

# Graph Neural Network-Based EEG Classification: A Survey

Dominik Klepl<sup>1</sup>, Min Wu<sup>2</sup>, *Senior Member, IEEE*, and Fei He<sup>3</sup>, *Senior Member, IEEE*

**Abstract**—Graph neural networks (GNN) are increasingly used to classify EEG for tasks such as emotion recognition, motor imagery and neurological diseases and disorders. A wide range of methods have been proposed to design GNN-based classifiers. Therefore, there is a need for a systematic review and categorisation of these approaches. We exhaustively search the published literature on this topic and derive several categories for comparison. These categories highlight the similarities and differences among the methods. The results suggest a prevalence of spectral graph convolutional layers over spatial. Additionally, we identify standard forms of node features, with the most popular being the raw EEG signal and differential entropy. Our results summarise the emerging trends in GNN-based approaches for EEG classification. Finally, we discuss several promising research directions, such as exploring the potential of transfer learning methods and appropriate modelling of cross-frequency interactions.

**Index Terms**—Graph neural network, classification, EEG, neuroscience, deep learning.

## I. INTRODUCTION

ELECTROENCEPHALOGRAPHY (EEG) is a non-invasive technique used for recording electrical brain activity with a wide range of applications in cognitive neuroscience [1], clinical diagnosis [2], [3], and brain-computer interfaces [4], [5]. However, analysing EEG signals poses several challenges, including a low signal-to-noise ratio, non-stationarity resulting from brain dynamics, and the multivariate nature of the signals [6], [7]. In this review, we focus on the classification of EEG, such as emotion recognition, motor imagery recognition or neurological disorders and diseases.

Traditional feature extraction methods for EEG classification, such as common spatial patterns [6], wavelet

transform [8], and Hilbert-Huang transform [9], have been commonly employed. These methods aim to extract meaningful features from EEG signals [10], [11], with key features like power spectral density (PSD) [7] to characterise brain states. However, relying on such manually defined features to train machine learning classifiers has several limitations. Subjectivity and biases in feature selection, along with time-consuming engineering and selection processes, limit scalability and generalisation [7], [12]. Automated feature extraction methods are needed to overcome these limitations, improve efficiency, reduce bias, and enhance classifier adaptability to different EEG datasets.

Deep learning architectures, such as convolutional neural networks (CNN) and long short-term memory (LSTM) networks, have also been explored for EEG analysis [13], [14]. However, they face challenges in effectively capturing the spatial dependencies between electrodes and handling the temporal dynamics of EEG signals [7]. Modelling the complex sequential and spatial relationships in EEG data is crucial for more accurate classification and analysis.

Network neuroscience offers an alternative approach to EEG modelling by framing the signals as a graph. The brain exhibits a complex network structure, with neurons forming connections and communicating with each other [15]. Analysing EEG data as a graph enables the study of network properties, including functional connectivity, providing insights into brain function and dysfunction [12], [16], [17]. Graph-based analysis facilitates the examination of network features, node importance, community structure, and information flow, offering insights into brain organisation and dynamics. Such graph-theory-based features were shown to be powerful predictive features for EEG classification [12], [17], [18], [19], [20], [21], [22]. However, these features have the same limitations as manually defined features based on traditional EEG analysis methods introduced above.

Graph neural networks (GNN) emerge as a powerful tool for modelling neurophysiological data [23], such as EEG, within the network neuroscience framework [7], [24]. GNNs are specifically designed to operate on graph-structured data. They can effectively leverage the spatial structure within EEG data to extract features, uncover patterns and make predictions based on the complex interactions between different electrodes. Designing GNN models for

Manuscript received 18 October 2023; revised 22 December 2023; accepted 12 January 2024. Date of publication 18 January 2024; date of current version 25 January 2024. This work was supported by the Agency for Science, Technology and Research (A\*STAR), AI, Analytics and Informatics (AI3) Horizontal Technology Program Office (HTPO) under Grant C211118015, and in part by the Engineering and Physical Sciences Research Council (EPSRC) under Grant EP/X020193/1. (Corresponding authors: Fei He; Min Wu.)

Dominik Klepl and Fei He are with the Centre for Computational Science and Mathematical Modelling, Coventry University, CV1 2JH Coventry, U.K. (e-mail: klepld@uni.coventry.ac.uk; fei.he@coventry.ac.uk).

Min Wu is with the Institute for Infocomm Research (I<sup>2</sup>R), Agency for Science, Technology and Research (A\*STAR), Singapore 138632 (e-mail: wumin@i2r.a-star.edu.sg).

Digital Object Identifier 10.1109/TNSRE.2024.3355750

EEG classification will likely improve classification tasks and potentially uncover new insights in neuroscience.

Motivated by the potential of GNNs and an increasing number of recent papers proposing GNN for various EEG classification tasks, there is an urgent need for a comprehensive review of GNN models for EEG classification. The main contributions of this paper include:

- Identifying emerging trends of GNN models tailored for EEG classification.
- Reviewing popular graph convolutional layers and their applicability to EEG data.
- Providing a unified overview of node feature and brain graph structure definitions in the context of EEG analysis.
- Examining techniques for transforming sets of node feature embeddings into a single graph embedding for graph classification tasks.

By addressing these essential aspects, this review paper will provide a comprehensive and in-depth analysis of the application of GNN models for EEG classification. The findings and insights gained from this review will serve as a resource to navigate this emerging field and identify promising future research directions.

## II. OVERVIEW OF GRAPH NEURAL NETWORKS

Graphs are widely used to capture complex relationships and dependencies in various domains, such as social networks, biological networks, and knowledge graphs. The problem of graph classification, which aims to assign a label to an entire graph, has gained attention in recent years. GNNs offer a promising solution to this problem by extending the concept of convolution from Euclidean inputs to graph-structured data. GNNs have been successfully applied in a wide range of fields, such as biology [23], bioinformatics [25], network neuroscience [26], chemistry [27], [28], drug design and discovery [29], [30], natural language processing [31], [32], recommendation systems [33], [34], traffic prediction [35], [36] and finance [37].

In graph classification problems, the input is a set of graphs, each with its own set of nodes, edges, and node features. Let  $G = (V, E, H)$  denote a featured graph, where  $V$  represents the set of nodes,  $E$  represents the set of edges connecting the nodes, and  $H$  represents the  $V \times D$  matrix of  $D$ -dimensional node features. In the case of EEG, the EEG channels are the nodes, and edges represent structural or functional connectivity between pairs of nodes. Each graph  $G$  is associated with a label  $y$ , indicating its class. The goal is to learn a function  $f(G) \rightarrow y$  that can predict the class label  $y$  given an input graph  $G$ . A general structure of a GNN model for EEG classification is presented in Fig 1.

Compared to other deep learning models, GNNs offer several advantages. First, GNNs were specifically designed for graph-structured inputs. This means that GNNs can adapt to irregularly structured inputs, i.e. graphs with varying numbers of nodes, compared to traditional deep learning, such as CNN, that require fixed-size inputs. Next, GNNs can simultaneously learn information from node features and the graph structure by accepting two inputs: node feature matrix and graph

structure. Such simultaneous integration is not possible with traditional deep learning methods.

Multiple types of GNNs have been well introduced in [38], [39]. In this survey, we briefly introduce the two main branches of GNNs, namely, spatial and spectral GNNs (Fig. 2). Other types of GNNs, such as attention GNNs [40], recurrent GNNs [41], and graph transformers [42], can be viewed as special cases of spatial GNNs, and thus we will not provide detailed discussion in this survey. Both spatial and spectral GNNs aim to extend the convolution mechanism to graph data. For a detailed review of their similarities and differences, see [43]. Moreover, for a comparison of different GNNs in terms of computational complexity, see [38].

