networks for eeg decoding and visualization," *Human Brain Mapping*, vol. 38, no. 11, pp. 5391–5420, 2017. DOI: https://doi.org/10.1002/hbm.23730. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hbm.23730. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.23730.

[59] A. Vilamala, K. H. Madsen, and L. K. Hansen, "Deep convolutional neural networks for interpretable analysis of eeg sleep stage scoring," in *2017 IEEE 27th International Workshop on Machine Learning for Signal Processing (MLSP)*, 2017, pp. 1–6. DOI: 10.1109/MLSP.2017.8168133.

[60] H. A. Siddiqa, Z. Tang, Y. Xu, *et al.*, "Single-channel eeg data analysis using a multi-branch cnn for neonatal sleep staging," *IEEE Access*, vol. 12, pp. 29 910–29 925, 2024. DOI: 10.1109/ACCESS.2024.3365570.

[61] S. K. Satapathy and D. Loganathan, "Automated classification of multi-class sleep stages classification using polysomnography signals: A nine- layer 1d-convolution neu-

ral network approach," *Multimedia Tools and Applications*, vol. 82, no. 6, pp. 8049–8091, Mar. 1, 2023. DOI: 10.1007/s11042-022-13195-2. [Online]. Available: https://doi.org/10.1007/s11042-022-13195-2.

[62] W. Chen, Y. Wang, Y. Ren, *et al.*, "An automated detection of epileptic seizures eeg using cnn classifier based on feature fusion with high accuracy," *BMC Medical Informatics and Decision Making*, vol. 23, no. 1, p. 96, May 22, 2023. DOI: 10.1186/s12911-023-02180-w. [Online]. Available: https://doi.org/10.1186/s12911-023-02180-w.

[63] A. A. Ein Shoka, M. M. Dessouky, A. El-Sayed, and E. El-Din Hemdan, "An efficient cnn based epileptic seizures detection framework using encrypted eeg signals for secure telemedicine applications," *Alexandria Engineering Journal*, vol. 65, pp. 399–412, 2023, ISSN: 1110-0168. DOI: https://doi.org/10.1016/j.aej.2022.10.014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1110016822006639.

[64] L. Joseph, S. Izudheen, U. Narayanana, A. R. Nair, and A. G. K A, "An efficient approach for eeg seizure detection using cnn with feature extraction," in *2024 7th International Conference on Circuit Power and Computing Technologies (ICCPCT)*, vol. 1, 2024, pp. 1924–1929. DOI: 10.1109/ICCPCT61902.2024.10673095.

[65] X. Wang, Y. Ma, J. Cammon, F. Fang, Y. Gao, and Y. Zhang, "Self-supervised eeg emotion recognition models based on cnn," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 1952–1962, 2023. DOI: 10.1109/TNSRE.2023.3263570.

[66] T. Song, W. Zheng, P. Song, and Z. Cui, "Eeg emotion recognition using dynamical graph convolutional neural networks," *IEEE Transactions on Affective Computing*, vol. 11, no. 3, pp. 532–541, 2020. DOI: 10.1109/TAFFC.2018.2817622.

[67] S. K. Jha, S. Suvvari, and M. Kumar, "Emotion recognition from electroencephalogram (eeg) signals using a multiple column convolutional neural network model," *SN Computer Science*, vol. 5, no. 2, p. 213, Jan. 20, 2024. DOI: 10.1007/s42979-023-02543-0. [Online]. Available: https://doi.org/10.1007/s42979-023-02543-0.

[68] N. Yao, H. Su, D. Li, *et al.*, "Eeg spatial projection and an improved 3d cnn with channel spatiotemporal joint attention mechanism for emotion recognition," *Signal, Image and Video Processing*, pp. 1–16, 2024.

[69] Y. Zhao, C. Dong, G. Zhang, *et al.*, "Eeg-based seizure detection using linear graph convolution network with focal loss," *Computer Methods and Programs in Biomedicine*, vol. 208, p. 106 277, 2021, ISSN: 0169-2607. DOI: https://doi.org/10.1016/j.cmpb.2021.106277. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169260721003515.

