



Lightweight Machine Learning with Brain Signals

Author:

Lou, Haowei

Publication Date:

2023

DOI:

<https://doi.org/10.26190/unsworks/24944>

License:

<https://creativecommons.org/licenses/by/4.0/>

Link to license to see what you are allowed to do with this resource.

Downloaded from <http://hdl.handle.net/1959.4/101237> in <https://unsworks.unsw.edu.au> on 2024-04-27

ORIGINALITY STATEMENT

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.

COPYRIGHT STATEMENT

I hereby grant the University of New South Wales or its agents a non-exclusive licence to archive and to make available (including to members of the public) my thesis or dissertation in whole or part in the University libraries in all forms of media, now or here after known. I acknowledge that I retain all intellectual property rights which subsist in my thesis or dissertation, such as copyright and patent rights, subject to applicable law. I also retain the right to use all or part of my thesis or dissertation in future works (such as articles or books).

For any substantial portions of copyright material used in this thesis, written permission for use has been obtained, or the copyright material is removed from the final public version of the thesis.

AUTHENTICITY STATEMENT

I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis.

UNSW is supportive of candidates publishing their research results during their candidature as detailed in the UNSW Thesis Examination Procedure.

Publications can be used in the candidate's thesis in lieu of a Chapter provided:

- The candidate contributed **greater than 50%** of the content in the publication and are the "primary author", i.e. they were responsible primarily for the planning, execution and preparation of the work for publication.
- The candidate has obtained approval to include the publication in their thesis in lieu of a Chapter from their Supervisor and Postgraduate Coordinator.
- The publication is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in the thesis.

The candidate has declared that **their thesis has publications - either published or submitted for publication - incorporated into it in lieu of a Chapter/s. Details of these publications are provided below.**

Publication Details #1

Full Title:	Less is More: Brain Functional Connectivity Empowered Generalisable Intention Classification with Task-relevant Channel Selection
Authors:	Haowei Lou., Zesheng Ye., Lina Yao. and Yu Zhang.
Journal or Book Name:	Transactions on Neural Systems and Rehabilitation Engineering
Volume/Page Numbers:	1888 - 1899
Date Accepted/Published:	06 March 2023
Status:	published
The Candidate's Contribution to the Work:	Propose Ideas, write code, conduct experiment, collect result, and write it on paper
Location of the work in the thesis and/or how the work is incorporated in the thesis:	Chapter 3

Candidate's Declaration



I confirm that where I have used a publication in lieu of a chapter, the listed publication(s) above meet(s) the requirements to be included in the thesis. I also declare that I have complied with the Thesis Examination Procedure.

Lightweight Machine Learning with Brain Signals

Haowei Lou

A thesis in fulfillment of the requirements for the degree of
Master of Philosophy



School of Computer Science and Engineering
Faculty of Engineering
The University of New South Wales

February 2023

Copyright Statement

‘I hereby grant the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all proprietary rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstract International (this is applicable to doctoral theses only).

I have either used no substantial portions of copyright material in my thesis or I have obtained permission to use copyright material; where permission has not been granted I have applied/will apply for a partial restriction of the digital copy of my thesis or dissertation.’

Haowei Lou
July 10, 2023

Authenticity Statement

‘I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis. No emendation of content has occurred and if there are any minor variations in formatting, they are the result of the conversion to digital format.’

Haowei Lou
July 10, 2023

Abstract

Electroencephalography (EEG) signals are gaining popularity in Brain-Computer Interface (BCI) systems and neural engineering applications thanks to their portability and availability. Inevitably, the sensory electrodes on the entire scalp would collect signals irrelevant to the particular BCI task, increasing the risks of overfitting in machine learning-based predictions.

While this issue is being addressed by scaling up the EEG datasets and handcrafting the complex predictive models, this also leads to increased computation costs. Moreover, the model trained for one set of subjects cannot easily be adapted to other sets due to inter-subject variability, which creates even higher over-fitting risks. Meanwhile, despite previous studies using either convolutional neural networks (CNNs) or graph neural networks (GNNs) to determine spatial correlations between brain regions, they fail to capture brain functional connectivity beyond physical proximity.

To this end, we propose 1) removing task-irrelevant noises instead of merely complicating models; 2) extracting subject-invariant discriminative EEG encodings, by taking functional connectivity into account; 3) navigating and training deep learning model with the most critical EEG channels; 4) detecting most similar EEG segments with target subject to reduce the cost of computation as well as inter-subject variability.

Specifically, we construct a task-adaptive graph representation of brain network based on topological functional connectivity rather than distance-based connections. Further, non-contributory EEG channels are excluded by selecting only functional regions relevant to the corresponding intention. Lastly, contributory EEG segments are detected by several similarity estimation metrics, we then evaluate and train our proposed framework upon detected EEG segments to compare the performance of different metrics in EEG BCI tasks.

We empirically show that our proposed approach, SIFT-EEG, outperforms state-of-the-art, with around 4% and 7% improvements over CNN-based and GNN-based models, on performing motor imagery predictions. Also, the task-adaptive channel selection demonstrates similar predictive performance with only 20% of raw EEG data. Moreover, the best-performed metric can achieve a high level of accuracy with less than 9% training data, suggesting a possible shift in direction for future works other than simply scaling up the model.

Acknowledgements

I wish to express my gratitude to all the people who have helped me during my studies. Thanks to their help, I have been able to overcome difficulties and complete my research project and this paper.

First of all, I would like to thank my primary academic supervisor, A/Prof. Lina Yao, for her support and guidance throughout my MPhil study. We first met in December 2019 during The CHALLENG Program and she has been extremely supportive, providing me with every resources I need and spending enormous time guiding me through my research. Thank you, A/Prof. Yao.

In addition, I would also like to thank Zesheng Ye, PhD candidate, for often discussing academic problems with me and guiding me to solve encountered difficulties. Thank you, Zesheng, for your valuable assistance in teaching me how to conduct the research and proofreading my paper.

Furthermore, the senior students in our group were essential in helping me throughout my journey. Thank you, Xiacong Chen and Dr. Chaoran Huang for all kinds of help to me in these years

Also, I would like to thank my partner, Yongyi Wu, for the love, care, and support she offers me during my tough times. As well as my kittens, September and October, for the happiness and loveliness they give me.

Lastly, I would like to conclude by expressing my gratitude to my parents, Fanggang Lou and Jie Du, for their love, support, and financial assistance. They have given me everything I need since my childhood, letting me freely choose the path I want to go and supporting me unconditionally.

Without them, I would not have been able to complete my study and get through this journey. I would like to express my sincere gratitude for your generous help in completing my studies. I wish all the best in your research, career and family!

Abbreviations

ACC Classification Accuracy

AGG Aggregation Function

APs Extracellular Action Potentials

AUC Area Under ROC Curve

BCI Brain Computer Interface

CNN Convolutional Neural Network

CV Computer Vision

DL Deep Learning

EEG Electroencephalogram

ELU Exponential Linear Unit activation function

ERP Event Related Potential

GCN Graph Convolutional Network

GIN Graph Isomorphism Network

GNN Graph Neural Network

LFPs Local Field Potentials

MEG Magnetoencephalography

MI Motor Imagery

ML Machine Learning

MLP Multilayer Perceptron

NLP Natural Language Processing

NN Neural Network

RECSYS Recommender System

RNN Recurrent Neural Network

SOTA State Of The Art

VEP Visually Evoked Potential

WL Weisfeiler Lehman Isomorphism

fMRI functional Magnetic Resonance Imaging

fNIRS functional Near Infrared Spectroscopy

Contents

1	Introduction	1
1.1	Background	2
1.1.1	Motor Imagery EEG Signals	2
1.1.2	Open Challenges	3
1.1.3	NN-based MI Recognition	3
1.1.4	Subject-Independent MI Recognition	4
1.1.5	Topological Pattern of Brain Activity	4
1.1.6	Subject-Specific Training Data Selection	6
1.1.7	Lightweight Machine Learning	6
1.2	Motivations	7
1.2.1	Contribution	8
1.3	Dissertation Organization	9
2	Literature Review	12
2.1	BCI Overview	12
2.2	Brain Signal Collection	13
2.2.1	Invasive Signal Acquisition	13
2.2.2	Non-invasive Signal Acquisition	14

2.3	Motor Imagery Intention Recognition	16
2.4	Functional Connectivity	17
2.5	State-of-the-Art NN-Based Approaches in MI-EEG Signal	18
2.5.1	CNN-based SOTA	19
2.5.2	RNN-based SOTA	20
2.5.3	Attention-based SOTA	21
2.5.4	Graph-based SOTA	21
2.5.5	Limitations	23
3	Functional Connectivity Empowered Intention Classification	25
3.1	Introduction	25
3.2	Preliminaries	27
3.2.1	Graph representation for EEG signals	27
3.2.2	Graph Neural Networks	28
3.3	Methodology	30
3.3.1	Overview	30
3.3.2	Functional Adjacency Matrix	31
3.3.3	Task-Adaptive Channel Selection	31
3.3.4	Temporal Embedding	34
3.3.5	Topological Embedding	34
3.3.6	Intention Classification	37
3.4	Empirical Studies	38
3.4.1	Experiment Setting	38
3.4.2	Result & Discussion	40

3.5	Conclusion	50
4	Utilization Towards Lightweight Machine Learning	56
4.1	Introduction	56
4.2	Preliminaries	59
4.2.1	Graph simulation for Brain network	59
4.2.2	1D & 2D CNN Embedding Block	60
4.2.3	Graph Neural Network	62
4.2.4	Generalization in Machine Learning	63
4.2.5	Similarity	65
4.3	Methodology	66
4.3.1	Overview	66
4.3.2	Functional Adjacency Matrix	67
4.3.3	Navigate similar EEG-segments	68
4.3.4	Temporal Embedding	71
4.3.5	Topological Embedding	72
4.3.6	Intention Classification	75
4.4	Experiment setting	76
4.4.1	Dataset	76
4.4.2	Implementation Detail	77
4.5	Result & Discussion	78
4.5.1	Overall Performance	79
4.5.2	Impact of 1D-CNN embedding block	81
4.5.3	Fewer trials - Data efficiency	83

4.5.4	Computationally & Storage efficiency	84
4.6	Conclusion	85
5	Conclusion	86
5.1	Conclusion	86
5.2	Future work	87
	Bibliography	89

List of Figures

2.1	General workflow of a BCI system	13
2.2	Non-invasive Brain Signal Acquisition devices with correspond readings	16
3.1	SIFT-EEG Architecture	27
3.2	Task-relevant channels and channel-wise connectivity	35
3.3	Hidden Feature Visualization	37
3.4	The brain topology maps acquired with various importance metrics under different ratios of subjects sampled from 105 subjects.	41
3.5	Comparison for training loss change when the number of training epochs increases.	41
3.6	Electrodes position in international 10-10 EEG signal acquisition device	43
3.7	Classification accuracy for top $t\%$ task-relevant channels selected by the proposed algorithm.	53
3.8	Relative elapsed train and inference time for SIFT-EEG under different channel selection ratios.	54
3.9	Comparison of computation efficiency for total time taken for processing arbitrarily sampled 400 EEG segments.	54
3.10	Impact of varying the number of training subjects	55
4.1	Comparison of feature extraction between 2D & 1D CNN layers . . .	62
4.2	1D-SIFT-EEG architecture	64

4.3	The visualization of hidden features with t-SNE. Green dots refer to left-hand imaginary motion, while orange dots indicate right-hand imaginary motion.	78
4.4	Comparison for training loss change when the number of training epochs increases.	79
4.5	Computational cost comparison for different methods	81
4.6	Storage cost comparison for different methods	83

List of Tables

2.1	Summary of non-invasive brain signals' characteristics	15
3.1	Overall performances for subject-independent evaluation. All results are obtained over 12 runs with mean \pm std.	52
3.2	Comparison of the classification accuracy for task-adaptive and random selection of 20% of original channels.	52
3.3	The statistically significant tests. We find that 20 out of 24 comparisons are significant ($p \leq 0.05$), with results shown in bold	52
3.4	Comparison of model performance with different channel selection strategies applied to SIFT-EEG	53
4.1	Architecture of the temporal embedding block, 1DCNN(\cdot)	67
4.2	Overall performances for subject-independent evaluation using EEG-MMIDB dataset. All results are obtained over 12 runs with mean \pm std.	70
4.3	Overall performances for subject-independent evaluation using BCI-COMPIV2A dataset. All results are obtained using Leave-One-Out Cross-Validation over 9 subjects with mean \pm std.	72
4.4	The statistically significant tests. We find that 29 out of 32 comparisons are significant ($p \leq 0.10$), with results shown in bold	74

4.5	Leave-One-Out Cross-Validation (LOOCV) evaluation to the performance of the BCICOMP IV2A dataset using nine subjects. The results of the evaluation are presented in the form of accuracy scores for each subject, which were calculated when that subject was in the testing set.	75
4.6	Comparison of performance when model is trained using the most similar EEG segments, as determined by a similarity metric.	75
4.7	Storage cost comparison for different methods	82

List of Algorithms

1	Task-relevant channel selection	33
2	Task-relevant edge formation algorithm	51

List of Publications

This dissertation contains parts of content of the following publications:

- [1] H. Lou, Z. Ye, L. Yao, and Y. Zhang, “Less is more: Brain functional connectivity empowered generalizable intention classification with task-relevant channel selection,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 1888–1899, 2023.

Chapter 1

Introduction

Electroencephalography (EEG)-based brain-computer inference (BCI) systems have enabled a variety of neurological tasks, such as motion intention recognition [1], emotion analysis [2] and brain disease detection [3]. The EEG-based BCI uses non-invasive scalp electrodes to record and further analyze electrical fluctuations that occur as a result of brain activity. A task that is of interest to this study is recognising Motor Imagery (MI), a cognitive process in which subjects imagine moving different parts of their bodies. There have been years of research into algorithms for detecting a particular MI, as well as implications for individuals with disabilities in a range of applications, including brain typing [4], mind-controlled wheelchairs [5] and prosthetic arm [6].

1. Introduction

1.1 Background

1.1.1 Motor Imagery EEG Signals

Motor Imagery is a cognitive process in which a subject imagines that he/she performs a movement without actually performing the movement [7]. Studies from neurological research indicate brain regions that have been engaged in actual movements will also be active during imagery movements [8, 9]. This lead to brain signals produced by the primary-cortex region being consistent regardless of whether the subject is executing or imaging motions.

Due to its wide potential field including human-computer interaction, and neurorehabilitation. Extensive research has been conducted in building BCI applications with MI-based EEG signals. Zhang, et.al. [4] propose a brain typing system that incorporates deep-learning modules to interpret raw signals and uses the output to type words in a virtual keyboard. Specifically, this work designed a joint convolutional-recurrent neural network that simultaneously learns high-level features from raw MI-EEG signals. Followed by an Autoencoder to eliminate noises from artifacts. Imran, et.al. [5] developed a mind-control wheelchair system to improve elder life quality. The system contains an EEG headset to capture brain signals. Followed by Arduino, a single-board microcontroller, to interpret acquired brain signals and control wheelchairs. Taha, et al. [6]'s prosthetic arm project fits 3D-printed Prosthetic limbs with users, the EEG system will record the user's brain waves when he is thinking of a certain action or implementing a facial expression. Then, converted to commands to control the arm.

1. Introduction

1.1.2 Open Challenges

The open challenge for implementing MI-based BCI systems comes from:

1. Human-generated thoughts are non-stationary, and generated signals are non-linear. How to design a system to find deeper insights from the human brain signal is challenging [10].
2. Raw EEG signals have high noise-to-signal ratio. It contains too much-unwanted information such as noise, artifacts, or interference [11].
3. The extensive training time required for model development and the high cost of collecting EEG datasets significantly increase the cost of developing a BCI system [11].

1.1.3 NN-based MI Recognition

Recent MI recognition research takes advantages of deep neural networks (NN) to extract discriminative representations from enormous amounts of data without the need for carefully-designed features. For instance, convolutional neural networks (CNNs) are widely adopted to extract spatial correlations between different sensory channels, whilst recurrent neural networks (RNNs) are promising for capturing temporal dependencies from raw EEG signals. A single CNN architecture consisting of two convolution blocks, dubbed EEGNet, was proposed by [12] for EEG classification across different paradigms. Zhang et al. [4] design a hybrid model using CNN and RNN to extract spatial and temporal features from EEG signals, paired with an Auto-encoder to eliminate artifacts. Zhang et al. [13] additionally transform

1. Introduction

raw EEG signals into a matrix-like form to explore correlations between physically adjacent sensory channels.

1.1.4 Subject-Independent MI Recognition

While achieving success in subject-dependent settings, they still suffer from subject-independent evaluation, in which training and test data are collected from different subjects. In fact, EEG signals may manifest different patterns even when subjects are performing the same cognitive task. It is necessary either to adapt to new subjects or to model subject-invariant features to address this issue. Fahimi et al. [14] combine general mental states with new subject’s data for personalized modeling. Chen et al. [15] reduce discrepancy between two subjects with adversarial training. However, each new subject must be adapted once, and negative transfer may also occur in view of noisy EEG signals [16]. Alternatively, another line of works seeks to identify patterns that are universal across many subjects. [17] incorporate self-attention [18] into a convolutional-recurrent model to explore concentrated temporal periods.

1.1.5 Topological Pattern of Brain Activity

Despite being dominant in learning spatial representations, CNNs are structured on dense and regular “grid”-like inputs, which limits their ability to identify non-Euclidean relationships. It is, however, true that EEG channels are non-Euclidean by nature, as are brain regions. Additionally, even with the same acquisition device, signals do not necessarily originate from the same location due to the variance between subjects. Zhang et al. [19, 20] represent EEG nodes as a graph to learn a

1. Introduction

topological-based positioning relationship, which appears to be less subject dependent than conventional representations. It brings to light recent interest on using graph neural networks (GNNs) as an alternative to CNNs obtain non-Euclidean representations from raw EEG input signals [21]. There are concerns raised by existing graph-based EEG representations despite the topological features showing robustness to inter-subject variability.

channel redundancy Common EEG acquisition devices measure brain activity from the entire scalp. A particular cognitive task will, however, not require the activation of all brain regions [22], and different tasks may even activate different regions [23]. It follows that redundancy would be included if all channels are used indiscriminately.

edge formation confusion still exists regarding how to properly depict inter-regional brain connections as graph edges. The graph representation of certain non-euclidean structures, such as brain networks, is shown to be more effective than using CNNs [21]. Still, the extensively studied distance-based edge formation [19, 20] does not account for dynamic functional connectivity between different MI intentions.

graph noise EEG recordings are bound to contain noise from both external and internal factors, implying raw EEG signals are likely to result in poor model generalization and over-fitting risks given GNNs' sensitivity to the quality of the input graph [24]. This makes it even more challenging to identify task-relevant patterns while invariant to subjects, as EEG datasets are practically impossible to scale up as much as other domains.

1. Introduction

1.1.6 Subject-Specific Training Data Selection

Increasing the size of training dataset brings plenty of benefits, include performance improvement, reduce over-fitting, and improve the robustness of NN-model. However, the performance growth that is brought by size of dataset decreases as the volume of data increases. Recent study [25] shows the relationship between performance of NN-model in CV tasks and the amount of training data is in logarithmic, instead of linear relationship. Resources needed to collect EEG signals are higher than other ML-tasks, as it requires dedicated signal acquisition devices and plenty of time for training participant and collecting the signal. Engineering practice, seeking performance improvement by scale up the size of training data, with millions of training data is not applicable in EEG-BCI tasks. Because the size of EEG dataset is restricted by the limited number of public-available training dataset and cost of collection. Given that the improvement brought by increasing the amount of data is no longer significant, and the cost for collecting new EEG data is expensive. It is worthwhile to try an engineering practice, to train a subject-specific model by identifying and training NN-model with training EEG segments that is similar with targeting subjects. We assume the data with high difference is full with noises from inter-subject variability. Excluding these data can not only makes NN-models learn more subject-specific features, but also reduce the amount of training time, which make the overall framework more cost-effective.

1.1.7 Lightweight Machine Learning

The portability of EEG systems, which is a key factor in real-world BCI applications, is contradicted by complexity and expense. This creates barriers for end-users

1. *Introduction*

who wish to use BCI applications in their everyday life. Consequently, some well-performed massive NN-models are not suitable for real BCI applications due to their resource-intensive nature.

Lightening the NN-model is a fundamental practice for many BCI applications, as it can significantly reduce the cost of the system. This allows the applications to be deployed on edge devices of the user, including cell phones, wheelchairs, and laptops.

In addition, lightweight neural network models are typically faster to train and can be deployed in a cost-effective and timely manner, thereby reducing the overall cost of the system and decreasing the barrier to using the application.

1.2 Motivations

In order to solve previously addressed limitations and open challenges in this area of research, we wish to develop a cost-effective brain computer inference (BCI) system that can let humans interact with the computer using their brain signals in a more lightweight and accurate manner. Firstly, the proposed framework should be reusable, generalized, and accurate, with good performance in subject-independent experimental settings, which indicates higher generalizability and lower calibration costs. Secondly, we need to find a way to navigate and eliminate unwanted sources of noise in the extracted features, thereby improving the quality of classification. Thirdly, we aim to minimize the size and inference cost as much as possible to make it easier to deploy on edge devices. Finally, an effective training methodology is required to reduce the amount of data needed in the training stage, thus reducing the cost of collecting data and training the system.