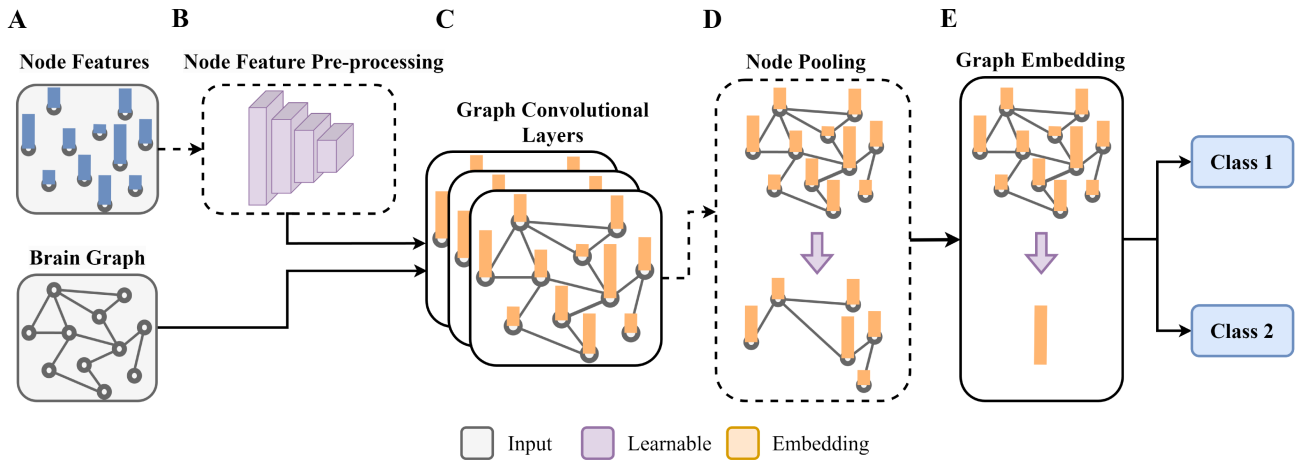
Spatial GNNs aggregate information from neighbouring nodes, similar to traditional convolution applied to image data aggregating information from adjacent pixels. Stacking multiple spatial GNN layers leads to information aggregation from various scales going from local to global patterns being captured in early and later layers, respectively. In contrast, spectral GNNs perform information aggregation in the graph frequency domain, with low-frequency and high-frequency components capturing global and local patterns, respectively. However, both approaches learn to capture local and global patterns within the graph, i.e. high and low-frequency information in the spectral domain. The advantage of spectral GNNs is their connection to graph signal processing, allowing for interpretation from the perspective of graph filters. However, spectral GNNs do not generalise well to large graphs since they depend on the eigendecomposition of graph Laplacian. In contrast, spatial GNNs can be applied to large graphs since they perform only local message-passing. On the other hand, spatial GNNs may be challenging to interpret and prone to overfitting because of over-smoothing, where embeddings of all nodes become similar.

### A. Spatial GNNs

Spatial GNNs directly operate on the graph structure via the adjacency matrix operator. Given a set of nodes and associated features, spatial GNNs perform neighbourhood aggregation to derive node embeddings. This process is referred to as message passing. Intuitively, nodes connected by edges should have similar node embeddings, i.e. local node similarity. Message passing implements this idea by updating node embeddings with aggregated information collected from the node's neighbourhood. Formally, the node update equation in  $l^{\text{th}}$  layer of spatial GNN with  $L$  layers is defined as follows:

$$h_i^{(l+1)} = \sigma \left( W_1^{(l)} h_i^{(l)} + \sum_{j \in \mathcal{N}(v_i)} W_2^{(l)} h_j^{(l)} e_{ji} \right), \quad (1)$$

where  $h_i$  is the node embedding vector, or when  $l = 1$ , this is the input node feature vector,  $\sigma$  is the activation function,  $\sum$  is the aggregation function,  $\mathcal{N}(v_i)$  is the neighbourhood of node  $v_i$ ,  $W \in \mathbb{R}^{d_1 \times d_2}$  is a learnable parameter matrix projecting node embeddings from input dimension  $d_1$  to hidden dimension  $d_2$  and  $e_{ji}$  is the edge weight ( $e_{ji} = 1$  for unweighted graphs).



**Fig. 1.** General architecture of a GNN model for classification of EEG. (A) The input to the model consists of node features and a possibly learnable brain graph structure. (B) Optionally, the node features can undergo pre-processing via a neural network. (C) Next, the node features are passed to a block of graph convolutional layers, where node embeddings are learned. (D) Then, a node pooling module can be utilised to coarsen the graph. Node pooling may contain learnable parameters as well. (E) Finally, the set of node embeddings forms a graph embedding, which can be used to predict the outcome.

A single spatial GNN layer aggregates information from the 1-hop neighbourhood. Thus, to increase the reception field of the model,  $L$  spatial GNN layers can be stacked to aggregate information from up to  $L$ -hop neighbourhoods. A disadvantage of spatial GNNs is the difficulty of training deep models with many layers. With an increasing number of layers, the node embeddings become increasingly smooth, i.e. variance among embeddings of all nodes decreases. This happens when the messages already contain aggregated information from the whole graph; continual message passing of such saturated messages leads to oversmoothing, i.e., all node embeddings becoming essentially identical.

### B. Spectral GNNs

Spectral GNNs can also be applied to EEG classification tasks by leveraging the spectral domain analysis of graph-structured data. The EEG graph is transformed into the spectral domain using the Graph Fourier Transform (GFT) and Graph Signal Processing (GSP) techniques. For a detailed review of spectral GNN methods, please refer to [44].

The graph spectrum is defined as the eigendecomposition of the graph Laplacian matrix. The GFT is then defined as  $\hat{\mathbf{H}} = \mathbf{U}^T \mathbf{H}$ , its inverse as  $\mathbf{H} = \mathbf{U} \hat{\mathbf{H}}$ , where  $\mathbf{U}$  is the orthonormal matrix of eigenvectors of the graph Laplacian  $\mathbf{L}$  and  $\mathbf{H} \in \mathbb{R}^{N \times D}$  is the matrix of node feature vectors with  $N$  and  $D$  being the number of nodes and dimensionality of node features, respectively. The graph Laplacian is defined as  $\mathbf{L} = \mathbf{D} - \mathbf{A}$ , but often the normalised version is preferred:  $\hat{\mathbf{L}} = \mathbf{I} - \mathbf{D}^{-1/2} \mathbf{A} \mathbf{D}^{-1/2}$  ( $\mathbf{A}$  and  $\mathbf{D}$  are the adjacency and degree matrices, respectively).

Spectral GNN is then typically defined as the convolution ( $*$ ) of a signal defined on graph  $\mathbf{H}$  and a spatial kernel  $g$  in the spectral domain, thus becoming an element-wise multiplication ( $\odot$ ):

$$\mathbf{H} * g = \mathbf{U} \left( \left( \mathbf{U}^T \mathbf{H} \right) \odot \left( \mathbf{U}^T g \right) \right). \quad (2)$$

Generally,  $\mathbf{U}^T g$  is defined as a learnable diagonal matrix  $\mathbf{G} = \text{diag}(g_1, \dots, g_V)$  spectral filter [44].

However, the full spectral graph convolution can be computationally expensive. A popular approximation is the Chebyshev GNN (ChebConv) [45], which performs localised spectral filtering on the graph. The node embedding update equation of a ChebConv is defined as:

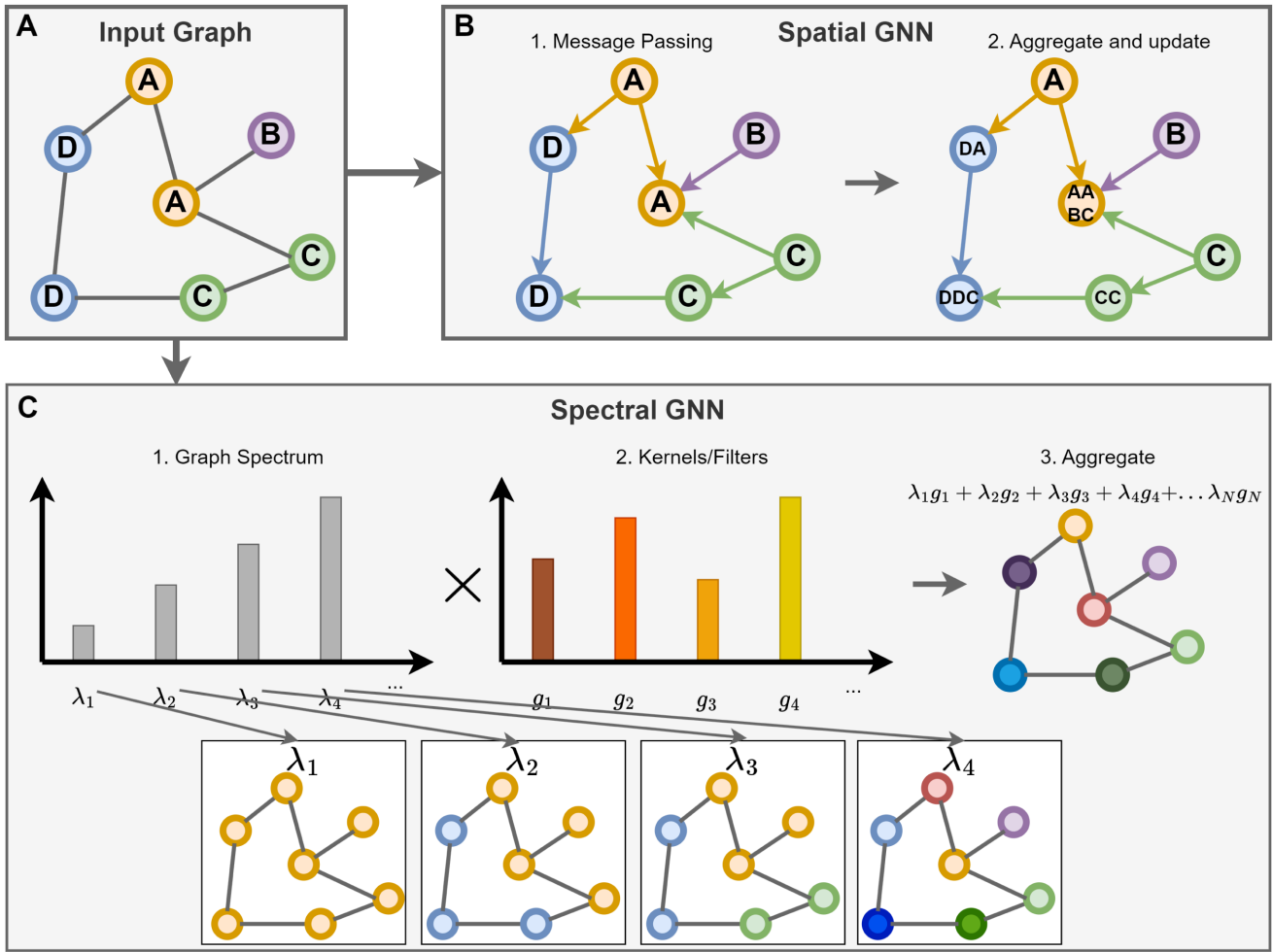
$$\mathbf{H} * g \approx \sum_{i=1}^K \Theta_i T_i(\hat{\mathbf{L}}'), \quad (3)$$