[70] S. S. Saboksayr, G. Mateos, and M. Cetin, "Eeg-based emotion classification using graph signal processing," in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 1065–1069. DOI: 10.1109/ICASSP39728.2021.9414342.

[71] T. Chen, L. Li, and X. Yuan, "A graph neural network with spatial attention for emotion analysis," *Cognitive Computation*, vol. 17, no. 1, p. 3, Nov. 21, 2024. DOI: 10.1007/s12559-024-10358-1. [Online]. Available: https://doi.org/10.1007/s12559-024-10358-1.

[72] S. Asadzadeh, T. Y. Rezaii, S. Beheshti, and S. Meshgini, "Accurate emotion recognition utilizing extracted eeg sources as graph neural network nodes," *Cognitive Computation*, vol. 15, no. 1, pp. 176–189, Jan. 1, 2023. DOI: 10.1007/s12559-022-10077-5. [Online]. Available: https://doi.org/10.1007/s12559-022-10077-5.

[73] A. Hajisafi, H. Lin, Y.-Y. Chiang, and C. Shahabi, "Dynamic gnns for precise seizure detection and classification from eeg data," in *Advances in Knowledge Discovery and Data Mining*, D.-N. Yang, X. Xie, V. S. Tseng, J. Pei, J.-W. Huang, and J. C.-W. Lin, Eds., Singapore: Springer Nature Singapore, 2024, pp. 207–220, ISBN: 978-981-97-2238-9.

[74] L. Tang and M. Zhao, "Epileptic seizure detection in neonatal eeg using a multi-band graph neural network model," *Applied Sciences*, vol. 14, no. 21, 2024, ISSN: 2076-3417. DOI: 10.3390/app14219712. [Online]. Available: https://www.mdpi.com/2076-3417/14/21/9712.

[75] S. Abadal, P. Galván, A. Mármol, *et al.*, "Graph neural networks for electroencephalogram analysis: Alzheimer's disease and epilepsy use cases," *Neural Networks*, vol. 181, p. 106 792, 2025, ISSN: 0893-6080. DOI: https://doi.org/10.1016/j.neunet.2024.106792. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0893608024007160.

[76] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: A Bradford Book, 2018, ISBN: 0262039249.

[77] A. D. Pelosi, N. Roth, T. Yehoshua, D. Itah, O. Braun Benyamin, and A. Dahan, "Personalized rehabilitation approach for reaching movement using reinforcement learning," *Scientific Reports*, vol. 14, no. 1, p. 17 675, Jul. 30, 2024. DOI: 10.1038/s41598-024-64514-6. [Online]. Available: https://doi.org/10.1038/s41598-024-64514-6.

[78] N. Sanghi, *Deep reinforcement learning with python*. Springer, 2021.

[79] S. Raschka, Y. H. Liu, and V. Mirjalili, *Machine Learning with PyTorch and Scikit-Learn: Develop machine learning and deep learning models with Python*. Packt Publishing Ltd, 2022.

[80] M. Orban, M. Elsamanty, K. Guo, S. Zhang, and H. Yang, "A review of brain activity and eeg-based brain–computer interfaces for rehabilitation application," *Bioengineering*, vol. 9, no. 12, 2022, ISSN: 2306-5354. DOI: 10.3390/bioengineering9120768. [Online]. Available: https://www.mdpi.com/2306-5354/9/12/768.

[81] A. Gonzalez-Santocildes, J.-I. Vazquez, and A. Eguiluz, "Enhancing robot behavior with eeg, reinforcement learning and beyond: A review of techniques in collaborative robotics," *Applied Sciences*, vol. 14, no. 14, 2024, ISSN: 2076-3417. DOI: 10.3390/app14146345. [Online]. Available: https://www.mdpi.com/2076-3417/14/14/6345.

