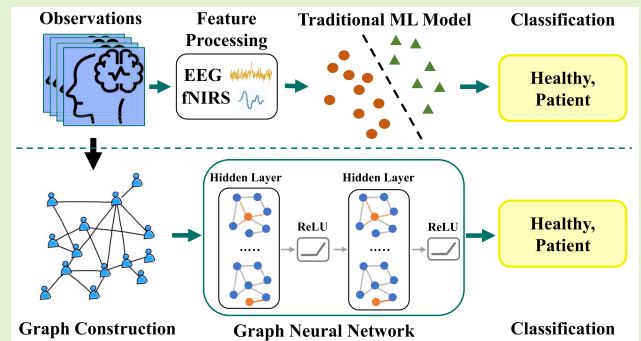


Topology-Aware Multimodal Fusion for Neural Dynamics Representation Learning and Classification

Neela Rahimi^{ID}, Graduate Student Member, IEEE, Chetan Kumar, John McLinden^{ID}, Sarah Ismail Hosni^{ID}, Seyyed Bahram Borgheai^{ID}, Yalda Shahriari^{ID}, Member, IEEE, and Ming Shao^{ID}, Member, IEEE

Abstract—Applications of multimodal neuroimaging techniques, including electroencephalography (EEG) and functional near-infrared spectroscopy (fNIRS) have gained prominence in recent years, and they are widely practiced in brain–computer interface (BCI) and neuro-pathological diagnosis applications. Most existing approaches assume observations are independent and identically distributed (i.i.d.), as shown in the top section of the right figure, yet ignore the difference among subjects. It has been challenging to model subject groups to maintain topological information (e.g., patient graphs) while fusing BCI signals for discriminant feature learning. In this article, we introduce a topology-aware graph-based multimodal fusion (TaGMF) framework to classify amyotrophic lateral sclerosis (ALS) and healthy subjects, illustrated in the lower section of the right image. Our framework is built on graph neural networks (GNNs) but with two unique contributions. First, a novel topology-aware graph (TaG) is proposed to model subject groups by considering: 1) intersubject; 2) intrasubject; and 3) intergroup relations. Second, the learned representation of EEG and fNIRS signals of each subject allows for explorations of different fusion strategies along with the TaGMF optimizations. Our analysis demonstrates the effectiveness of our graph-based fusion approach in multimodal classification by achieving a 22.6% performance improvement over classical approaches.

Index Terms—Brain–computer interface (BCI), data fusion, graph neural network (GNN), multimodal neuroimaging.



I. INTRODUCTION

RESEARCH efforts have been actively pursuing multiple measurements to support effective decoding of neural activities to improve brain–computer interface (BCI) systems [1], [2] and have developed a variety of techniques, including electroencephalography (EEG), and

functional near-infrared spectroscopy (fNIRS) [3], [4] for studying brain signals. Most studies have conducted unimodal investigations, using either EEG or fNIRS to learn and detect neural responses [4], [5]. However, each of these modalities can capture limited information regarding brain functions due to the underlying technical constraints and complex nature of neural processing in the brain [2], [6], [7].

Applications of multimodal fusion have gained prominence in recent years and are widely practiced in BCI and neuro-pathological diagnosis applications [8], [9]. In this regard, integrating EEG and fNIRS signals provides two cost-effective sources of information: electrical activities and hemodynamic responses of the brain from EEG and fNIRS, respectively [10], [11]. Earlier studies with EEG-fNIRS fusion for classification reported improved performance over a single modality [1], [8], [12], [13]. These fusion approaches usually follow independent and identically distributed (i.i.d.) assumptions and use either decision-level [14], [15] or feature-level [12], [16], [17], [18], [19] fusion strategy to integrate features and classify observations independently.

Manuscript received 4 April 2024; accepted 30 April 2024. Date of publication 17 May 2024; date of current version 1 July 2024. This work was supported by the National Science Foundation under Grant NSF-1913492, Grant NSF-2024418, and Grant NSF-2225818. The associate editor coordinating the review of this article and approving it for publication was Dr. Geethu Joseph. (Neela Rahimi and Chetan Kumar contributed equally to this work.) (Corresponding author: Neela Rahimi.)

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the University of Rhode Island Institutional Review Board.

Neela Rahimi, Chetan Kumar, and Ming Shao are with the Department of Computer and Information Science, University of Massachusetts Dartmouth, Dartmouth, MA 02747 USA (e-mail: nrahimi@umassd.edu; ckumar@umassd.edu; mshao@umassd.edu).

John McLinden, Sarah Ismail Hosni, Seyyed Bahram Borgheai, and Yalda Shahriari are with the Department of Electrical, Computer, and Biomedical Engineering, University of Rhode Island, Kingstown, RI 02881 USA (e-mail: john_mclinden@uri.edu; sarah_hosni@uri.edu; borgheai@uri.edu; yalda_shahriari@uri.edu).

Digital Object Identifier 10.1109/JSEN.2024.3400006

Recent explorations have opened new frontiers in modeling brain signals using graph-based methods, which have achieved appealing performance in analyzing, classifying, and interpreting neurological disorders and disease prediction in cases such as Alzheimer's and autism [10], [20], depression [21], and anxiety [22]. In addition, brain signals can be modeled by a graph according to regional connectivity to represent the underlying networks in brain responses [23], [24]. Moreover, graphs have the potential to represent topology within larger populations where nodes (i.e., individuals) are connected based on their similarities [20] or to provide knowledge bases for recognized challenges such as depression detection [25]. While graph-based models offer insights into neural dynamics through regional connectivity, our study diverges by focusing on an end-to-end representation learning framework for EEG and fNIRS signals using graph neural networks (GNNs) [26] for groups of subjects. This decision stems from the challenges of interpreting functional connectivity in fNIRS due to extracerebral hemodynamic interactions, which can confound true neural activity, especially without short channels to mitigate these effects [27]. By prioritizing the integration of EEG and fNIRS data for representation learning, our work seeks to harness the complementary strengths of these modalities, thus contributing novel methodological advancements to neuroimaging research beyond the traditional functional connectivity approach.

As an extension to previous works, we also explore graph-guided BCI fusion for groups of subjects (i.e., healthy/patient) and their classification. The local smoothness assumption in graphs allows us to model different subjects and groups that are usually non-i.i.d. for representation learning. The framework is built on GNNs for representation learning, while the learned features can be verified through various classifiers. GNN approaches generally consider node features and their interaction as edges to aggregate node features over local neighborhoods and pursue optimal representations layer by layer [28]. The efficacy of GNNs hinges on accurately reflecting the topology among subjects and observations in the graph. Assuming there are multiple observations per subject, when building the graph for GNN, several distinct relations would be considered, including: 1) inter-subject; 2) intrasubject; and 3) intergroup relations. In particular, intrasubject relations indicate a strong correlation among observations of the same subject. However, to our knowledge, these concepts have not been thoroughly explored in previous research.

We combine modal integration and cross-subject information within a single framework. The motivation behind this dual approach is rooted in our aim to develop a comprehensive strategy that not only enhances the integration of multimodal neuroimaging data but also addresses the inherent variability and distributional differences across subjects—a common challenge in neuroscientific studies. By incorporating these two aspects into a unified framework, we strive to highlight the synergistic benefits of handling modality fusion and cross-subject variability concurrently. This integration allows for a more holistic improvement in classification and representation learning performance, leveraging the complementary

information from multiple modalities while accounting for the unique characteristics of individual subjects' data.

To that end, this article introduces topology-aware graph-based multimodal-fusion (TaGMF), a learnable GNN model for feature extraction, achieved by leveraging the union of three dedicated subgraphs. Compared to existing works, TaGMF explores subject relations over networks beyond the scope of single-subject neural dynamics modeling to pursue better representations, along with the network optimization is the single and multimodal fusion strategy at different levels to account for multimodal data from subjects. In particular, we investigate two fusion strategies, namely *early* and *late* fusion, and compare them with the single modality data in classification tasks. To demonstrate our TaGMF approach, we extensively evaluate the collected EEG and fNIRS data from healthy controls (HCs) and subjects with amyotrophic lateral sclerosis (ALS).

Our contribution can be summarized as follows.

- 1) A novel TaGMF framework has been developed to model subject groups and intra- and intersubject relations of different observations to learn discriminant node representations for downstream tasks.
- 2) Explorations of fusion approaches with multimodal measures provide a basis for understanding complementary features and improving the performance of learning tasks.
- 3) Extensive evaluations on a collected multimodal EEG-fNIRS dataset demonstrate that our proposed TaGMF outperforms conventional machine learning models.

II. RELATED WORK

Machine learning algorithms have achieved remarkable improvements in neuroimaging and signal processing. This section reviews the methods closely related to our work, narrowing it down to: 1) graph-based learning approaches in neuroimaging studies and 2) multimodal EEG-fNIRS fusion.

A. Graph Neural Networks

In clinical studies, graphs are explored broadly under two scenarios: 1) cross-subject graphs and 2) single-subject graphs. Cross-subject graphs usually model patient networks in computer-aided diagnosis (CADx) applications. In contrast, single-subject graphs aim to present inherent brain neural connections as nodes and edges for a single subject [29].