1. Introduction

1.2.1 Contribution

Targeting each of these concerns, our study presents a **Subject-Independent** MI classification model using brain **F**unctional **C**onnect**I**ty (SIFT-EEG). For *redundant channel*, functional connectivity statistically contributes to identifying the brain regions involved when performing a certain cognitive task. For *edge definition*, graph adjacency is dynamically determined by functional connectivity, which filters out weak associations and identifies task-relevant active brain regions based on their importance in the functional connectivity graph. This is followed by performing self-attentive temporal convolution to extract discriminative task-adaptive temporal embeddings that mitigates *graph noise* contained in raw EEG signals. For *brain topological pattern extraction*, we derive robust topological embeddings from the task-adaptive temporal graphs of different subjects using a Graph Isomorphic Network (GIN) that reliably detects equivalent graphs [26]. For *training data selection*, we investigate the effectiveness of four similarity metrics in identifying similar training EEG segments, which could lead to a possible engineering practice for practical EEG-based BCI applications. Finally, for *lightweight machine learning*, we try our best to balance the performance and cost of our proposed framework. In order to reduce the computational time and size of proposed NN-model, and make it more lightweight.

- We propose a subject-independent MI prediction model built upon functional topological adjacency, which further takes into account self-attentive temporal convolution and graph isomorphism, thus capturing task-adaptive but subject-invariant EEG embeddings.
- We present a data-driven channel selection algorithm based on active brain regions, which can exclude non-contributory channels and reduce the impact

1. Introduction

of task-irrelevant noises.

- We evaluate SIFT-EEG on a large-scale EEG-based MI dataset, demonstrating its effectiveness for subject-independent classification, with competitive performances against state-of-the-arts even using 20% of raw data.
- We propose an improved version of the SIFT-EEG framework, the 1DSIFT-EEG framework, which further improves the performance in subject-independent MI classification. The framework is more improved when there are small groups of training subjects.
- We empirically compare the effectiveness of four similarity estimation metrics, which can be used to navigate EEG signals with high similarity and makes models remain well-performance even with fewer training data.
- We utilize the computational and storage cost of our proposed deep-learning framework, 1D-SIFTEEG with SIFT-EEG framework. Our result shows the improved version is more lightweight and cost-effective, making it easier to deploy on edge devices.

1.3 Dissertation Organization

This chapter briefly introduces current research gaps and the motivation for this research. The following chapters are organised as follows:

Chapter 2 Literature Review Research into Brain-Computer Interfaces (BCI) began in the 1970s. The emergence of Neural Networks (NN) and Deep Learning (DL) techniques has enabled computers to learn and analyze various data types,

1. Introduction

such as Computer Vision (CV), Natural Language Processing (NLP) and Recommender Systems (RecSYS). In recent years, numerous powerful Machine Learning (ML)-based BCI models have been developed. However, most of these models have millions of parameters, which necessitate high computational costs for classification. In Chapter 2, we provide a summary and comparison of existing approaches for BCI systems, along with their advantages and limitations, followed by unresolved issues and research gaps.

Chapter 3 Functional Connectivity Empowered Intention Classification

Electroencephalography (EEG) signals are gaining popularity in Brain-Computer Interface (BCI) thanks to their portability and availability. Inevitably, the sensory electrodes on the entire scalp would collect signals irrelevant to the particular BCI task, increasing the risks of overfitting in machine learning-based predictions. While this issue is being addressed by scaling up the EEG datasets and handcrafting the complex predictive models, which also leads to increased computation costs. Moreover, the model trained for one set of subjects cannot easily be adapted to other sets due to inter-subject variability, which creates even higher over-fitting risks. Meanwhile, despite previous studies using either convolutional neural networks (CNNs) or graph neural networks (GNNs) to determine spatial correlations between brain regions, they fail to capture brain functional connectivity beyond physical proximity. In Chapter 3, we propose 1) removing task-irrelevant noises instead of merely complicating models; 2) extracting subject-invariant discriminative EEG encodings, by taking functional connectivity into account. Specifically, we construct a task-adaptive graph representation of brain network based on topological functional connectivity rather than distance-based connections. Further, non-contributory EEG channels are excluded by selecting only functional regions relevant to the corre-

1. Introduction

sponding intention. We empirically show that the proposed approach outperforms state-of-the-art, with around 1% and 4% improvements over CNN-based and GNN-based models, on performing MI predictions.

Chapter 4 Utilization Towards Lightweight Machine Learning Our recently proposed framework SIFTEEG, discussed in Chapter 3, has been found to be computationally and storage expensive. Furthermore, our results demonstrate that SIFT-EEG does not perform well on datasets with a small number of subjects. In Chapter 4, we proposed an improved version of SIFTEEG, 1D-SIFTEEG, which has lower complexity and performs better on datasets with fewer subjects. We also investigated different similarity estimation metrics and proposed an engineering practice to reduce the cost of the framework by selecting similar training data to the target subject. Our results show that the overall framework achieves a satisfactory level of classification accuracy even when using less than 10% of the training data compared with traditional approaches.

Chapter 5 Conclusion We summarize and highlight the content of this dissertation in Chapter 5, and discuss the limitation and future direction of research.

Chapter 2

Literature Review

2.1 BCI Overview

A Brain-Computer Interface (BCI) system is composed of several components, including brain signal acquisition, signal analysis and processing, and control of external devices. Figure 2.1 illustrates the general workflow of a BCI system. Initially, signal acquisition devices record brain signals from the scalp, which are then sent to a computer for further analysis. The computer then pre-processes the raw signal to reduce noise caused by environmental and artifact effects (e.g., environmental noise, eye blink, and fatigue), followed by a feature extraction process to extract discriminative features from the brain signals. The extracted features are then passed to a classifier to distinguish types of brain activity. Finally, the output of the classifier is transformed into machine-recognizable commands, which are used to control external devices and provide feedback to the brain in the real world.

2. Literature Review

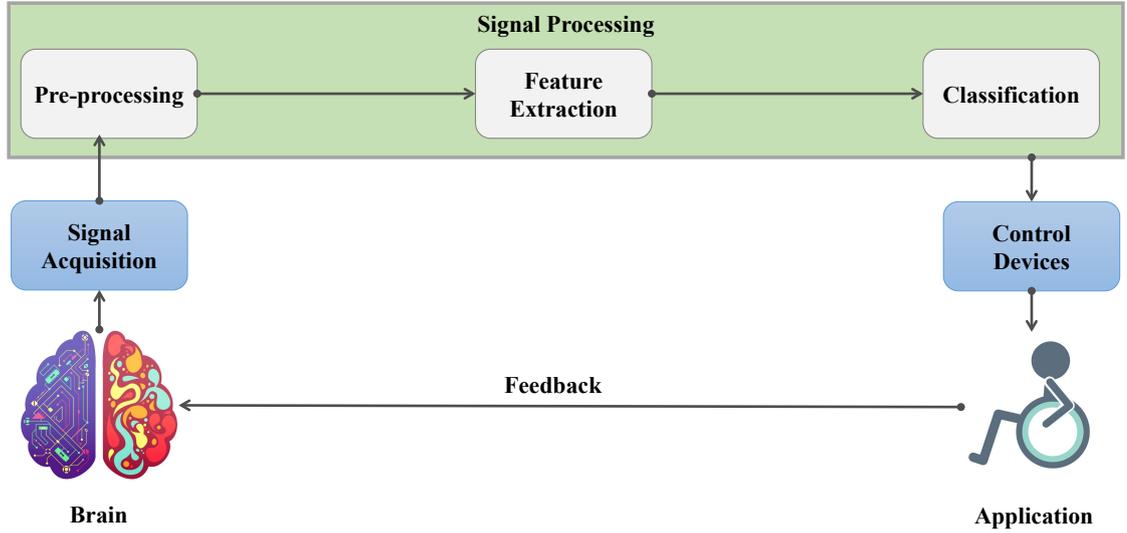


Figure 2.1: General workflow of a BCI system

2.2 Brain Signal Collection

Signal acquisition devices for Brain-Computer Interface (BCI) systems can be classified as either invasive or non-invasive, depending on the location of the signal detector. Invasive devices offer the benefit of collecting high-quality, low-signal-to-noise ratio signals, whereas non-invasive devices are more competitive in terms of portability and low clinical risk.

2.2.1 Invasive Signal Acquisition

The current techniques enable the interface of electric neuronal activity in vivo, ranging from intracellular potentials, extracellular action potentials (APs) to local field potentials (LFPs), through the use of invasive brain signal acquisition devices implanted into the brain to measure brain activity [27].

2. Literature Review

Implanted detectors are situated close to the source signal, enabling them to directly collect signals from the cortex of the brain [28]. This reduces the influence of environmental noise on the collected signals, thus potentially increasing the information transfer rate [29]. Consequently, the collected signals are more accurate and generally have superior performance compared to non-invasive signal acquisition approaches.

The low user acceptance of invasive signal acquisition devices, due to safety concerns during the process of neurosurgery and the implant, which is much lower than that of non-invasive devices, limits their application [30].

2.2.2 Non-invasive Signal Acquisition

Non-invasive signal acquisition devices place a detector near the scalp without contact with the brain itself. Various methods are employed to capture brain signals, including the measurement of electrical activity, magnetic fields from the scalp, and alterations in cerebral blood flow via functional magnetic resonance imaging (fMRI).

Brain signals are subject to interference from a variety of sources in the course of their travel through the skull, scalp, and hair to be detected [28]. Researchers have found that electric fields produced by neurons decay exponentially with increasing distance, resulting in the collected signals from non-invasive devices containing a large amount of redundant and noisy information, rendering them less accurate than those collected by implanted detectors.

Common non-invasive signal acquisition techniques utilized in the fields of rehabilitation and medical care include:

2. Literature Review

- Electroencephalography (EEG) involves the placement of small metallic detectors (electrodes) close to the scalp in order to monitor brain state by recording brain electrical activity [31].
- Functional Magnetic Resonance Imaging (fMRI) is a neuroimaging technique that produces detailed imaging of brain state through measurements of changes in brain blood-volume and/or oxygen level using magnetic fields [32].
- Functional Near-Infrared Spectroscopy (fNIRS) is a neuroimaging technique that can be used to produce brain state imaging. It utilizes near-infrared light to measure changes in oxygenation of the blood in the brain [33].
- Magnetoencephalography (MEG) is a technique that seeks to measure the magnetic fields generated by the brain in order to detect and record its activity [34].

Figure 2.2 and Table 2.1 in [35] provide visual and tabular representations, respectively, of the appearance and characteristics of various non-invasive brain signal acquisition devices.

Signals	EEG	fMRI	fNIRS	MEG
Noise-to-Signal Ratio	<i>low</i>	<i>middle</i>	<i>low</i>	<i>low</i>
Portability	<i>high</i>	<i>low</i>	<i>high</i>	<i>low</i>
Cost	<i>low</i>	<i>high</i>	<i>low</i>	<i>high</i>

Table 2.1: Summary of non-invasive brain signals' characteristics

This study primarily focuses on developing lightweight machine learning to process the most cost-effective and widely-accepted EEG signals, taking into account the portability and low cost nature of EEG based signal, as well as the recent release of consumer-level EEG systems from [36, 37, 38].

2. Literature Review

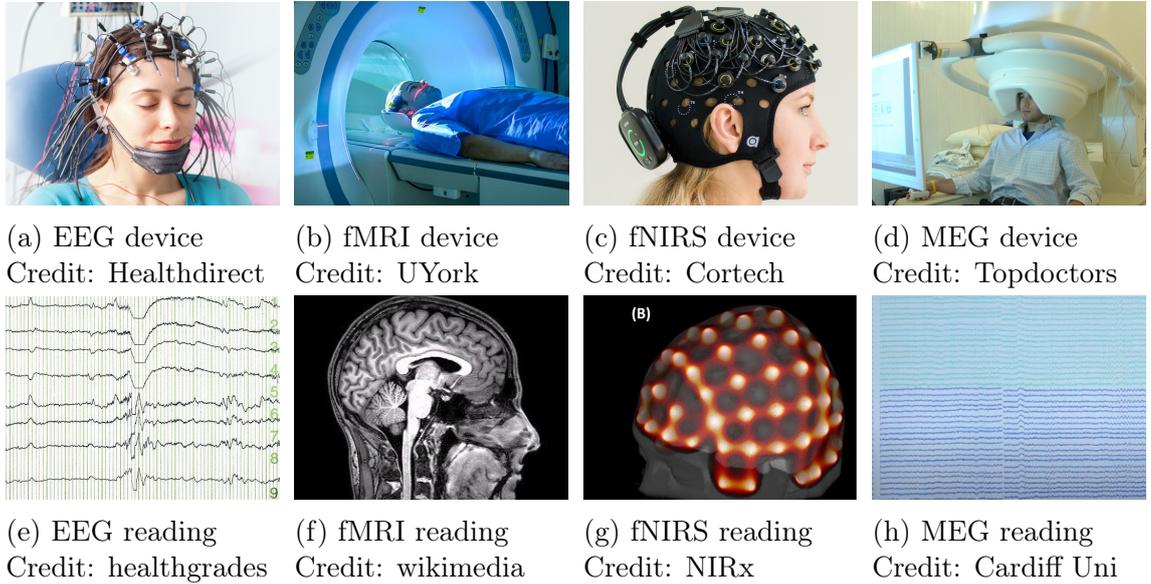


Figure 2.2: Non-invasive Brain Signal Acquisition devices with correspond readings

2.3 Motor Imagery Intention Recognition

Since the 1970s, a considerable number of researchers have been interested in EEG and have discovered numerous applications of EEG signals. Recent studies have demonstrated that EEG signals can be used to analyse various types of brain activity, such as sleep state [39], emotion recognition [40], fatigue detection [41], motion intention [42], brain disease [3, 43], and biometric identification [44].

Motor Imagery (MI) is a cognitive process in which individuals mentally simulate the movement of body parts in preparation for physical activity [45]. MI has been proposed as an intuitive mapping for Brain-Computer Interface (BCI) commands, as the tasks associated with MI are closely related to the natural production of muscle movement commands. This intuitive mapping increases the usability of BCI systems while reducing the mental effort required for operation, making it an advantageous alternative to other BCI paradigms such as Visually Evoked Potential (VEP) or Event-Related Potential (ERP). Taking into consideration the benefits of MI signals,

2. Literature Review

we decide to utilize them as a source in our Brain-Computer Interface (BCI) system, thereby enhancing the performance of Neural Network (NN) models in analyzing MI-based Electroencephalography (EEG) signals.

2.4 Functional Connectivity

Functional connectivity is a measure of the degree of communication between different regions of the brain and how well they are connected. Investigating the functional connectivity of the human brain can provide greater insight into its workings and lead to more effective feature extraction. There are various methods for quantifying functional connectivity, which can be divided into non-directed and directed connections. Non-directed metrics measure the association between signals from different brain regions, without determining the direction of influence. Directed metrics, on the other hand, attempt to establish statistical causation to describe the inter-regional associations [46]. Subsequently, this section will introduce two of the most commonly used estimation metrics for functional connectivity.

Pearson Coefficient is a non-directional measure of functional connectivity that disregards the temporal characteristics of data and treats it as a set of random variables. It quantifies the statistical features of two sets of random variables, X and Y , and determines the correlation strength between them. The Pearson Coefficient (r) can be calculated as follows:

$$r = \frac{\sum(x_i - \bar{X})(y_i - \bar{Y})}{\sqrt{\sum(x_i - \bar{X})^2 \sum(y_i - \bar{Y})^2}}$$

2. Literature Review

\bar{X} refers to the mean of random variable X .

Granger Causality is a direct measurement of functional connectivity that utilizes linear regression models or non-parametric spectral matrix factorization to produce an estimation of directed interactions between brain regions [47].

2.5 State-of-the-Art NN-Based Approaches in MI-EEG Signal

The emergence of Neural Networks (NN) and Deep Learning (DL) techniques has enabled computers to autonomously learn and analyze a variety of data. In comparison to traditional methods, feature extraction and intention classification are two distinct modules that necessitate distinct algorithms for each. The remarkable capabilities of the deep learning model in feature extraction and classification have enabled the integration of the feature extractor and classifier's functionality into a single, end-to-end framework.

Intra-subject evaluation also referred to as subject-dependent evaluation, involves the use of EEG segments from the same subject for both training and testing. This approach disregards the individual differences in brain patterns among subjects, thus resulting in a model that is specific to a single subject.

Cross-subject evaluation approach, which combines EEG segments from different subjects and splits training and testing sets, fails to consider the fact that each individual subject has distinct brain patterns. Consequently, the model evaluated using this approach is only applicable to a specific group of subjects, rather than

2. Literature Review

providing a general conclusion for the entire human population.

Evaluation that is not subject-dependent occurs when EEG segments used for training and testing are taken from two distinct groups of subjects, thus allowing the studied neural network model to have increased generalizability across different populations. This renders the entire framework more adaptive to practical applications.

In this section, we present a series of baseline and state-of-the-art (SOTA) deep learning-based frameworks for resolving electroencephalography-MI brain-computer interface (EEG-MI BCI) tasks, followed by an analysis of their strengths and weaknesses.

2.5.1 CNN-based SOTA

Convolutional Neural Network (CNN) proposed by LeCun for the processing of image, speech, and time-series data [48]. Like traditional Neural Networks, CNNs are composed of learnable weights and biases. However, they possess a distinct architecture that is specifically tailored to image data, allowing them to capture the relationship between pixels in an image. Recent frameworks typically employ Convolutional Neural Network (CNN) filters to extract features from raw brain signals. However, there are still a number of frameworks in their early stages that are composed solely of CNN filters.

Lawhern et al. [12] developed EEGNet, a widely accepted baseline deep learning model, in 2018. This convolutional neural network (CNN) was designed to be small and efficient, while still providing good classification performance. Lawhern et al. reported that EEGNet achieved an accuracy of approximately 70% under intra-subject evaluation and near 40% under cross-subject evaluation.

2. Literature Review

2.5.2 RNN-based SOTA

The human brain is one of the most complex and dynamic systems in the world, with its state constantly changing and thus impacting observed signals. A limitation of the conventional Convolutional Neural Network (CNN) framework is its inability to analyse the brain state over a prolonged period of time, as it is limited to the time taken for each Electroencephalography (EEG) segment. Recurrent Neural Networks (RNNs) can be employed to address this issue, as they are capable of recognizing sequential characteristics, thus allowing for the analysis of historical signals and the capture of long-term associations.

Zhang et al. [4] proposed a practical brain typing system, which utilizes a hybrid deep learning model to decode brain signals and convert them to machine commands for typing purposes. This model incorporates Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to learn temporal and spatial dependency features from raw Electroencephalography (EEG) signals. Zhang et al. reported that their proposed method can achieve approximately 95 % accuracy under intra-subject evaluation.

Zhang et al. [17] proposed CRAM, an end-to-end deep learning model which deploys convolutional neural networks (CNNs), recurrent neural networks (RNNs) and attentive mechanisms. This model was evaluated under a subject-independent setting, achieving an accuracy of approximately 60 % in four-class classification.

Other works [20, 13] also make use of Recurrent Neural Networks (RNNs) in their frameworks to capture sequential associations. To ensure a comprehensive comparison, we have added some of the state-of-the-art frameworks to our comparison model list.

2. Literature Review

2.5.3 Attention-based SOTA

The attention mechanism is a technique that enables a model to make more accurate predictions by allowing the model to focus on the most pertinent aspects of an input. In electroencephalography (EEG) studies, an attentive layer is typically deployed to identify the most discriminating EEG slices across a sequence of segments or the most significant channels. This is accomplished by first calculating a weighted score for each element; once the most important components are identified, they are amplified by multiplying them by their respective weighted scores. This serves to increase the weight of more important elements in the classification process, while minimizing the contribution of less essential elements (e.g., discriminative time slices, electrodes, and edges).

Zhang et al. proposed frameworks CRAM [17], GHAM [19], and GCRAM [20], which all incorporate an attentive module in the framework. Their experimental results demonstrate that the inclusion of an attentive module can effectively enhance the performance of the framework.

2.5.4 Graph-based SOTA

Brain activity is characterized by the collective efforts of multiple brain regions [49], which renders the brain more akin to a dynamic graph structure than a static image. Electroencephalography (EEG) signal acquisition devices can detect signals from a brain region, analogous to vertices in a graph. Functional connectivity is used to describe the relationships between different brain regions, analogous to edges in a graph. By transforming EEG readings into a graph with inter-regional associations, it is possible to better simulate the brain network and extract more relevant features,

2. Literature Review

which can assist the classifier in making a more accurate classification.

While CNN-based approaches [13, 4, 17] assume a Euclidean structure of EEG electrodes to capture inter-regional correlations, these works ignore the natural geometry of brain structure and connections between different regions beyond their immediate vicinity. In contrast, graph-based brain representations appear to better reflect the non-Euclidean nature of the human scalp [21], and also encode subject-invariant positioning priors of electrodes [19, 20] to the model. However, their graph construction cannot represent dynamic functional connectivity that adapts to different MI tasks, since the edges therein are fixed in terms of the distances between nodes.

Graph Neural Networks (GNNs) are a generalised type of neural network designed to extract topological features from graph-structured data. This has led to an increased interest in GNNs from both academia and industry. While Convolutional Neural Network (CNN) models have been successful in extracting features in Euclidean space, their inability to be applied to data in non-Euclidean space limits their effectiveness in extracting topological features.