[82] A. Xavier Fidêncio, C. Klaes, and I. Iossifidis, "Error-related potentials in reinforcement learning-based brain-machine interfaces," *Frontiers in Human Neuroscience*, vol. 16, 2022, ISSN: 1662-5161. DOI: 10.3389/fnhum.2022.806517. [Online]. Available: https://www.frontiersin.org/journals/human-neuroscience/articles/10.3389/fnhum.2022.806517.

[83] M. A. Lebedev and M. A. Nicolelis, "Brain–machine interfaces: Past, present and future," *TRENDS in Neurosciences*, vol. 29, no. 9, pp. 536–546, 2006.

[84] D. L. Schomer and F. L. Da Silva, *Niedermeyer's electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 2012.

[85] J. Hubbard, A. Kikumoto, and U. Mayr, "Eeg decoding reveals the strength and temporal dynamics of goal-relevant representations," *Scientific reports*, vol. 9, no. 1, p. 9051, 2019.

[86] D. J. McFarland, L. A. Miner, T. M. Vaughan, and J. R. Wolpaw, "Mu and beta rhythm topographies during motor imagery and actual movements," *Brain topography*, vol. 12, pp. 177–186, 2000.

[87] M. Jeannerod, "The representing brain: Neural correlates of motor intention and imagery," *Behavioral and Brain sciences*, vol. 17, no. 2, pp. 187–202, 1994.

[88] A. Biasiucci, R. Leeb, I. Iturrate, *et al.*, "Brain-actuated functional electrical stimulation elicits lasting arm motor recovery after stroke," *Nature communications*, vol. 9, no. 1, p. 2421, 2018.

[89] C. Farabet, C. Couprie, L. Najman, and Y. LeCun, "Learning hierarchical features for scene labeling," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 8, pp. 1915–1929, 2012.

[90] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[91] Y. LeCun, K. Kavukcuoglu, and C. Farabet, "Convolutional networks and applications in vision," in *Proceedings of 2010 IEEE international symposium on circuits and systems*, IEEE, 2010, pp. 253–256.

[92] Z. Zhang, P. Cui, and W. Zhu, "Deep learning on graphs: A survey," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 1, pp. 249–270, 2020.

[93] R. Zhang, Z. Wang, F. Yang, and Y. Liu, "Recognizing the level of organizational commitment based on deep learning methods and eeg," in *ITM Web of Conferences*, EDP Sciences, vol. 47, 2022, p. 02 044.

[94] M. Jia, W. Liu, J. Duan, *et al.*, "Efficient graph convolutional networks for seizure prediction using scalp eeg," *Frontiers in Neuroscience*, vol. 16, p. 967 116, 2022.

[95] N. Wagh and Y. Varatharajah, "Eeg-gcnn: Augmenting electroencephalogram-based neurological disease diagnosis using a domain-guided graph convolutional neural network," in *Machine Learning for Health*, PMLR, 2020, pp. 367–378.

[96] G. Bao, K. Yang, L. Tong, *et al.*, "Linking multi-layer dynamical gcn with style-based recalibration cnn for eeg-based emotion recognition," *Frontiers in Neurorobotics*, vol. 16, p. 834 952, 2022.

[97] W. Ma, C. Wang, X. Sun, X. Lin, and Y. Wang, "A double-branch graph convolutional network based on individual differences weakening for motor imagery eeg classification," *Biomedical Signal Processing and Control*, vol. 84, p. 104 684, 2023.

[98] L. Meng, J. Hu, Y. Deng, and Y. Hu, "Electrical status epilepticus during sleep electroencephalogram waveform identification and analysis based on a graph convolutional neural network," *Biomedical Signal Processing and Control*, vol. 77, p. 103 788, 2022.

[99] N. Khaleghi, T. Y. Rezaii, S. Beheshti, and S. Meshgini, "Developing an efficient functional connectivity-based geometric deep network for automatic eeg-based visual decoding," *Biomedical Signal Processing and Control*, vol. 80, p. 104 221, 2023.