where  $\Theta \in \mathbb{R}^{K \times d \times d}$  are learnable parameters,  $T_i(\hat{\mathbf{L}}') = 2T_{i-1}(\hat{\mathbf{L}}') - T_{i-2}(\hat{\mathbf{L}}')$ ,  $T_1(\hat{\mathbf{L}}') = \mathbf{H}$ ,  $T_2(\hat{\mathbf{L}}') = \hat{\mathbf{L}}' \mathbf{H}$ , and  $\hat{\mathbf{L}}' = \frac{2\hat{\mathbf{L}}}{\lambda_{max}} - \mathbf{I}$  ( $\lambda_{max}$  is the largest eigenvalue of  $\hat{\mathbf{L}}$ , often approximated as  $\lambda_{max} = 2$ ). The  $K$  parameter controls the size of the Chebyshev filter.

However, spectral GNNs are limited to input graphs with a fixed number of nodes. This is because of the explicit use of the graph Laplacian. This is in contrast to spatial GNNs, which do not rely on explicitly materialising the adjacency matrix.

### III. SURVEY RESULTS

This survey is based on a review of 63 articles. These articles were selected by title and abstract screening from a search on Google Scholar and ScienceDirect queried on November 1st, 2022. The search query for collecting the articles was defined as: (“Graph neural network” OR “Graph convolutional network”) AND (“Electroencephalography” OR “EEG”). Both peer-reviewed articles and preprints were included. All types of EEG classification tasks were included. We summarise the various types of EEG classification tasks identified in the surveyed papers in Fig 3. The most common classification tasks are emotion recognition, epilepsy diagnosis and detection and motor imagery. However, the type of classification task should have a relatively minor effect on the GNN architecture design. Thus, we do not analyse and discuss



**Fig. 2.** Illustration of core mechanisms of spatial and spectral GNNs. **A)** An undirected featured graph is given as an example input graph with node features shown as node labels and colours. **B)** Spatial GNNs operate in the graph domain directly using message passing to update node embeddings. 1) Messages, i.e. transformed node features or embeddings, are sent along edges. For simplicity, we show only one direction of the flow of messages. 2) The collected messages at each node are aggregated using a permutation-invariant function and are fused with the original node embedding to form an updated node embedding. Thus, one spatial GNN layer results in node embeddings containing information about the 1-hop neighbourhood of a given node. Thus,  $L$  layers are required for node embeddings to access the information from the  $L$ -hop neighbourhood. **C)** In contrast, spectral GNNs operate in the graph spectral domain. 1) Node features are treated as signals on top of a graph and are deconstructed into graph frequencies given by the eigendecomposition of the graph Laplacian. Graph frequencies can be interpreted as variations of the signals. 2) The contribution of each graph frequency is weighted by the set of learnable kernels  $G$  that effectively function as graph filters. 3) Node embeddings are then obtained by aggregating the filtered graph frequencies and transforming them back to the spatial graph domain. Thus, full spectral GNNs can access information from  $N$ -hop neighbourhoods where  $N$  is the number of nodes of a given graph. However, in practice, approximations such as Chebyshev graph convolution restrict this to the chosen hop size.

this in detail. Instead, we survey the various GNN-based methods for EEG classification, intending to systematically categorise the types of GNN modules and identify emerging trends in this field independent of the specific classification task.

In the remaining portion of this paper, we report the categories of comparisons we identified in the surveyed papers. These are based on the different modules of the proposed GNN-based models. Specifically, these are:

- Definition of brain graph structure
- Type of node features
- Type of graph convolutional layer
- Node feature preprocessing
- Node pooling mechanisms
- Formation of graph embedding from the set of node embeddings

The following sections will provide further details on these categories, and the paper will conclude by discussing trends and proposing plausible directions for future research.

#### IV. DEFINITION OF BRAIN GRAPH STRUCTURE

The first part of the input to a GNN model is the brain graph structure inferred from the EEG data itself (Fig. 1A). We summarise the methods for defining the brain graphs in Table I. These methods can be generally categorised as learnable or pre-defined.

An alternative categorisation of the brain graph structures is the functional (FC) and the “structural” connectivity (SC). Generally, SC graphs are pre-defined, whereas FC graphs can be both pre-defined and learnable. SC in the classical sense of physical connections between brain regions is not possible to obtain using EEG signals since these are recorded at the scalp

TABLE I  
OVERVIEW OF METHODS FOR OBTAINING THE BRAIN GRAPH STRUCTURE

Method	Learnable	Pre-defined	Papers
Distance between electrode positions	✗	✓	[46]–[64]
Functional connectivity measure	✗	✓	[7], [47], [49], [50], [55], [59], [61], [65]–[86]
Manually defined	✗	✓	[46], [53], [54], [57], [58], [62], [87]
Shared learnable mask	✓	✗	[46], [63], [69], [72], [82], [88]–[92]
Feature similarity	✓	✗	[56], [63], [72], [87], [92]–[94]
Feature distance	✓	✗	[64], [95]–[97]
Transformer-style attention	✓	✗	[98], [99]
Concatenation attention	✓	✗	[81], [100]
Dense projection	✓	✗	[101]–[103]
LSTM-based	✓	✗	[104]
Multiple/Combined graph definitions	-	-	[47], [49], [53], [54], [57]–[59], [61]–[64], [67], [69], [72], [79], [79], [81], [82], [87], [92], [102]

TABLE II  
OVERVIEW OF METHODS IN DEFINING THE INPUT NODE FEATURES

Method	Time domain	Frequency domain	Graph Domain	Papers
Differential entropy	✓	✗	✗	[46], [51], [53], [55], [57], [71], [72], [75], [78], [81], [82], [87], [89], [90], [92], [93], [93], [95]–[99], [101], [102]
Raw signal	✓	✗	✗	[48], [52], [56], [60], [62]–[64], [66], [67], [69], [72], [76], [77], [79], [80], [83], [84], [86], [91], [94], [100], [104]–[106]
Fourier Transform	✗	✓	✗	[50], [88], [107]
PSD/Band Power	✗	✓	✗	[7], [49], [55], [57]–[59], [61], [78], [85], [90]
Graph theory metrics	✗	✗	✓	[65], [73]
Descriptive statistics	✓	✗	✗	[47], [58], [61]

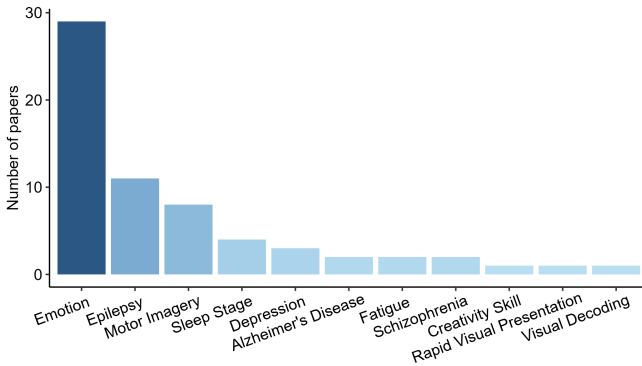


Fig. 3. Classification tasks presented in the current EEG-GNN literature.

surface. Instead, we use the term to describe methods that construct brain graphs based on the physical distance between EEG electrodes. In contrast, FC refers to pairwise statistical relationships between EEG signals.