Demir et al. experimentally evaluated several GNN-based approaches, namely GIN0 [26], GraphSAGE [50] and EEG-GAT [51], under the same framework [21]. However, this study connected the vertices using a complete graph or based on electrode alignment, which was not capable of capturing the dynamic features.

Wu et al. proposed a framework [52] that contains a data-driven algorithm to select emotion-relevant critical subnetworks automatically to explore universal emotionally relevant functional connectivity patterns among different subjects. This algorithm obtains emotionally relevant subnetworks by averaging all test subjects' functional connectivity matrices, followed by proportional thresholding to filter out weak con-

2. Literature Review

nections and merging of the emotionally relevant subnetwork for all other edges to construct the final emotional-relevant critical subnetwork. This algorithm has been utilized to capture the dynamic functional connectivity and to make the algorithm more adaptive to MI tasks, as detailed in Chapter 3, methodology section.

2.5.5 Limitations

Recent research into MI recognition leverages the powerful capabilities of deep neural networks (NNs) to extract features from large amounts of signals. However, these studies mainly focus on extracting features solely from the temporal domains, without taking into account the functional connectivity. The high intersubject variability easily leads to trained model overfits training data. Resulting in most SOTA methods not perform well in extracting subject-invariant features and making accurate subject-independent classifications. To address this issue, there has been a need for research into analyzing brain signals from a neurological perspective, such as simulating the brain's state when generating signals, and extract features from the brain-topological domain.

Additionally, EEG signals have a high noise-to-signal ratio (NTS), which necessitates the use of additional operations/parameters in neural network models to distinguish noise from useful information. Many effective methods have been proposed by recent studies to remove environmental noises and human artifacts. However, there has been a lack of research into how to train models only with effective data in order to accelerate training and inference time, such as excluding non-contributory EEG channels, channel-wise interference, and non-contributory EEG segments.

In order to make BCI systems more adaptive to new users, fine-tuning a pre-trained

2. Literature Review

deep learning model using signals that have a similar distribution to the user's signals is highly effective in improving the end-user experience. By sacrificing some of the subject independence, the model gains the capacity to analyze the user's personalized features with higher classification accuracy, thus enhancing the user experience in the application. Despite the advances in MI recognition, how to identify similar signals is still an open problem. There is a lack of research in determining the most appropriate method for estimating the similarity between EEG signals.

Chapter 3

Functional Connectivity Empowered Intention Classification

3.1 Introduction

Chapter 3 presents the architecture of our proposed DL framework, SIFT-EEG, that takes task-adaptive functional connectivity into account when recognizing EEG-MI signals. Specifically, it extracts temporal dynamic features from raw EEG signals using CNN temporal encoders and combines them with functional connectivity to construct graph representations of the brain network. Then, it employs GNN topological encoders to navigate subject-invariant discriminative patterns inside the brain network and extract neurology topological features to make accurate classifications. We also present our proposed task-adaptive channel selection algorithm in this chapter, the purpose of this algorithm is to exclude task-irrelevant functional brain regions that are non-contributory in recognizing corresponding intentions. Empir-

3. *Functional Connectivity Empowered Intention Classification*

ical comparisons for the proposed DL framework and algorithm are shown in this chapter, we can expect around 1% and 11% improvements in state-of-the-art CNN-based and GNN-based models. Furthermore, the task-adaptive channel selection reaches similar predictive performance with only 20% of raw EEG data, suggesting a possible shift in direction for future works other than simply scaling up the model. The findings of this chapter have been organized into the paper, *Less is More: Brain Functional Connectivity Empowered Generalisable Intention Classification with Task-relevant Channel Selection*, and published in *IEEE Transactions on Neural Systems and Rehabilitation Engineering* (TNSRE).

Key contributions for Chapter 3 are:

- We propose a subject-independent MI prediction model built upon functional topological adjacency, which further takes into account self-attentive temporal convolution and graph isomorphism, thus capturing task-adaptive but subject-invariant EEG embeddings.
- We present a data-driven channel selection algorithm based on active brain regions, which can exclude non-contributory channels and reduce the impact of task-irrelevant noises.
- We evaluate SIFT-EEG on a large-scale EEG-based MI dataset, demonstrating its effectiveness for subject-independent classification, with competitive performances against state-of-the-art even using 20% of raw data.

3. Functional Connectivity Empowered Intention Classification

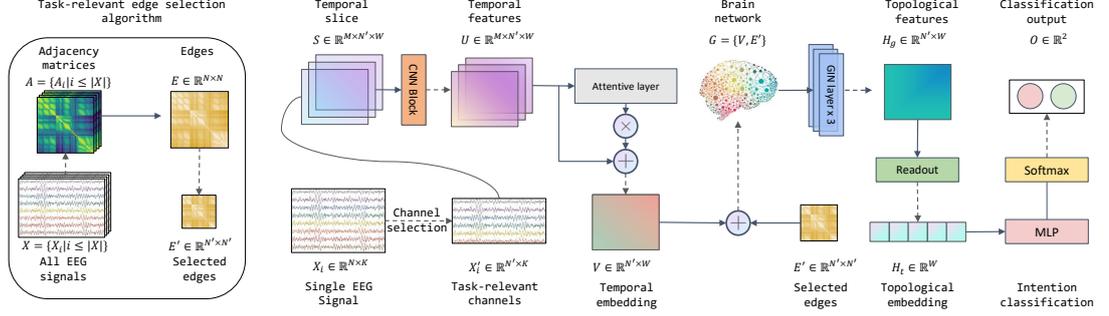


Figure 3.1: Overview of the **Subject-Independent MI classification model using brain Functional ConnecTivity (SIFT-EEG)** model. We first crop raw EEG signal into a sequence of time slices using the slide window technique; adopt CNN to extract temporal features and a self-attentive module to search for the most discriminative temporal slice; then we combine nodes in the temporal embedding with functional connectivity to generate the graph representation of EEG signals; select three layers of Graph Isomorphism Network (GIN) to extract topological features and lastly, the extracted topological embedding are classified to different motion intention using a fully connected network with a softmax activation function.

3.2 Preliminaries

3.2.1 Graph representation for EEG signals

A graph $G = \langle V, E, W \rangle$ is an abstract structure in non-euclidean space, composed of nodes V , edges E , and weights W . Thus, the brain network topology can be represented by graphs. The nodes $\{v_i\}_{i=1:n} \in V$ refer to EEG scalp electrodes located on specific brain regions, with a total electrode count of n . Each edge $e_{ij} \in E$ represents the inter-regional relation between electrode pairs (v_i, v_j) . Accordingly, the pairwise edge weight $w_{ij} \in W$ indicates the strength of each relation.

Graph adjacency A takes the form of a 2D matrix $R^{N \times N}$, where N denotes the number of nodes. One can be either a weighted or an unweighted graph. For unweighted adjacency, $A_{ij} = 1$ represents a pair of connected nodes while $A_{ij} = 0$ means there is no connection between them. A weighted adjacency has an additional

3. Functional Connectivity Empowered Intention Classification

attribute to the strength of relationship by setting $A_{ij} = w_{ij}$. Degree $d(v_i)$ measures a node’s centrality by the number of edges connecting to it, indicating the node importance within the graph, calculated by adding up the i -th row of A .

Edge definitions for EEG-based graph representations are yet to be provided. The literature mostly uses either complete graph [21], i.e., each pair of nodes are connected, or distance-based connections [19, 20], that is, two nodes connected if their physical distance between electrodes is lower than a pre-defined threshold.

3.2.2 Graph Neural Networks

Graph Neural Networks (GNNs) are a family of generalized neural networks excelling at analyzing graph-structured data.

Graph Convolutional Network (GCN) [53] simplify spectral graph convolutions to the framework of neural networks, which defines a GCN operator to exchange node-wise information through edge connections, along with a layer-wise propagation rule that updates hidden node features. Throughout the K -layer propagation process, the nodes receive the averaged features from their neighbors via a mean aggregation, and the shape of the graph structure remains the same at the next layer.

$$h_v^{(k+1)} = \sigma (W \cdot \text{MEAN} \{h_v^{(k)} \cup h_n^{(k)} \mid n \in \mathcal{N}(v)\}) \quad (3.1)$$

where $h_v^{(k+1)}$ is node v ’s hidden feature at k -th GCN layer, W are model parameters and σ is non-linear activation. $\text{READOUT}(\cdot)$ function is further applied after propagating the last layer to extract the graph-level embedding,

$$h_G = \text{READOUT} (\{h_v^{(K)} \mid v \in V\}) \quad (3.2)$$

3. Functional Connectivity Empowered Intention Classification

GraphSAGE [50] interpret GNNs from a spatial perspective and generalize the mean-aggregator in GCN to a wider range of operators,

$$h_v^{(k+1)} = \sigma \left(W \cdot \text{AGG} \{ h_v^{(k)} \cup h_n^{(k)} \mid n \in \mathcal{N}(v) \} \right) \quad (3.3)$$

where W are model parameters and $\text{AGG}(\cdot)$ refers to a permutation invariant function such as min/max/mean pooling.

The recent Graph Isomorphism Network (GIN) [26] is derived from the Weisfeiler-Lehman (WL) isomorphism test for checking if graphs are topologically identical, proving that GIN is as powerful as the WL test when $\text{AGGREGATE}(\cdot)$ and $\text{READOUT}(\cdot)$ are permutation invariant and injective, which produces more discriminative embeddings than other GNNs variants in graph-level classification.

Assume the function $f : \mathcal{X} \rightarrow \mathbb{R}^n$ for any countable node feature space \mathcal{X} . Then for infinitely many choices of ϵ including all irrational numbers, such that any function g can be decomposed into $g(c, X) = \phi((1 + \epsilon) \cdot f(c) + \sum_{x \in X} f(x))$ for some function ϕ , where $c \in X$ and $X \subset \mathcal{X}$. A multilayer perceptron (MLP) with more than one hidden layer can be used to approximate injective function [54], so the node embedding h_v becomes distinguishable, using the GINConv operator defined by

$$h_v^{(k+1)} = \text{MLP}^{(k)} \left((1 + \epsilon^{(k)}) h_v^{(k)} + \sum_{n \in \mathcal{N}(v)} h_n^{(k)} \right) \quad (3.4)$$

where the $\text{AGG}(\cdot)$ of neighbor nodes is implemented as a summation to impose injective mapping.

3.3 Methodology

3.3.1 Overview

The overall structure of SIFT-EEG is presented in Fig. 3.1. We now describe the goal of EEG-based Motor Imagery (MI) classification and overview the key steps of our approach. Given an EEG segment $X \in \mathbb{R}^{N \times K}$ collected while a subject was performing a MI task, we aim to estimate the associated specific intention Y by training a predictive model that performs supervised classification under the subject-independent setting, meaning that model training and evaluation are conducted on two disjoint groups of subjects. $K = T \times f$ is the number of time points within a segment, N is the number of electrodes, T is the recording duration, and f is the sampling frequency.

Let X be an EEG segment and Y be the intention, our approach approximates the mapping $f : X \rightarrow Y$, parameterized by a Neural Network $\hat{Y} = f_{\Theta}(X)$, by involving the following five steps:

- 1) Calculating adjacency $A \in \mathbb{R}^{N \times N}$ across the readings of all sensory channels X based on functional connectivity;
- 2) Selecting the top- N' strongest channels $X' \in \mathbb{R}^{N' \times K}$ tailored to the MI task and identifying task-relevant edges $E' \in \mathbb{R}^{N' \times N'}$;
- 3) Extracting temporal node embedding $V \in \mathbb{R}^{N' \times \omega}$, by summarizing the features of each temporal slice $S_m \in \mathbb{R}^{N' \times \omega}$, where $S = \{S_m\}_{m=1}^M$ results from sliding window applied on raw data of task-relevant channels X' ;
- 4) Generating topological graph embedding $H \in \mathbb{R}^h$ from the brain functional

3. Functional Connectivity Empowered Intention Classification

network $G = \langle V, E' \rangle$ constructed by temporal node embeddings and task-relevant edges;

- 5) Predicting the intention $f(X) = \hat{Y}$ of an EEG segment from the graph-level topological embedding.

3.3.2 Functional Adjacency Matrix

We first calculate the adjacency of an EEG-based brain network with functional connectivity. We prefer this approach to previous distance-based or complete graph connections which either ignore topological relationships or are very computationally intensive if too many nodes are involved. Functional connectivity defines the statistical dependencies among temporal signals, commonly measured as the similarity between two brain regions by using the Pearson coefficient:

$$P(i, j) = \frac{\text{Cov}(i, j)}{\text{Var}(i)\text{Var}(j)} \quad (3.5)$$

where $\text{Cov}(i, j)$ is the covariance of measurement readings between sensory electrodes i and j , $\text{Var}(i)$ denotes the standard deviation of i -th channel readings throughout K timesteps.

This results in a complete functional adjacency $A \in \mathbb{R}^{N \times N}$ with N being the number of electrodes. The following section discusses how we further reduce nodes and edges by only keeping task-relevant channels.

3.3.3 Task-Adaptive Channel Selection

The activation of different brain regions varies with brain activity [55], indicating that some regions may not respond as actively to certain tasks as others. However,

3. Functional Connectivity Empowered Intention Classification

raw EEG data collected from electrodes distributed throughout the scalp will thus always contain information irrelevant to a specific MI task. The presence of such redundancy increases the computational burden and over-fitting risks.

This has led to the demand for filtering out task-irrelevant channels, yet most efforts are directed at manually solving complex optimization problems [56], which even requires extensive domain expertise[57]. The problem may, however, require an automatic solution in some cases, such as cross-subject analysis [58]. Instead, we propose a simple yet effective data-driven channel selection strategy in Algorithm. 1 upon node **importance** within all EEG channels, representing the activation intensity of corresponding brain regions. In particular, we suggest two metrics for measuring the nodes' importance $\Lambda \in \mathbb{R}^N$.

Degree-based Importance.

The degree of a node indicates its centrality within a brain network as well as the implication at the graph-level. Within this metric, we simply define the nodes' importance W as the number of incoming edges by looking up the channel-wise adjacency matrix A . Then, we apply a proportional threshold T^* to preserve channels N' adapting to a specific intention by descending nodes upon the importance Λ . The above steps are applied for each EEG segment X in the dataset.

Strength-based Importance.

We assume the channels with high correlation strength across different MI tasks are more active than those less correlated. Say we have EEG segments of C tasks with N_c for each, being performed by multiple subjects. We compute the absolute

3. Functional Connectivity Empowered Intention Classification

ALGORITHM 1: Task-relevant channel selection

Require: raw EEG segments set \mathcal{X} , intention set C , importance weight Λ and threshold T^*

Ensure: EEG segments set \mathcal{X}' of task-relevant channels

1: $\mathcal{X}' \leftarrow \emptyset$

2: **for** $X \in \mathcal{X}$ **do**

3: Sort all channels by importance scores S in descending order, and store the indices into I

$$I \leftarrow \text{argsort}(W)$$

4: Select the most active channels with threshold T^*

$$X' \leftarrow X_i, \quad \forall i < T^* \text{ and } i \in I$$

5: Preserve readings of the most active channels X' and append them to the task-relevant EEG set

$$\mathcal{X}' \leftarrow \mathcal{X}' \cup X'$$

6: **end for**

Pearson coefficient between the same channel's N_c readings of every two intentions $c_i, c_j \in C$ across all intentions. Then the task-relevant strength of each channel is obtained by averaging its $C - 1$ coefficient values across all intention pairs. We define this strength of all channels as the importance weight Λ for segment X . Lastly, we apply a threshold T^* to preserve task-relevant channels N' by descending nodes upon their task-relevant strength.

3. Functional Connectivity Empowered Intention Classification

3.3.4 Temporal Embedding

There might be different concentration periods between subjects during MI, leading to different temporal properties even if two subjects were performing the same task [17]. Thus, our objective is to identify the most discriminative period within each EEG segment of different subjects, which we refer to as the subject-invariant temporal embedding.

Specifically, we first crop the EEG segment $X' \in \mathbb{R}^{N' \times K}$ into M temporal slices $S \in \mathbb{R}^{M \times N' \times \omega}$ using a sliding window of size ω , after selecting the strongest N' channels. We then create initial temporal features U_m of each temporal slice $S_m \in \mathbb{R}^{N' \times \omega}$ by applying a 3×3 convolution kernel followed by an elu non-linear activation,

$$U_m = \text{elu}(\text{conv}(S_m)), \quad \text{for } m = 1, \dots, M \quad (3.6)$$

where U_m retains the same tensor shape as input slice S_m . Next, we use self-attention [18] to capture slice-wise correlations and adapt temporal weights to their temporal features, leading to a temporal embedding $V \in \mathbb{R}^{N' \times W}$ that summarizes the EEG segment,

$$V = \sum_{m=1}^M \frac{\exp(H_m^\top W_m)}{\sum_{j=1}^M \exp(H_j^\top W_j)} U_m \quad (3.7)$$

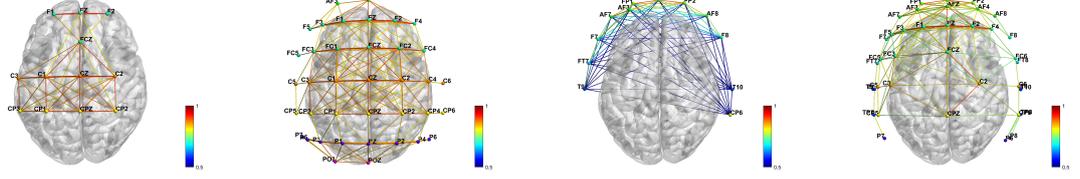
with

$$H_m = W_h U_m + b_h \quad (3.8)$$

where W_m , W_h and b_h are learnable parameters. The temporal embedding V aggregates all of M slices, taking into account each slice's importance, which derives the most discriminative representation upon input values.

3.3.5 Topological Embedding

3. Functional Connectivity Empowered Intention Classification



(a) Degree sampling 12 nodes with top 100% of edges (b) Degree sampling 38 nodes with top 30% of edges (c) Strength sampling 12 nodes with top 100% of edges (d) Strength sampling 38 nodes with top 30% of edges

Figure 3.2: Selected task-relevant channels and channel-wise connectivity. We present part of the edges with the highest correlation for readability purposes. The graph is visualized using BrainNet Viewer [59].

The brain activity is coordinated by multiple brain regions [49]. These inter-regional correlations are potentially beneficial when incorporated into predictive models. While CNN-based methods [13, 4, 17] assume Euclidean-structure of EEG electrodes, these works ignore the natural geometry of brain structure and connections between different regions beyond their immediate vicinity. As opposed, graph-based brain representations appear to reflect better the non-Euclidean nature of human’s scalp [21], but also encode the subject-invariant positioning priors of electrodes [19, 20] to the model. Nevertheless, their graph constructions cannot encode dynamic functional connectivity that adapts to different MI tasks, since the edges therein are fixed in terms of the distances between nodes.

Task-Adaptive Edge Formation.

In a brain network, the correctly illustrated edges could model how each brain region associates with other active regions during a MI task. Meanwhile, channel-wise correlations across MI intentions do not remain constant as aforementioned. Having identified the task-relevant channels N' , we now formulate the task-relevant edges, Algorithm 2, $E' \in \mathbb{R}^{N' \times N'}$ to leverage strong associations tailored to the task [52],

3. Functional Connectivity Empowered Intention Classification

in the following stages:

1. *Constructing*: Calculate channel-wise correlation adjacency $A_c \in \mathbb{R}^{L \times S \times N \times N}$ of all subjects for each intention $c \in C$, using the Pearson coefficient, where L denotes the number of labels, S is the number of subjects, and N represents the number of channels/electrodes.
2. *Generalizing*: Find the generalized connectivity across all subjects by averaging the adjacency matrices of all subjects for each intention c .
3. *Thresholding*: Preserve critical connections that exceed the threshold T^e for each intention c .
4. *Merging*: Derive task-relevant edges E' by merging the critical connections of all C intentions.

Algorithm 3 shows the pseudocode to derive a purely functional adjacency E' . Moreover, considering the temporal embedding V results from the task-relevant channels X' , we construct the task-adaptive brain network $G = \langle V, E' \rangle$ to obtain topological embedding.

Embedding Computation.

We adopt a L -layer GIN to investigate the topological embedding of a given brain network G , as it shows promise for graph-level classification [26, 60]. The graph-level topological embedding $H \in \mathbb{R}^W$ after L -layers' propagation is produced by

$$H = \text{READOUT} \{h_v^{(L)} \mid v \in V\} \quad (3.9)$$

3. Functional Connectivity Empowered Intention Classification

with

$$h_v^{(l)} = \text{MLP}^{(l-1)} \left((1 + \epsilon^{(l-1)})h_v^{(l-1)} + \sum_{n \in N(v)} h_n^{(l-1)} \right) \quad (3.10)$$

where W is the embedding dimension. We set MLP with 2 hidden layers, $\epsilon = 0$, and $k = 3$ in practice. While READOUT(\cdot) function can be either non-injective aggregations (e.g., pooling) or injective mappings (e.g., MLP(\cdot)), it is believed that the latter will yield a more discriminative graph embedding, as shown by [26]. Our empirical studies experiment with different setups to verify this argument in EEG-based applications.

