Exploring Non-Euclidean Approaches: A Comprehensive Survey on Graph-Based Techniques for EEG Signal Analysis

Harish C. Bhandari [1], Yagya R. Pandeya [2], Kanhaiya Jha [1], Sudan Jha [2,3], and Sultan Ahmad [4,5,]*

[1] Department of Mathematics, School of Science, Kathmandu University, Dhulikhel, Nepal
[2] Department of Computer Science and Engineering, School of Engineering, Kathmandu University, Dhulikhel, Nepal
[3] School of Computer Science and Engineering, Lovely Professional University, Phagwara, India
[4] Department of Computer Science, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Saudi Arabia
[5] Department of Computer Science and Engineering, Chandigarh University, Punjab, India
Email: harish.bhandari@ku.edu.np (H.C.B.); yagya.pandeya@ku.edu.np(Y.R.P.); jhakn@ku.edu.np(K.J.); sudan.jha@ku.edu.np (S.J.); s.alisher@psau.edu.sa (S.A.)
*Corresponding author

Abstract—**Electroencephalogram (EEG) signals are widely applied in emotion recognition, sentiment analysis, disease classification, sleep disorder identification, and fatigue detection. Recent research has highlighted the active exploration of neurological disease analysis using EEG signals. Various machine learning and deep learning techniques, using feature-based and Euclidean approaches, have been employed to analyse these EEG signals. However, non-Euclidean approaches have proven more effective than Euclidean methods in EEG signal research. This superiority may stem from the nonlinear and dynamic characteristics of EEG signals, intricate interplay among brain regions, and resilience to common EEG signal noise. Unfortunately, limited studies on the graph representation of EEG signals exist due to constraints such as insufficient datasets, unavailable source code, and the complexity of graph representation. Hence, we aim to conduct a survey on various graph representation techniques, graph neural networks, existing methods, and available resources for EEG signal analysis using the non-Euclidean approach. In addition, visibility graph-based methods have been applied to single-channel EEG signals, while graph neural networks have been shown to have promising outcomes in multichannel EEG signal analysis. Thus, the survey concluded that the non-Euclidean approach uses a graph to map more with the brain structure than with the Euclidean structure. Additionally, the inclusion of visibility graphs in multichannel EEG signals and graph neural networks would justify the robustness of the non-Euclidean approach in EEG signal analysis.**

Keywords—**electroencephalogram signals, graph representation, graph neural network, intelligent processing, deep analytics**

I. Introduction

According to the World Health Organization, approximately one billion people worldwide have experienced Neurological Disorders (NDs), and a study [1] indicated that one out of every six individuals has suffered from some form of ND. The incidence of neurological disorders has been massively increasing; in 2022, the Neuroepidemiology Editorial [2] reported that every one in three people globally is suffering from NDs. Electroencephalography (EEG) [3, 4] is the most versatile, economical, nondestructive, painless, and side effect-free solution for studying brain function and diagnosing NDs. EEG signals are obtained by placing electrodes on the scalp, which detect the electrical activity of neurons in the brain. These signals are then transformed and stored in an external device such as a computer, generating a graph of brain activity over time. EEG signals are valuable for studying a variety of brain processes, such as perception, attention, emotion, memory, and motor control. Moreover, EEG signals can be used to diagnose conditions such as epilepsy, sleep disorders, and brain injuries.

Electroencephalography (EEG) is a powerful tool for revealing novel insights into brain function and dynamics. By detecting intricate patterns, identifying biomarkers, exploring brain connectivity, and facilitating real-time monitoring, EEG contributes significantly to advancing our understanding of brain structure, function, and dynamics. This not only advances neuroscience research but also enhances clinical diagnosis and informs therapeutic interventions. Moreover, the classical visibility graph approach, particularly its effectiveness in single-channel EEG signal analysis, underscores its potential. With technological advancements enabling the incorporation of multiple signals for visibility graphs and the application of data-driven approaches, the field presents exciting opportunities for future research. These advancements align with our conclusions, highlighting the profound implications of EEG findings in bridging the gap between new results and established domain knowledge.

Manuscript received January 25, 2024; revised February 26, 2024; accepted March 26, 2024; published October 8, 2024.

However, identifying NDs from nonlinear temporal signal frequencies has been a meticulous task [5]. Experts usually require hours of EEG signals to detect abnormal signals from epileptic seizures, despite the seizure duration of abnormal EEG signals being only a matter of seconds [6]. This may explain why half of patients appear normal to the naked eye of an expert, and the consequences can be irreversible. Using GNNs for EEG signal processing is a suitable approach for comprehending the nonlinear and non-Euclidean nature of signals. The non-Euclidean data can be represented as graphs or networks, where the distance between two nodes is not a Euclidean distance but a graph-theoretic distance such as the shortest path.