1) *Cross-Subject Graph*: The success of GNNs in social networks and recommendation systems has motivated graph-based algorithms in patient networks [30] where subjects are treated as nodes and their interaction as edges [24]. Parisot et al. [31] was the first to employ GNN in the medical domain for brain analysis in a semi-supervised fashion. Later, they enhanced the population graph in disease diagnosis by using meta-features [20]. Nonetheless, the graph quality largely affects the performance of GNN. To that end, [32], [33], [34], [35] recommended using multigraph fusions to enhance the learning process by incorporating latent attributes in the population graph, and an optimal graph, based on relevancy to medical features selected by statistical tests. While

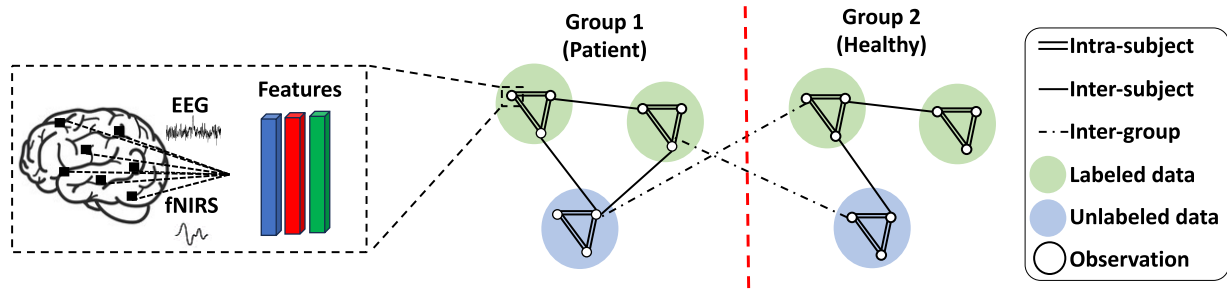


Fig. 1. Overview of the proposed TaGMF framework. The left side represents the collection of EEG and fNIRS brain signals and subsequent feature extraction. The center part provides a visual description of graph structures with three types of edges: 1) intrasubject edges connect observations within the same subject; 2) intersubject edges between different subjects provide insights into shared patterns across individuals; and 3) intergroup edges connect subjects across healthy and patient groups highlighting group-level similarities.

this ensemble strategy improved the performance, it introduced additional computation in latent attribute discovery and efforts in fusion layer design. To alleviate these challenges, Wang et al. [36] conducted feature selection for clinical and genomic data to build two different matrices. Zhan et al. [37] proposed to use multiple graphs for heterogeneous features with an optimization module to adaptively optimize the graph weights. Cosmo et al. [30] demonstrated the existence of a single optimal graph by incorporating multiple features embedded in Euclidean space. Zhong et al. [38] proposed EEG-based emotion recognition using regularized GNN (RGNN) to handle cross-subject EEG variations and noisy labels.

2) Single-Subject Graph: GNN with single-subject graphs mostly concentrated on capturing richer features from brain signals and eliminating noise to achieve more refined results. Research in [5] proposed an EEG-based hierarchy graph convolution network to extract the information between adjacent electrodes for emotion recognition. Yin et al. [39] used GNN to extract graph domain features for emotion valence and arousal recognition and classification. Multiview graph convolutional network (GCN) [40] proposed a graph embedding of brain networks to capture brain topological information and eliminate noisy and spurious connections for autism spectrum disorder classification. Faskowitz et al. [41] introduced an edge-centric perspective on brain networks, highlighting the significance of connections in understanding the architecture of cerebral cortex interactions. Furthermore, the work by Ismail and Karwowski [42] explored innovative approaches to modeling functional brain networks using graph theory and covariance matrices, offering insights into the complex interplay of neural connections. Similarly, the comprehensive review by Dragomir and Omurtag [43] on graph theoretical approaches in neuroimaging consolidates the current understanding and applications of these methods in elucidating the brain's functional architecture. While extensive research has leveraged single modalities independently, studies exploring the integration of multimodal signals, particularly EEG and fNIRS, through covariance matrices for single-subject connectivity analysis remain an emerging area. Unlike existing works, this article also proposes a general graph-based multimodal fusion framework for groups of subjects. In particular, we focus on developing new graph construction methods to incorporate various relations among subjects for neural dynamics representation learning.

B. Multimodal Brain EEG-fNIRS Fusion

Multimodal fusion [1] has been employed in many tasks, including mental analysis, emotion measurement, motor control, clinical evaluation, rehabilitation, and perception assessment [44]. Fused brain signals demonstrated superiority through shallow learning models, such as support vector machines (SVMs), random forest [44], [45], [46], linear discriminant analysis (LDA) [47], and deep models, including deep neural networks (DNNs) [7], [48], [49], [50], recurrent neural network (RNN) with long short-term memory (LSTM) [51], [52], [53], convolutional neural network (CNN) [22], [54], [55] and GNN [20], [23]. Nonetheless, fusing high dimensional EEG-fNIRS neuroimaging data is still a challenge [2]. Common issues include weak generalization and overfitting given limited data in clinical research [56]. In this regard, Lin et al. [8] conducted the correlation analysis to select the most correlated signal channels, and Yin et al. [57] adopted joint mutual information for feature optimization. Al-Shargie et al. [17] developed a fusion technique to incorporate temporal properties of EEG and spatial features from fNIRS, and later used canonical correlation analysis (CCA) as a linear mixing model that maximizes the covariance between EEG and fNIRS [15]. Khan and Hasan [13] applied multiresolution singular value decomposition (MSVD) to perform the feature-based fusion. First, statistical features are extracted from the fNIRS data and discrete wavelet transform (DWT) features from the EEG data are normalized, which are then decomposed into sub-bands using MSVD. Recently, Deligani et al. [2] proposed an approach to optimizing the complementary features of EEG-fNIRS using label information at the feature level through a cross-validation process. Qiu et al. [58] combined the time- and frequency-domain features of EEG and fNIRS data and used the ASO algorithm for feature selection to remove the information redundancy caused by multidomain features.

In this work, we explore EEG-fNIRS feature extraction and fusion through a novel TaGMF framework. In particular, we compare early and late fusion strategies to understand the roles of TaGMF in neuroimaging data representation learning.

III. METHODOLOGY

Fig. 1 illustrates a visual overview of the TaGMF method, integrating EEG and fNIRS data within a multimodal graph

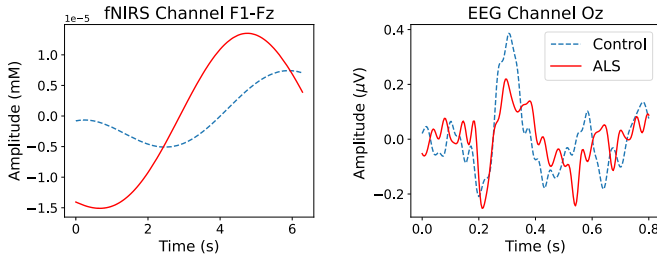


Fig. 2. Visualization of ALS versus control signals.

representation. The left side represents the collection of EEG and fNIRS brain signals and subsequent feature extraction. The center part provides a visual description of graph structures with three types of edges: 1) intrasubject edges connect observation within the same subject; 2) intersubject edges between different subjects to provide insights into shared patterns across individuals; and 3) intergroup edges connect subjects across healthy and patient groups highlighting group-level similarities. TaGMF leverages the graph structure to extract meaningful features which are used to learn discriminating patterns essential for the downstream classification task.

A. Data Analysis

1) *Participants*: Data were collected by recruiting a total of eighteen subjects (nine subjects with ALS and nine HCs). Seven out of nine ALS subjects were males with an average age of 56.8 years. Participants with ALS had ALS revised functional rating scale (ALSFRS-R) scores of 0, 4, 4, 23, 22, 39, 41, 33, and 26 with a mean score of 21.3 ± 15.5 on a 48-point scale. Four out of nine age-matched HCs were males with an average age of 60.7 years. Age-matched control participants reported no known history of visual, mental, or substance-related disorders which could possibly influence the results or their performance during data collection. Data collection was performed at the University of Rhode Island (URI) with Institutional Review Board (IRB) approval and written consent was obtained from all the subjects or their caregivers [2].

2) *Acquisition*: EEG and fNIRS signals were recorded simultaneously using a single cap mounted with both EEG electrodes and fNIRS optodes. A g.USBamp amplifier (g.tec Medical Tech., Schiedlberg, Austria) was used to record the EEG data at a sampling rate of 256 Hz. A NIRScout system (NIRx Inc.) with two NIR lights (760 and 850 nm wavelengths) was used to record fNIRS data at a sampling rate of 7.81 Hz. This setup allowed us to capture the complementary electro-hemodynamic characteristics of neural responses with minimal interference. Data was recorded using 16 EEG and 16 fNIRS channels. EEG channels were placed at AF3*, AF4*, F1*, Fz*, F2*, T7, Cz, T8, P7, P3, Pz, P4, P8, PO7, PO8, and Oz which covers all prefrontal, frontal, central, parietal, temporal and occipital areas. fNIRS channels consisted of eight emitters and seven detectors where emitters were placed at Fpz, AF3, AF4, F3, Fz, F4, CP5, and CP6, and the detectors were located at Fp1, Fp2, AFz, F1, F2, P5, and P6 covering the prefrontal and frontal areas. More details can be found in [2]. Fig. 2 shows ALS versus control signals, averaged

over all subjects using the F1-Fz channel of fNIRS and Oz channel of EEG as the most representative channels.