SC graph is pre-defined such that electrodes are connected by an edge in the following way:

$$e_{ij} = \begin{cases} 1 \text{ or } 1/d_{ij}, & \text{if } d_{ij} \leq t \\ 0, & \text{otherwise} \end{cases}, \quad (4)$$

where  $e_{ij}$  is the edge weight connecting nodes  $i$  and  $j$ ,  $d_{ij}$  is a measure of distance between EEG electrodes, and  $t$  is a manually defined threshold controlling the graph sparsity.

Such an approach offers several advantages. First, the SC graph is insensitive to any noise effects of EEG recording since it is independent of the actual signals. Second, all data samples share an identical graph structure, provided the same EEG montage was utilised during the recording. This

offers explainability advantages when combined with spectral GNN since the graph frequency components defined by the eigenvectors of graph Laplacian are fixed. On the other hand, the SC graph is limited to short-range relationships. Thus, it might not accurately represent the underlying brain network. Some papers propose to overcome this limitation by manually inserting global [53], [56], [57], [58], [62] or inter-hemispheric edges [46], [54], [87].

In contrast, an FC graph can be obtained from either classical FC measures (FC measure in Table I or learnable methods (e.g. feature concatenation/distance and attention methods in Table I)). We refer to all of these methods as FC because they all measure the degree of interaction between two nodes, thus falling within the traditional definition of FC. Unlike SC, the FC graph is unique for each data sample and can contain both short- and long-range edges. On the other hand, since it is derived directly from EEG signals, it might be sensitive to noise.

Learnable FC based on node feature distance or feature concatenation are generally computed as:

$$e_{ij} = \theta_1(|h_i - h_j|) \text{ and} \quad (5)$$

$$e_{ij} = \theta_2(h_i \parallel h_j), \quad (6)$$

respectively, where  $\theta_1(\cdot)$  and  $\theta_2(\cdot)$  are neural networks with input-output dimensions of  $\mathbb{R} : d \rightarrow 1$  and  $\mathbb{R} : 2 \times d \rightarrow 1$ , respectively;  $|\cdot|$  denotes absolute value;  $\parallel$  denotes concatenation and  $h_i$  is the node feature/embedding of node  $i$ . We discuss the attention-based graphs together with the types of graph convolutional layers in Section VI and thus skip these methods in this section.

Special cases of brain graph definition are the shared-mask methods. These methods defined a matrix of learnable parameters with the same shape as the adjacency matrix of

TABLE III  
OVERVIEW OF NODE FEATURE PRE-PROCESSING  
BEFORE GNN LAYERS

Method	Trained separately	Papers
1D CNN		[48], [53], [60], [93], [94], [100], [104], [106]
Feature-wise attention weighting	×	[73]
bidirectional LSTM	✓	[76]
Temporal CNN	×	[56], [91], [94]
WaveletCNN	×	[91]
SincCNN	×	[60]
MLP	×	[85]
CNN Feature Extractor	✓	[64]

the input graphs that acts as a mask/filter by multiplying it with the adjacency matrix. This learnable matrix is a part of the model. Thus, the same mask is applied to all input graphs. However, a shared mask limits the size of the input graphs, i.e. the number of nodes must remain fixed so that the adjacency matrix can be multiplied with the shared mask.

In the current stage, which method should be preferred for brain graph classification tasks is unclear. Some authors attempt to avoid this issue by combining multiple methods. However, we instead suggest that the researchers carefully consider each of the presented methods in the context of the given classification task, as each method poses its unique set of strengths and weaknesses.

## V. NODE FEATURE DEFINITIONS

The second part of the input to a GNN model is the node feature matrix (Fig. 1A). We summarise the various definitions of node features in Table II. We categorise these definitions based on which domain they are computed, i.e. time, frequency and graph domains.

The time-domain methods are the most commonly used in the current literature. In particular, these are the differential entropy (DE) and raw signal methods. The popularity of DE is given by the fact that many of the open EEG datasets include this feature, such as the SEED [108] emotion recognition dataset. DE describes the complexity of a continuous variable and is defined as:

$$DE(X) = - \int_X f(x) \log(f(x)) dx \quad (7)$$

where  $X$  is a random continuous variable and  $f(x)$  is the probability density function.

Many papers define the node feature as the raw EEG signal. However, the raw signal can be too long for a GNN to process effectively. Thus, it is often coupled with node feature pre-processing module and spatio-temporal GNNs (See V-A and VI, respectively) to either reduce the dimensionality or to extract the temporal patterns contained within the signal effectively. An alternative to the raw signal node feature is descriptive statistics, such as mean, median or standard deviation.

Frequency-domain node features are usually defined as the Fourier frequency components obtained by the Fourier transform or PSD. Both of these methods aim to quantify

the strength of various frequency components within the EEG signal. An advantage of these representations is their relatively low dimensionality compared to the raw signal described previously.

Finally, graph-theoretical features can be utilised to describe the nodes, e.g. mean node weight [65] and betweenness centrality [65], [73]. A severe limitation of this method is that the graph structure needs to be defined prior to node feature extraction. Thus, this node feature type is incompatible with learnable brain graph methods.

## A. Node Feature Preprocessing

An optional next step after node features construction is some kind of node feature pre-processing module (NFP) (Fig. 1B). We summarise the types of NFPs in Table III.

Most of the NFPs are integrated within the GNN architecture, thus allowing the model to be trained in an end-to-end manner. The exceptions are methods that utilise a pre-trained feature extraction neural network implemented as a bidirectional LSTM [76] or a CNN [64].

The surveyed NFPs are all based on a neural network. In most cases, these are variants of a CNN and a multilayer perceptron (MLP). These modules aim to (1) reduce the dimensionality of the node features and (2) enhance the node features, including potentially suppressing noise or redundant information.

## VI. TYPE OF GRAPH CONVOLUTIONAL LAYER

A core part of a GNN model are the graph convolutional layers (GCN) (Fig. 1C). We summarise the utilised types of GCNs in Table IV. We further categorise them based on the type of GNN as introduced in Section II, i.e. spatial, spectral. Additionally, we add the temporal category, which is not a type of standalone GCN layer but must be combined with spatial or spectral GCN.

Interestingly, ChebConv is used in the majority of the surveyed papers (counting both ChebConv and spectral spatio-temporal GNN in Table IV). Since EEG typically uses 128 electrodes in high-density montages, the size of the brain graphs is relatively small. In such cases, even a full spectral GNN would not be too computationally expensive for EEG classification. Therefore, it remains unclear why many authors opt for the ChebConv approximation of spectral GNN. We speculate that the influence of classical signal processing tools in EEG analysis might also serve as a sufficient argument for using spectral GNNs for EEG classification.

On the other hand, the other half of the surveyed papers experiment with a wide range of spatial GNNs. The (simplified) GCN is a popular method amongst these, which is equivalent to a 1st-order ChebConv ( $K = 1$ ). A special case of spatial GNN is the graph attention network (GAT). GAT allows for adjusting the graph by re-weighting the edges using an attention mechanism. Generally, the attention mechanism for computing the new softmax-normalised edge weight  $e_{ij}$  is defined as follows:

$$e_{i,j} = \frac{\exp(\sigma(\mathbf{w}^\top [\mathbf{W}h_i \parallel \mathbf{W}h_j]))}{\sum_{k \in \mathcal{N}(i)} \exp(\sigma(\mathbf{w}^\top [\mathbf{W}h_i \parallel \mathbf{W}h_k]))}, \quad (8)$$

TABLE IV  
OVERVIEW OF GRAPH CONVOLUTIONAL LAYERS

Method	Spatial	Spectral	Temporal	Papers
Graph Isomorphism Network	✓	✗	✗	[48], [65], [79]
(Simplified) Graph Convolution Network	✓	✗	✗	[7], [46], [53], [54], [56], [58], [61], [70], [72], [75], [83], [89], [106]
Chebyshev Graph Convolution	✗	✓	✗	[49], [51], [55], [57], [59], [66], [67], [69], [71], [74], [76]–[78], [80], [82], [85], [90], [97], [99], [104]
Graph Attention Network	✓	✗	✗	[60], [62], [73], [84], [88], [94], [98], [100]
Diffusion recurrent gated	✗	✓	✗	[50]
Spatio-temporal GNN (Spectral)	✗	✓	✓	[52], [81], [86], [95], [96], [107]
Spatio-temporal GNN (Spatial)	✓	✗	✓	[63], [105]
Powers of Adjacency Matrix GNN	✓	✗	✗	[101], [102]
GraphSAGE	✓	✗	✗	[48], [107]
Spectral GNN	✗	✓	✗	[87], [93]
B-Spline Kernel GCN	✓	✗	✗	[47]
Residual GCN	✓	✗	✗	[91]
Multibranch architectures	-	-	-	[58], [63], [80], [92], [97], [100]

TABLE V  
OVERVIEW OF NODE POOLING MECHANISMS

Method	Learnable	Papers
TopK	✓	[62], [67]
Hierarchical tree pooling	✓	[65]
SortPool	✓	[48]
EdgePool	✓	[48]
SAGPool	✓	[48], [54]
Set2Set	✓	[48]
Manual Clustering	✗	[101], [102]
Graclus Clustering	✗	[77], [80]

where  $w$  and  $W$  are the learnable parameters of the model,  $\sigma$  is an activation function,  $h$  is the node feature vector/embedding, and  $N(i)$  is the set of nodes connected to node  $i$ . The resulting edge weights can then be passed to Equation 1.