GNNs are a type of deep learning model specifically designed to process data represented in the form of graphs. In a graph, nodes signify entities, and edges represent the relationships or connections between them. The primary advantage of GNNs is their ability to handle data with intricate relationships, such as social networks, molecular structures, and brain connectivity networks. Thus, GNNs are well suited for capturing the spatial and temporal relationships between different brain regions and reducing the dimensionality of EEG data. Moreover, GNNs are robust to noise, artifacts, and other sources of interference that often affect EEG signals. Therefore, over the last few years, machine learning on graphs has gained much attention and achieved remarkable success in various fields, including computer vision, particle physics, medical imaging, social sciences, chemistry, and recommendation systems. The success of machine learning in these fields can be attributed to the availability of graph-structured data and the effective use of GNNs.

GNNs are a type of neural network designed to handle graph-structured data and are considered an extension of Convolutional Neural Networks (CNNs) from Euclidean to non-Euclidean data. GNNs are capable of processing information from nonlinear signals that are supported on graphs and leveraging topologically distant information in a nonlinear manner [7, 8]. GNNs transmit messages to neighboring nodes via the edges of graphs and aggregate them to generate node embeddings, which can be utilized for tasks such as node classification, edge prediction, and graph classification [9].

Previous researches [10–13] on EEG signals has focused primarily on feature extraction strategies, including statistical feature extraction methods [14], EEG waveform analysis [15], data collection and processing techniques, and software for handling EEG data [16, 17]. Survey papers have mostly concentrated on algorithms for selecting channels to extract features [18], different processing methods, acquisition types, and applications of EEG signals [19]. Lotte *et al.* [20] and Hosseini *et al.* [21] conducted reviews of classification algorithms for EEG signal analysis, and some papers have reviewed machine learning methods [22–24] for EEG signal analysis, such as Support Vector Machines (SVMs), k-means methods, Artificial Neural Networks (ANNs), linear classifiers, and deep learningmethods [25, 26], e.g., Deep Neural Networks (DNNs), Recurrent Neural Networks (RNNs), and CNNs, using Euclidean data processing approaches.

However, to the best of our knowledge, there has been no analytical study on non-Euclidean data processing approaches in EEG signal analysis using machine learning or deep learning technology. This paper aims to fill that gap by surveying the methods developed between 2012 and 2023 for EEG signal processing using graph-based neural networks. This paper presents a complete pipeline, from the basic graph representation to the implementation of the GNN. The paper also discusses the proposed methods, dataset, source code, challenges, and future directions for research in EEG signal processing using GNNs.

The paper is structured as follows. Section II explains the materials and methods, and Section III describes the graph representation techniques. Section IV presents the frameworks for graph neural networks, which have been used for EEG signal analysis. In Section V, brief EEG signal analysis approaches are illustrated with a comparative study. Section VI describes the dataset, and Section VII describes the challenges and future research scopes. Finally, Section VIII presents concluding remarks.

II. Materials and Methods

For this work, we selected various survey/review papers, journals, conference proceedings and public websites from 2012 to 2023 through the Google Scholar (https://scholar.google.com/) and PubMed (https://pubmed.ncbi.nlm.nih.gov/) search engines. The choice of search engines is not random. Google Scholar's extensive coverage, including books, articles, and conference papers, coupled with its user-friendly interface, makes it an invaluable resource. Additionally, its citation tracking feature aids in assessing the impact and significance of research papers. On the other hand, PubMed's specialization in biomedical and life sciences literature, along with its provision of direct links to full-text articles, streamlines the process of accessing pertinent research materials, particularly for studies in the field of EEG signal analysis. Initially, we searched for papers with the key phrase "EEG signals and graph neural networks". During the process, many EEG signal analysis techniques involve visibility graphs, so we slightly changed the strategy and collected the papers with the following three key phrases:

- Visibility graph and EEG signal analysis
- EEG signal and graph neural networks
- Graph learning on EEG

Using the phrase "visibility graph and EEG signal", we considered the articles after the year 2012, and for the remaining two phrases, only the articles after 2016 were considered to be graph neural network-based learning after the work of the graph convolution network paper [27]. Initially, 299 articles were identified using specific keywords and databases. Of these, 94 articles were inaccessible either publicly or through third-party websites. Additionally, 81 papers were eliminated because their titles did not align with the scope of the study. Further screening based on abstracts resulted in the exclusion of an additional 56 papers, leaving a refined selection for analysis. After removing duplications among the selected articles, only 53 were considered for this work. Table I provides information about the screening methods and the selected number of articles.

TABLE I. ARTICLE SELECTION AND SCREENING CRITERIA

Screening Methods	Visibility graph and EEG signal analysis	EEG signals and graph neural networks	Graph learning on EEG
Total collected articles	100	109	90
Not freely available	22	37	35
Abstract reading	20	22	14
Selected	18	32	18

III. GRAPH REPRESENTATION

A graph is a kind of non-Euclidean data structure that shows the relationships between entities; such entities are known as nodes, and relationships are the edges of the graph. Formally, a graph can be defined as $G = (V, E)$, where $V = \{v_1, v_2, \dots, v_n\}$ is the set of n vertices and $E = \{e_1, e_2, \dots, e_m\}$ is the set of m edges. If the edges on the graph are directed from one node to another, then the resulting graph is the directed graph. Additionally, if the strengths of the relationships between nodes are not the same, such graphs are weighted graphs. The graph defined above is an undirected and unweighted graph.