3) *Preprocessing and Raw Feature Extraction*: For EEG data, each group, that is, HC and ALS, has nine participants. Each participant has two runs and each run has 14 trials. Therefore, in total, we recorded $9 \times 2 \times 14 = 252$ number of observations/samples for each group. EEG signals were bandpass filtered 0.3–35 Hz to remove baseline drift and out-of-band artifacts and further, it was visually inspected for any outliers. Segmentation for task/event-related analysis was conducted by isolating specific windows poststimulus presentation to ensure that the extracted features pertain directly to the tasks/events. The EEG spectral features were obtained by decomposing the data into spectrograms and averaged into four different frequency bands, namely delta (1–3 Hz), theta (4–7 Hz), alpha (8–12 Hz), and beta (13–30 Hz) to obtain four different features. Thus, in total, we obtained $16 \times 4 = 64$ spectral features where 16 is the number of channels, and 4 is the number of frequency bands. EEG temporal features were acquired using five event-related potential (ERP) features corresponding to three maximum and two minimum peaks respectively. Therefore, we obtained $16 \times 4 = 64$ EEG temporal features where 16 is the number of channels, and 4 is the number of ERP components. In this study, we consider EEG spectral as *EEG Power* and EEG temporal as *EEG ERP*.

For fNIRS data, the segmentation approach mirrors that of the EEG, with an equivalent number of observations/samples per group. Bandpass filter was applied to fNIRS data at 0.01–0.2 Hz to remove physiological noises produced by respiratory and cardiac activities [59]. Subsequently, using the modified Beer-Lambert Law [60] on the raw optical intensity data, we extracted oxy-hemoglobin (HbO₂) concentration changes. This process specifically pinpointed hemodynamic responses to the tasks/events by analyzing 0–6 s poststimulus windows for each channel, ensuring the extracted features accurately reflected task/event-induced hemodynamic activity. In total, fNIRS data contain $16 \times 2 = 32$ features where 16 is the number of channels and 2 is the number of feature types.

4) *Experimental Protocol*: Subjects participated in a visuo-mental test based on the commonly used visual oddball paradigm with a mathematical task during data collection. This paradigm is fully described in [61]. These tasks provoke both electrical and hemodynamic responses relevant to visual oddball simulations and mental arithmetic operations.

5) *Training/Validation Split*: Raw EEG and fNIRS features are normalized using the unit norm and split into training and test sets. Data are split into five folds where training and test sets do not contain observations from the same subject. We have 504 observations of 18 subjects (9 HC and 9 ALS). Out of nine HC subject observations, seven subjects and their $7 \times 28 = 196$ observations are used for training, and the remaining two subjects and their $2 \times 28 = 56$ observations are for testing. The same data split strategy is applied to the ALS group.

B. TaG-Based Multimodal Fusion

In this section, we will elaborate on our TaGMF framework for classifying disease states using the collected data, as

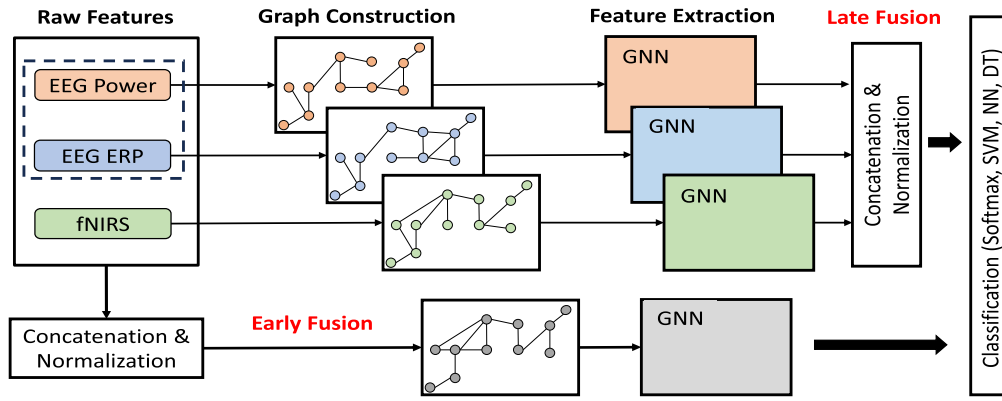


Fig. 3. Pipeline of our TaGMF framework. Early fusion entails the concatenation of raw features and their TaG construction before feeding them to the GNN model. Late fusion performs TaG construction and feature extraction through the GNN models first, followed by the concatenation of features. Classification is then conducted for both approaches.

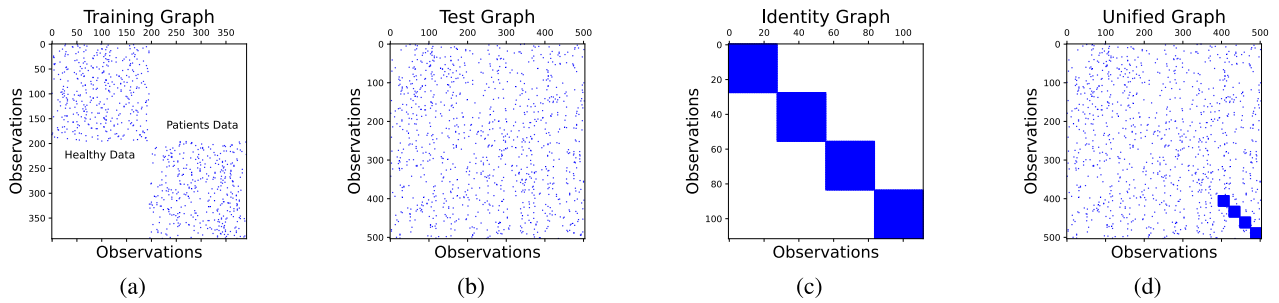


Fig. 4. Visualizations of (a) training graph, (b) test graph, (c) identity graph, and (d) unified graph on EEG Power data.

illustrated in Fig. 3. TaGMF is based on classic GNN models but tailored to accommodate multimodal data for groups of subjects. First, we propose a topology-aware graph (TaG) to present pairwise relations among observations precisely. Second, two fusion strategies, that is, *early* and *late*, are explored to account for common and unique topological information of each modality.

1) *TaG Construction*: GCN [62] entails graphs in training and inference; however, built-in graphs are not always available. To that end, we develop a general graph construction criteria suitable for neural dynamics modeling of subject groups, termed “TaG.” TaG is essentially a nonparametric and knowledge-driven graph motivated by the following. First, we apply the k -nearest-neighbor (KNN) method to explore observations and their neighbors to build a nondirectional KNN graph where only neighbors (measured by similarities) are connected. Second, knowledge about each subject and their observations can contribute additional connections to the KNN graph, especially for test data without label information. For example, observations of the same subject can be connected. Third, most GNNs are trained in a transductive fashion, meaning the graph incorporates all the observations, including training (with labels) and test data (without labels), that are intertwined during the model training. Our TaG includes three sub-graphs to accommodate different facets of data and craft different ways of building connections, as explained below.

a) *Training graph G_1* : The training graph is built for both HC and ALS training data using: 1) label information and

2) connections learned by the KNN graph. First, data of the same label are connected, yielding two fully connected graphs for HC and ALS observations, respectively. Second, we further apply KNN criteria to the fully connected graphs to produce the training graph. This is equivalent to the intersection of the fully connected HC/ALS graph and their individual KNN graphs. An example EEG Power training graph is shown in Fig. 4(a). The graph is sparse in general due to the intersection with the sparse KNN graph.

b) *Test graph G_2* : Test data without labels can only apply the KNN criterion to create connections for observations. Note that these connections may yield test-test data connections or test-train data connections that may propagate labels from training to test data through the graph. Unlike the training graph, the test graph may also introduce unwanted connections between HC and ALS observations, primarily due to their high similarities. An example test graph is shown in Fig. 4(b).

c) *Identity graph G_3* : Observations of the same subject share the label and thus should connect to each other. This is particularly useful for test graphs with limited discriminant information and leads to fully connected sub-blocks, as shown in Fig. 4(c). The formulated “clique” benefits discriminant representation learning and classification as observations of the same subject will be treated as a group at the test. This is essentially similar to the majority voting criterion determined by predictions of a group test data, thus providing potentially better performance.

d) *Unified graph G* : The union of three sub-graphs: $G = G_1 \cup G_2 \cup G_3$ enables the integration of nonparametric KNN

graph and knowledge of subject groups. An example unified graph for the same EEG Power data is shown in Fig. 4(d).