Next, the spatio-temporal GNNs were tested for EEG classification in several instances. A spatio-temporal block consists of one GCN layer and one 1D-CNN applied temporally. This structure allows the model to extract both spatial (i.e. graph) and temporal patterns. There are both spatial and spectral variants of spatio-temporal GNN, and there is no indication as to which one should be preferred as no comparative study exists to date.

Finally, several papers adopt multi-branch architectures. These methods utilise multiple GCN layers applied in parallel to allow the model to focus on various aspects (also views) of the input graph. An example of such a model utilises two-branch GNN to learn from both FC- and SC-based brain graph structure [63]. Alternatively, the individual frequency bands of EEG signals can be used to construct various graph views [85].

## VII. NODE POOLING MECHANISMS

In some instances, reducing the number of nodes in the graph might be desirable. This can be achieved with a node pooling module (Fig. 1D). We summarise the node pooling modules utilised in the surveyed papers in Table V.

There are both learnable and non-learnable node pooling modules in the literature. Please see the corresponding papers for a detailed description of these methods (Table V). Node pooling modules remain a relatively unexplored topic in the EEG-GNN classification models. Node pooling can (1) remove

redundant nodes, (2) reduce the size of the graph embedding in a setting where the concatenation of node embeddings forms it, and (3) aid in the explainability of the model by identifying node importance with respect to the classification task.

## VIII. FROM NODE EMBEDDINGS TO GRAPH EMBEDDING

The output of the graph convolutions is a set of learned node embeddings. Node embeddings in this form are suitable for tasks such as node classification and link prediction. However, for graph classification, the set of node features needs to be transformed into a unified graph representation (Fig. 1E). We summarise the methods for this transformation in Table VI.

The most straightforward method to form a graph embedding is to simply concatenate the node features. This approach poses a few limitations. First, the resulting graph embedding grows with the number of nodes. Thus, the classification layer requires a large number of parameters. Second, all input graphs need to have the same number of nodes, limiting the model’s generalisation to other datasets. Finally, such an approach is likely to include redundant or duplicated information in the graph embedding since GNN produces node embeddings by aggregating information from neighbouring nodes.

A readout function is one of the methods to form a graph embedding that addresses these issues. A readout forms the embedding by passing the node features through a permutation-invariant function. A general definition of a readout to obtain graph embedding of a graph  $G_i$  from a set of  $V$  node embeddings  $H = [h_1, \dots, h_V]$  is given by:

$$G_i = \sum_{k=1}^V h_k, \quad (9)$$

where  $\sum$  can be any permutation-invariant function. In the surveyed papers, these functions were sum, average and maximum. A few papers also experiment with attention-weighted sum to attenuate the role of unimportant nodes within the graph embedding [88]. An interesting alternative is to apply CNN-style average or maximum pooling node-wise [105].

TABLE VI  
OVERVIEW OF METHODS FOR THE FORMATION OF GRAPH EMBEDDING FROM A SET OF NODE EMBEDDINGS

Method	Learnable	Papers
Sum readout	✗	[46], [65], [82]
Average readout	✗	[49], [54], [61], [62], [72], [85], [107]
Maximum readout	✗	[7], [54], [62], [76], [106]
Concatenate node embeddings	✗	[47], [55]–[60], [64], [66], [67], [70], [74], [77], [80], [81], [86], [87], [90]–[93], [98], [100]–[102], [104]
CNN-like Average/Maximum Pooling	✗	[83], [105]
SortPool	✓	[68]
Attention weighted	✓	[63], [88], [97]
CNN	✓	[52], [69], [78]
LSTM	✓	[51], [99]
Capsule Network	✓	[73]
Transformer	✓	[89]
Bidirectional LSTM	✓	[83], [84], [100]

Alternatively, researchers explored various neural network models to obtain graph embeddings, such as CNN [52], [69], [78], (bi-)LSTM [51], [83], [84], [99], [100], transformer [89] and capsule networks [73]. Additionally, graph pooling methods, such as DiffPool [109], SAGPool [110], iPool [111], TAP [112] and HierCorrPool [113] can be used for this purpose.

## IX. DISCUSSION

Despite most of the surveyed papers being relatively recent, a wide range of GNN-based methods has already been proposed to classify EEG signals in a diverse set of tasks, such as emotion recognition, brain-computer interfaces, and psychological and neurodegenerative disorders and diseases (Fig 3). This recent rise in the popularity of GNN models for EEG might be attributed to (1) the development of new GNN methods and (2) advances in network neuroscience inspired an extension of this framework to deep learning. GNNs offer unique advantages over other deep learning methods. This is mainly the possibility of modelling multivariate time series and interactions among them with a single GNN model, which is not possible with CNN or recurrent networks. Additionally, patterns learned by GNNs can readily be interpreted in the context of network neuroscience, thus enabling a wide range of avenues for model explainability.

This survey categorises the proposed GNN models in terms of their inputs and modules. Specifically, these are brain graph structure, node features and their preprocessing, GCN layers, node pooling mechanisms, and formation of graph embeddings. This categorisation allows us to provide a quick and simple overview of the different methods presented in the EEG-GNN literature, appreciate the current state of the art in this field and identify promising future directions.

### A. Limitations of Surveyed Papers

Surprisingly, we have identified the least variety and innovation in the category of GCN layers (Table IV). A significant proportion of the surveyed papers utilise either ChebConv or “vanilla” spatial GCN. This might be due to the relative novelty of the EEG-GNN field, and thus, many papers explore other areas of model design, such as node features and brain graph definitions. A few papers seem to successfully

experiment with more complex types of GCN layers [47], [50], [91] and multi-branch architectures [58], [63], [80], [92], [97], [100].

A major limitation of most surveyed papers is the lack of generalisability to external datasets that might use a different number of EEG signals. This is caused by (1) the use of ChebConv and (2) forming graph embedding by node feature concatenation [47], [55], [56], [57], [58], [59], [60], [64], [66], [67], [70], [74], [77], [80], [81], [86], [87], [90], [91], [92], [93], [98], [100], [101], [102], [104]. (1) can be addressed by utilising spatial GCN layers as suggested above, and (2) can be solved by using a readout function or a suitable node pooling mechanism, which coarsens the graph to a fixed number of nodes. Additionally, there is a general lack of transfer learning experiments for EEG-GNN models, which might be a promising direction for future research.

Finally, we have identified an interesting gap in EEG-GNN research: the lack of utilising frequency band information in a more complex way. A few papers train separate models for each frequency band in isolation [46], [47], [65]. Alternatively, they propose concatenating the graph embeddings generated from the frequency-band-GNN branches [52], [87], [101].

### B. Future Directions

Several promising directions can be identified in the rapidly evolving landscape of EEG-GNN research. First, a comprehensive comparison of the various GCN layers (e.g. spatial GNN, ChebConv, GAT and graph transformer) with respect to their influence on classification performance should be carried out to address this crucial design question in a systematic manner.

Second, enhancing the generalisability of models by addressing issues related to the varying number of EEG signals/electrodes and exploring transfer learning approaches can open new avenues for research. For instance, pre-trained GNN models on cheap-to-obtain large datasets, such as open databases for emotion recognition or BCI applications, would allow the application of complex GNN architectures to problems with limited data availability due to the high costs or small populations (e.g. clinical data, rare diseases and disorders). Focusing on these issues would likely improve the



generalisability of the models when evaluated on a diverse set of EEG datasets and different classification tasks.

Lastly, the rich frequency information of EEG signals should be explored more. For instance, we suggest a plausible utility of integrating cross-frequency coupling (CFC) approaches into EEG-GNN models. There is growing evidence in the literature concerning the advanced brain functions (e.g. learning, memory) enabled by CFC [114]. Thus, integrating findings from neuroscience research into the EEG-GNN design promises both performance and explainability gains.

### C. Limitations of Our Survey

It is worth noting that this paper does not follow a systematic review methodology; therefore, we do not assert that our findings are exhaustive. Instead, our objective is to offer a succinct and cohesive overview of the current research on EEG-GNN models to facilitate the development of innovative approaches and assist researchers new to this field.