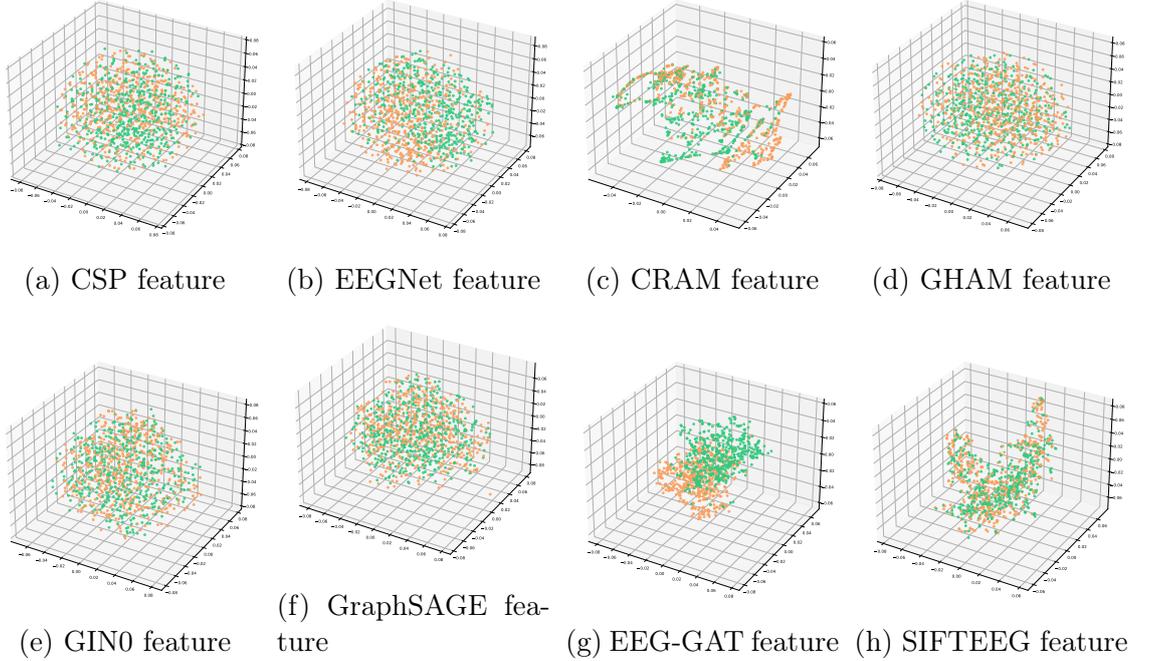


Figure 3.3: The visualization of hidden features with t-SNE. Green dots refer to left-hand imaginary motion, while orange dots indicate right-hand imaginary motion.

3.3.6 Intention Classification

We take the topological embedding H of an EEG segment X to perform intention prediction, with a softmax function to estimate the probabilities of each possible

3. Functional Connectivity Empowered Intention Classification

intention,

$$P_c = \text{softmax}(W_o H + b_o), \quad \text{for } c = 1, \dots, C \quad (3.11)$$

where W_o and b_o are trainable parameters. The predicted intention is thus given by $\hat{Y} = \arg \max_c P_c$.

3.4 Empirical Studies

3.4.1 Experiment Setting

Dataset.

In this work, we evaluate the performance of the proposed SIFT-EEG on a widely studied large-scale cross-subject EEG-based MI dataset eegmddb (EEG motor movement/imagery database) downloaded from Physionet database [61]. The dataset was collected using BCI2000 [36] containing 64 electrodes with the sampling frequency of 160Hz from 109 healthy subjects. Following the conventions [19, 20], we remove subjects #88, #89, #92, #100 from the dataset due to consecutively resting states. Within three sessions of MI tasks, EEG signals were recorded when subjects were executing left/right fist open and closed imagery. There are approximately fifteen 3.1-second segments of imaginary left/right-hand movement in each session; each subject performed three sessions. We randomly select 90 subjects and use their EEG segments as the training set, whereas the remaining 15 subjects are used as the test set. There are no disjoint subjects in the training and testing set, which ensures the evaluation is conducted in a subject-independent setting. We experiment with twelve different train/test splits to reduce randomness in the results. We fix a unique random seed for each split that specifies the training and testing

3. Functional Connectivity Empowered Intention Classification

subjects. Noticeably, the last split includes all untested subjects from the previous eleven splits to ensure each subject is tested at least once.

Preprocessing.

In line with most related works, raw EEG data is normalized with a z-score, calculated by subtracting the mean value and dividing by the standard deviation of training samples X . We do not further apply filters or data augmentation on raw EEG signals. We use temporal data as model input.

Baselines.

We reproduce several baselines for subject-independent MI classification, including both traditional and DNN-based approaches. We adopt the Common spatial pattern and linear discriminant analysis (CSP+LDA) [62] as the representative traditional means. The DNN-based approaches are further divided into the CNN and GNN families.

Specifically, the CNN-based methods include EEGNet [12], CRAM [17], GHAM [20]. The GNN members GIN0 [26], GraphSAGE [50] and EEG-GAT [51] apply different GNNs under the same framework [21]. In addition, we implement a two-layer RNN as a vanilla DNN baseline to model temporal data.

Implementation Detail.

Each input EEG segment $X \in \mathbb{R}^{64 \times 496}$ contains 496 timesteps with 64 channels. For temporal embedding, we apply the slide window technique with window size of 400

3. *Functional Connectivity Empowered Intention Classification*

and a step size is 10. Hence, the input signal contains ten temporal slices, where each slice has the shape of $[64,400]$ (i.e., $N = 64, \omega = 400$). All the models are implemented with PyTorch¹ and trained and trained on an NVidia 3060-Ti GPU in a fully-supervised manner. We use cross-entropy as the objective function and optimize model parameters using Adam with a learning rate of 0.001. The training batch size is 500. Each model is trained for 120 epochs, and the dropout probability is set to 0.5 to avoid over-fitting. We implement all GNN-related components using Torch_Geometric².

3.4.2 Result & Discussion

The proposed SIFT-EEG is empirically compared with a range of baselines, focusing on three research questions.

- 1) Does the modeling of temporal dependencies benefit spatial/topological correlations and model performance?
- 2) Does task-adaptive dynamic connectivity outperform distance-based connectivity for topological learning?
- 3) Does task-relevant channel selection with topological features still achieve competitive performance?

The evaluation metrics include classification Accuracy and the Area Under ROC-Curve(ROC-AUC). All models are trained and evaluated with the same setting for fair comparison.

¹<https://pytorch.org>

²<https://pytorch-geometric.readthedocs.io>

3. Functional Connectivity Empowered Intention Classification

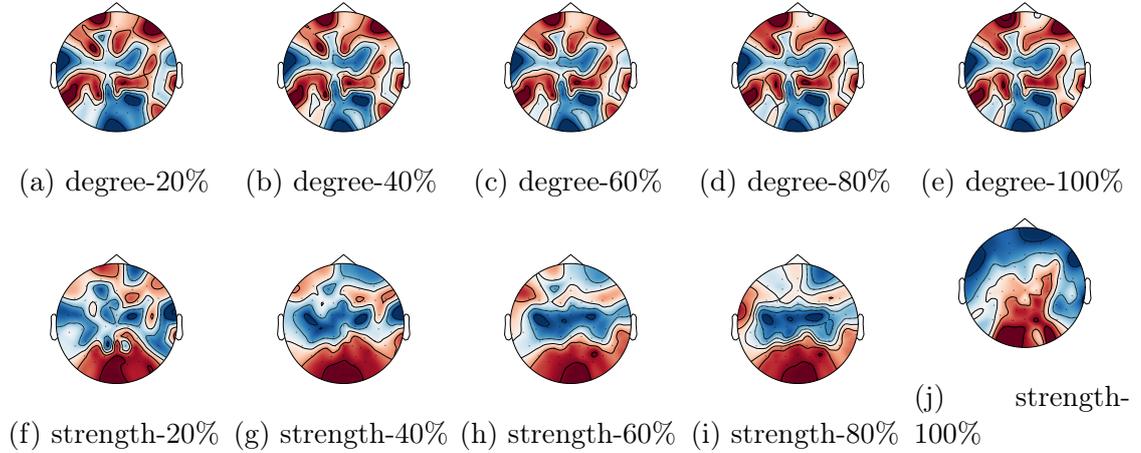


Figure 3.4: The brain topology maps acquired with various importance metrics under different ratios of subjects sampled from 105 subjects.

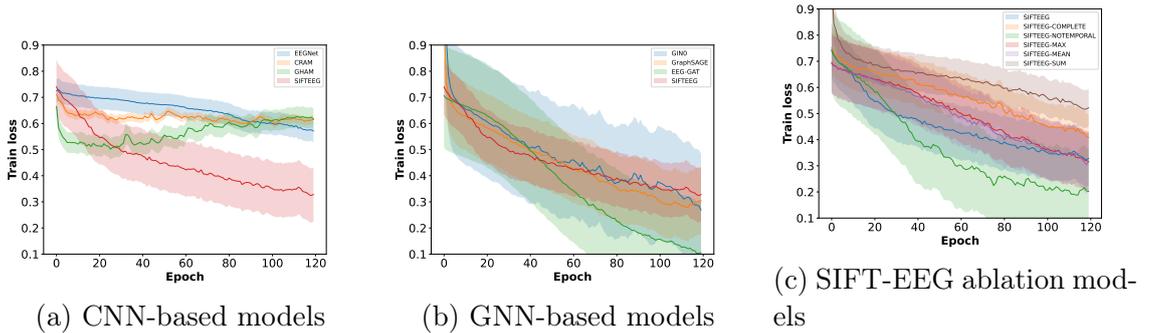


Figure 3.5: Comparison for training loss change when the number of training epochs increases.

Overall Performance.

Tab. 3.1 reports the model performance of all approaches on both metrics. Observe that SIFT-EEG outperforms all baseline models, with 1.13% and 14.68% accuracy improvements, as well as ROC-AUC gains of 1.09% and 15.44%, compared to the best performing CNN-based state-of-the-art (GHAM) and GNN-based state-of-the-art (EEG-GAT), respectively. Meanwhile, the top performers in both DNN families prove to be more accurate than CSP+LDA. In addition, all DNN baselines (except RNN) include mechanisms to represent the spatial correlation between EEG chan-

3. Functional Connectivity Empowered Intention Classification

nels, providing better results than vanilla RNN. It follows that such representations should be incorporated into modeling.

Fig. 3.5 depicts how the training loss changes with the number of training epochs increase. GNN-based methods generally have lower training losses and faster convergence rates compared with CNN-based methods, which suggests the capability of topological features in task-specific predictions. However, their test performances are worse than CNN-based methods, showing a tendency to overfit the training data. The reason may be that GNN-based methods extract topological features directly from raw EEG signals. Recall that our evaluations take place in a subject-independent setting. In this case, the variances in temporal patterns between subjects, i.e., graph noises, may cause them to perform inconsistently [63, 64]. Conversely, CNN-based methods explicitly handle temporal correlations before looking at the spatial domain. For instance, CRAM locates discriminative temporal features adaptively for different subjects by using attention. This eases the burden on the spatial feature extractor and generalizes the model to new subjects, albeit fitting these models takes longer. The proposed SIFT-EEG combines the merits of both, leveraging flexible topological features of the brain signal, as well as subject-adaptive temporal features that reduce task-irrelevant noise.

In addition, we perform statistically significant tests to evaluate the model performance improvements of SIFT-EEG over baselines. We use the pairwise t-test, assuming the pairwise difference is significant if p -value is less than 0.05. The results are reported in Tab. 3.3, where statistically significant differences are **bolded**.

3. Functional Connectivity Empowered Intention Classification

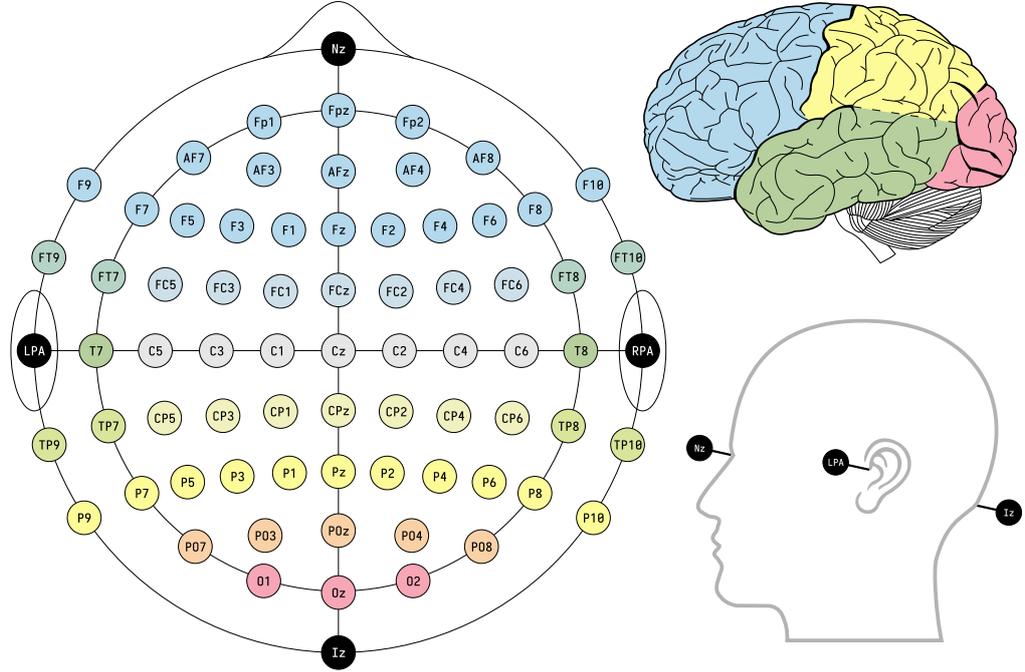


Figure 3.6: Electrodes position in international 10-10 EEG signal acquisition device

Impact of Temporal Embedding.

Now we analyze how temporal embedding contributes to EEG-based MI classification and answer the first research question. EEGNet design a CNN block to extract temporal features, while CRAM and GHAM adopt self-attention on a recurrent network to capture long-term temporal dependency and adaptive subject-specific patterns, thus improving model performance in subject-independent experiments. In contrast, GIN0 and GraphSAGE take raw EEG signals directly as input without learning temporal features. This leads to much lower predictive results than their CNN-based counterpart. EEG-GAT, on the other hand, attempts to extract temporal information with a 2D temporal convolution operator, which derives around 32% and 24% improvement over GIN0 and GraphSAGE, even when all three approaches are developed under a similar predictive framework. We also examine the impact of temporal embedding within SIFT-EEG. Denoted as SIFT-EEG(w/o t), a

3. *Functional Connectivity Empowered Intention Classification*

SIFT-EEG variant without temporal embedding is compared with full SIFT-EEG to reveal its effect. There is an increase of 29.01% and 28.96% in the results of accuracy and ROC-AUC, respectively. It is thus clear from the comparisons of all three groups that modeling temporal dependency in conjunction with spatial/topological embedding is essential.

Comparison of Topological Feature.

The second research question investigates whether the proposed task-adaptive dynamic connectivity facilitates topological learning. We first identify the graph connection of each model in comparison. GHAM uses distance-based graph representations of EEG channels, but not GNN for representation learning. All three GNN baselines define complete graph connections to yield the best results. SIFT-EEG chooses functional connectivity and bypasses the position limits.

GHAM extends CRAM with a graph definition of input that leads to an increase in 1.5% on accuracy and 0.7% on ROC-AUC to CRAM, implying the merit of non-euclidean assumption imposed by graph representation. Among the three GNN baselines, EEG-GAT shows better performance than the other two. A possible explanation is, GAT could benefit from well-designed attention in large and noisy graphs [65], suggesting GAT seems to win out over GIN in a complete graph of the brain network (as per their framework). Figure. 3.3 presents the hidden features extracted by different methods, features extracted by the SIFT-EEG framework for each class tend to become more disjoint with another class, and possess stronger discriminative power.

We further replace dynamic connectivity with complete connection in SIFT-EEG

3. *Functional Connectivity Empowered Intention Classification*

and observe a clear performance drop as in SIFT-EEG (w/o d). The GIN applied to SIFT-EEG aggregates all nodes indiscriminately under complete connections. This leads to the conclusion that SIFT-EEG needs to be defined with task-adaptive edges.

Injective Mapping Benefits SIFT-EEG.

We additionally test with different READOUT functions in SIFT-EEG, including three pooling-based variants SIFT-EEG (Avg), SIFT-EEG (Sum), SIFT-EEG (Max), referring to mean-, sum-, and max-pooling. The performance also drops noticeably. As a reminder, SIFT-EEG emphasizes a) graph-level prediction instead of node-level prediction; and b) structural information of the graph (brain network). Our results empirically align with the theoretical findings of [26], i.e., MLP-based injective mapping increases capacity over its non-injective READOUT counterpart.

Analysis of Task-Adaptive Channel Selection.

Our third research question examines task-adaptive channel selection from four perspectives. First, we visualize and compare the effects of two importance metrics. We next apply task-adaptive channel selection to CRAM and GHAM to examine its applicability. We evaluate its effectiveness against three additional channel selection approaches using SIFT-EEG. Finally, we discuss the model elapsed time results with different channel selection ratios.

Qualitativeness Fig. 3.2 showcases the positions of task-relevant channels selected by two different importance metrics. The nodes resulting from degree-based importance are primarily located around the central sulcus, while those derived by

3. *Functional Connectivity Empowered Intention Classification*

strength-based importance clusters in the frontal lobe correspond to the functional area for motor control in the study of neuroscience [66].

Applicability In addition to SIFT-EEG, the performance of two other models, CRAM and GHAM, is compared with task-adaptive and random selection by keeping 20% original channels. For comparison, CRAM makes no graph assumptions, GHAM represents the input with a hard-ruled graph definition, whilst SIFT-EEG dynamically determines the graph representation. As seen by Table 3.2, the accuracy of all models for task-adaptive selection is similar to (with CRAM and GHAM) or even slightly higher (with SIFT-EEG) than all channels when only 20% are used, whereas random strategy returns lower performance. Hence, our task-adaptive channel selection could benefit all models in general. Still, it works best when coupled with dynamic functional graph connectivity.

Effectiveness Moreover, we compare the proposed task-adaptive selection with two additional channel selection strategies. Shan et al. [67] identify subject-specific channels by finding channels with strong correlations to the central channel. In reproducing this strategy, we select the channels with an average correlation ≥ 0.7 , except for the reference channels C3/C4/Cz. Mattioli et al. [68] reduce the number of channels required by segmenting motion functional regions and producing regions of interest. In our experiments, we select channels located in the motor cortex region for this method. Table 3.4 reports the least number of channels required to achieve an accuracy greater than 60%, for each of these strategies applied to SIFT-EEG. The task-adaptive selection with strength-based importance is the top performer with the fewest channels required. The degree-based metric, however, appears to be relatively ineffective. It might be the case that nodes are unequally distributed

3. *Functional Connectivity Empowered Intention Classification*

throughout the scalp. According to Fig. 3.6, the sensory nodes near edge areas (e.g., AF7, O1 and P10) have fewer neighbors than those in central areas such as Cz, Cpz and Fcz. Meanwhile, the distance between nodes affects both the number of incoming edges and the strength of the connection. This eventually biases the estimation when using the degree-based importance metric.

On the other hand, the strength-based task-adaptive selection does not only report higher accuracy but with fewer channels, even compared to using channels from the widely-recognized motor cortex region [68]. The results indicate that task-adaptive channel selection can help encode the most discriminative EEG embedding with the fewest channels, even in the absence of extensive domain knowledge.

Efficiency Our final step is to examine the model performance and running efficiency when varying the number of channels and edges. Fig. 3.7 and Fig. 3.8 show accuracy and relative elapsed time, respectively. The random selection consistently results in degraded performance as available channels decrease. In contrast, both task-adaptive selections perform better than random selection. The strength-based selection maintains the highest consistency of performance, regardless of the number of channels. Meanwhile, both training time and inference time are trending downward with fewer channels. In an interesting twist, running with 20% channels would cost slightly more time than the case of 40%, which, however, would lead to rebounded classification accuracy.

In addition, we compare the computational efficiency of SIFT-EEG with baselines when all 64 channels are used, as shown in Fig. 3.9. We evaluate how long it takes to handle 400 arbitrarily sampled EEG segments with 400 forward passes with batch size 100, for both training and inference. The proposed SIFT-EEG takes 3.4% less

3. Functional Connectivity Empowered Intention Classification

training time than EEG-GAT, the strongest GNN baseline whilst improving the predictive performance by 14.68%. Moreover, SIFT-EEG consumes 5 fewer but 10 more seconds than CRAM and GHAM, respectively. The higher recognition accuracy of these three models comes at the expense of being more costly during inference. Nevertheless, we note that SIFT-EEG can improve its efficiency with our task-relevant channel selection, denoted by SIFT-EEG (w/c) in Fig. 3.9. At the time it runs with top-40% task-relevant channels, which improves training and inference efficiency by more than 50% and 75%, respectively. This proves SIFT-EEG to be faster than most baselines while gaining substantial performance benefits.