2) **TaGMF Learning:** Let $X \in \mathbb{R}^{N \times d}$ be the BCI features, including both training and test data, where N is the total number of observations, d is the dimension of raw features, and $G \in \mathbb{R}^{N \times N}$ is the built TaG graph. TaGMF aims to learn the GNN parameters W_l at layer l in training. The forward pass proceeds layer by layer through the function: $Y_l = g(Y_{l-1}, G)$ where Y_l is the feature at layer l , and g is a ReLU activation function. In particular, each layer aggregates the node features via graph convolutions and passes to the next layer as follows:

$$Y_l = g(D^{-1/2} \hat{G} D^{-1/2} Y_{l-1} W_{l-1}) \quad (1)$$

where $\hat{G} = G + I_N$ is the modified adjacency matrix with self-loops through adding the identity matrix $I_N \in \mathbb{R}^{N \times N}$, and $D_{ii} = \sum_j G_{ij}$ is a diagonal matrix indicating the degree of each node. The component $D^{-1/2} \hat{G} D^{-1/2}$ is also recognized as the normalized graph Laplacian in spectral methods such as spectral clustering to maintain a good balance among different clusters. Here $Y_0 = X$ uses the raw features as the input, and the final output Z is produced through a softmax function

$$Z = \text{softmax}(D^{-1/2} \hat{G} D^{-1/2} Y_L W_L) \quad (2)$$

where Z is the normalized probabilistic output for classification purposes.

3) **TaGMF-Based Feature Extraction:** As opposed to the conventional usage of GNN as a classifier, we treat TaGMF as a representation learning model to extract discriminant features. This allows us to explore different classifiers with the learned TaGMF features and usually achieves better performance when the number of training samples is limited. To extract TaGMF features, we first train TaGMF as before, then pass the data through the learned GNN model and extract the features from the second to last layer. The learned features for training and test data will be used to train various classifiers.

4) **Fusion Strategies:** In TaGMF, we propose two fusion mechanisms: *early fusion* and *late fusion*. In early fusion, raw features from multimodal data are concatenated to form new vectors based on which a common TaG can be learned. Then TaG and raw features are passed to GNN for training and feature extraction. In late fusion, however, separate GNN models are learned for each modality to extract features. The GNN features of each modality are then concatenated. The early and late fusion approaches are illustrated in Fig. 3.

C. Comparisons With Classical Methods

We elaborate on baselines by considering different features and classifiers used by classical approaches. First, TaGMF features are compared with raw features and feature selection approaches such as lasso regularization capable of selecting the most significant and nonredundant sparse features [63], [64]. Second, various classifiers are explored, including the Softmax classifier, SVM, nearest neighbor (NN), and decision tree (DT). In parallel, both unimodal and multimodal features

are compared in classification tasks. This provides a comprehensive comparison between TaGMF and many other classical approaches.

IV. EXPERIMENTAL RESULTS

A. Raw Features Versus TaGMF Features

Fig. 5(a)–(d) compares the raw features extracted directly from the original signals and proposed TaGMF features. It can be observed that TaGMF features yield higher accuracy than raw features in all three classifiers, and our proposed approach improves the performance by approximately 22%, 15%, and 19% with classifiers SVM, NN, and DT, respectively as shown in Fig. 5(d). It is noteworthy that SVM outperformed NN and DT in terms of classification accuracy.

B. Single Versus Multimodality

Fig. 5 evaluates TaGMF features under unimodal and multimodal setups, applying late fusion. As hypothesized, classification using multimodal features extracted by TaGMF demonstrated improved performance [Fig. 5(d)] compared to each unimodal signal [Fig. 5(a)–(c)], achieving 15.9%, 19.7%, and 22.6% improvement with the three classifiers: SVM, NN, and DT, respectively. SVM consistently exhibited enhanced performance in all modalities. The results also indicated that TaGMF features work better with SVM than the default classifier Softmax applied in conventional GNNs.

C. Early and Late Fusion Strategies

Fig. 6 compares early and late fusion for multimodal TaGMF features followed by conventional classifiers. As shown in the figure, the late fusion approach outperforms early fusion by 9.29%, 14.9%, and 11.7% in SVM, NN, and DT. While both approaches aim to harness the complementary strengths of EEG and fNIRS modalities, our analysis reveals a distinct advantage in favor of late fusion. This strategy's superiority is attributed to its capacity to preserve and optimally leverage modality-specific features up to the decision-making stage. Late fusion allows for independent and tailored optimization of each modality's data representations. Theoretically, late fusion can require additional computation due to separate processing streams that are merged at a later stage, in the context of our study, the quantitative increase in processing time is approximately 1.5-fold compared to early fusion and 1.2 times higher than memory consumption in early fusion. This insight into the comparative effectiveness of fusion strategies underscores the value of late fusion in achieving a more nuanced and effective representation of neuroimaging data.

1) **Roles of Early and Late Fusion in Representation Learning:** Early Fusion integrates raw data or features from EEG and fNIRS modalities at an initial stage. This approach tests the hypothesis that combined feature space, when processed holistically, could reveal new insights into the neural mechanisms under study. However, it may also obscure modality-specific patterns. Late Fusion maintains the distinction between modalities until the decision level, allowing for independent optimization of representation learning for each

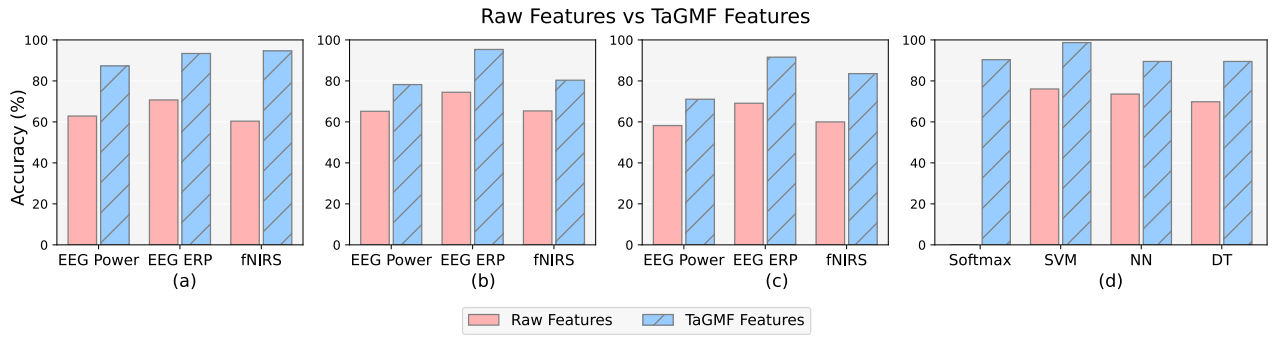


Fig. 5. Accuracy comparisons between raw features and TaGMF features in late fusion strategy for unimodal data using three classifiers. (a) SVM, (b) NN, and (c) DT. (d) TaGMF features of multimodal data using four classifiers, including Softmax, SVM, NN, and DT.

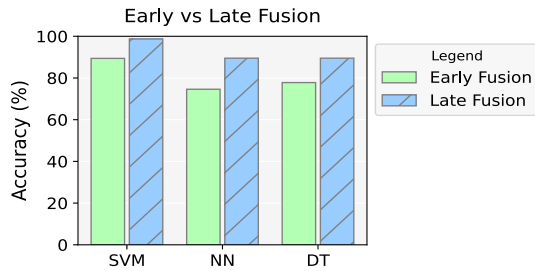


Fig. 6. Performance comparison of multimodal data in early and late fusion strategies using three classifiers: SVM, NN, and DT.

modality. This strategy acknowledges the unique contributions of EEG and fNIRS to understanding neural activity, ensuring that the richness of modality-specific information is preserved and effectively utilized.

D. Why KNN Algorithm For Graph Construction

In selecting the KNN method for our study, we aimed to establish connections between different nodes in the absence of ground truth or explicit physical knowledge. There are two popular options for constructing graphs for vectorized data, including: 1) θ -neighborhood graph and 2) KNN graph, and the latter one shows empirically better performance in graph clustering and representation learning [65]. Compared to other graph construction approaches, the KNN graph is a nonparametric modeling that does not explicitly rely on the data distribution. This is particularly critical and helpful when the size of the dataset is small or data is not distributed in Gaussian. While there is no built-in graph for our patient/subject based on EEG-fNIRS data, KNN allows us to consider multidimensional data in a more natural and data-driven manner suitable for GNN in the next step. The neighborhood information, especially the nonparametric local connections, can be propagated in GNN layer by layer. The use of KNN in this context aligns well with our goal to leverage GNNs for our analysis.