[100] Y. Hou, S. Jia, X. Lun, *et al.*, "Gcns-net: A graph convolutional neural network approach for decoding time-resolved eeg motor imagery signals," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 6, pp. 7312–7323, 2024. DOI: 10.1109/TNNLS.2022.3202569.

[101] X. Lin, J. Chen, W. Ma, W. Tang, and Y. Wang, "Eeg emotion recognition using improved graph neural network with channel selection," *Computer Methods and Programs in Biomedicine*, vol. 231, p. 107 380, 2023, ISSN: 0169-2607. DOI: https://doi.org/10.1016/j.cmpb.2023.107380. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169260723000470.

[102] T. Song, W. Zheng, P. Song, and Z. Cui, "Eeg emotion recognition using dynamical graph convolutional neural networks," *IEEE Transactions on Affective Computing*, vol. 11, no. 3, pp. 532–541, 2018.

[103] T. Chen, Y. Sui, X. Chen, A. Zhang, and Z. Wang, "A unified lottery ticket hypothesis for graph neural networks," in *International conference on machine learning*, PMLR, 2021, pp. 1695–1706.

[104] S. Taran, V. Bajaj, D. Sharma, S. Siuly, and A. Sengur, "Features based on analytic imf for classifying motor imagery eeg signals in bci applications," *Measurement*, vol. 116, pp. 68–76, 2018, ISSN: 0263-2241. DOI: https://doi.org/10.1016/j.measurement.2017.10.067. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0263224117307029.

[105] T. Yu, J. Xiao, F. Wang, *et al.*, "Enhanced motor imagery training using a hybrid bci with feedback," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 7, pp. 1706–1717, 2015. DOI: 10.1109/TBME.2015.2402283.

[106] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Filter bank common spatial pattern (fbcsp) in brain-computer interface," in *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, 2008, pp. 2390–2397. DOI: 10.1109/IJCNN.2008.4634130.

[107] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, *et al.*, "Deep learning with convolutional neural networks for eeg decoding and visualization," *Human Brain Mapping*, vol. 38, no. 11, pp. 5391–5420, 2017. DOI: https://doi.org/10.1002/hbm.23730. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hbm.23730. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.23730.

[108] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "Eegnet: A compact convolutional neural network for eeg-based brain–computer interfaces," *Journal of Neural Engineering*, vol. 15, no. 5, p. 056 013, Jul. 2018. DOI: 10.1088/1741-2552/aace8c. [Online]. Available: https://dx.doi.org/10.1088/1741-2552/aace8c.

[109] D. Zhang, L. Yao, K. Chen, S. Wang, X. Chang, and Y. Liu, "Making sense of spatio-temporal preserving representations for eeg-based human intention recognition," *IEEE Transactions on Cybernetics*, vol. 50, no. 7, pp. 3033–3044, 2020. DOI: 10.1109/TCYB.2019.2905157.

[110] S. Sartipi and M. Cetin, "Subject-independent deep architecture for eeg-based motor imagery classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 32, pp. 718–727, 2024. DOI: 10.1109/TNSRE.2024.3360194.

[111] Z. Chen, R. Yang, M. Huang, F. Li, G. Lu, and Z. Wang, "Eegprogress: A fast and lightweight progressive convolution architecture for eeg classification," *Computers in Biology and Medicine*, vol. 169, p. 107 901, 2024, ISSN: 0010-4825. DOI: https://doi.org/10.1016/j.compbiomed.2023.107901. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0010482523013665.

[112] O. Ali, M. Saif-ur-Rehman, T. Glasmachers, I. Iossifidis, and C. Klaes, "Contranet: A hybrid network for improving the classification of eeg and emg signals with limited training data," *Computers in Biology and Medicine*, vol. 168, p. 107 649, 2024, ISSN: 0010-4825. DOI: https://doi.org/10.1016/j.compbiomed.2023.107649. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0010482523011149.

[113] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and S. Y. Philip, "A comprehensive survey on graph neural networks," *IEEE transactions on neural networks and learning systems*, vol. 32, no. 1, pp. 4–24, 2020.