One of the major parts of EEG-GNN models we omit in this survey is the model explainability. We suggest that a survey paper is not well suited for comprehensively covering this aspect of research. Instead, we suggest a comparative experimental study to be better suited to explore the various explainability options of GNN explainability. However, to maintain the comprehensiveness of this survey, we list the papers that report the use of certain methods of model explainability: [24], [50], [55], [89], [105], [106].

## X. CONCLUSION

In conclusion, this survey examined the current research on EEG-GNN models for classifying EEG signals. Various GNN-based methods have been proposed for tasks such as emotion recognition, brain-computer interfaces, and psychological and neurodegenerative disorders. The surveyed papers were categorised based on inputs and modules, including brain graph structure, node features, GCN layers, node pooling mechanisms, and graph embeddings.

GNNs offer a unique method for analysing and classifying EEG in the graph domain, thus allowing the exploitation of complex spatial information in brain networks that other neural networks do not. Additionally, GNNs can be easily extended with CNN and recurrent network-based modules at various stages of the GNN architecture, such as for node feature pre-processing, node embedding post-processing and graph embedding formation.

However, limitations and areas for improvement were identified. There is a lack of variety and innovation in GCN layers, with many papers utilising ChebConv or “simple” spatial GCN without clear justification. Generalisability to external datasets with varying numbers of EEG electrodes is limited. Transfer learning experiments and integration of cross-frequency coupling approaches are potential future research to enhance the performance and explainability of GNN.

## REFERENCES

- [1] S. Morales and M. E. Bowers, “Time-frequency analysis methods and their application in developmental EEG data,” *Develop. Cognit. Neurosci.*, vol. 54, Apr. 2022, Art. no. 101067.
- [2] S. J. M. Smith, “EEG in the diagnosis, classification, and management of patients with epilepsy,” *J. Neurol., Neurosurgery Psychiatry*, vol. 76, no. suppl\_2, pp. ii2–ii7, Jun. 2005.
- [3] S. K. Loo and R. A. Barkley, “Clinical utility of EEG in attention deficit hyperactivity disorder,” *Appl. Neuropsychology*, vol. 12, no. 2, pp. 64–76, Jun. 2005, doi: [10.1207/s15324826an1202\\_2](https://doi.org/10.1207/s15324826an1202_2).
- [4] D. McFarland and J. Wolpaw, “EEG-based brain–computer interfaces,” *Current Opinion Biomed. Eng.*, vol. 4, pp. 194–200, Jul. 2017.
- [5] F. Lotte et al., “A review of classification algorithms for EEG-based brain–computer interfaces: A 10 year update,” *J. Neural Eng.*, vol. 15, no. 3, Apr. 2018, Art. no. 031005.
- [6] W. Samek, C. Vidaurre, K.-R. Müller, and M. Kawanabe, “Stationary common spatial patterns for brain–computer interfacing,” *J. Neural Eng.*, vol. 9, no. 2, Apr. 2012, Art. no. 026013.
- [7] D. Klepl, F. He, M. Wu, D. J. Blackburn, and P. Sarrigiannis, “EEG-based graph neural network classification of Alzheimer’s disease: An empirical evaluation of functional connectivity methods,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 2651–2660, 2022.
- [8] M. Murugappan, N. Ramachandran, and Y. Sazali, “Classification of human emotion from EEG using discrete wavelet transform,” *J. Biomed. Sci. Eng.*, vol. 3, no. 4, pp. 390–396, 2010.
- [9] R. J. Oweis and E. W. Abdulhay, “Seizure classification in EEG signals utilizing Hilbert–Huang transform,” *Biomed. Eng. Online*, vol. 10, no. 1, pp. 1–15, 2011.
- [10] M.-P. Hosseini, A. Hosseini, and K. Ahi, “A review on machine learning for EEG signal processing in bioengineering,” *IEEE Rev. Biomed. Eng.*, vol. 14, pp. 204–218, 2021.
- [11] K. Rasheed et al., “Machine learning for predicting epileptic seizures using EEG signals: A review,” *IEEE Rev. Biomed. Eng.*, vol. 14, pp. 139–155, 2021.
- [12] D. Klepl, F. He, M. Wu, D. J. Blackburn, and P. G. Sarrigiannis, “Cross-frequency multilayer network analysis with bispectrum-based functional connectivity: A study of Alzheimer’s disease,” *Neuroscience*, vol. 521, pp. 77–88, Jun. 2023.
- [13] 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,” *J. Neural Eng.*, vol. 15, no. 5, Oct. 2018, Art. no. 056013.
- [14] R. G. Hefron, B. J. Borghetti, J. C. Christensen, and C. M. S. Kabbani, “Deep long short-term memory structures model temporal dependencies improving cognitive workload estimation,” *Pattern Recognit. Lett.*, vol. 94, pp. 96–104, Jul. 2017.
- [15] D. S. Bassett and O. Sporns, “Network neuroscience,” *Nature Neurosci.*, vol. 20, no. 3, pp. 353–364, Feb. 2017.
- [16] R. Cao et al., “EEG functional connectivity underlying emotional valance and arousal using minimum spanning trees,” *Frontiers Neurosci.*, vol. 14, p. 355, May 2020.
- [17] A. T. Adebisi and K. C. Veluvolu, “Brain network analysis for the discrimination of dementia disorders using electrophysiology signals: A systematic review,” *Frontiers Aging Neurosci.*, vol. 15, Mar. 2023, Art. no. 1039496.
- [18] B. Kılıç and S. Aydın, “Classification of contrasting discrete emotional states indicated by EEG based graph theoretical network measures,” *Neuroinformatics*, vol. 20, no. 4, pp. 863–877, Oct. 2022.
- [19] M. Jalili, “Graph theoretical analysis of Alzheimer’s disease: Discrimination of AD patients from healthy subjects,” *Inf. Sci.*, vol. 384, pp. 145–156, Apr. 2017.
- [20] F. Hasanzadeh, M. Mohebbi, and R. Rostami, “Graph theory analysis of directed functional brain networks in major depressive disorder based on EEG signal,” *J. Neural Eng.*, vol. 17, no. 2, Mar. 2020, Art. no. 026010.
- [21] S. Nobukawa, T. Yamanishi, S. Kasakawa, H. Nishimura, M. Kikuchi, and T. Takahashi, “Classification methods based on complexity and synchronization of electroencephalography signals in Alzheimer’s disease,” *Frontiers Psychiatry*, vol. 11, p. 255, Apr. 2020.
- [22] S. Supriya, S. Siuly, H. Wang, and Y. Zhang, “Epilepsy detection from EEG using complex network techniques: A review,” *IEEE Rev. Biomed. Eng.*, vol. 16, pp. 292–306, 2023.
- [23] R. Li et al., “Graph signal processing, graph neural network and graph learning on biological data: A systematic review,” *IEEE Rev. Biomed. Eng.*, vol. 16, pp. 109–135, 2023.
- [24] D. Klepl, F. He, M. Wu, D. J. Blackburn, and P. G. Sarrigiannis, “Adaptive gated graph convolutional network for explainable diagnosis of Alzheimer’s disease using EEG data,” 2023, *arXiv:2304.05874*.