To demonstrate the efficacy of the KNN graph in our problem, we compare the KNN graph with the θ -neighborhood graph in Fig. 7. We adopt cosine similarity as the metric for θ -neighborhood graph construction, and any other similarity metric should work. θ -neighborhood graph will connect two nodes if their pairwise similarity is larger than θ , and disconnect them otherwise. All other setups are identical in this

experiment. We have used different numbers of neighbors (k) and similarity thresholds (θ) to connect the nodes to make a fair comparison. In Fig. 7, we explore $k = [3, 8, 16]$ and $\theta = [0.9, 0.7, 0.5]$ and found that a smaller k or larger θ empirically performs better. It demonstrated that a sparse graph is preferred by GNN in our problem. Overall, the KNN graphs perform better than θ -neighborhood graphs, especially with smaller k .

E. Comparison With Feature Extraction Methods

In this section, we compare TaGMF with established feature extraction approaches, including: 1) 1D-CNN; 2) common spatial pattern (CSP); and 3) filter bank CSP (FBCSP) [66]. CSP is widely applied for feature extraction for EEG signal analysis in motor imagery and BCI applications [67], [68]. FBCSP, a notable extension of CSP, effectively enhances feature extraction by utilizing filter banks to capture a broader range of signal characteristics [69]. As shown in Fig. 8, the proposed TaGMF outperforms all three methods. Notably, fused features boost the performance in all methods except for FBCSP. TaGMF excels in feature extraction from fNIRS data, likely due to its graph-based methodology that leverages the spatial details present in fNIRS data. This allows TaGMF to outperform other methods, effectively harnessing the spatially rich information of cortical hemodynamics to enhance feature extraction.

F. Parameters Analysis

A few hyperparameters are explored to verify their impacts on the proposed model, including: 1) the number of neighbors k in building the KNN graph; 2) the number of selected features in lasso regularization; and 3) the dimensionality of extracted TaGMF features.

1) *Number of Neighbors for KNN Graph*: Table I shows the impacts of different values of k on the KNN graph. Note the same value of k is applied to the three sub-graphs when building the unified graph. We conduct experiments by setting $k = 3, 8, 16, 32, 64$. As the value of k increases, the number of connections increases, too, along with the density and complexity of the graph structure. We can identify a decrease in performance with larger k in all cases.

2) *Lasso Regularization for Feature Selection*: Fig. 9 visualizes the performance by lasso feature selection, which is

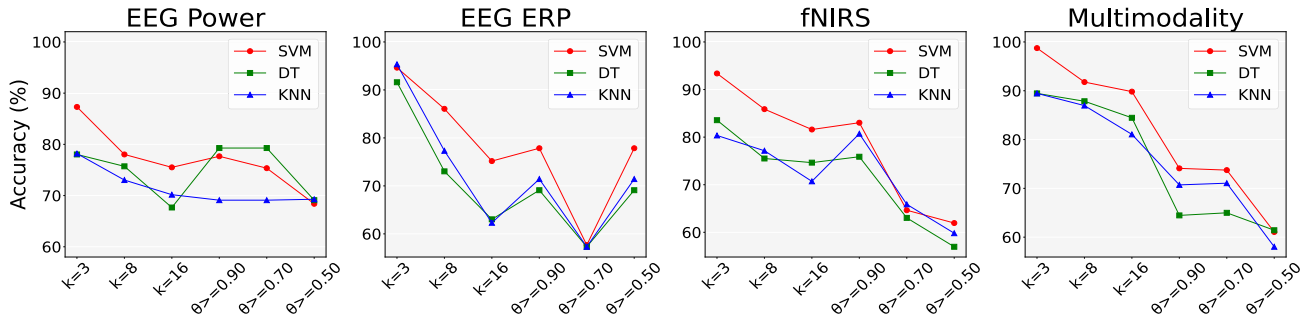


Fig. 7. Comparisons between KNN graph and θ -neighborhood graphs. Best seen in color.

TABLE I

CLASSIFICATION ACCURACY (%) ACHIEVED WITH DIFFERENT k USED IN KNN GRAPH CONSTRUCTION. THE TAGMF FEATURES AND SVM, DT, AND NN CLASSIFIERS ARE APPLIED FOR UNIMODAL AND MULTIMODAL DATA. (a) EEG POWER. (b) EEG ERP. (c) fNIRS. (d) MULTIMODALITY

(a)				(b)				(c)				(d)			
k	SVM	DT	KNN	k	SVM	DT	KNN	k	SVM	DT	KNN	k	SVM	DT	KNN
3	87.31	71.06	78.20	3	94.64	91.60	95.35	3	93.38	83.56	80.35	3	98.74	89.46	89.46
8	78.03	75.71	73.03	8	86.06	73.03	77.31	8	85.88	75.71	77.14	8	91.78	87.85	86.96
16	75.51	67.67	70.17	16	75.17	63.03	62.31	16	81.60	74.64	70.71	16	89.82	84.46	81.06
32	65.88	64.81	65.71	32	61.06	63.21	57.13	32	72.67	69.63	67.85	32	72.31	72.67	65.89
64	58.56	60.89	59.28	64	46.96	54.99	52.85	64	67.49	65.71	66.60	64	65.70	62.31	61.78

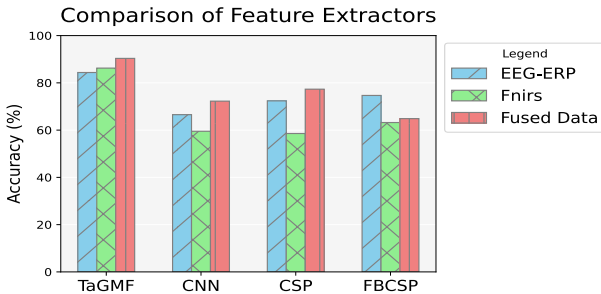


Fig. 8. Comparisons with 1D-CNN, CSP, and FBCSP. All methods follow the same experimental protocol. Note 1D-CNN is trained end-to-end, and multimodal features are fused before being fed into the 1D-CNN model. For TaGMF, CSP, and FBCSP, SVM classifiers are applied with optimized kernels and parameters.

calculated with features selected using lasso regularization with a fixed step size across different classifiers. When it comes to EEG Power, adding more contributing features has a negative impact on the accuracy of SVM and DT. However, it does improve the performance of all classifiers in EEG ERP. For fNIRS data, NN and DT experience performance declines beyond 80% of selected features, while SVM's performance remains relatively stable with a maximum decline of 1.25%. The use of multimodal signals demonstrates a notable increase in performance as more features are added, reaching a peak of 43% and declining afterward. On average, the SVM method surpasses both NN and DT by 2.5% and 6.3%, respectively. Furthermore, compared to the best unimodal performance (EEG ERP), SVM and NN improve the classification accuracy by 3.5% and 0.9%, respectively, while DT experiences a slight drop of 0.4%.

3) *Dimensionality of TaGMF Features*: Table II explores the impacts of dimensionality of TaGMF features on classification. The dimensionality is changed by setting different numbers

TABLE II

IMPACTS OF DIFFERENT DIMENSIONS OF TAGMF FEATURES ON UNIMODAL AND MULTIMODAL CLASSIFICATION ACCURACY (%)

Dimension	EEG Power	EEG ERP	fNIRS	Multimodality
8	85.53	90.17	88.56	98.40
16	84.99	95.00	91.06	98.60
32	87.31	95.00	93.38	98.70
64	91.96	94.64	96.96	93.75

of hidden nodes in TaGMF. The experiment starts with a default value of 16 and further increases it to 32 and 64. It can be observed that dimensionality 32 provides higher accuracy than others in unimodal and multimodal experiments. However, the margin becomes less significant in multimodal experiments.

G. Multilevel Graph Construction

Table III explores the efficacy of each sub-graph considered in TaG, including G(a): KNN graph for all observations without using any labels; G(b): graph with connections among observations of the same label in training data; G(c): graph with connections among observations of the same subject in the test data. At each level of graph construction, we add topological information from the previous graph. In general, adding more sub-graphs leads to improved performance for both unimodal and multimodal data. Comparing the KNN graph incorporating label information only (i.e., row a,b), we observe notable improvements in classification accuracy for SVM, NN, and DT classifiers. The most striking improvement occurs in multimodal data, when identity connections (c) are added, resulting in remarkable accuracy improvement of 30.1% for SVM, 18.9% for DT, and 18.4% for NN.

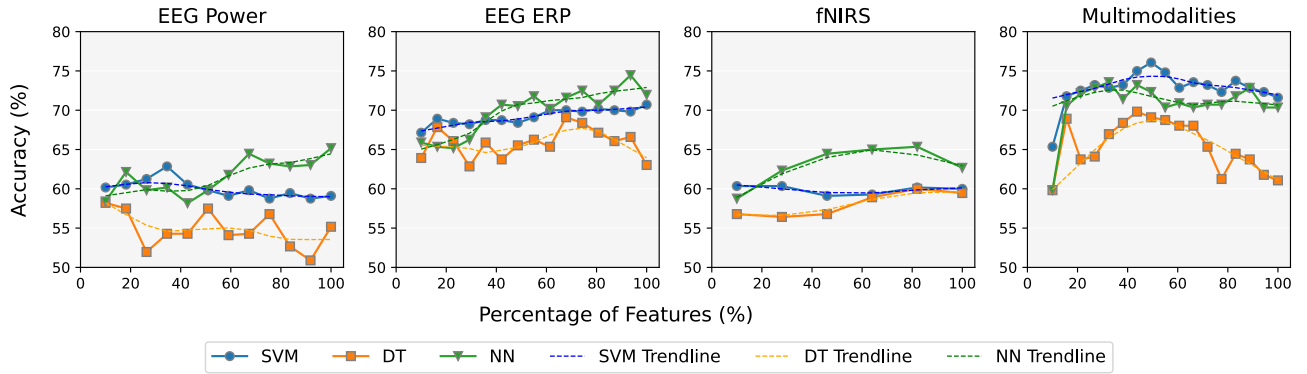


Fig. 9. Impacts of raw features selected using lasso regularization for unimodality and multimodality classification.