[114] J. Bruna, W. Zaremba, A. Szlam, and Y. LeCun, "Spectral networks and locally connected networks on graphs," *arXiv preprint arXiv:1312.6203*, 2013.

[115] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," *Advances in neural information processing systems*, vol. 29, 2016.

[116] R. Levie, F. Monti, X. Bresson, and M. M. Bronstein, "Cayleynets: Graph convolutional neural networks with complex rational spectral filters," *IEEE Transactions on Signal Processing*, vol. 67, no. 1, pp. 97–109, 2018.

[117] W. Hamilton, Z. Ying, and J. Leskovec, "Inductive representation learning on large graphs," *Advances in neural information processing systems*, vol. 30, 2017.

[118] F. Monti, D. Boscaini, J. Masci, E. Rodola, J. Svoboda, and M. M. Bronstein, "Geometric deep learning on graphs and manifolds using mixture model cnns," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 5115–5124.

[119] M. Niepert, M. Ahmed, and K. Kutzkov, "Learning convolutional neural networks for graphs," in *International conference on machine learning*, PMLR, 2016, pp. 2014–2023.

[120] H. Gao, Z. Wang, and S. Ji, "Large-scale learnable graph convolutional networks," in *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, 2018, pp. 1416–1424.

[121] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE signal processing magazine*, vol. 30, no. 3, pp. 83–98, 2013.

[122] D. Zeng, K. Huang, C. Xu, H. Shen, and Z. Chen, "Hierarchy graph convolution network and tree classification for epileptic detection on electroencephalography signals," *IEEE transactions on cognitive and developmental systems*, vol. 13, no. 4, pp. 955–968, 2020.

[123] D. Zhang, L. Yao, K. Chen, S. Wang, P. D. Haghighi, and C. Sullivan, "A graph-based hierarchical attention model for movement intention detection from eeg signals," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 11, pp. 2247–2253, 2019. DOI: 10.1109/TNSRE.2019.2943362.

[124] B. Sun, Z. Liu, Z. Wu, C. Mu, and T. Li, "Graph convolution neural network based end-to-end channel selection and classification for motor imagery brain–computer interfaces," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 9, pp. 9314–9324, 2023. DOI: 10.1109/TII.2022.3227736.

[125] H. Liu, D. Yang, X. Liu, *et al.*, "Todynet: Temporal dynamic graph neural network for multivariate time series classification," *Information Sciences*, vol. 677, p. 120 914, 2024, ISSN: 0020-0255. DOI: https://doi.org/10.1016/j.ins.2024.120914. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0020025524008284.

[126] Z. L. Li, G. W. Zhang, J. Yu, and L. Y. Xu, "Dynamic graph structure learning for multivariate time series forecasting," *Pattern Recognition*, vol. 138, p. 109 423, 2023, ISSN: 0031-3203. DOI: https://doi.org/10.1016/j.patcog.2023.109423. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0031320323001243.

[127] A. L. Goldberger, L. A. Amaral, L. Glass, *et al.*, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals," *circulation*, vol. 101, no. 23, e215–e220, 2000.

[128] S. P. van den Broek, F. Reinders, M. Donderwinkel, and M. Peters, "Volume conduction effects in eeg and meg," *Electroencephalography and clinical neurophysiology*, vol. 106, no. 6, pp. 522–534, 1998.

[129] C. Tallon-Baudry and O. Bertrand, "Oscillatory gamma activity in humans and its role in object representation," *Trends in cognitive sciences*, vol. 3, no. 4, pp. 151–162, 1999.

[130] A. D. Wilkie, "A stochastic investment model for actuarial use," *Transactions of the Faculty of Actuaries*, vol. 39, pp. 341–403, 1984.

[131] W. Chen, B. Koo, Y. Wang, *et al.*, "Using a stochastic economic scenario generator to analyse uncertain superannuation and retirement outcomes," *Annals of Actuarial Science*, vol. 15, no. 3, pp. 549–566, 2021. DOI: 10.1017/S1748499520000305.