Visualization of Dynamic Functional Connectivity

Having learned the task-adaptive dynamic functional connectivity topology, we compare it with a fixed functional connectivity topology originating from the pre-motor, supplementary, and primary motor area [69]. We visualize the active brain region across certain proportions of 105 subjects upon performing MI tasks, according to two important metrics, as in Fig. 3.4. The degree-based importance is associated with activities in the frontal and parietal lobes, whereas strength-based importance seems to target the occipital lobes mainly. There are overlaps between the dynamic active region and the fixed motor region regardless of the number of subjects used. Moreover, our dynamic brain topology shows that active regions within each importance metric are distributed similarly across subjects. Increasing the number of subjects converges the distribution to an almost identical pattern, implying that the statistical characteristics of the dynamic topology can be improved by having more samples. Still, it shows robustness despite different sample sizes.

3. Functional Connectivity Empowered Intention Classification

Impact of Varying Training Subjects

Furthermore, we investigate how different numbers of training subjects affect classification accuracy. Our experiment begins with fifty subjects used for training, i.e., the training and test subjects are close to a 1:1 ratio. Following, the training set expands by five subjects per ratio record until 100 subjects have been included. For each train/test ratio, we perform cross-validation to split the train and test sets, ensuring each subject has been tested at least once. We report performance improvements across GNN family methods with varying numbers of training subjects, using fifty as the baseline and the mean and standard deviation for each ratio, as illustrated in Fig. 3.10. Fig. 3.10 illustrates a strong correlation is observed between the number of subjects and classification accuracy. Increasing the number of train subjects from fifty to seventy almost linearly improves classification accuracy. This may suggest that SIFT-EEG learns more subject-independent features as it is trained on more subjects. However, the improvement becomes negligible as the training set size increases, when there are seventy to eighty-five subjects. Observations show that the overall performance continues to improve beyond eighty-five training subjects. The presence of more training subjects may increase the likelihood of testing subjects exhibiting similar patterns to those in the training set, as well as improved predictive performance. We also include the performance changes of other GNN-based methods. Whereas all methods demonstrate performance improvements, SIFT-EEG benefits the most from the use of a greater number of training subjects, as this may facilitate the extraction of subject-independent patterns among the population. In contrast, other methods show fluctuating results and irregular patterns while underperforming SIFT-EEG consistently, mostly because neither of these methods bakes subject-independent features into the representation. Interesting to note that EEG-GAT reports considerably higher variances despite being the best-performing GNN

3. Functional Connectivity Empowered Intention Classification

member.

3.5 Conclusion

This chapter focused on task-adaptive modeling of brain networks with functional connectivity. We conducted extensive experiments with a large-scale EEG dataset to demonstrate that the proposed predictive model outperforms the state-of-the-art for MI classification in the subject-independent setting. Moreover, our investigation indicated that task-adaptive region selection produces similar predictive performance with only 20% of raw EEG data, with a considerable reduction in computation cost during model training and deployment.

Our future work may lend itself to the data scarcity problem in EEG-related research by examining other MI datasets with fewer subjects. We will examine how to apply the proposed subject-independent channel-selection methods more efficiently in situations where training subjects are limited.

3. Functional Connectivity Empowered Intention Classification

ALGORITHM 2: Task-relevant edge formation algorithm

Require: EEG segments set \mathcal{X} , adjacency matrices \mathcal{A} , intention set C , edge threshold T^e

1: **for** c in C **do**

2: Average the matrices A_c of all EEG segments within the same intention c

$$A_c = \text{mean}_{X \rightarrow c}(A), \forall A \in \mathcal{A}, \forall X \in \mathcal{X}$$

3: Sort edges E_{A_c} in A_c based on the absolute value of weights Λ in descending order

$$E_c = \text{sort}(\text{abs}(E_{A_c}))$$

4: Derive critical edges corresponding to the strongest associations using edge threshold T^e

$$E_c = \text{indices}(E_c(0 : t \times |E_c|))$$

5: **end for**

6: Merge critical edges together to produce task-relevant edges

$$E = \text{union}_{c \in C}(E_c)$$

3. Functional Connectivity Empowered Intention Classification

Table 3.1: Overall performances for subject-independent evaluation. All results are obtained over 12 runs with mean \pm std.

	Method	Accuracy	ROC-AUC
Baselines	CSP+LDA	64.73 \pm 1.85 _(+20.79%)	0.6471 \pm 0.019 _(+21.62%)
	RNN	54.77 \pm 3.47 _(+42.76%)	0.5374 \pm 0.010 _(+46.45%)
	EEGNet	72.29 \pm 8.10 _(+8.17%)	0.7202 \pm 0.082 _(+9.28%)
	CRAM	75.65 \pm 3.12 _(+3.35%)	0.7643 \pm 0.041 _(+2.97%)
	GHAM	77.31 \pm 3.49 _(+1.13%)	0.7785 \pm 0.035 _(+1.09%)
	GIN0	54.47 \pm 2.75 _(+43.55%)	0.5449 \pm 0.027 _(+44.43%)
	GraphSAGE	58.42 \pm 3.86 _(+33.84%)	0.5845 \pm 0.039 _(+34.66%)
Ablation	EEG-GAT	68.18 \pm 4.30 _(+14.68%)	0.6817 \pm 0.076 _(+15.44%)
	SIFT-EEG (w/o d)	61.87 \pm 1.72 _(+26.38%)	0.6187 \pm 0.017 _(+27.21%)
	SIFT-EEG (w/o t)	64.10 \pm 3.95 _(+21.97%)	0.6454 \pm 0.036 _(+21.95%)
	SIFT-EEG (Avg)	61.51 \pm 2.94 _(+27.11%)	0.6089 \pm 0.018 _(+29.26%)
	SIFT-EEG (Sum)	62.91 \pm 4.42 _(+24.30%)	0.6159 \pm 0.034 _(+27.79%)
	SIFT-EEG (Max)	64.10 \pm 6.44 _(+21.99%)	0.6365 \pm 0.056 _(+23.64%)
	SIFT-EEG	78.19 \pm 3.42	0.7870 \pm 0.035

Table 3.2: Comparison of the classification accuracy for task-adaptive and random selection of 20% of original channels.

	Model	Random	Task-Adaptive	All Channels
	CRAM	55.11	77.05	75.93
	GHAM	59.55	76.96	77.04
	SIFT-EEG	53.11	79.81	77.93

Table 3.3: The statistically significant tests. We find that 20 out of 24 comparisons are significant ($p \leq 0.05$), with results shown in **bold**.

Metric	Method					
Accuracy P-value	CSP+LDA	EEGNet	CRAM	HAM	GIN0	GraphSAGE
	0.00	0.03	0.07	0.54	0.00	0.00
ROC-AUC P-value	EEG-GAT	SIFTEEG (Avg)	SIFTEEG (Sum)	SIFTEEG (Max)	SIFTEEG (w/o d)	SIFTEEG (w/o t)
	0.00	0.00	0.00	0.00	0.00	0.00
Accuracy P-value	CSP+LDA	EEGNet	CRAM	HAM	GIN0	GraphSAGE
	0.00	0.02	0.16	0.56	0.00	0.00
ROC-AUC P-value	EEG-GAT	SIFTEEG (Avg)	SIFTEEG (Sum)	SIFTEEG (Max)	SIFTEEG (w/o d)	SIFTEEG (w/o t)
	0.00	0.00	0.00	0.00	0.00	0.00

3. Functional Connectivity Empowered Intention Classification

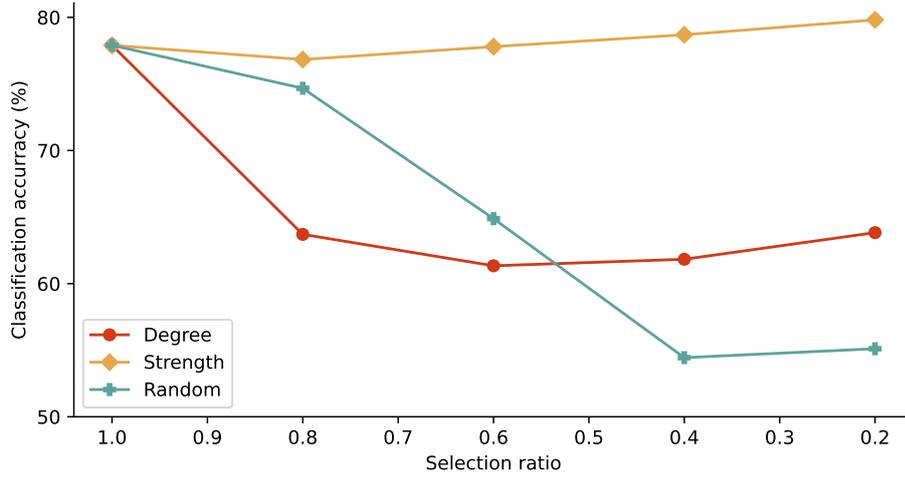


Figure 3.7: Classification accuracy for top t% task-relevant channels selected by the proposed algorithm.

Table 3.4: Comparison of model performance with different channel selection strategies applied to SIFT-EEG

Strategy	Accuracy	#Channels	Percent
Original	78.19 ± 0.034	64	100%
Random	62.00 ± 0.010	39	61%
Subject-specific [67]	75.89 ± 0.038	33	52%
Motor cortex region [68]	76.60 ± 0.019	18	28%
Task-adaptive (Degree)	63.83 ± 0.013	13	20%
Task-adaptive (Strength)	79.81 ± 0.045	13	20%

3. Functional Connectivity Empowered Intention Classification

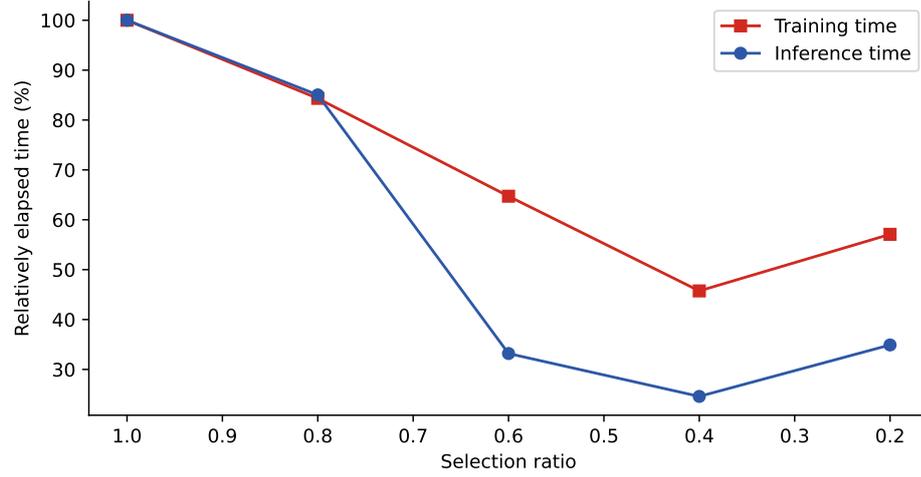
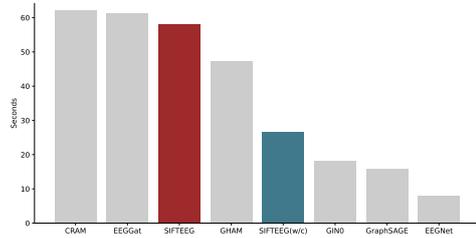
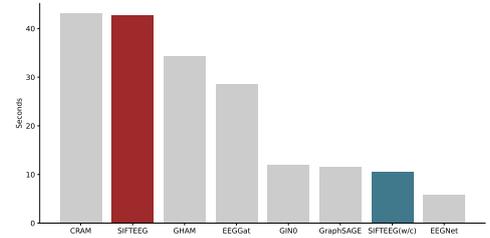


Figure 3.8: Relative elapsed train and inference time for SIFT-EEG under different channel selection ratios.



(a) Elapsed training time for 400 iterations.



(b) Elapsed inference time for 400 iterations.

Figure 3.9: Comparison of computation efficiency for total time taken for processing arbitrarily sampled 400 EEG segments.

3. Functional Connectivity Empowered Intention Classification

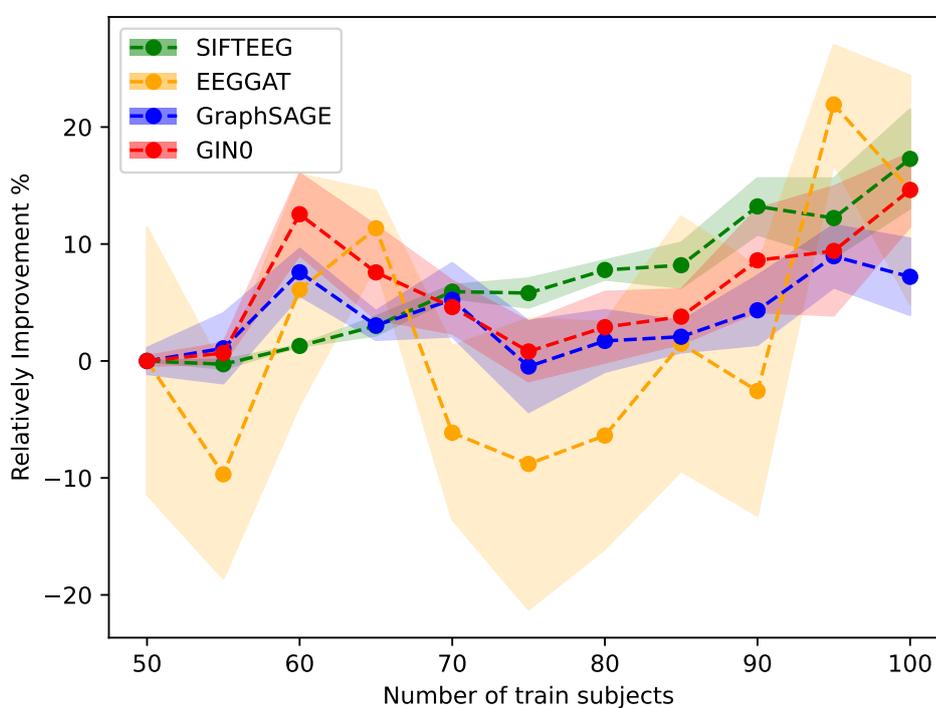


Figure 3.10: Impact of varying the number of training subjects. Every subject is included in the test set at least once for each train/test ratio. The results are reported with mean \pm std values of the relatively improved classification accuracy within each method itself.

Chapter 4

Utilization Towards Lightweight Machine Learning

4.1 Introduction

Chapter 3 proposed framework SIFT-EEG takes advantage of extracting subject-invariant features from the neurological topology domain. SIFT-EEG adopts 2D-CNN kernels to extract temporal dynamic features from raw EEG signals, as these kernels are widely chosen by the majority of SOTA studies [17, 19, 20]. Traditional CV tasks leverage the translation equivalence property of CNN ensuring that the object remains detectable after it translates position in the image. It helps navigate repetitive elements in images and makes them more detectable by DL-models [70]. EEG signals typically consist of time series signals from multiple channels. SOTA methods leverage 2DCNN to extract both temporal repetitive and channel-wise repetitive patterns to find useful patterns for classification. However,

4. Utilization Towards Lightweight Machine Learning

the SIFT-EEG framework has already used functional connectivity to estimate the correlational association between different channels and exclude those weak associations. Keep using 2DCNNs as the temporal encoder will lead to each channel's temporal dynamic features aggregating information from other channels, thus making it difficult to accurately estimate the state of its corresponding brain region. In Chapter 4, we have evaluated SIFT-EEG's performance on another dataset with fewer training subjects. Experimental results show that SIFT-EEG has a more than 10% classification accuracy reduction, whereas the CNN-based SOTA, CRAM [17] only reduces less than 3%. This phenomenon highlights SIFT-EEG's limitation in learning subject-invariant features with limited subjects and raises our concerns about SIFT-EEG's generalization ability. It also has the disadvantage of being computationally and storage intensive, thus necessitating the need to find an improved version of SIFT-EEG with reduced complexity in order to make it more cost-effective.

Subject-specific features are also essential for real-world BCI applications, as they enable DL models to understand users' brain activity and classify their intentions in a more personalized and accurate manner. However, training DL models is a costly and time-consuming process, especially due to the high labor cost of collecting EEG signals and labeling training data. It is too expensive and not practical to collect training data and train a personalized DL-model for every user from scratch. Therefore, learning features from EEG signals that share similar distribution with user's brain signals is essential for DL-models to capture more personalized features. There are a high number of similarity estimation metrics that can be used to calculate the similarity scores between pairs of EEG signals, such as distance, correlation, and hidden features. Despite this, the effectiveness of different similarity estimation metrics in navigating valuable EEG signals that contributory for DL-models to capture subject-specific features is yet to be explored. The outcome of such research

4. Utilization Towards Lightweight Machine Learning

is especially valuable for industrial communities, as they could adopt EEG signals selected by the most effective metric to make their DL-models more adaptive in recognizing users' personalized features, hence improving user experience.

Targeting each of these concerns, our study presents an improved version of SIFT-EEG, **Subject-Independent MI classification model using brain Functional Connectivity with One Dimensional CNN encoding block** (1D-SIFTEEG) illustrated in Figure. 4.2. On top of that, we heuristically select four similarity estimation metrics to navigate EEG signals similar to signals produced by targeting users. Then, training the DL model from scratch with top-K% EEG signals with the highest similarity scores. We then explore the effectiveness of these similarity estimation metrics in capturing subject-specific features by comparing the trained DL-model's performance in recognizing the target user's brain signal. Key contributions for Chapter 4 are:

- We propose an improved version of the SIFT-EEG framework, the 1D-SIFT-EEG framework, which further improves the performance in subject-independent motor imaginary classification. The framework is more improved when there are small groups of training subjects.
- We empirically compare the effectiveness of four similarity estimation metrics, which can be used to navigate EEG signals with high similarity and makes models remain well-performance even with fewer training data.
- We utilize the computational and storage cost of our proposed deep-learning framework, 1D-SIFTEEG with SIFT-EEG framework. Our result shows the improved version is more lightweight and cost-effective, making it easier to deploy on edge devices.

4.2 Preliminaries

4.2.1 Graph simulation for Brain network

EEG signals are collected from electrodes (channels in following contexts) that are distributed across the scalp, and each channel corresponds to a specific functional region inside the brain. The brain generates a variety of electrical signals through the collaboration of different functional areas. The electrical signals are received by neurons, which will be used to control the body to perform various activities [71, 22].

Functional regions can be easily navigated by placing channels closely with the region; however, the association across different regions is relatively hard to be captured. Some techniques, such as fMRI [72], can capture the association by measuring blood flow occurring with brain activity using a magnetic resonance imaging technique. However, it is not applicable to EEG-Based devices as it can only detect electric potential differences.

Graph $G = \langle V, E, W \rangle$ is an abstract structure in non-euclidean space that is made up of nodes V , edges E , and weights W . This means that the brain network topology can be represented by graphs. The nodes $\{v_i\}_{i=1:n} \in V$ refer to EEG scalp channels located on specific brain regions. There is a total channel count of n . Each edge $e_{ij} \in E$ represents the inter-regional relation between channel pairs (v_i, v_j) . The pairwise edge weight $w_{ij} \in W$ indicates the strength of each relation.

Each edge $e_{ij} \in E$ represents the inter-regional relation between channel pairs (v_i, v_j) . Accordingly, the pairwise edge weight $w_{ij} \in W$ indicates the strength of each relation.

4. Utilization Towards Lightweight Machine Learning

Graph adjacency A takes the form of a 2D matrix $R^{N \times N}$, where N denotes the number of nodes. One can be either a weighted or an unweighted graph. For unweighted adjacency, $A_{ij} = 1$ represents a pair of nodes are connected while $A_{ij} = 0$ means there is no connection between. A weighted adjacency has an additional attribute to the strength of relationship by setting $A_{ij} = w_{ij}$. Degree $d(v_i)$ measures a node's centrality by the number of edges connecting to it, indicating the importance of node within the graph, calculated by adding up the i -th row of A .

Edge definitions for EEG-based graph representations are yet to be provided. Literature mostly use either complete graph [21], i.e., each pair of nodes are connected, or distance-based connections [19, 20], that is, two nodes connected if their physical distance between channels is lower than a pre-defined threshold.

4.2.2 1D & 2D CNN Embedding Block

Convolutional Neural Network (CNN) is one of the most powerful and commonly used deep learning models due to its outstanding ability in feature extraction [73]. Each CNN layer consists of several CNN filters/kernels (mathematically a matrix) to extract features from the input.

Two dimensional convolutional neural network (2D-CNN) is the most well-known approach to resolving tasks such as image classifications, embedding, and object detection. The kernel inside 2D-CNN layer k is a 2D matrix, that move both horizontally and vertically over input data X . The input can be seen as another, larger 2D matrix compared with the kernel. By splitting the data into pieces of sub-region, and let kernel will perform dot product with each of the sub-regions. Each of output from dot-product will capture the correlation inside that sub-regions; After combin-

4. Utilization Towards Lightweight Machine Learning

ing all output together, we will get a new matrix, O , which contains aggregated features from input data.

One dimensional convolutional neural network (1D-CNN) is more advantageous in some applications such as anomaly detection, signal processing, and speech recognition [74]. Additionally, 1D-CNN requires also takes advantage in terms of computational complexity. To process an $N \times N$ image with a $K \times K$ kernel, 2D-CNN will have a complexity of $O(N^2K^2)$ and 1D-CNN will only have a complexity of $O(NK)$. The kernel inside the 1D-CNN layer k is a 1D matrix and only captures the feature in the temporal domain.