TABLE III

CLASSIFICATION ACCURACY (%) THROUGH DIFFERENT GRAPHS AND THEIR COMBINATIONS. G(a): KNN GRAPH FOR ALL OBSERVATIONS WITHOUT USING ANY LABELS; G(b): GRAPH WITH CONNECTIONS AMONG OBSERVATIONS OF THE SAME LABEL IN TRAINING DATA. G(c): GRAPH WITH CONNECTIONS AMONG OBSERVATIONS OF THE SAME SUBJECT IN THE TEST DATA.

(a) EEG POWER. (b) EEG ERP. (c) fNIRS. (d) MULTIMODALITY

(a)				(b)				(c)				(d)			
G	SVM	DT	KNN	G	SVM	DT	KNN	G	SVM	DT	KNN	G	SVM	DT	KNN
a	58.56	59.99	62.85	a	69.10	68.03	62.85	a	68.03	58.56	58.92	a	69.81	66.60	68.74
a,b	64.63	61.24	65.35	a,b	68.03	70.16	71.06	a,b	63.74	64.46	63.21	a,b	68.56	70.53	71.06
a,b,c	87.30	71.10	78.20	a,b,c	93.40	91.60	95.35	a,b,c	94.60	83.60	80.35	a,b,c	98.70	89.50	89.50

V. DISCUSSION

This study makes a significant contribution to the field of BCIs and neuropathological diagnosis, achieved through the innovative fusion of EEG and fNIRS brain signals using graph-based modeling. TaGMF's integration of inter-subject, intra-subject, and inter-group relations offers a deeper insight into brain response networks. The key practical implications of our research are as follows.

A. Feature Extraction of Multimodal Data

Firstly, the TaGMF framework captures spatio-temporal features from EEG-fNIRS data, offering a more comprehensive understanding of brain activity, as shown by the substantial improvement in classification accuracy. These findings aligned with [19], [20], [70] highlighted GNN's ability to capture inter-dependencies in multimodal data. Secondly, this work showcases improved diagnosis of neuropathological conditions. By effectively capturing the complex interplay of different brain activities, this approach could aid in the early detection of conditions such as autism spectrum disorders. Third, the flexibility of our graph-based modeling approach allows for customization to specific BCI applications.

B. Early and Late Fusion Strategies

The late fusion approach indicated significant performance enhancements in SVM, NN, and DT classifiers (Fig. 6). Integrating multimodal information at a later stage, the classifier benefits from discriminant unimodal features and the synergistic effects of their combined representation. The synergistic properties of EEG-fNIRS have also been experimented with in many works, including [71], [72] leading to higher diagnosis accuracy. Differently, early fusion, as shown in [73], excels when data modalities exhibit lower heterogeneity, effectively

merging in the temporal domain. In our problem, late fusion allows GNNs to maintain independent topology information of each modality and keep each in representation to enable better fusion in the next stage. Late fusion proved more suitable in our experiments, allowing each modality's strengths to be independently optimized through GNN before integration. Late fusion's superior performance can be attributed to its approach to learning the representations (features). Late fusion enables a more comprehensive and interpretable model of neural dynamics. This aligns with our observation that maintaining modality-specific representations until the final decision stage allows for a more targeted and effective use of neuroimaging data for classification tasks.

C. Number of k for KNN Graph

Our analysis of the k value in KNN graphs Fig. 1 showed an inverse relationship with classification performance. The experiments revealed an inverse correlation between the value of k and classification performance. Larger values of k allow for broader connections, forming redundant, possibly incorrect links among subjects of HC and ALS groups, which negatively impacted information propagation. Smaller values of k exclude such redundant connections, and we found that $k = 3$ captures the optimal associations with neighboring nodes. Similar experiments in [20] revealed $k = 3$ and $k = 4$ yielded the highest accuracy for *phenotypic* and *complete* graphs.

D. Lasso Regularization for Feature Selection

The number of selected features in lasso regularization influences the discriminative power of the extracted features. Fig. 9 illustrates that the optimal point varied among different modalities and classifiers. SVM classifier exhibited the highest accuracy in multimodal classification, followed by NN,

highlighting the complementary properties between EEG and fNIRS features achieved from data fusion. The DT classifier did not exhibit the highest performance on fused features; instead, it achieved enhanced results on EEG ERP data. This outcome could be attributed to the possibility of a suboptimal selection of feature sets by this classifier. Generally, classification performance improves as the number of selected features increases until a certain threshold is reached. After that point, performance begins to deteriorate due to overfitting, as aligned with the results in [74] and [75] experiments.

E. Dimensionality of GNN

The dimensionality of TaGMF features impacts both expressiveness and generalizability. A high-dimensional feature space captures fine-grained details but may introduce noise and additional computations. In contrast, a low-dimensional feature space may lead to information loss and limited representational capacity. Our analysis revealed that feature dimension significantly affects classification performance [70], [76], [77]. In unimodality, larger feature dimensions enhance performance, but in multimodal settings, feature dimensions differences are less pronounced (Table II). This is attributed to multimodal fusion effectively condensing discriminative information in a compact space, unlike unimodal setups. Thus, with fewer features, the multimodal data still exhibit strong performance.

F. Multilevel Graph Construction

The multilevel graph analysis provides valuable insights into the significance of graph construction in improving classification performance. Each level of graph construction, incorporating KNN, label information, and identity connections, contributed to enhanced classification performance. Notably, incorporating identity connections led to remarkable improvements. Moreover, the consistent improvement across modalities highlights the generalizability and robustness of the approach, making it a promising method for multimodal data analysis.

G. Comparison With Feature Extraction Methods

The performance of TaGMF in feature extraction from fNIRS data as compared to other methods, depicted in Fig. 7, suggests that graph-based methods are particularly more adept at capturing spatial properties. This inherent strength could explain why TaGMF outperforms other feature extraction methods, especially given that fNIRS data is spatially informative, reflecting cortical hemodynamic responses. The graph-based approach of TaGMF facilitates more effective utilization of these spatial characteristics, leading to improved feature extraction and, consequently, classification performance.

The implications of our findings are significant, particularly in the context of ALS classification. The accuracy improvement of approximately 22.63% achieved by our proposed fusion strategy over raw feature showcases the potential clinical utility of the TaGMF framework. More accurate classification of ALS patients and healthy individuals can

aid in early detection, monitoring disease progression, and optimizing treatment strategies. Despite the promising results, our study has some limitations. Firstly our experiments were conducted using specific datasets, and the generalizability of the TaGMF framework to other datasets warrants further investigation. Additionally, the sample size of our study may impact the generalizability of the findings. In the future, we plan to augment our current dataset and evaluate the proposed model on other multisubject datasets.

VI. CONCLUSION

In conclusion, our study introduced a novel deep graph-based framework for hybrid EEG-fNIRS signal analysis and subject group classification. By building a graph to represent multimodal brain signals and leveraging intra-subject, inter-subject, and inter-group relations, our proposed model learned discriminant feature representations, resulting in enhanced classification performance over unimodal signals. Through extensive experiments, we demonstrated the superiority of our graph-based feature fusion approach, achieving significant performance enhancements of 22.6% over unimodal data. Our findings emphasized the potential of graph-based methods in capturing complex relationships and improving classification accuracy in visuo-mental multimodal brain signal analysis.