[132] A. Minney, Z. Zhu, Y. Guo, *et al.*, "Using the pension multiple to measure retirement outcomes," *Finance Research Letters*, vol. 49, p. 103 149, 2022.

[133] S. Sahin, A. Cairns, and T. Kleinow, "Revisiting the wilkie investment model," in *18th International AFIR Colloquium, Rome*, 2008, pp. 1–24.

[134] S. Scheidegger and I. Bilionis, "Machine learning for high-dimensional dynamic stochastic economies," *Journal of Computational Science*, vol. 33, pp. 68–82, 2019.

[135] S. Moshiri and N. Cameron, "Neural network versus econometric models in forecasting inflation," *Journal of forecasting*, vol. 19, no. 3, pp. 201–217, 2000.

[136] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE transactions on neural networks*, vol. 20, no. 1, pp. 61–80, 2008.

[137] D. Marshall, D. Coyle, S. Wilson, and M. Callaghan, "Games, gameplay, and bci: The state of the art," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 5, no. 2, pp. 82–99, 2013. DOI: 10.1109/TCIAIG.2013.2263555.

[138] M. Bamdad, H. Zarshenas, and M. A. Auais, "Application of bci systems in neurorehabilitation: A scoping review," *Disability and Rehabilitation: Assistive Technology*, vol. 10, no. 5, pp. 355–364, 2015, PMID: 25560222. DOI: 10.3109/17483107.2014.961569. eprint: https://doi.org/10.3109/17483107.2014.961569. [Online]. Available: https://doi.org/10.3109/17483107.2014.961569.

[139] M.-V. Drăgoi, I. Nisipeanu, A. Frimu, *et al.*, "Real-time home automation system using bci technology," *Biomimetics*, vol. 9, no. 10, 2024, ISSN: 2313-7673. DOI: 10.3390/biomimetics9100594. [Online]. Available: https://www.mdpi.com/2313-7673/9/10/594.

[140] A. Saibene, H. Ghaemi, and E. Dagdevir, "Deep learning in motor imagery eeg signal decoding: A systematic review," *Neurocomputing*, vol. 610, p. 128 577, 2024, ISSN: 0925-2312. DOI: https://doi.org/10.1016/j.neucom.2024.128577. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231224013481.

[141] O. Abdel-Hamid, A.-r. Mohamed, H. Jiang, L. Deng, G. Penn, and D. Yu, "Convolutional neural networks for speech recognition," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 10, pp. 1533–1545, 2014. DOI: 10.1109/TASLP.2014.2339736.

[142] W. Wang and J. Gang, "Application of convolutional neural network in natural language processing," in *2018 International Conference on Information Systems and Computer Aided Education (ICISCAE)*, 2018, pp. 64–70. DOI: 10.1109/ICISCAE.2018.8666928.

[143] M. Asfand-e-yar, Q. Hashir, A. A. Shah, H. A. M. Malik, A. Alourani, and W. Khalil, "Multimodal cnn-ddi: Using multimodal cnn for drug to drug interaction associated events," *Scientific Reports*, vol. 14, no. 1, p. 4076, Feb. 19, 2024. DOI: 10.1038/s41598-024-54409-x. [Online]. Available: https://doi.org/10.1038/s41598-024-54409-x.

[144] H. W. Aung, J. J. Li, B. Shi, Y. An, and S. W. Su, "Eeg_glt-net: Optimising eeg graphs for real-time motor imagery signals classification," *Biomedical Signal Processing and Control*, vol. 104, p. 107 458, 2025, ISSN: 1746-8094. DOI: https://doi.org/10.1016/j.bspc.2024.107458. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1746809424015167.

[145] M. Tangermann, K.-R. Müller, A. Aertsen, *et al.*, "Review of the bci competition iv," *Frontiers in Neuroscience*, vol. 6, 2012, ISSN: 1662-453X. DOI: 10.3389/fnins.2012.00055. [Online]. Available: https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2012.00055.