Many frameworks in EEG-based BCI adopt 2D-CNN as an encoding layer to extract information from raw signals. However, the effectiveness of this approach is doubtful since recognizing EEG signals are time series classification. Each segment of EEG signal acquired by signal collection devices $X \in \mathbb{R}^{N \times W}$, contains N time-series data from each of channels and W time points depending on the recording time. Directly applying 2D-CNN on raw signals will not only capture the temporal correlation but also the channel-wise correlation when it moves vertically. Move CNN kernel vertically will collect no valuable information other than how the readings from each channel are ordered. This will make the extracted features contains redundant and irrelevant information, which can impact the performance of the overall framework. Fig. 4.1 shows the difference between 1D-CNN and 2D-CNN, we use the blend color to illustrate the features with combined channels collected by 2D-CNN kernels.

4. Utilization Towards Lightweight Machine Learning



Figure 4.1: Comparison of feature extraction between 2D & 1D CNN layers

4.2.3 Graph Neural Network

Graph Neural Networks (GNNs) are a family of generalised neural networks excelling at analysing graph-structured data.

Graph Convolutional Network (GCN) [53] simplify spectral graph convolutions to the framework of Neural Networks, which defines a GCN operator to exchange node-wise information through edge connections, along with a layer-wise propagation rule that updates hidden node features. Throughout the K -layer propagation process, the nodes receives the averaged features from its neighbours via a mean-aggregation, and the shape of the graph structure remains the same at the next layer.

$$h_v^{(k+1)} = \sigma (W \cdot \text{MEAN} \{h_v^{(k)} \cup h_n^{(k)} \mid n \in \mathcal{N}(v)\}) \quad (4.1)$$

where $h_v^{(k+1)}$ is node v 's hidden feature at k -th GCN layer, W are model parameters and σ is non-linear activation. $\text{READOUT}(\cdot)$ function is further applied after propagating the last layer to extract the graph-level embedding,

$$h_G = \text{READOUT} (\{h_v^{(K)} \mid v \in V\}) \quad (4.2)$$

GraphSAGE [50] interpret GNNs from a spatial perspective and generalize the mean-

4. Utilization Towards Lightweight Machine Learning

aggregator in GCN to a wider range of operators,

$$h_v^{(k+1)} = \sigma \left(W \cdot \text{AGG} \{ h_v^{(k)} \cup h_n^{(k)} \mid n \in \mathcal{N}(v) \} \right) \quad (4.3)$$

where W are model parameters and $\text{AGG}(\cdot)$ refers to a permutation invariant function such as min/max/mean pooling.

The recent Graph Isomorphism Network (GIN) [26] is derived from the Weisfeiler-Lehman (WL) isomorphism test for checking if graphs are topologically identical, proving that GIN are as powerful as the WL test when $\text{AGGREGATE}(\cdot)$ and $\text{READOUT}(\cdot)$ are permutation invariant and injective, which produces more discriminative embeddings than other GNNs variants in graph-level classification.

Assume the function $f : \mathcal{X} \rightarrow \mathbb{R}^n$ for any countable node feature space \mathcal{X} . Then for infinitely many choices of ϵ including all irrational number, such that any function g can be decomposed into $g(c, X) = \phi((1 + \epsilon) \cdot f(c) + \sum_{x \in X} f(x))$ for some function ϕ , where $c \in X$ and $X \subset \mathcal{X}$. A multilayer perceptron (MLP) with more than one hidden layer can be used to approximate injective function [54], so the node embedding h_v becomes distinguishable, using GINConv operator defined by

$$h_v^{(k+1)} = \text{MLP}^{(k)} \left((1 + \epsilon^{(k)})h_v^{(k)} + \sum_{n \in \mathcal{N}(v)} h_n^{(k)} \right) \quad (4.4)$$

where the $\text{AGG}(\cdot)$ of neighbor nodes is implemented as summation to impose injective mapping.

4.2.4 Generalization in Machine Learning

Generalization in machine learning refers to the ability of a trained model can be applied to unseen data that is different from training data with low errors. It can

4. Utilization Towards Lightweight Machine Learning

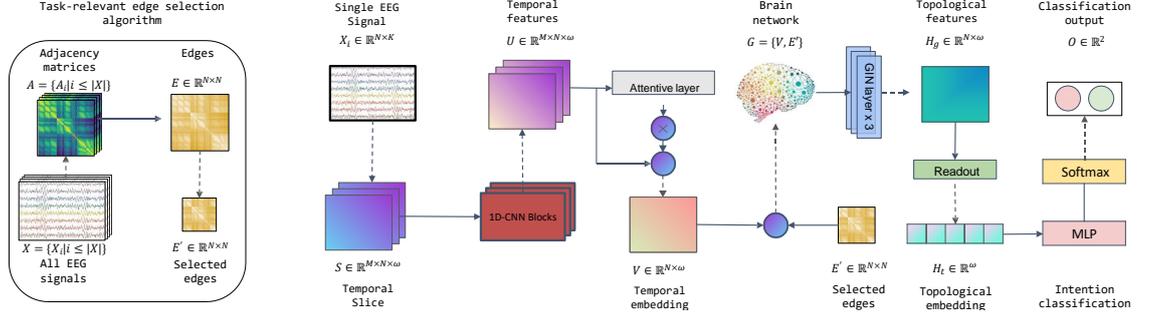


Figure 4.2: Overview of the **1-Dimensional Subject-Independent MI** classification model using brain **F**unctional **C**onnectivity (**1D-SIFT-EEG**) model. We first crop raw EEG signal into a sequence of the temporal slice using slide window technique; adopt 1D-CNN to extract temporal features and the self-attentive module to search most discriminative temporal slice; then we combine nodes in temporal embedding with functional connectivity to generate the graph representation of EEG signals; select three layer of graph isomorphism network(GIN) to extract topological features and lastly the extracted topological embedding are classified to different motion intention using a fully connected network with softmax activation function.

be also seen as the wellness of a model’s ability to extract useful features from raw data. Generalization ability is a challenging aspect in EEG-based BCI systems since the EEG signal is non-stationary [75]. This means that the statistical characteristics of the EEG signal change over time, which results in the trained model being more likely to overfit the training data. The difference between each person results in a further difference between training and testing samples under the subject-independent setting.

There are many common approaches to improve a model’s generalization ability in machine learning, such as: using more training data, applying data augmentation, early stopping the training process, or using regularization techniques [76]. However, collecting/augmenting more EEG data will cause additional costs and hyper-parameter tuning is not easily transferable to other experiments.

Network reduction (simplifying the model) becomes the most practical and cost-

4. Utilization Towards Lightweight Machine Learning

effective way to avoid over-fitting. By reducing the complexity of the classification by eliminating less meaningful or irrelevant data, it is possible to prevent overfitting and improve classification accuracy [76, 77].

4.2.5 Similarity

The estimation of similarity is a crucial step in machine learning. Different similarity estimation metrics will significantly influence the selected training EEG segments and hence, impact the performance of the machine learning model. In this study, we select four different similarity metrics to estimate the similarity and investigate their effectiveness in navigating the most relevant EEG segments.

Distance-based similarity is one of the easiest-to-understand similarity metrics. It measures the similarity between two data points in euclidean space based on the distance between them. The more similar the data points are, the closer they are to each other.

Correlation-based similarity measures how two sets of variables are correlated. Since each EEG segment contains readings from N channels. In order to make a fair comparison, we have to calculate the correlation for each channel using the Pearson coefficient and sum all channel's correlations together to obtain segment-level correlational similarity.

Autoencoder-based similarity can find similar items from a dataset with help from the deep learning model Autoencoder [78]. An autoencoder can effectively encode data to a lower-dimension space, which can effectively compress the data and learn features. Autoencoder-based approach has been widely investigated and deployed in many applications include recommender system [79], computer vision [80] and

4. Utilization Towards Lightweight Machine Learning

natural language processing [81].

Hidden feature-based similarity takes similar ideas with autoencoder-based similarity. It will first encode raw EEG segments to a lower-dimension space, and find EEG segments that have similar hidden features in the lower-dimension space. The only difference between the autoencoder-based and hidden feature-based approach is it uses a pre-defined deep learning model for encoding rather than autoencoder.

4.3 Methodology

4.3.1 Overview

We now describe the goal of EEG-based Motor Imagery (MI) classification and key steps of our framework. Given an EEG segment $X \in \mathbb{R}^{N \times K}$ collected while a subject was performing a MI task, we aim to estimate the associated specific intention Y by training a predictive model that performs supervised classification under the subject-independent setting, meaning that model training and evaluation are conducted on two disjoint groups of subjects. $K = T \times f$ is the number of time points within a segment, N is the number of channels, T is the recording duration, and f is the sampling frequency. Let X be an EEG segment and Y be the intention, our approach approximates the mapping $f : X \rightarrow Y$, parameterized by a Neural Network $\hat{Y} = f_{\Theta}(X)$, by involving the following five steps:

- 1) Calculating adjacency $A \in \mathbb{R}^{N \times N}$ across the readings of all sensory channels X based on functional connectivity;
- 2) Selecting the top- N' most relevant EEG segments $X' \subset X$, $X' \in \mathbb{R}^{N \times K}$ tailored

4. Utilization Towards Lightweight Machine Learning

Layer	Kernel	Stride	Shape
Input			(N, 400)
Conv1D	20	2	(N, 192)
Conv1D	20	2	(N, 87)
Conv1D	6	1	(N, 80)
AvgPool1d	3	2	(N, 40)
Conv1D	6	1	(N, 35)

Table 4.1: Architecture of the temporal embedding block, 1DCNN(\cdot)

to the MI task and identifying task-relevant edges $E \in \mathbb{R}^{N \times N}$;

- 3) Extracting temporal node embedding $V \in \mathbb{R}^{N \times W}$, by summarizing the features of each temporal slice $S_m \in \mathbb{R}^{N \times W}$, where $S = \{S_m\}_{m=1}^M$ results from sliding window applied on raw EEG-segments X ;
- 4) Generating topological graph embedding $H \in \mathbb{R}^h$ from the brain functional network $G = \langle V, E \rangle$ constructed by temporal node embeddings and task-relevant edges;
- 5) Predicting the intention $f(X) = \hat{Y}$ of an EEG segment from the graph-level topological embedding.

4.3.2 Functional Adjacency Matrix

We first calculate the adjacency of EEG-based brain network with functional connectivity. We prefer this approach to previous distance-based or complete graph connections that either disregard topological relationships or are very computationally intensive if too many nodes are involved. Functional connectivity defines the statistical dependencies among temporal signals, commonly measured as the simi-

4. Utilization Towards Lightweight Machine Learning

ilarity between two brain regions by using the Pearson coefficient:

$$P(i, j) = \frac{\text{Cov}(i, j)}{\sqrt{\text{Var}(i)\text{Var}(j)}} \quad (4.5)$$

where $\text{Cov}(i, j)$ is the covariance of measurement readings between sensory channels i and j , $\text{Var}(i)$ denotes the standard deviation of i -th channel readings throughout K timesteps.

This results in a complete functional adjacency $A \in \mathbb{R}^{N \times N}$ with N being the number of channels. The following section discusses how we further reduce nodes and edges by only keeping task-relevant channels.

4.3.3 Navigate similar EEG-segments

In order to find the most practical way in finding similar EEG segments, we select four commonly used metrics to make the evaluation. Firstly, we split the original dataset into source X_S and target X_T domain. Then, we calculate the similarity between each EEG segment in the source and target domain. After that, we sort EEG segments in the source domain based on the calculated similarity metric in ascending order and select some of the most similar segments. In the end, we train the model using selected segments and evaluate the performance in the target domain.

Distance-based metric

To calculate the Distance-based similarity S_D , we first need to flat the EEG segments, $X_i \in [N, K] \Rightarrow [N \times K]$. For each EEG segment in source domain X_{S_i} , we calculate the distance for all T segments in the target domain and summarize

4. Utilization Towards Lightweight Machine Learning

all distances together to obtain distance-based similarity S_{d_i} for EEG segment i in the source domain. Unlike distance, a high correlation strength symbolizes strong associations, which means EEG segments are strongly correlated. Thus, we have to sort the obtained metrics in descending order to find the most similar segment.

$$S_{D_i} = \sum_{j=1}^T \text{abs}(X_{S_i} - X_{T_j})$$

Correlation-based metric

The correlation-based similarity S_C needs to calculate the channel-level correlation strength using the Pearson coefficient in the EEG segments $X \in [N, K]$. Then, summarize all channel-level together to obtain segment-level correlation. The correlation-based similarity S_{c_i} for EEG segment i is finally obtained by summarizing all segment-level correlation in target domain X_T together.

$$S_{C_i} = \sum_{j=1}^T \sum_{k=1}^N \text{Pearson}(X_{S_i}^k, X_{T_j}^k)$$

Autoencoder-based metric

Autoencoder-based similarity S_A requires a fully trained Autoencoder model. We follow the same architecture with the recently proposed autoencoder framework for EEG signals [82] and trained the model using our training data. We encode the raw EEG segments into lower-dimensional hidden embedding using the well-trained autoencoder, $H = \text{AE}(X)$. We then calculate the distance between the hidden embedding to estimate the similarity between EEG segments in the source and

4. Utilization Towards Lightweight Machine Learning

Table 4.2: Overall performances for subject-independent evaluation using EEGM-MIDB dataset. All results are obtained over 12 runs with mean \pm std.

Method	Accuracy	ROC-AUC
CSP+LDA	64.73 \pm 1.85 _(+24.32%)	0.6471 \pm 0.019 _(+24.40%)
EEGNet	72.29 \pm 8.10 _(+11.31%)	0.7202 \pm 0.082 _(+11.77%)
CRAM	75.65 \pm 3.12 _(+6.37%)	0.7643 \pm 0.041 _(+5.32%)
GHAM	77.31 \pm 3.49 _(+4.09%)	0.7785 \pm 0.035 _(+3.40%)
GIN0	54.47 \pm 2.75 _(+47.73%)	0.5449 \pm 0.027 _(+47.73%)
GraphSAGE	58.42 \pm 3.86 _(+37.74%)	0.5845 \pm 0.039 _(+37.72%)
EEG-GAT	68.18 \pm 4.30 _(+18.02%)	0.6817 \pm 0.076 _(+18.09%)
SIFT-EEG	78.19 \pm 3.42 _(+2.91%)	0.7870 \pm 0.035 _(+2.28%)
1D-SIFTEEG	80.47 \pm 2.61	0.8050 \pm 0.026

target domain.

$$S_{A_i} = \sum_{j=1}^T \text{abs}(H_{S_i} - H_{T_j})$$

Hidden Feature-based metric

Hidden Feature-based similarity S_H requires a fully trained deep learning model in analysing EEG signals. We use our recently proposed SIFT-EEG as the encoding model and trained the model using our training data. This model can be replaced with any deep learning models as long as it does not underfit the training data. We encode the raw EEG segments into lower-dimensional hidden embedding using SIFT-EEG, $H = \text{SIFTEEG}(X)$. We then calculate the distance between the hidden embedding using similar setting in Autoencoder-based similarity.

$$S_{H_i} = \sum_{j=1}^T \text{abs}(H_{S_i} - H_{T_j})$$

4.3.4 Temporal Embedding

The concentration periods between subjects during MI may differ, leading to different temporal properties even if two subjects were performing the same task. [17]. Our objective is to identify the most discriminatory period within each EEG segment of different subjects, which we refer to as the subject-invariant temporal embedding.

More specifically, we first crop the EEG segment $X \in \mathbb{R}^{N \times K}$ into M temporal slices $S \in \mathbb{R}^{M \times N \times W}$ using a sliding window of size W . We then create initial temporal features U_m of each temporal slice $S_m \in \mathbb{R}^{N \times W}$ by applying four 1D convolution layers, each followed by an elu non-linear activation and a 1d Batch normalization regularization function. Table. 4.1 shows the architecture for the embedding block. The temporal embedding U_m extracted by the block $1\text{DCNN}(\cdot)$ is smaller than the input slice S_m .

$$U_m = 1\text{DCNN}(S_m), \quad \text{for } m = 1, \dots, M \quad (4.6)$$

Next, we use self-attention [18] to capture slice-wise correlations and adapt temporal weights to their temporal features, leading to a temporal embedding $V \in \mathbb{R}^{N \times W}$ that summarizes the EEG segment,

$$V = \sum_{m=1}^M \frac{\exp(H_m^\top W_m)}{\sum_{j=1}^M \exp(H_j^\top W_j)} U_m \quad (4.7)$$

with

$$H_m = W_h U_m + b_h \quad (4.8)$$

where W_m , W_h and b_h are learnable parameters. The temporal embedding V aggregates all of M slices with taking into account each slice’s importance, which derives the most discriminative representation upon input values.

4. Utilization Towards Lightweight Machine Learning

Table 4.3: Overall performances for subject-independent evaluation using BCICOM-PIV2A dataset. All results are obtained using Leave-One-Out Cross-Validation over 9 subjects with mean \pm std.

Method	Accuracy	ROC-AUC
CSP+LDA	$50.52 \pm 4.64_{(+51.07\%)}$	$0.5050 \pm 0.046_{(+36.89\%)}$
EEGNet	$65.93 \pm 11.12_{(15.76\%)}$	$0.6170 \pm 0.096_{(12.04\%)}$
CRAM	$74.14 \pm 5.66_{(2.94\%)}$	$0.6633 \pm 0.047_{(4.22\%)}$
GHAM	$64.90 \pm 1.99_{(17.60\%)}$	$0.5371 \pm 0.040_{(28.71\%)}$
GIN0	$56.02 \pm 2.78_{(36.24\%)}$	$0.5068 \pm 0.040_{(36.62\%)}$
GraphSAGE	$51.04 \pm 1.37_{(49.53\%)}$	$0.5000 \pm 0.000_{(38.26\%)}$
EEG-GAT	$55.43 \pm 2.29_{(37.69\%)}$	$0.5239 \pm 0.020_{(31.95\%)}$
SIFTEEG	$62.39 \pm 1.69_{(22.33\%)}$	$0.4954 \pm 0.025_{(39.54\%)}$
1D-SIFTEEG	76.32 ± 5.29	0.6913 ± 0.061

4.3.5 Topological Embedding

Brain activity involves the joint efforts of multiple brain regions [49]. While CNN-based approaches [13, 4, 17] assume Euclidean-structure of EEG channels to capture such inter-regional correlations, these works ignore the natural geometry of brain structure and connections between different regions beyond their immediate vicinity. As opposed, graph-based brain representations appear to better reflect the non-Euclidean nature of human’s scalp [21], but also encode the subject-invariant positioning priors of channels [19, 20] to the model. Nevertheless, their graph construction cannot represent dynamic functional connectivity that adapts to different MI tasks, since the edges therein are fixed in terms of the distances between nodes.

Task-Adaptive Edge Formation.

In the brain network, correctly illustrated edges could model how each brain region associates with other active regions during a MI task. Meanwhile, channel-wise

4. Utilization Towards Lightweight Machine Learning

correlations across MI intentions do not remain constant as aforementioned. Having identified the task-relevant channels N , we now formulate the task-relevant edges $E \in \mathbb{R}^{N \times N}$ to leverage strong associations tailored to the task [52], in the following stages:

1. *Constructing*: Calculate channel-wise correlation adjacency $A_c \in \mathbb{R}^{L \times S \times N \times N}$ of all subjects for each intention $c \in C$, using the Pearson coefficient, where L denotes the number of labels, S is the number of subjects, and N represents the number of channels.
2. *Generalizing*: Find the generalized connectivity across all subjects by averaging the adjacency matrices of all subjects for each intention c .
3. *Thresholding*: Preserve critical connections that exceed the threshold T^e for each intention c .
4. *Merging*: Derive task-relevant edges E' by merging the critical connections of all C intentions.

Algorithm 3 shows the pseudocode to derive a purely functional adjacency E . Moreover, consider the temporal embedding V results from the raw EEG segments X , we construct the task-adaptive brain network $G = \langle V, E \rangle$ to obtain topological embedding.

Embedding Computation.

We adopt a L -layer GIN to investigate topological embedding of a given brain network G , as it shows promise for graph-level classification [26, 60]. The graph-level

4. Utilization Towards Lightweight Machine Learning

Table 4.4: The statistically significant tests. We find that 29 out of 32 comparisons are significant ($p \leq 0.10$), with results shown in **bold**.