ACKNOWLEDGMENT

The authors would like to thank the participants who took part in this study, without whom this study would not have been possible. They declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

REFERENCES

- [1] J. Sui, T. Adali, Q. Yu, J. Chen, and V. D. Calhoun, "A review of multivariate methods for multimodal fusion of brain imaging data," *J. Neurosci. Methods*, vol. 204, no. 1, pp. 68–81, Feb. 2012.
- [2] R. J. Deligani, S. B. Borgheai, J. McLinden, and Y. Shahriari, "Multimodal fusion of EEG-fNIRS: A mutual information-based hybrid classification framework," *Biomed. Opt. Exp.*, vol. 12, no. 3, p. 1635, 2021.
- [3] S. Harrison and D. Hartley, "Shedding light on the human auditory cortex: A review of the advances in near infrared spectroscopy (NIRS)," *Rep. Med. Imag.*, vol. 12, pp. 31–42, Oct. 2019.
- [4] R. Li, S. Li, J. Roh, C. Wang, and Y. Zhang, "Multimodal neuroimaging using concurrent EEG/fNIRS for poststroke recovery assessment: An exploratory study," *Neurorehabil. Neural Repair*, vol. 34, no. 12, pp. 1099–1110, Dec. 2020.
- [5] F. Zheng, B. Hu, S. Zhang, Y. Li, and X. Zheng, "EEG emotion recognition based on hierarchy graph convolution network," in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, Dec. 2021, pp. 1628–1632.
- [6] S. B. Borgheai et al., "Enhancing communication for people in late-stage ALS using an fNIRS-based BCI system," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 5, pp. 1198–1207, May 2020.
- [7] W. Yan et al., "Deep learning in neuroimaging: Promises and challenges," *IEEE Signal Process. Mag.*, vol. 39, no. 2, pp. 87–98, Mar. 2022.
- [8] X. Lin, L. Sai, and Z. Yuan, "Detecting concealed information with fused electroencephalography and functional near-infrared spectroscopy," *Neuroscience*, vol. 386, pp. 284–294, Aug. 2018.
- [9] F. Al-Shargie, "Assessment of mental stress among undergraduate students using novel fusion method on EEG and fNIRS features," in *Proc. Int. Conf. Educ. Neurosci.*, Abu Dhabi, United Arab Emirates, Mar. 2019, doi: 10.3389/conf.fnhum.2019.229.00021.

- [10] R. Li, T. Nguyen, T. Potter, and Y. Zhang, "Dynamic cortical connectivity alterations associated with Alzheimer's disease: An EEG and fNIRS integration study," *NeuroImage, Clin.*, vol. 21, 2019, Art. no. 101622.
- [11] S. Ahn and S. C. Jun, "Corrigendum: Multi-modal integration of EEG-fNIRS for brain-computer interfaces—Current limitations and future directions," *Frontiers Hum. Neurosci.*, vol. 15, p. 34, Feb. 2021.
- [12] T. Nguyen, S. Ahn, H. Jang, S. C. Jun, and J. G. Kim, "Utilization of a combined EEG/fNIRS system to predict driver drowsiness," *Sci. Rep.*, vol. 7, no. 1, pp. 1–10, 2017.
- [13] M. U. Khan and M. A. H. Hasan, "Hybrid EEG-fNIRS BCI fusion using multi-resolution singular value decomposition (MSVD)," *Frontiers Hum. Neurosci.*, vol. 14, Dec. 2020, Art. no. 599802.
- [14] S. Fazli et al., "Enhanced performance by a hybrid NIRS-EEG brain computer interface," *NeuroImage*, vol. 59, no. 1, pp. 519–529, Jan. 2012.
- [15] F. Al-Shargie, T. B. Tang, and M. Kiguchi, "Assessment of mental stress effects on prefrontal cortical activities using canonical correlation analysis: An fNIRS-EEG study," *Biomed. Opt. Exp.*, vol. 8, no. 5, p. 2583, 2017.
- [16] A. P. Buccino, H. O. Keles, and A. Omurtag, "Hybrid EEG-fNIRS asynchronous brain-computer interface for multiple motor tasks," *PLoS ONE*, vol. 11, no. 1, Jan. 2016, Art. no. e0146610.
- [17] F. Al-Shargie, M. Kiguchi, N. Badruddin, S. C. Dass, A. F. M. Hani, and T. B. Tang, "Mental stress assessment using simultaneous measurement of EEG and fNIRS," *Biomed. Opt. Exp.*, vol. 7, no. 10, pp. 3882–3898, Oct. 2016.
- [18] M. Saadati, J. Nelson, and H. Ayaz, "Convolutional neural network for hybrid fNIRS-EEG mental workload classification," in *Proc. Int. Conf. Appl. Hum. Factors Ergonom. Cham, Switzerland: Springer*, 2019, pp. 221–232.
- [19] J. Zhang, X. Zhang, G. Chen, and Q. Zhao, "Granger-causality-based multi-frequency band EEG graph feature extraction and fusion for emotion recognition," *Brain Sci.*, vol. 12, no. 12, p. 1649, Dec. 2022.
- [20] S. Parisot et al., "Disease prediction using graph convolutional networks: Application to autism spectrum disorder and Alzheimer's disease," *Med. Image Anal.*, vol. 48, pp. 117–130, Aug. 2018.
- [21] W. Zheng, L. Yan, C. Gou, and F.-Y. Wang, "Graph attention model embedded with multi-modal knowledge for depression detection," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2020, pp. 1–6.
- [22] Y. Xie et al., "Anxiety and depression diagnosis method based on brain networks and convolutional neural networks," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 1503–1506.
- [23] X. Li et al., "BrainGNN: Interpretable brain graph neural network for fMRI analysis," *Med. Image Anal.*, vol. 74, Dec. 2021, Art. no. 102233.
- [24] K. Gopinath, C. Desrosiers, and H. Lombaert, "Graph convolutions on spectral embeddings for cortical surface parcellation," *Med. Image Anal.*, vol. 54, pp. 297–305, May 2019.
- [25] M. Ghorbani, A. Kazi, M. Soleymani Baghshah, H. R. Rabiee, and N. Navab, "RA-GCN: Graph convolutional network for disease prediction problems with imbalanced data," *Med. Image Anal.*, vol. 75, Jan. 2022, Art. no. 102272.
- [26] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE Trans. Neural Netw.*, vol. 20, no. 1, pp. 61–80, Jan. 2008.
- [27] M. A. Yücel et al., "Best practices for fNIRS publications," *Neurophotonics*, vol. 8, no. 1, Jan. 2021, Art. no. 012101.
- [28] G. Li, M. Müller, B. Ghanem, and V. Koltun, "Training graph neural networks with 1000 layers," in *Proc. Int. Conf. Mach. Learn.*, in Proceedings of Machine Learning Research, M. Meila and T. Zhang, Eds., 1000, pp. 6437–6449. [Online]. Available: <https://proceedings.mlr.press/v139/li21o.html>
- [29] M. Wang et al., "Hierarchical structured sparse learning for schizophrenia identification," *Neuroinformatics*, vol. 18, no. 1, pp. 43–57, Jan. 2020.
- [30] L. Cosmo, A. Kazi, S.-A. Ahmadi, N. Navab, and M. Bronstein, "Latent-graph learning for disease prediction," 2020, *arXiv:2003.13620*.
- [31] S. Parisot et al., "Spectral graph convolutions for population-based disease prediction," in *Medical Image Computing and Computer Assisted Intervention 2017*, M. Descoteaux, L. Maier-Hein, A. Franz, P. Jannin, D. L. Collins, S. Duchesne, Eds. Cham, Switzerland: Springer, 2017, pp. 177–185.
- [32] A. Kazi et al., "Graph convolution based attention model for personalized disease prediction," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. Cham, Switzerland: Springer*, 2019, pp. 122–130.
- [33] A. Kazi, S. A. Krishna, S. Shekarforoush, K. Kortuem, S. Albarqouni, and N. Navab, "Self-attention equipped graph convolutions for disease prediction," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 1896–1899.
- [34] S. Deng, S. Wang, H. Rangwala, L. Wang, and Y. Ning, "Cola-GNN: Cross-location attention based graph neural networks for long-term ILI prediction," in *Proc. 29th ACM Int. Conf. Inf. Knowl. Manage.*, Oct. 2020, pp. 245–254.
- [35] H. Lu and S. Uddin, "A weighted patient network-based framework for predicting chronic diseases using graph neural networks," *Sci. Rep.*, vol. 11, no. 1, pp. 1–12, Nov. 2021.
- [36] C. Wang et al., "A cancer survival prediction method based on graph convolutional network," *IEEE Trans. Nanobiosci.*, vol. 19, no. 1, pp. 117–126, Jan. 2020.
- [37] K. Zhan, X. Chang, J. Guan, L. Chen, Z. Ma, and Y. Yang, "Adaptive structure discovery for multimedia analysis using multiple features," *IEEE Trans. Cybern.*, vol. 49, no. 5, pp. 1826–1834, May 2019.
- [38] P. Zhong, D. Wang, and C. Miao, "EEG-based emotion recognition using regularized graph neural networks," *IEEE Trans. Affect. Comput.*, vol. 13, no. 3, pp. 1290–1301, Jul. 2022.
- [39] Y. Yin, X. Zheng, B. Hu, Y. Zhang, and X. Cui, "EEG emotion recognition using fusion model of graph convolutional neural networks and LSTM," *Appl. Soft Comput.*, vol. 100, Mar. 2021, Art. no. 106954.
- [40] G. Wen, P. Cao, H. Bao, W. Yang, T. Zheng, and O. Zaiane, "MVS-GCN: A prior brain structure learning-guided multi-view graph convolution network for autism spectrum disorder diagnosis," *Comput. Biol. Med.*, vol. 142, Jan. 2022, Art. no. 105239.
- [41] J. Faskowitz, F. Z. Esfahlani, Y. Jo, O. Sporns, and R. F. Betzel, "Edge-centric functional network representations of human cerebral cortex reveal overlapping system-level architecture," *Nature Neurosci.*, vol. 23, no. 12, pp. 1644–1654, Dec. 2020.
- [42] L. E. Ismail and W. Karwowski, "A graph theory-based modeling of functional brain connectivity based on EEG: A systematic review in the context of neuroergonomics," *IEEE Access*, vol. 8, pp. 155103–155135, 2020.
- [43] A. Dragomir and A. Omurtag, "Brain's networks and their functional significance in cognition," in *Handbook of Neuroengineering*. Cham, Switzerland: Springer, 2021, pp. 1–30.
- [44] S. Ge et al., "A brain-computer interface based on a few-channel EEG-fNIRS bimodal system," *IEEE Access*, vol. 5, pp. 208–218, 2017.
- [45] S. Uddin, A. Khan, M. E. Hossain, and M. A. Moni, "Comparing different supervised machine learning algorithms for disease prediction," *BMC Med. Informat. Decis. Making*, vol. 19, no. 1, pp. 1–16, Dec. 2019.
- [46] S. Mohan, C. Thirumalai, and G. Srivastava, "Effective heart disease prediction using hybrid machine learning techniques," *IEEE Access*, vol. 7, pp. 81542–81554, 2019.
- [47] A. Alhudhaif, "An effective classification framework for brain-computer interface system design based on combining of fNIRS and EEG signals," *PeerJ Comput. Sci.*, vol. 7, p. e537, May 2021.
- [48] A. M. Chiarelli, P. Croce, A. Merla, and F. Zappasodi, "Deep learning for hybrid EEG-fNIRS brain-computer interface: Application to motor imagery classification," *J. Neural Eng.*, vol. 15, no. 3, Jun. 2018, Art. no. 036028.
- [49] H. Ghonchi, M. Fateh, V. Abolghasemi, S. Ferdowsi, and M. Rezvani, "Deep recurrent-convolutional neural network for classification of simultaneous EEG-fNIRS signals," *IET Signal Process.*, vol. 14, no. 3, pp. 142–153, May 2020.
- [50] S. Kim, D. Y. Shin, T. Kim, S. Lee, J. K. Hyun, and S.-M. Park, "Enhanced recognition of amputated wrist and hand movements by deep learning method using multimodal fusion of electromyography and electroencephalography," *Sensors*, vol. 22, no. 2, p. 680, Jan. 2022.
- [51] P. Sirpal, A. Kassab, P. Pouliot, and D. K. Nguyen, "fNIRS improves seizure detection in multimodal EEG-fNIRS recordings," *J. Biomed. Opt.*, vol. 24, no. 5, p. 1, Feb. 2019.
- [52] Md. H. R. Rabbani and S. Md. R. Islam, "Multimodal decision fusion of EEG and fNIRS signals," in *Proc. 5th Int. Conf. Electr. Eng. Inf. Commun. Technol. (ICEEICT)*, Nov. 2021, pp. 1–6.
- [53] Y. Duan, X. Liu, and Y. Lian, "Progress on hybrid EEG-fNIRS system and its application," in *Proc. 2nd Int. Symp. Artif. Intell. Med. Sci.*, Oct. 2021, pp. 1–6.
- [54] T. K. K. Ho, I. Kim, Y. Jeon, J.-I. Song, and J. Gwak, "An EEG-fNIRS hybridization technique in the multi-class classification of Alzheimer's disease facilitated by machine learning," in *Proc. Korean Soc. Comput. Inf. Conf.*, 2021, pp. 305–307.