- [25] X.-M. Zhang, L. Liang, L. Liu, and M.-J. Tang, "Graph neural networks and their current applications in bioinformatics," *Frontiers Genet.*, vol. 12, pp. 1–22, Jul. 2021.
- [26] A. Bessadok, M. A. Mahjoub, and I. Rekik, "Graph neural networks in network neuroscience," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 5, pp. 5833–5848, May 2023.
- [27] O. Wieder et al., "A compact review of molecular property prediction with graph neural networks," *Drug Discovery Today, Technol.*, vol. 37, pp. 1–12, Dec. 2020.
- [28] P. Reiser et al., "Graph neural networks for materials science and chemistry," *Commun. Mater.*, vol. 3, no. 1, pp. 1–18, Nov. 2022.
- [29] J. Xiong, Z. Xiong, K. Chen, H. Jiang, and M. Zheng, "Graph neural networks for automated *de novo* drug design," *Drug Discovery Today*, vol. 26, no. 6, pp. 1382–1393, Jun. 2021.
- [30] M. Sun, S. Zhao, C. Gilvary, O. Elemento, J. Zhou, and F. Wang, "Graph convolutional networks for computational drug development and discovery," *Briefings Bioinf.*, vol. 21, no. 3, pp. 919–935, May 2020.
- [31] M. Malekzadeh, P. Hajibabaei, M. Heidari, S. Zad, O. Uzuner, and J. H. Jones, "Review of graph neural network in text classification," in *Proc. IEEE 12th Annu. Ubiquitous Comput., Electron. Mobile Commun. Conf. (UEMCON)*, Dec. 2021, pp. 0084–0091.
- [32] L. Wu et al., "Graph neural networks for natural language processing: A survey," *Found. Trends Mach. Learn.*, vol. 16, no. 2, pp. 119–328, 2023.
- [33] C. Gao et al., "A survey of graph neural networks for recommender systems: Challenges, methods, and directions," *ACM Trans. Recommender Syst.*, vol. 1, no. 1, pp. 3:1–3:51, Mar. 2023.
- [34] S. Wu, F. Sun, W. Zhang, X. Xie, and B. Cui, "Graph neural networks in recommender systems: A survey," *ACM Comput. Surveys*, vol. 55, no. 5, pp. 97:1–97:37, Dec. 2022.
- [35] W. Jiang and J. Luo, "Graph neural network for traffic forecasting: A survey," *Exp. Syst. Appl.*, vol. 207, Nov. 2022, Art. no. 117921.
- [36] M. Lv, Z. Hong, L. Chen, T. Chen, T. Zhu, and S. Ji, "Temporal multi-graph convolutional network for traffic flow prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 6, pp. 3337–3348, Jun. 2021.
- [37] J. Wang, S. Zhang, Y. Xiao, and R. Song, "A review on graph neural network methods in financial applications," 2021, *arXiv:2111.15367*.
- [38] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 1, pp. 4–24, Jan. 2021.
- [39] J. Zhou et al., "Graph neural networks: A review of methods and applications," *AI Open*, vol. 1, pp. 57–81, Jan. 2020.
- [40] P. Veličković et al., "Graph attention networks," 2017, *arXiv:1710.10903*.
- [41] Y. Seo, M. Defferrard, P. Vandergheynst, and X. Bresson, "Structured sequence modeling with graph convolutional recurrent networks," in *Neural Information Processing (Lecture Notes in Computer Science)*, L. Cheng, A. C. S. Leung, and S. Ozawa, Eds. Cham, Switzerland: Springer, 2018, pp. 362–373.
- [42] Y. Shi, Z. Huang, S. Feng, H. Zhong, W. Wang, and Y. Sun, "Masked label prediction: Unified message passing model for semi-supervised classification," 2020, *arXiv:2009.03509*.
- [43] Z. Chen et al., "Bridging the gap between spatial and spectral domains: A survey on graph neural networks," 2020, *arXiv:2002.11867*.
- [44] D. Bo, X. Wang, Y. Liu, Y. Fang, Y. Li, and C. Shi, "A survey on spectral graph neural networks," 2023, *arXiv:2302.05631*.
- [45] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," in *Advances in Neural Information Processing Systems*, vol. 29. Red Hook, NY, USA: Curran Associates, 2016.
- [46] 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.
- [47] Z. Lin, T. Qiu, P. Liu, L. Zhang, S. Zhang, and Z. Mu, "Fatigue driving recognition based on deep learning and graph neural network," *Biomed. Signal Process. Control*, vol. 68, Jul. 2021, Art. no. 102598.
- [48] A. Demir, T. Koike-Akino, Y. Wang, M. Haruna, and D. Erdogmus, "EEG-GNN: Graph neural networks for classification of electroencephalogram (EEG) signals," in *Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Nov. 2021, pp. 1061–1067.
- [49] K. Raiesi, M. Khazaei, P. Croce, G. Tamburro, S. Comani, and F. Zappasodi, "A graph convolutional neural network for the automated detection of seizures in the neonatal EEG," *Comput. Methods Programs Biomed.*, vol. 222, Jul. 2022, Art. no. 106950.
- [50] S. Tang et al., "Self-supervised graph neural networks for improved electroencephalographic seizure analysis," in *Proc. Int. Conf. Learn. Represent.*, 2021, pp. 1–23.
- [51] 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.
- [52] B. Sun, H. Zhang, Z. Wu, Y. Zhang, and T. Li, "Adaptive spatiotemporal graph convolutional networks for motor imagery classification," *IEEE Signal Process. Lett.*, vol. 28, pp. 219–223, 2021.
- [53] G. Du et al., "A multi-dimensional graph convolution network for EEG emotion recognition," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–11, 2022.
- [54] T. Xu, W. Dang, J. Wang, and Y. Zhou, "DAGAM: A domain adversarial graph attention model for subject independent EEG-based emotion recognition," 2022, *arXiv:2202.12948*.
- [55] S. Jang, S.-E. Moon, and J.-S. Lee, "EEG-based video identification using graph signal modeling and graph convolutional neural network," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2018, pp. 3066–3070.
- [56] Y. Ding, N. Robinson, C. Tong, Q. Zeng, and C. Guan, "LGGNet: Learning from local-global-graph representations for brain-computer interface," 2021, *arXiv:2105.02786*.
- [57] 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.
- [58] D. Zeng, K. Huang, C. Xu, H. Shen, and Z. Chen, "Hierarchy graph convolution network and tree classification for epileptic detection on electroencephalography signals," *IEEE Trans. Cognit. Develop. Syst.*, vol. 13, no. 4, pp. 955–968, Dec. 2021.
- [59] J. Jia, B. Zhang, H. Lv, Z. Xu, S. Hu, and H. Li, "CR-GCN: Channel-relationships-based graph convolutional network for EEG emotion recognition," *Brain Sci.*, vol. 12, no. 8, p. 987, Jul. 2022.
- [60] D. Priyasad, T. Fernando, S. Denman, S. Sridharan, and C. Fookes, "Affect recognition from scalp-EEG using channel-wise encoder networks coupled with geometric deep learning and multi-channel feature fusion," *Knowl.-Based Syst.*, vol. 250, Aug. 2022, Art. no. 109038.
- [61] M. Jia et al., "Efficient graph convolutional networks for seizure prediction using scalp EEG," *Frontiers Neurosci.*, vol. 16, Aug. 2022, Art. no. 967116.
- [62] T. Chen, Y. Guo, S. Hao, and R. Hong, "Exploring self-attention graph pooling with EEG-based topological structure and soft label for depression detection," *IEEE Trans. Affect. Comput.*, vol. 13, no. 4, pp. 2106–2118, Oct. 2022.
- [63] M. Li, H. Chen, and Z. Cheng, "An attention-guided spatiotemporal graph convolutional network for sleep stage classification," *Life*, vol. 12, no. 5, p. 622, Apr. 2022.
- [64] Z. Jia et al., "Multi-view spatial-temporal graph convolutional networks with domain generalization for sleep stage classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1977–1986, 2021.
- [65] H. Liu, J. Zhang, Q. Liu, and J. Cao, "Minimum spanning tree based graph neural network for emotion classification using EEG," *Neural Netw.*, vol. 145, pp. 308–318, Jan. 2022.
- [66] Y. Kwak, W.-J. Song, and S.-E. Kim, "Graph neural network with multilevel feature fusion for EEG based brain-computer interface," in *Proc. IEEE Int. Conf. Consum. Electron. Asia (ICCE-Asia)*, Nov. 2020, pp. 1–3.
- [67] Q. Chang et al., "Classification of first-episode schizophrenia, chronic schizophrenia and healthy control based on brain network of mismatch negativity by graph neural network," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1784–1794, 2021.
- [68] J. Bi, F. Wang, X. Yan, J. Ping, and Y. Wen, "Multi-domain fusion deep graph convolution neural network for EEG emotion recognition," *Neural Comput. Appl.*, vol. 34, no. 24, pp. 22241–22255, Dec. 2022.
- [69] S. Asadzadeh, T. Y. Rezaii, S. Beheshti, and S. Meshgini, "Accurate emotion recognition using Bayesian model based EEG sources as dynamic graph convolutional neural network nodes," *Sci. Rep.*, vol. 12, no. 1, pp. 1–14, Jun. 2022.
- [70] Y. Zhao et al., "EEG-based seizure detection using linear graph convolution network with focal loss," *Comput. Methods Programs Biomed.*, vol. 208, Sep. 2021, Art. no. 106277.