Dataset	Metric	Method			
EEGMMIDB	Accuracy P-value	CSP+LDA	EEGNet	CRAM	GHAM
		0.00	0.00	0.00	0.02
	ROC-AUC P-value	GIN0	GraphSAGE	EEG-GAT	SIFTEEG
		0.00	0.00	0.00	0.08
	Accuracy P-value	CSP+LDA	EEGNet	CRAM	GHAM
		0.00	0.00	0.01	0.05
ROC-AUC P-value	GIN0	GraphSAGE	EEG-GAT	SIFTEEG	
	0.00	0.00	0.00	0.17	
BCICOMP1V2a	Accuracy P-value	CSP+LDA	EEGNet	CRAM	GHAM
		0.00	0.02	0.47	0.00
	ROC-AUC P-value	GIN0	GraphSAGE	EEG-GAT	SIFTEEG
		0.00	0.00	0.00	0.00
	Accuracy P-value	CSP+LDA	EEGNet	CRAM	GHAM
		0.00	0.09	0.33	0.00
ROC-AUC P-value	GIN0	GraphSAGE	EEG-GAT	SIFTEEG	
	0.00	0.00	0.00	0.00	

topological embedding $H \in \mathbb{R}^W$ after L -layers propagation is produced by

$$H = \text{READOUT} \{h_v^{(L)} \mid v \in V\} \quad (4.9)$$

with

$$h_v^{(l)} = \text{MLP}^{(l-1)} \left((1 + \epsilon^{(l-1)})h_v^{(l-1)} + \sum_{n \in N(v)} h_n^{(l-1)} \right) \quad (4.10)$$

where W is the embedding dimension. We set MLP with 2 hidden layers, $\epsilon = 0$, and $k = 3$ in practice. While READOUT(\cdot) function can be either non-injective aggregations (e.g., pooling) or injective mappings (e.g., MLP(\cdot)), the latter leads to more discriminative graph-level embedding, shown by [26]. Our empirical studies experiment with different setups to verify this argument in EEG-based applications.

4. Utilization Towards Lightweight Machine Learning

Method	1	2	3	4	5	6	7	8	9
EEGNet	66.67	61.03	53.28	55.04	67.44	68.14	90.23	59.09	72.41
GGRAM	71.74	69.12	86.13	72.09	74.42	67.26	75.19	79.55	74.14
GHAM	64.49	65.44	62.77	66.67	67.44	61.95	66.92	62.88	65.52
GNNEEG	55.07	53.68	61.31	58.14	54.26	53.98	58.65	53.03	56.03
GraphSAGE	50.00	50.74	50.36	51.94	51.16	50.44	50.38	50.00	54.31
EEGGat	57.97	55.88	54.74	50.39	54.26	55.75	57.89	56.82	55.17
SIFTEEG	61.59	62.50	58.39	62.79	62.02	63.72	63.91	63.64	62.93
1D-SIFTEEG	76.81	72.06	84.67	75.19	77.52	69.91	75.94	84.09	70.69

Table 4.5: Leave-One-Out Cross-Validation (LOOCV) evaluation to the performance of the BCICOMP IV2A dataset using nine subjects. The results of the evaluation are presented in the form of accuracy scores for each subject, which were calculated when that subject was in the testing set.

Metric	1	3	5	7	9	Mean
Distance	53.65 ± 8.49	55.32 ± 9.68	55.24 ± 8.77	56.28 ± 9.36	56.78 ± 9.58	55.45
Correlation	62.41 ± 7.86	66.09 ± 7.28	66.90 ± 7.49	68.31 ± 7.46	67.35 ± 6.42	66.22
Autoencoder	62.04 ± 7.20	63.69 ± 6.95	65.32 ± 6.40	65.76 ± 6.56	65.64 ± 6.84	64.49
Hidden feature	60.28 ± 5.64	65.95 ± 6.48	67.25 ± 6.62	66.61 ± 7.10	67.75 ± 6.35	65.57
Overall	59.60 ± 3.53	62.76 ± 4.40	63.68 ± 4.93	64.24 ± 4.69	64.38 ± 4.46	62.93

Table 4.6: Comparison of performance when model is trained using the most similar EEG segments, as determined by a similarity metric.

4.3.6 Intention Classification

We take the topological embedding H of EEG segment X to perform intention prediction, with a softmax function to estimate the probabilities of each possible intentions,

$$P_c = \text{softmax}(W_o H + b_o), \quad \text{for } c = 1, \dots, C \quad (4.11)$$

where W_o and b_o are trainable parameters. The predicted intention is thus given by $\hat{Y} = \arg \max_c(P_c)$.

4.4 Experiment setting

4.4.1 Dataset

In this work, we evaluate the performance of the proposed method on widely studied large-scale cross-subject EEG-based MI dataset eegmmidb (EEG motor movement/imagery database) downloaded from the Physionet database and a relatively small-scaled dataset with fewer number of subjects namely BCICOMP IV2a (BCI Competition IV dataset-2a) downloaded from BCI Competition.

EEGMMIDB was collected using BCI2000 [33] containing 64 channels with sampling frequency of 160Hz from 109 healthy subjects. Following the conventions [14], [15], we remove subjects 88, 89, 92, 100 from the dataset due to consecutively resting states. Within three sessions of MI tasks, EEG signals were recorded when subjects were executing left/right fist open and closed imagery. There are approximately fifteen 3.1-second segments of imaginary left/right-hand movement in each session; each subject performed three sessions. We randomly select 90 subjects and use their EEG segments as training set, whereas the remain 15 subjects are used as test set. There are no disjoint subjects in the training and testing set, which ensures the evaluation is conducted under a subject independent setting. We experiment with twelve different train/test splits to reduce randomness in the results. We fix a unique random seed for each split that specifies the training and testing subjects. Noticeably, the last split includes all untested subjects from the previous eleven splits to ensure each subject is tested at least once.

BCICOMP IV2a was collected using [DeviceName] containing 22 channels with sampling frequency of 250Hz from 9 healthy subjects. Within six sessions of MI tasks,

4. Utilization Towards Lightweight Machine Learning

EEG signals were recorded when subjects were executing left/right fist open and closed imagery. There are approximately twenty-four 7.5-seconds segments of imaginary left/right-hand movement in each session; each subject performed six sessions. Due to this dataset only containing small number of subjects, we chose to use Leave-One-Out cross validator in evaluation. We select EEG segments from 8 subjects as training set, and the remaining 1 subject are used as testing set. There are no disjoint subjects in the training and testing set, which ensures the evaluation is conducted in a subject-independent setting. We experiment with each one of 9 subjects and take the mean value as result to make fair comparison of performance.

4.4.2 Implementation Detail

Each EEG segment in EEGMMIDB dataset $X \in \mathbb{R}^{64 \times 496}$ contains 496 timesteps with 64 channels. For temporal embedding, we apply the slide window technique with window size of 400 and the step size is 10. Hence, the input signal contains ten temporal slices, where each slice has the shape of [64,400] (i.e., $N = 64, W = 400$).

Segments in BCICOMPIV2a dataset $X \in \mathbb{R}^{22 \times 1875}$ contains 1875 timesteps with 22 channels. We apply slide window technique with same window size and the step size is 50. In the end, each sliced signal contain thirty temporal slices, and each slice has the shape of [22,400] (i.e., $N = 22, W = 400$).

All the models are implemented with PyTorch¹ and trained on an NVidia TITAN X in a fully-supervised manner. We use cross-entropy as the objective function and optimize model parameters using Adam with a learning rate of 0.001. The training batch size is 500. Each model in EEGMMIDB and BCICOMPIV2a is trained for

¹<https://pytorch.org>

4. Utilization Towards Lightweight Machine Learning

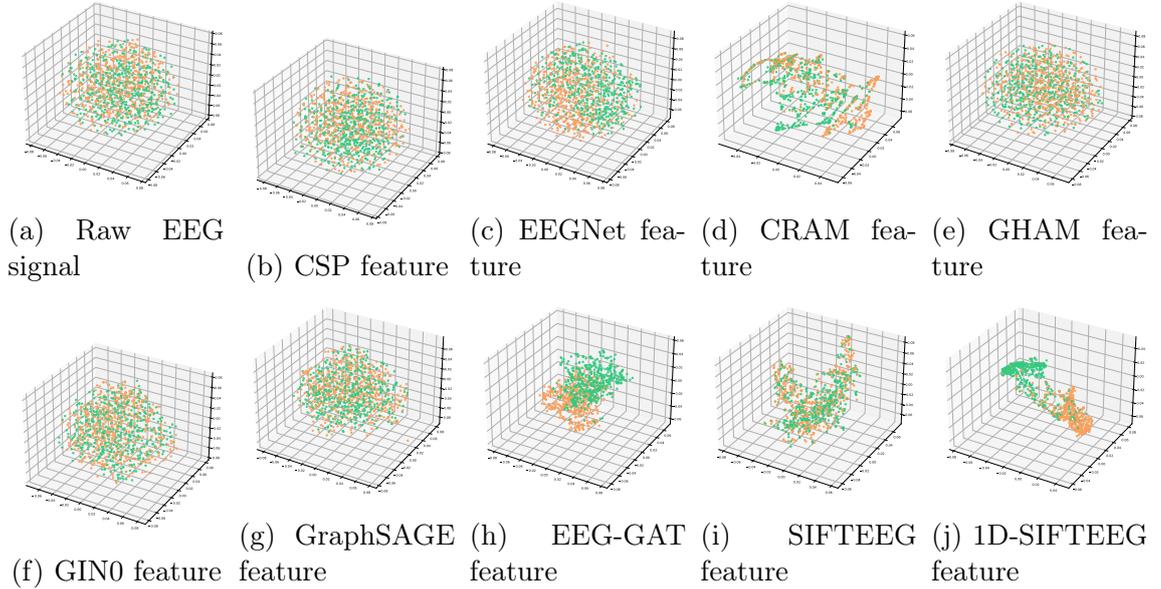


Figure 4.3: The visualization of hidden features with t-SNE. Green dots refer to left-hand imaginary motion, while orange dots indicate right-hand imaginary motion.

120 and 500 epochs respectively, and the dropout probability is set to 0.5 to avoid over-fitting. We implement all GNN-related components using `Torch_Geometric`².

4.5 Result & Discussion

The proposed 1D-SIFT-EEG is empirically compared with a range of baselines across two datasets, focusing on three research questions.

- 1) Does the proposed model with 1D-CNN embedding block benefit model’s generalization ability and performance?
- 2) What is the most efficient similarity metric to utilize in order to minimize the computational expense of the training phase and detect analogous EEG

²<https://pytorch-geometric.readthedocs.io>

4. Utilization Towards Lightweight Machine Learning

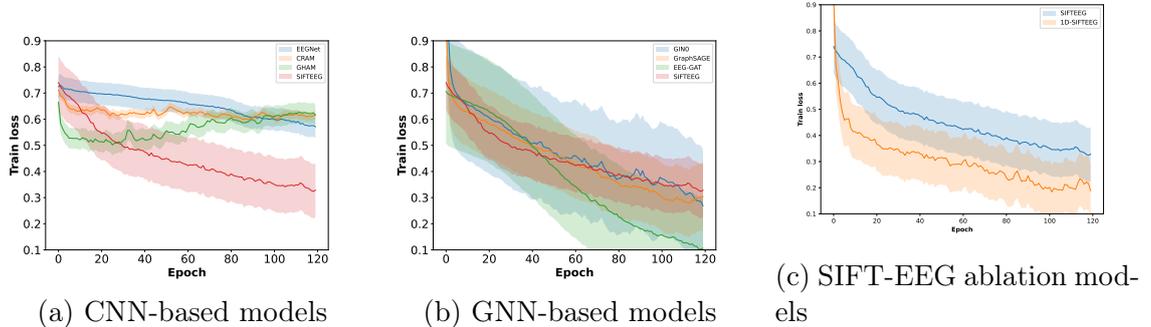


Figure 4.4: Comparison for training loss change when the number of training epochs increases.

segments to enhance performance?

- 3) Does the proposed model demonstrate a lightweight and computationally cost-effective comparison to the previous version (SIFT-EEG)?

The evaluation metrics include classification Accuracy and the Area Under ROC-Curve(ROC-AUC). All models are trained and evaluated with the same setting for a fair comparison.

4.5.1 Overall Performance

We compared our proposed 1D-SIFT-EEG framework with several state-of-the-art models, including both traditional and DNN-based approaches. We adopt the Common spatial pattern and linear discriminant analysis (CSP+LDA) [62] as the representative traditional means. Specifically, the CNN-based methods include EEG-Net [12], CRAM [17], GHAM [20]. The GNN members GIN0 [26], GraphSAGE [50] and EEG-GAT [51] apply different GNNs under the same framework [21]. Plus our recently proposed framework (SIFT-EEG), respectively. In addition, all baselines

4. Utilization Towards Lightweight Machine Learning

include mechanisms to represent the spatial correlation between EEG channels. It follows that such representations should be incorporated into modelling.

Table 4.2 reports the model performance of all approaches on both metrics for the EEGMMIDB dataset. It can be expected that 1D-SIFT-EEG will outperform all baseline models, with 4.09%, 18.02%, and 2.91% improvement in accuracy, as well as 3.40%, 18.09%, and 2.28% improvement on ROC-AUC in comparison with the best-performing CNN-based, GNN-based, and recently proposed SIFT-EEG methods, respectively.

Table 4.3 presents the model performance of all approaches for the BCICOMPIV2a dataset, and Table 4.5 contains subject-specific evaluation when each one of nine subjects is the testing set. It is observed that 1D-SIFTEEG outperforms the best-performed CNN-based, GNN-based and recently proposed SIFT-EEG methods by 2.94%, 26.24%, and 22.33% on the accuracy, and 4.22%, 31.95%, and 39.54% on ROC-AUC, respectively.

Fig. 4.4 depicts how the training loss changes with the number of training epochs increase. GNN-based methods generally have lower training losses and faster convergence rates compared with CNN-based methods, which suggests the capability of topological features in task-specific predictions. However, their test performances are worse than CNN-based methods, showing a tendency to overfit the training data. The reason may be that GNN-based methods extract topological features directly from raw EEG signals. Recall that our evaluations take place in a subject-independent setting. In this case, the variances in temporal patterns between subjects, i.e., graph noises, may cause them to perform inconsistently [63, 64]. Conversely, CNN-based methods explicitly handle temporal correlations before looking at the spatial domain. For instance, CRAM locates discriminative temporal features

4. Utilization Towards Lightweight Machine Learning

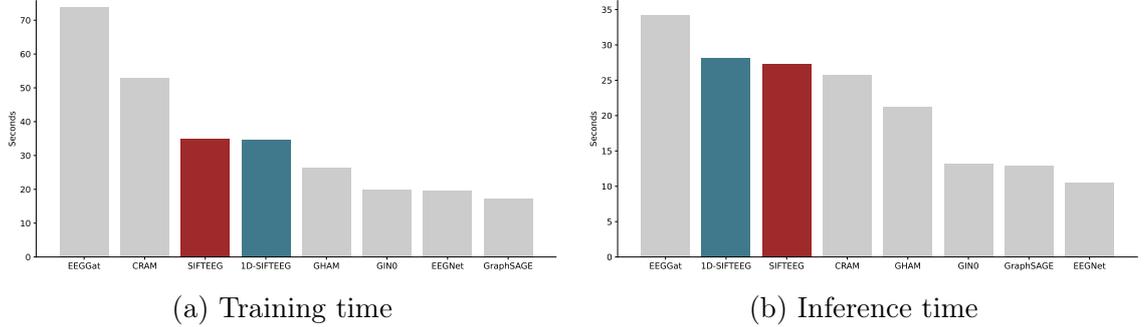


Figure 4.5: Computational cost comparison for different methods

adaptively for different subjects by using attention. This eases the burden on the spatial feature extractor and generalizes the model to new subjects, albeit fitting these models takes longer. The proposed SIFT-EEG combines the merits of both, leveraging flexible topological features of the brain signal, as well as subject-adaptive temporal features that reduce task-irrelevant noise.

In addition, we perform statistically significant tests to evaluate the model performance improvements of SIFT-EEG over baselines. We use the pairwise t-test, assuming the pairwise difference is significant if p -value is less than 0.10. The results are reported in Tab. 4.4, where statistically significant differences are **bolded**.

4.5.2 Impact of 1D-CNN embedding block

An analysis of the contribution of the 1D-CNN embedding block to EEG-based MI classification is conducted to answer the first research question. Our previous study in SIFT-EEG demonstrated that temporal embedding is advantageous for extracting subject-invariant features, which is further corroborated by the experimental results presented in Tables 4.2 and 4.3. Models that incorporated CNN embedding generally outperformed pure GNN models that fed the raw signals into the neural network.

4. Utilization Towards Lightweight Machine Learning

The experiments in Table 4.2 demonstrate that the SIFTEEG framework can extract subject-invariant features when the number of training subjects is large (≥ 90 subjects). Conversely, as indicated in Table 4.3, when the number of subjects in training is small (≤ 10 subjects), the performance of this framework is significantly diminished, with accuracy even lower than that of all CNN-based models. This phenomenon may be attributed to interference between channels during the feature extraction process due to the 2D-CNN kernel.

To address this issue, we employed a 1D-CNN kernel, which is capable of extracting temporal correlations while excluding inter-channel interference. The results presented in Tables 4.2&4.3 demonstrate that the 1D-SIFTEEG model, which applies the 1D-CNN kernel, outperforms other models in subject-independent experiments, thus confirming the efficacy of the 1D-CNN kernel in obtaining subject-invariant embedding. Hidden features extracted by the 1D-SIFTEEG framework are also more distinct compares with features extracted by the SIFTEEG framework, as illustrated in Figure 4.3.

It is evident from the comparisons of the three groups that temporal embedding is essential, and the temporal embedding extracted from the 1D-CNN block is more subject-invariant than the features extracted by the 2D-CNN kernels.

Method	Parameters (M)	Storage (MB)
SIFT-EEG	2.92	11.13
1D-SIFT-EEG	0.47	1.78

Table 4.7: Storage cost comparison for different methods

4. Utilization Towards Lightweight Machine Learning

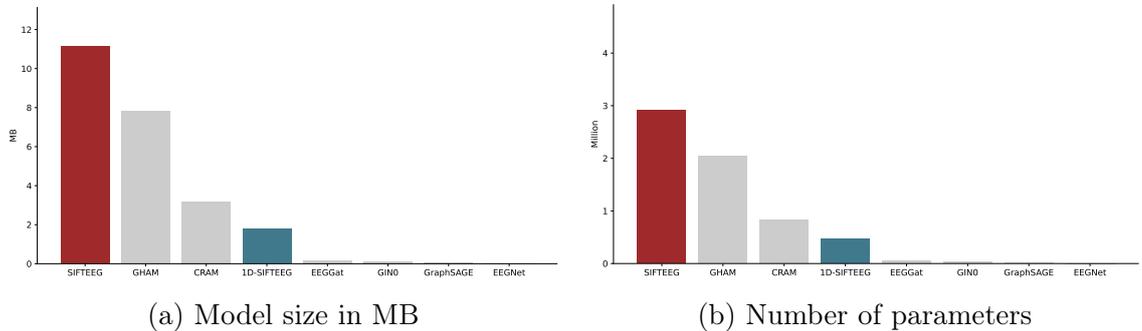


Figure 4.6: Storage cost comparison for different methods

4.5.3 Fewer trials - Data efficiency

The effectiveness of each similarity estimation matrix was evaluated using EEGM-MIDB dataset, given its higher volume of trainable subjects that could be employed for evaluation. The Leave-One-Out cross-validation was utilized and the mean value was taken in order to make a fair comparison.

The similarity score for each EEG segment in the training set was calculated, and the TopK most similar segments were selected as training data, where $K = |X_T| \times 1, 3, 5, 7, 9$, which is roughly equivalent to the top 0.96%, 2.88%, 4.81%, 6.73%, 8.65% most similar segments of the target domain X_T collected from the source domain X_S .

Table 4.6 presents a comparison of the classification accuracy when the topK training EEG segments are selected using each similarity metric.

From the table, it is evident that the correlation-based similarity metric is the most successful metric in comparison to the other metrics. The distance-based metric has the poorest performance, which can be attributed to the presence of a high degree of noise in the raw EEG Segments, thus making the training segments selected biased.

4. Utilization Towards Lightweight Machine Learning

The performance of the Hidden feature metric is superior to that of the Autoencoder metric in the learning process. The Hidden feature is designed to classify EEG segments and extract features that can be used for classification, thus making the overall selected EEG segments more discriminative and enabling a more accurate classification. Conversely, the Autoencoder is intended to produce an output that is identical to the input EEG segments, which lacks discriminative features, resulting in a poorer performance than that of the Hidden feature.

The overall performance growth is more evident when K increases from 1 to 7; however, the marginal growth has a noticeable decrease when K is between 7 and 9. This could be attributed to the large discrepancy between the lower-ranked train segments and the test segments. Thus, the introduction of more data does not bring more improvements in classification accuracy.

4.5.4 Computationally & Storage efficiency

We generated 400 EEG segments, each with a shape of [22,1875], and trained the model using the generated segments for 100 Epochs with GPU acceleration provided by an NVidia TITAN X. We then recorded the amount of time taken to process all generated segments for training and inference in order to compare the overall computational cost for each model. The results are presented in Fig. 4.5.

When compared to other models, the computational resources consumed by 1D-SIFTEEG during the training phase are moderate. Additionally, it requires a slightly lower amount of time than the SIFT-EEG method. Furthermore, the computational resources required in the inference stage are reduced in comparison to the training stage, and 1D-SIFTEEG consumes slightly higher time than SIFT-EEG in this stage.

4. Utilization Towards Lightweight Machine Learning

Table 4.7 and Figure. 4.6 present the number of trainable parameters and the size of the model for each model. Notably, 1D-SIFTEEG requires only approximately 2MB of storage space and its total number of trainable parameters is less than half-million, which is only roughly 16% of the space required for SIFT-EEG. This is much less than the space required by the previous version, and the most accurate method in CNN-family (GHAM), making 1D-SIFTEEG more suitable for deployment on edge devices.