[146] S. Sakhavi, C. Guan, and S. Yan, "Learning temporal information for brain-computer interface using convolutional neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 11, pp. 5619–5629, 2018. DOI: 10.1109/TNNLS.2018.2789927.

[147] D. Zhao, F. Tang, B. Si, and X. Feng, "Learning joint space–time–frequency features for eeg decoding on small labeled data," *Neural Networks*, vol. 114, pp. 67–77, 2019, ISSN: 0893-6080. DOI: https://doi.org/10.1016/j.neunet.2019.02.009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0893608019300711.

[148] T. M. Ingolfsson, M. Hersche, X. Wang, N. Kobayashi, L. Cavigelli, and L. Benini, *Eeg-tcnet: An accurate temporal convolutional network for embedded motor-imagery brain-machine interfaces*, 2020. arXiv: 2006.00622 [eess.SP]. [Online]. Available: https://arxiv.org/abs/2006.00622.

[149] M.-a. Li and Z.-w. Ruan, "A novel decoding method for motor imagery tasks with 4d data representation and 3d convolutional neural networks," *Journal of Neural Engineering*, vol. 18, no. 4, p. 046 029, Apr. 2021. DOI: 10.1088/1741-2552/abf68b. [Online]. Available: https://dx.doi.org/10.1088/1741-2552/abf68b.

[150] Y. Song, Q. Zheng, B. Liu, and X. Gao, "Eeg conformer: Convolutional transformer for eeg decoding and visualization," eng, *IEEE transactions on neural systems and rehabilitation engineering*, vol. 31, pp. 710–719, 2023, ISSN: 1534-4320.

[151] X. Li, X. Tang, S. Qiu, X. Deng, H. Wang, and Y. Tian, "Subdomain adversarial network for motor imagery eeg classification using graph data," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 8, no. 1, pp. 327–336, 2024. DOI: 10.1109/TETCI.2023.3301385.

[152] A. Salami, J. Andreu-Perez, and H. Gillmeister, "Eeg-itnet: An explainable inception temporal convolutional network for motor imagery classification," *IEEE Access*, vol. 10, pp. 36 672–36 685, 2022. DOI: 10.1109/ACCESS.2022.3161489.

[153] C. Gao, W. Liu, and X. Yang, "Convolutional neural network and riemannian geometry hybrid approach for motor imagery classification," *Neurocomputing*, vol. 507, pp. 180–190, 2022, ISSN: 0925-2312. DOI: https://doi.org/10.1016/j.neucom.2022.08.024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231222009985.

[154] Z. Miao, M. Zhao, X. Zhang, and D. Ming, "Lmda-net:a lightweight multi-dimensional attention network for general eeg-based brain-computer interfaces and interpretability," *NeuroImage*, vol. 276, p. 120 209, 2023, ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2023.120209. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811923003609.

[155] J. Wang, L. Yao, and Y. Wang, "Ifnet: An interactive frequency convolutional neural network for enhancing motor imagery decoding from eeg," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 1900–1911, 2023. DOI: 10.1109/TNSRE.2023.3257319.

[156] J. Cao, L. Yang, P. G. Sarrigiannis, D. Blackburn, and Y. Zhao, "Dementia classification using a graph neural network on imaging of effective brain connectivity," *Computers in Biology and Medicine*, vol. 168, p. 107 701, 2024, ISSN: 0010-4825. DOI: https://doi.org/10.1016/j.compbiomed.2023.107701. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0010482523011666.

[157] Z. Wang, C. Hu, W. Liu, X. Zhou, and X. Zhao, "Eeg-based high-performance depression state recognition," *Frontiers in Neuroscience*, vol. 17, 2024, ISSN: 1662-453X. DOI: 10.3389/fnins.2023.1301214. [Online]. Available: https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2023.1301214.