- [71] J. Chen, H. Qian, and X. Gong, "Bayesian graph neural networks for EEG-based emotion recognition," in *Clinical Image-Based Procedures, Distributed and Collaborative Learning, Artificial Intelligence for Combating COVID-19 and Secure and Privacy-Preserving Machine Learning*. Cham, Switzerland: Springer, 2021, pp. 24–33.
- [72] R. Li, Y. Wang, and B.-L. Lu, "A multi-domain adaptive graph convolutional network for EEG-based emotion recognition," in *Proc. 29th ACM Int. Conf. Multimedia*, Oct. 2021, pp. 5565–5573.
- [73] S. Ghosh, A. Konar, and A. K. Nagar, "Decoding subjective creativity skill from visuo-spatial reasoning ability using capsule graph neural network," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2021, pp. 1–8.
- [74] N. Khaleghi, T. Y. Rezaii, S. Beheshti, and S. Meshgini, "Developing an efficient functional connectivity-based geometric deep network for automatic EEG-based visual decoding," *Biomed. Signal Process. Control*, vol. 80, Feb. 2023, Art. no. 104221.
- [75] D. Wang et al., "Identification of depression with a semi-supervised GCN based on EEG data," in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, Dec. 2021, pp. 2338–2345.
- [76] Y. Hou et al., "Deep feature mining via attention-based BiLSTM-GCN for human motor imagery recognition," 2020, *arXiv:2005.00777*.
- [77] Y. Hou et al., "GCNs-net: A graph convolutional neural network approach for decoding time-resolved EEG motor imagery signals," *IEEE Trans. Neural Netw. Learn. Syst.*, pp. 1–12, 2022.
- [78] W. Kong, M. Qiu, M. Li, X. Jin, and L. Zhu, "Causal graph convolutional neural network for emotion recognition," *IEEE Trans. Cognit. Develop. Syst.*, vol. 15, pp. 1686–1693, 2022.
- [79] T. Tao, L. Guo, Q. He, H. Zhang, and L. Xu, "Seizure detection by brain-connectivity analysis using dynamic graph isomorphism network," in *Proc. 44th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2022, pp. 2302–2305.
- [80] B. Wang et al., "Graph learning with co-teaching for EEG-based motor imagery recognition," *IEEE Trans. Cognit. Develop. Syst.*, vol. 15, pp. 1722–1731, 2022.
- [81] Y. Li, N. Zhong, D. Taniar, and H. Zhang, "MutualGraphNet: A novel model for motor imagery classification," 2021, *arXiv:2109.04361*.
- [82] G. Bao et al., "Linking multi-layer dynamical GCN with style-based recalibration CNN for EEG-based emotion recognition," *Frontiers Neuroinformatics*, vol. 16, pp. 1–12, Feb. 2022.
- [83] L. Feng, C. Cheng, M. Zhao, H. Deng, and Y. Zhang, "EEG-based emotion recognition using spatial-temporal graph convolutional LSTM with attention mechanism," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 11, pp. 5406–5417, Nov. 2022.
- [84] J. He, J. Cui, G. Zhang, M. Xue, D. Chu, and Y. Zhao, "Spatial-temporal seizure detection with graph attention network and bi-directional LSTM architecture," *Biomed. Signal Process. Control*, vol. 78, Oct. 2022, Art. no. 103908.
- [85] X. Sun et al., "A novel complex network-based graph convolutional network in major depressive disorder detection," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–8, 2022.
- [86] X. Shan, J. Cao, S. Huo, L. Chen, P. G. Sarrigiannis, and Y. Zhao, "Spatial-temporal graph convolutional network for Alzheimer classification based on brain functional connectivity imaging of electroencephalogram," *Human Brain Mapping*, vol. 43, no. 17, pp. 5194–5209, Dec. 2022.
- [87] H. Li, X. Zhang, and Y. Xia, "EEG emotion recognition based on dynamically organized graph neural network," in *Proc. Int. Conf. Multimedia Model*. Cham, Switzerland: Springer, 2022, pp. 344–355.
- [88] C. Li et al., "EEG emotion recognition based on self-attention dynamic graph neural networks," in *Proc. 44th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2022, pp. 292–296.
- [89] M. Sun, W. Cui, S. Yu, H. Han, B. Hu, and Y. Li, "A dual-branch dynamic graph convolution based adaptive TransFormer feature fusion network for EEG emotion recognition," *IEEE Trans. Affect. Comput.*, vol. 13, no. 4, pp. 2218–2228, Oct. 2022.
- [90] T. Zhang, X. Wang, X. Xu, and C. L. P. Chen, "GCB-net: Graph convolutional broad network and its application in emotion recognition," *IEEE Trans. Affect. Comput.*, vol. 13, no. 1, pp. 379–388, Jan. 2022.
- [91] Y. Li, Y. Liu, Y.-Z. Guo, X.-F. Liao, B. Hu, and T. Yu, "Spatio-temporal-spectral hierarchical graph convolutional network with semisupervised active learning for patient-specific seizure prediction," *IEEE Trans. Cybern.*, vol. 52, no. 11, pp. 12189–12204, Nov. 2022.
- [92] Y. Xue, W. Zheng, Y. Zong, H. Chang, and X. Jiang, "Adaptive hierarchical graph convolutional network for EEG emotion recognition," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2022, pp. 1–8.
- [93] J. Li, S. Li, J. Pan, and F. Wang, "Cross-subject EEG emotion recognition with self-organized graph neural network," *Frontiers Neurosci.*, vol. 15, p. 689, Jun. 2021.
- [94] J. Lu, Y. Tian, S. Wang, M. Sheng, and X. Zheng, "PearNet: A Pearson correlation-based graph attention network for sleep stage recognition," 2022, *arXiv:2209.13645*.
- [95] Z. Jia et al., "GraphSleepNet: Adaptive spatial-temporal graph convolutional networks for sleep stage classification," in *Proc. IJCAI*, 2020, pp. 1324–1330.
- [96] T. Wierciński, M. Rock, R. Zwierzycki, T. Zawadzka, and M. Zawadzki, "Emotion recognition from physiological channels using graph neural network," *Sensors*, vol. 22, no. 8, p. 2980, Apr. 2022.
- [97] H. Zeng et al., "Siam-GCAN: A Siamese graph convolutional attention network for EEG emotion recognition," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–9, 2022.
- [98] Y. Zhu, K. Gan, and Z. Yin, "Locally temporal-spatial pattern learning with graph attention mechanism for EEG-based emotion recognition," 2022, *arXiv:2208.11087*.
- [99] X. Li, W. Zheng, Y. Zong, H. Chang, and C. Lu, "Attention-based spatio-temporal graphic LSTM for EEG emotion recognition," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2021, pp. 1–8.
- [100] X. Li, J. Li, Y. Zhang, and P. Tiwari, "Emotion recognition from multi-channel EEG data through a dual-pipeline graph attention network," in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, Dec. 2021, pp. 3642–3647.
- [101] T. Song, S. Liu, W. Zheng, Y. Zong, and Z. Cui, "Instance-adaptive graph for EEG emotion recognition," in *Proc. AAAI Conf. Artif. Intell.*, vol. 34, no. 3, 2020, pp. 2701–2708.
- [102] T. Song et al., "Variational instance-adaptive graph for EEG emotion recognition," *IEEE Trans. Affect. Comput.*, vol. 14, no. 1, pp. 343–356, Jan. 2023.
- [103] Q. Lian, Y. Qi, G. Pan, and Y. Wang, "Learning graph in graph convolutional neural networks for robust seizure prediction," *J. Neural Eng.*, vol. 17, no. 3, Jun. 2020, Art. no. 035004.
- [104] T. Dissanayake, T. Fernando, S. Denman, S. Sridharan, and C. Fookes, "Geometric deep learning for subject independent epileptic seizure prediction using scalp EEG signals," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 2, pp. 527–538, Feb. 2022.
- [105] X. Li, B. Qian, J. Wei, A. Li, X. Liu, and Q. Zheng, "Classify EEG and reveal latent graph structure with spatio-temporal graph convolutional neural network," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Nov. 2019, pp. 389–398.
- [106] M. Zhdanov, S. Steinmann, and N. Hoffmann, "Investigating brain connectivity with graph neural networks and GNNExplainer," in *Proc. 26th Int. Conf. Pattern Recognit. (ICPR)*, Aug. 2022, pp. 5155–5161.
- [107] I. C. Covert et al., "Temporal graph convolutional networks for automatic seizure detection," in *Proc. Mach. Learn. Healthcare Conf.*, 2019, pp. 160–180.
- [108] W.-L. Zheng and B.-L. Lu, "Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks," *IEEE Trans. Auto. Mental Develop.*, vol. 7, no. 3, pp. 162–175, Sep. 2015.
- [109] Z. Ying et al., "Hierarchical graph representation learning with differentiable pooling," in *Advances in Neural Information Processing Systems*, vol. 31. Red Hook, NY, USA: Curran Associates, 2018.
- [110] J. Lee, I. Lee, and J. Kang, "Self-attention graph pooling," in *Proc. 36th Int. Conf. Mach. Learn.*, Jun. 2019, pp. 3734–3743.
- [111] X. Gao, W. Dai, C. Li, H. Xiong, and P. Frossard, "IPool—Information-based pooling in hierarchical graph neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 9, pp. 5032–5044, Sep. 2022.
- [112] H. Gao, Y. Liu, and S. Ji, "Topology-aware graph pooling networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 12, pp. 4512–4518, Dec. 2021.
- [113] Y. Wang, M. Wu, X. Li, L. Xie, and Z. Chen, "Multivariate time series representation learning via hierarchical correlation pooling boosted graph neural network," *IEEE Trans. Artif. Intell.*, vol. 5, pp. 1–13, 2023.
- [114] V. Jirsa and V. Müller, "Cross-frequency coupling in real and virtual brain networks," *Frontiers Comput. Neurosci.*, vol. 7, pp. 1–25, Sep. 2013.