- [55] M. Ramirez, S. Kaheh, M. A. Khalil, and K. George, "Application of convolutional neural network for classification of consumer preference from hybrid EEG and fNIRS signals," in *Proc. IEEE 12th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Jan. 2022, pp. 1024–1028.
- [56] G. Brown, A. Pocock, M.-J. Zhao, and M. Luján, "Conditional likelihood maximisation: A unifying framework for information theoretic feature selection," *J. Mach. Learn. Res.*, vol. 13, no. 1, pp. 27–66, Jan. 2012.
- [57] X. Yin et al., "A hybrid BCI based on EEG and fNIRS signals improves the performance of decoding motor imagery of both force and speed of hand clenching," *J. Neural Eng.*, vol. 12, no. 3, Jun. 2015, Art. no. 036004.
- [58] L. Qiu, Y. Zhong, Z. He, and J. Pan, "Improved classification performance of EEG-fNIRS multimodal brain-computer interface based on multi-domain features and multi-level progressive learning," *Frontiers Hum. Neurosci.*, vol. 16, 2022, Art. no. 973959.
- [59] F. Scarpa, S. Cutini, P. Scatturin, R. Dell'Acqua, and G. Sparacino, "Bayesian filtering of human brain hemodynamic activity elicited by visual short-term maintenance recorded through functional near-infrared spectroscopy (fNIRS)," *Opt. Exp.*, vol. 18, no. 25, p. 26550, 2010.
- [60] L. Kocsis, P. Herman, and A. Eke, "The modified Beer-Lambert law revisited," *Phys. Med. Biol.*, vol. 51, no. 5, pp. N91–N98, Mar. 2006.
- [61] S. B. Borgheai et al., "Multimodal exploration of non-motor neural functions in ALS patients using simultaneous EEG-fNIRS recording," *J. Neural Eng.*, vol. 16, no. 6, Nov. 2019, Art. no. 066036.
- [62] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," 2016, *arXiv:1609.02907*.
- [63] V. Fonti and E. Belitser, "Feature selection using Lasso," *VU Amsterdam Res. Paper Bus. Anal.*, vol. 30, pp. 1–25, Mar. 2017.
- [64] H. Zhang, J. Wang, Z. Sun, J. M. Zurada, and N. R. Pal, "Feature selection for neural networks using group lasso regularization," *IEEE Trans. Knowl. Data Eng.*, vol. 32, no. 4, pp. 659–673, Apr. 2020.
- [65] U. von Luxburg, "A tutorial on spectral clustering. Statistics and computing," in *Data Structures and Algorithms (cs. DS): Machine Learning*, 2007, pp. 395–416.
- [66] K. Keng Ang, Z. Yang Chin, H. Zhang, and C. Guan, "Filter bank common spatial pattern (FBCSP) in brain-computer interface," in *Proc. IEEE Int. Joint Conf. Neural Netw., IEEE World Congr. Comput. Intell.*, Jun. 2008, pp. 2390–2397.
- [67] R. Ameri, A. Pouyan, and V. Abolghasemi, "EEG signal classification based on sparse representation in brain computer interface applications," in *Proc. 22nd Iranian Conf. Biomed. Eng. (ICBME)*, Iran, Nov. 2015, pp. 21–24.
- [68] H. Wang and W. Zheng, "Local temporal common spatial patterns for robust single-trial EEG classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 16, no. 2, pp. 131–139, Apr. 2008.
- [69] S. Akuthota, K. Rajkumar, and J. Ravichander, "EEG based motor imagery BCI using four class iterative filtering & four class filter bank common spatial pattern," in *Proc. Int. Conf. Adv. Electron., Commun., Comput. Intell. Inf. Syst. (ICAECIS)*, Apr. 2023, pp. 429–434.
- [70] S. Zhu et al., "A graph-based feature extraction algorithm towards a robust data fusion framework for brain-computer interfaces," in *Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Nov. 2021, pp. 878–881.
- [71] Q. He, L. Feng, G. Jiang, and P. Xie, "Multimodal multitask neural network for motor imagery classification with EEG and fNIRS signals," *IEEE Sensors J.*, vol. 22, no. 21, pp. 20695–20706, Nov. 2022.
- [72] W.-C. Su et al., "Simultaneous multimodal fNIRS-EEG recordings reveal new insights in neural activity during motor execution, observation, and imagery," *Sci. Rep.*, vol. 13, no. 1, p. 5151, Mar. 2023.
- [73] Y. Li, X. Zhang, and D. Ming, "Early-stage fusion of EEG and fNIRS improves classification of motor imagery," *Frontiers Neurosci.*, vol. 16, Jan. 2023, Art. no. 1062889.
- [74] S. M. I. Hosni et al., "A graph-based nonlinear dynamic characterization of motor imagery toward an enhanced hybrid BCI," *Neuroinformatics*, vol. 20, no. 4, pp. 1169–1189, Oct. 2022.
- [75] M. Meng, L. Dai, Q. She, Y. Ma, and W. Kong, "Crossing time windows optimization based on mutual information for hybrid BCI," *Math. Biosci. Eng.*, vol. 18, no. 6, pp. 7919–7935, 2021, doi: [10.3934/mbe.2021392](https://doi.org/10.3934/mbe.2021392).
- [76] A. Zafar et al., "A hybrid GCN and filter-based framework for channel and feature selection: An fNIRS-BCI study," *Int. J. Intell. Syst.*, vol. 2023, pp. 1–14, Mar. 2023.
- [77] L. Zhao, R. Liu, S. Li, X. Wang, and D. Bao, "Spatio-temporal variable structure graph neural network for EEG data classification," in *Proc. 6th Int. Symp. Auto. Syst. (ISAS)*, Jun. 2023, pp. 1–6.