[158] W. Cappelletti, Y. Xie, and P. Frossard, "LEARNING SELF-SUPERVISED DYNAMIC NETWORKS FOR SEIZURE ANALYSIS," in *ICLR 2024 Workshop on Learning from Time Series For Health*, 2024. [Online]. Available: https://openreview.net/forum?id=S523wIxRTe.

[159] Y. Ding, N. Robinson, C. Tong, Q. Zeng, and C. Guan, "Lggnet: Learning from local-global-graph representations for brain–computer interface," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 7, pp. 9773–9786, 2024. DOI: 10.1109/TNNLS.2023.3236635.

[160] M. A. Abbasi, H. F. Abbasi, M. Z. Aziz, W. Haider, Z. Fan, and X. Yu, "A novel precisely designed compact convolutional eeg classifier for motor imagery classification," *Signal, Image and Video Processing*, vol. 18, no. 4, pp. 3243–3254, Jun. 1, 2024. DOI: 10.1007/s11760-023-02986-1. [Online]. Available: https://doi.org/10.1007/s11760-023-02986-1.

[161] H. Li, H. Ji, J. Yu, *et al.*, "A sequential learning model with gnn for eeg-emg-based stroke rehabilitation bci," *Frontiers in Neuroscience*, vol. 17, 2023, ISSN: 1662-453X. DOI: 10.3389/fnins.2023.1125230. [Online]. Available: https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2023.1125230.

[162] J. Janisch, T. Pevný, and V. Lisý, *Classification with costly features using deep reinforcement learning*, 2018. arXiv: 1711.07364 [cs.AI]. [Online]. Available: https://arxiv.org/abs/1711.07364.

[163] B. Roe, *MiniBooNE particle identification*, UCI Machine Learning Repository, DOI: https://doi.org/10.24432/C5QC87, 2005.

[164] P. Cortez and A. Morais, *Forest Fires*, UCI Machine Learning Repository, DOI: https://doi.org/10.24 2007.

[165] A. Krizhevsky and G. Hinton, *Learning multiple layers of features from tiny images*, 2009.

[166] P. Cortez, A. Cerdeira, F. Almeida, T. Matos, and J. Reis, *Wine Quality*, UCI Machine Learning Repository, DOI: https://doi.org/10.24432/C56S3T, 2009.

[167] C. Song, C. Chen, Y. Li, and X. Wu, "Deep reinforcement learning apply in electromyography data classification," in *2018 IEEE International Conference on Cyborg and Bionic Systems (CBS)*, 2018, pp. 505–510. DOI: 10.1109/CBS.2018.8612213.

[168] D. Dua and C. Graff, *Uci machine learning repository*, http://archive.ics.uci.edu/ml, Accessed: 2024-11-08, 2017.

[169] C. Gao, W. Liu, and X. Yang, "Convolutional neural network and riemannian geometry hybrid approach for motor imagery classification," *Neurocomputing*, vol. 507, pp. 180–190, 2022, ISSN: 0925-2312. DOI: https://doi.org/10.1016/j.neucom.2022.08.024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231222009985.

[170] Y. Song, Q. Zheng, B. Liu, and X. Gao, "Eeg conformer: Convolutional transformer for eeg decoding and visualization," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 710–719, 2023. DOI: 10.1109/TNSRE.2022.3230250.

[171] Z. Miao, M. Zhao, X. Zhang, and D. Ming, "Lmda-net:a lightweight multi-dimensional attention network for general eeg-based brain-computer interfaces and interpretability," *NeuroImage*, vol. 276, p. 120 209, 2023, ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2023.120209. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811923003609.

[172] J. Wang, L. Yao, and Y. Wang, "Ifnet: An interactive frequency convolutional neural network for enhancing motor imagery decoding from eeg," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 1900–1911, 2023. DOI: 10.1109/TNSRE.2023.3257319.

[173] A. Salami, J. Andreu-Perez, and H. Gillmeister, "Eeg-itnet: An explainable inception temporal convolutional network for motor imagery classification," *IEEE Access*, vol. 10, pp. 36 672–36 685, 2022. DOI: 10.1109/ACCESS.2022.3161489.