The findings of this study demonstrate that 1D-SIFTEEG is more lightweight and cost-effective than the previous version, SIFT-EEG, thus responding to the third research question.

4.6 Conclusion

We propose the 1D-SIFT-EEG framework, which builds on the SIFT-EEG framework and further improves the performance in subject-independent motor imaginary classification. Secondly, we compare different similarity estimation approaches and demonstrate that correlation-based similarity metrics can effectively navigate EEG signals with high similarity, which makes the model remain well-performing even with very little data. Lastly, our proposed deep-learning model, 1D-SIFTEEG is more lightweight, making it easier to deploy on edge devices.

Chapter 5

Conclusion

5.1 Conclusion

This research project proposed a lightweight, subject-independent and task-adaptive modelling of brain networks empowered by functional connectivity.

Unlike previous approaches, who construct brain network using a fixed functional connectivity topology originating from the pre-motor, supplementary, and primary motor area that is responsible for motion control in Neuroscience. Our proposed framework is more task-adaptive and dynamic to the given MI-task. The proposed data-driven Algorithm 1 is able to navigate channels that are contributory to providing more discriminative information.

Extensive experiments were conducted with both large-scale and small-scale EEG datasets to demonstrate that the proposed predictive model outperformed the state-of-the-art for Motor Imagery classification in the subject-independent setting. Table 3.1 demonstrates our proposed DL-frameworks, SIFT-EEG, which is effective

5. Conclusion

in making subject-independent classification by extracting subject-independent features. Furthermore, the combination of temporal and topological encoders is more capable of subject-independent feature extraction than other methods. Additionally, 1D-SIFT-EEG, which incorporates One-Dimensional CNN into the temporal encoder, has higher discriminative ability and is proven to be more effective in dealing with datasets with fewer subjects. Furthermore, our investigation indicated that task-adaptive region selection produced similar predictive performance with only 20% of raw EEG data, resulting in a considerable reduction in computation cost during model training and deployment. Table.3.2 shows the classification accuracy when using the top 20% most relevant channels to train different NN-models. Accuracy of all models for task-adaptive selection is similar to (with CRAM and GHAM) or even slightly higher (with SIFT-EEG) than all channels. Additionally, an improved version of the SIFT-EEG framework, 1DSIFT-EEG, was proposed to further improve the performance in subject-independent motor imaginary classification. Different similarity estimation approaches were compared and it was demonstrated that correlation-based similarity metrics could effectively navigate EEG signals with high similarity, which enabled the model to remain well-performing even with very little data. Lastly, the proposed deep-learning model, 1D-SIFTEEG, was found to be lightweight and efficient, making it easier to deploy on edge devices.

5.2 Future work

There are still many limitations to this study, such as the fact that most experiments are conducted under binary classification; the model is trained and evaluated purely based on EEG MI-based tasks; and all experiments are conducted in an off-line setting, with its online performance not being well-investigated and compared with

5. Conclusion

other methods.

There are also several opportunities for future research based on this study. Examples include: applying transfer learning or domain adaption technique with the SIFT-EEG framework so that we can adapt the pre-trained system to new users and learn their subject-specific features; incorporating the proposed model into a live BCI system to evaluate its performance in production environments; exploring whether the SIFT-EEG framework can act as a pre-trained model and be fine-tuned to address different BCI paradigms such as emotion analysis, brain diseases detection or bio-authentication.

Bibliography

- [1] W. Chen, S. Wang, X. Zhang, L. Yao, L. Yue, B. Qian, and X. Li, “Eeg-based motion intention recognition via multi-task rnns,” in *ICDM*, pp. 279–287, SIAM, 2018.
- [2] S. M. Alarcao and M. J. Fonseca, “Emotions recognition using eeg signals: A survey,” *IEEE Transactions on Affective Computing*, vol. 10, no. 3, pp. 374–393, 2017.
- [3] J. Jeong, “Eeg dynamics in patients with alzheimer’s disease,” *Clinical neurophysiology*, vol. 115, no. 7, pp. 1490–1505, 2004.
- [4] X. Zhang, L. Yao, Q. Z. Sheng, S. S. Kanhere, T. Gu, and D. Zhang, “Converting your thoughts to texts: Enabling brain typing via deep feature learning of eeg signals,” in *2018 IEEE international conference on pervasive computing and communications (PerCom)*, pp. 1–10, IEEE, 2018.
- [5] I. A. Mirza, A. Tripathy, S. Chopra, M. D’Sa, K. Rajagopalan, A. D’Souza, and N. Sharma, “Mind-controlled wheelchair using an eeg headset and arduino microcontroller,” in *ICTSD*, pp. 1–5, IEEE, 2015.
- [6] T. Beyrouthy, S. K. Al Kork, J. A. Korbane, and A. Abdulmonem, “Eeg mind controlled smart prosthetic arm,” in *EmergiTech*, pp. 404–409, IEEE, 2016.
- [7] T. Mulder, “Motor imagery and action observation: cognitive tools for rehabilitation,” *Journal of neural transmission*, vol. 114, pp. 1265–1278, 2007.
- [8] M. Hallett, J. Fieldman, L. G. Cohen, N. Sadato, and A. Pascual-Leone, “Involvement of primary motor cortex in motor imagery and mental practice,” *Behavioral and Brain Sciences*, vol. 17, no. 2, pp. 210–210, 1994.
- [9] A. Sirigu, L. Cohen, J. Duhamel, B. Pillon, B. Dubois, Y. Agid, and C. Pierrot-Deseilligny, “Congruent unilateral impairments for real and imagined hand movements,” *Neuroreport*, vol. 6, no. 7, pp. 997–1001, 1995.

Bibliography

- [10] M. F. Mridha, S. C. Das, M. M. Kabir, A. A. Lima, M. R. Islam, and Y. Watanobe, “Brain-computer interface: Advancement and challenges,” *Sensors*, vol. 21, no. 17, p. 5746, 2021.
- [11] M. Rashid, N. Sulaiman, A. PP Abdul Majeed, R. M. Musa, B. S. Bari, S. Khatun, *et al.*, “Current status, challenges, and possible solutions of eeg-based brain-computer interface: a comprehensive review,” *Frontiers in neuro-robotics*, p. 25, 2020.
- [12] 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. 056013, 2018.
- [13] D. Zhang, L. Yao, X. Zhang, S. Wang, W. Chen, R. Boots, and B. Benatallah, “Cascade and parallel convolutional recurrent neural networks on eeg-based intention recognition for brain computer interface,” in *Proceedings of the aaai conference on artificial intelligence*, vol. 32, 2018.
- [14] F. Fahimi, Z. Zhang, W. B. Goh, T.-S. Lee, K. K. Ang, and C. Guan, “Inter-subject transfer learning with an end-to-end deep convolutional neural network for eeg-based bci,” *Journal of neural engineering*, vol. 16, no. 2, p. 026007, 2019.
- [15] K. Chen, L. Yao, D. Zhang, X. Chang, G. Long, and S. Wang, “Distributionally robust semi-supervised learning for people-centric sensing,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, pp. 3321–3328, 2019.
- [16] Y.-P. Lin and T.-P. Jung, “Improving eeg-based emotion classification using conditional transfer learning,” *Frontiers in human neuroscience*, vol. 11, p. 334, 2017.
- [17] D. Zhang, L. Yao, K. Chen, and J. Monaghan, “A convolutional recurrent attention model for subject-independent eeg signal analysis,” *IEEE Signal Processing Letters*, vol. 26, no. 5, pp. 715–719, 2019.
- [18] P. Shaw, J. Uszkoreit, and A. Vaswani, “Self-attention with relative position representations,” *arXiv preprint arXiv:1803.02155*, 2018.
- [19] D. Zhang, L. Yao, K. Chen, S. Wang, P. D. Haghghi, 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.
- [20] D. Zhang, K. Chen, D. Jian, and L. Yao, “Motor imagery classification via temporal attention cues of graph embedded eeg signals,” *IEEE journal of biomedical and health informatics*, vol. 24, no. 9, pp. 2570–2579, 2020.

Bibliography

- [21] A. Demir, T. Koike-Akino, Y. Wang, M. Haruna, and D. Erdogmus, “Eeg-gnn: Graph neural networks for classification of electroencephalogram (eeg) signals,” in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 1061–1067, IEEE, 2021.
- [22] J. N. Sanes, J. P. Donoghue, *et al.*, “Plasticity and primary motor cortex,” *Annual review of neuroscience*, vol. 23, no. 1, pp. 393–415, 2000.
- [23] A. Bechara, H. Damasio, and A. R. Damasio, “Emotion, decision making and the orbitofrontal cortex,” *Cerebral cortex*, vol. 10, no. 3, pp. 295–307, 2000.
- [24] D. Luo, W. Cheng, W. Yu, B. Zong, J. Ni, H. Chen, and X. Zhang, “Learning to drop: Robust graph neural network via topological denoising,” in *WSDM*, pp. 779–787, 2021.
- [25] C. Sun, A. Shrivastava, S. Singh, and A. Gupta, “Revisiting unreasonable effectiveness of data in deep learning era,” in *Proceedings of the IEEE international conference on computer vision*, pp. 843–852, 2017.
- [26] K. Xu, W. Hu, J. Leskovec, and S. Jegelka, “How powerful are graph neural networks?,” in *ICLR*, 2019.
- [27] L. R. Hochberg, M. D. Serruya, G. M. Friehs, J. A. Mukand, M. Saleh, A. H. Caplan, A. Branner, D. Chen, R. D. Penn, and J. P. Donoghue, “Neuronal ensemble control of prosthetic devices by a human with tetraplegia,” *Nature*, vol. 442, no. 7099, pp. 164–171, 2006.
- [28] S. Waldert, “Invasive vs. non-invasive neuronal signals for brain-machine interfaces: will one prevail?,” *Frontiers in neuroscience*, vol. 10, p. 295, 2016.
- [29] S. Waldert, T. Pistohl, C. Braun, T. Ball, A. Aertsen, and C. Mehring, “A review on directional information in neural signals for brain-machine interfaces,” *Journal of Physiology-Paris*, vol. 103, no. 3-5, pp. 244–254, 2009.
- [30] C. H. Blabe, V. Gilja, C. A. Chestek, K. V. Shenoy, K. D. Anderson, and J. M. Henderson, “Assessment of brain-machine interfaces from the perspective of people with paralysis,” *Journal of neural engineering*, vol. 12, no. 4, p. 043002, 2015.
- [31] M. Teplan *et al.*, “Fundamentals of eeg measurement,” *Measurement science review*, vol. 2, no. 2, pp. 1–11, 2002.
- [32] N. K. Logothetis, “What we can do and what we cannot do with fmri,” *Nature*, vol. 453, no. 7197, pp. 869–878, 2008.

Bibliography

- [33] M. Ferrari and V. Quaresima, “A brief review on the history of human functional near-infrared spectroscopy (fnirs) development and fields of application,” *Neuroimage*, vol. 63, no. 2, pp. 921–935, 2012.
- [34] J. Vrba and S. E. Robinson, “Signal processing in magnetoencephalography,” *Methods*, vol. 25, no. 2, pp. 249–271, 2001.
- [35] X. Zhang, L. Yao, X. Wang, J. Monaghan, D. Mcalpine, and Y. Zhang, “A survey on deep learning-based non-invasive brain signals: recent advances and new frontiers,” *Journal of neural engineering*, vol. 18, no. 3, p. 031002, 2021.
- [36] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, “Bci2000: a general-purpose brain-computer interface (bci) system,” *IEEE Transactions on biomedical engineering*, vol. 51, no. 6, pp. 1034–1043, 2004.
- [37] M. Duvinage, T. Castermans, M. Petieau, T. Hoellinger, G. Cheron, and T. Dutoit, “Performance of the emotiv epoc headset for p300-based applications,” *Biomedical engineering online*, vol. 12, no. 1, pp. 1–15, 2013.
- [38] P. Durka, R. Kuś, J. Żygierewicz, M. Michalska, P. Milanowski, M. Łabęcki, T. Sputeck, D. Laszuk, A. Duszyk, and M. Kruszyński, “User-centered design of brain-computer interfaces: Openbci. pl and bci appliance,” *Bulletin of the Polish Academy of Sciences. Technical Sciences*, vol. 60, no. 3, pp. 427–431, 2012.
- [39] K. Šušmáková, “Human sleep and sleep eeg,” *Measurement science review*, vol. 4, no. 2, pp. 59–74, 2004.
- [40] D. O. Bos *et al.*, “Eeg-based emotion recognition,” *The influence of visual and auditory stimuli*, vol. 56, no. 3, pp. 1–17, 2006.
- [41] B. T. Jap, S. Lal, P. Fischer, and E. Bekiaris, “Using eeg spectral components to assess algorithms for detecting fatigue,” *Expert Systems with Applications*, vol. 36, no. 2, pp. 2352–2359, 2009.
- [42] H. Altaheri, G. Muhammad, M. Alsulaiman, S. U. Amin, G. A. Altuwaijri, W. Abdul, M. A. Bencherif, and M. Faisal, “Deep learning techniques for classification of electroencephalogram (eeg) motor imagery (mi) signals: a review,” *Neural Computing and Applications*, pp. 1–42, 2021.
- [43] R. Soikkeli, J. Partanen, H. Soininen, A. Pääkkönen, and P. Riekkinen Sr, “Slowing of eeg in parkinson’s disease,” *Electroencephalography and clinical neurophysiology*, vol. 79, no. 3, pp. 159–165, 1991.

Bibliography

- [44] R. Paranjape, J. Mahovsky, L. Benedicenti, and Z. Koles, “The electroencephalogram as a biometric,” in *Canadian Conference on Electrical and Computer Engineering 2001. Conference Proceedings (Cat. No. 01TH8555)*, vol. 2, pp. 1363–1366, IEEE, 2001.
- [45] A. M. Batula, H. Ayaz, and Y. E. Kim, “Evaluating a four-class motor-imagery-based optical brain-computer interface,” in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 2000–2003, IEEE, 2014.
- [46] A. M. Bastos and J.-M. Schoffelen, “A tutorial review of functional connectivity analysis methods and their interpretational pitfalls,” *Frontiers in systems neuroscience*, vol. 9, p. 175, 2016.
- [47] C. W. Granger, “Investigating causal relations by econometric models and cross-spectral methods,” *Econometrica: journal of the Econometric Society*, pp. 424–438, 1969.
- [48] Y. LeCun, Y. Bengio, *et al.*, “Convolutional networks for images, speech, and time series,” *The handbook of brain theory and neural networks*, vol. 3361, no. 10, p. 1995, 1995.
- [49] D. S. Bassett and O. Sporns, “Network neuroscience,” *Nature neuroscience*, vol. 20, no. 3, pp. 353–364, 2017.
- [50] W. Hamilton, Z. Ying, and J. Leskovec, “Inductive representation learning on large graphs,” *Advances in neural information processing systems*, vol. 30, 2017.
- [51] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, “Graph attention networks,” *arXiv preprint arXiv:1710.10903*, 2017.
- [52] X. Wu, W.-L. Zheng, Z. Li, and B.-L. Lu, “Investigating eeg-based functional connectivity patterns for multimodal emotion recognition,” *Journal of neural engineering*, vol. 19, no. 1, p. 016012, 2022.
- [53] T. N. Kipf and M. Welling, “Semi-supervised classification with graph convolutional networks,” *arXiv preprint arXiv:1609.02907*, 2016.
- [54] K. Hornik, “Approximation capabilities of multilayer feedforward networks,” *Neural networks*, vol. 4, no. 2, pp. 251–257, 1991.
- [55] M. C. Tate, G. Herbet, S. Moritz-Gasser, J. E. Tate, and H. Duffau, “Probabilistic map of critical functional regions of the human cerebral cortex: Broca’s area revisited,” *Brain*, vol. 137, no. 10, pp. 2773–2782, 2014.

Bibliography

- [56] T. Chugh, K. Sindhya, J. Hakanen, and K. Miettinen, “A survey on handling computationally expensive multiobjective optimization problems with evolutionary algorithms,” *Soft Computing*, vol. 23, no. 9, pp. 3137–3166, 2019.
- [57] L. A. Moctezuma and M. Molinas, “Classification of low-density eeg for epileptic seizures by energy and fractal features based on emd,” *Journal of biomedical research*, vol. 34, no. 3, p. 180, 2020.
- [58] L. A. Moctezuma and M. Molinas, “Eeg channel-selection method for epileptic-seizure classification based on multi-objective optimization,” *Frontiers in neuroscience*, vol. 14, p. 593, 2020.
- [59] M. Xia, J. Wang, and Y. He, “Brainnet viewer: a network visualization tool for human brain connectomics,” *PloS one*, vol. 8, no. 7, p. e68910, 2013.
- [60] B.-H. Kim and J. C. Ye, “Understanding graph isomorphism network for rs-fmri functional connectivity analysis,” *Frontiers in neuroscience*, p. 630, 2020.
- [61] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P. C. Ivanov, R. Mark, J. Mietus, G. Moody, C. Peng, and H. Stanley, “Components of a new research resource for complex physiologic signals,” *PhysioBank, PhysioToolkit, and Physionet*, 2000.
- [62] S.-L. Wu, C.-W. Wu, N. R. Pal, C.-Y. Chen, S.-A. Chen, and C.-T. Lin, “Common spatial pattern and linear discriminant analysis for motor imagery classification,” in *2013 IEEE Symposium on Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB)*, pp. 146–151, IEEE, 2013.
- [63] Y. Rong, W. Huang, T. Xu, and J. Huang, “Dropedge: Towards deep graph convolutional networks on node classification,” in *International Conference on Learning Representations*, 2020.
- [64] E. Dai, W. Jin, H. Liu, and S. Wang, “Towards robust graph neural networks for noisy graphs with sparse labels,” in *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*, pp. 181–191, 2022.
- [65] B. Knyazev, G. W. Taylor, and M. Amer, “Understanding attention and generalization in graph neural networks,” *Advances in neural information processing systems*, vol. 32, 2019.
- [66] M. Bear, B. Connors, and M. A. Paradiso, *Neuroscience: Exploring the Brain, Enhanced Edition: Exploring the Brain*. Jones & Bartlett Learning, 2020.
- [67] H. Shan, H. Xu, S. Zhu, and B. He, “A novel channel selection method for optimal classification in different motor imagery bci paradigms,” *Biomedical engineering online*, vol. 14, no. 1, pp. 1–18, 2015.

Bibliography

- [68] F. Mattioli, C. Porcaro, and G. Baldassarre, “A 1d cnn for high accuracy classification and transfer learning in motor imagery eeg-based brain-computer interface,” *Journal of Neural Engineering*, vol. 18, no. 6, p. 066053, 2022.
- [69] S. Sreeja, J. Rabha, D. Samanta, P. Mitra, and M. Sarma, “Classification of motor imagery based eeg signals using sparsity approach,” in *International Conference on Intelligent Human Computer Interaction*, pp. 47–59, Springer, Cham, 2017.
- [70] L. Lettry, M. Perdoch, K. Vanhoey, and L. Van Gool, “Repeated pattern detection using cnn activations,” in *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 47–55, IEEE, 2017.
- [71] N. A. Lassen, D. H. Ingvar, and E. Skinhøj, “Brain function and blood flow,” *Scientific American*, vol. 239, no. 4, pp. 62–71, 1978.
- [72] D. J. Heeger and D. Ress, “What does fmri tell us about neuronal activity?,” *Nature reviews neuroscience*, vol. 3, no. 2, pp. 142–151, 2002.
- [73] S. Albawi, T. A. Mohammed, and S. Al-Zawi, “Understanding of a convolutional neural network,” in *2017 international conference on engineering and technology (ICET)*, pp. 1–6, Ieee, 2017.
- [74] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, and D. J. Inman, “1d convolutional neural networks and applications: A survey,” *Mechanical systems and signal processing*, vol. 151, p. 107398, 2021.
- [75] W. Klonowski, “Everything you wanted to ask about eeg but were afraid to get the right answer,” *Nonlinear biomedical physics*, vol. 3, no. 1, pp. 1–5, 2009.
- [76] X. Ying, “An overview of overfitting and its solutions,” in *Journal of physics: Conference series*, vol. 1168, p. 022022, IOP Publishing, 2019.
- [77] M. M. Bejani and M. Ghatee, “A systematic review on overfitting control in shallow and deep neural networks,” *Artificial Intelligence Review*, vol. 54, no. 8, pp. 6391–6438, 2021.
- [78] D. Bank, N. Koenigstein, and R. Giryes, “Autoencoders,” *arXiv preprint arXiv:2003.05991*, 2020.
- [79] G. Zhang, Y. Liu, and X. Jin, “A survey of autoencoder-based recommender systems,” *Frontiers of Computer Science*, vol. 14, no. 2, pp. 430–450, 2020.
- [80] K. G. Lore, A. Akintayo, and S. Sarkar, “Llnet: A deep autoencoder approach to natural low-light image enhancement,” *Pattern Recognition*, vol. 61, pp. 650–662, 2017.

Bibliography

- [81] C.-Y. Liou, W.-C. Cheng, J.-W. Liou, and D.-R. Liou, "Autoencoder for words," *Neurocomputing*, vol. 139, pp. 84–96, 2014.
- [82] T. Wen and Z. Zhang, "Deep convolution neural network and autoencoders-based unsupervised feature learning of eeg signals," *IEEE Access*, vol. 6, pp. 25399–25410, 2018.