The graph density D of an undirected graph $G = (V, E)$ is given by the ratio of the number of edges to the maximum possible number of edges in the graph. That is,

$$D = \frac{|E|}{\frac{|V|(|V|-1)}{2}} = \frac{2|E|}{|V|(|V|-1)}$$

If $D = 0$, the graph is an empty graph or discrete graph, and if $\mathrm{D} = 1$, the graph is a fully connected or complete graph.

Based on this density, graphs can be classified as sparse graphs or dense graphs, i.e., if $D <$ threshold, the graph is a sparse graph; otherwise, it is a dense graph. The common rule of thumb is that the threshold is usually equal to 0.5. An EEG signal is a complex temporal signal from multiple electrodes on the scalp. To reduce the complexity of graph feature engineering, some have attempted to reduce the number of nodes or electrodes, while others have attempted to reduce the connectivity between EEG channels based on the notion of graph density.

In the real-world context, different graph variants are common in practice: dynamic graphs [28, 29], signed graphs [30], hypergraphs [31, 32], heterogeneous graphs [33, 34], bipartite graphs [35, 36], and visibility graphs [37]. The details of all those graphs are not within the scope of this paper, so we focus only on the graphs that have been used in practice for EEG signal analysis.

A. Temporal Representation

Visibility graphs can be used to transform time series data into scale-free graphs, and these graphs exhibit long-term temporal dependencies and chaotic properties of time series data [38]. Since the introduction of visibility graphs by Lacasa [33] in 2008, visibility graphs have been used for EEG signal analysis. Therefore, this section describes the chronological development of EEG signal analysis based on the concept of visibility graphs.

1) Visibility graph

For EEG signal analysis, the structure of such a nonlinear and nonstationary signal is crucial. However, the visual structure of large temporal sequences is often difficult for human experts to observe. Thus, it is essential to map such a temporal sequence into a domain that preserves the structure of the sequence with minimum cost and minimum information loss. Since the introduction of a visibility graph in 2008 by Lacasa [37], many researchers have implemented this concept in EEG signal analysis.

For a temporal sequence $x = [x_1, x_2, \dots, x_N]$ of N data points, a Visibility Graph (VG) of x is a graph $G = (V, E)$ where $V = \{1, 2, \dots, N\}$ and

$$E = \left\{(t_i, t_j): x_k - x_j < \frac{t_k - t_i}{t_j - t_i}(x_j - x_i), when\ t_i < t_k < t_j\right\}$$

The VG of temporal points is the graph whose nodes are the temporal points, and the edges are formed between these temporal points if a straight line can be drawn in the bar chart of the temporal sequence without the blockage of middle bars.

To the left of Fig. 1 is the bar chart of the temporal data $x = [10, 5, 3, 7, 10, 5, 4, 8]$, and the right figure is the bar with the corresponding visibility graph.

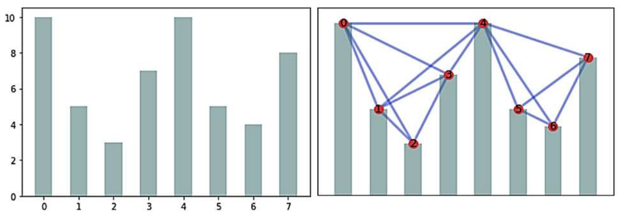

Fig. 1. A bar chart and the visibility graph on top of the bar chart.

EEGs are highly fluctuating, and a measure of such fluctuations is fractality. Ahmadlou *et al.* [39] used the notion of fractality to diagnose autism spectrum disorder with the help of VG. In the expression of the probability distribution of nodes in the VG, $P(k) = k^{-r}$, where k is the degree of the node and r is the fractality measure. This fractality can be measured as the slope of a least square line of $\log_2(P(k))$ against $\log_2\left(\frac{1}{k}\right)$.

Ji *et al.* [40] used the degree distribution of nodes of VGs obtained from EEG signals of normal people as well as those experiencing occupational stress. They showed that their degree distributions differ according to occupational stress. Similarly, Hao *et al.* [41] used VG signals to classify EEG signals into ictal and interictal signals.

VGs can be large, and in real life, working with large VGs is not preferable because of their high complexity. To reduce (sparsify) the connection and still preserve the structure of the temporal sequence, Luque *et al.* [42]

introduced the concept of a Horizontal Visibility Graph (HVG).

2) Horizontal visibility graph

For the temporal sequence $x = [x_1, x_2, \dots, x_N]$, the HVG of x is denoted by $HVG(x) = (V, E)$, where the set of vertices is exactly the same as that of VG and

$$E = \{(t_i, t_j): x_i, x_j > x_k \; when \; 1 \leq i < k < j \leq N\}$$

Zhu *et al.* [43] applied the concept of the VG similarity score using the Graph Synchronization (GS) score as a feature of the Horizontal Visibility Directed Graph (HVDG) to classify sleep stages from EEG signals. The synchronization score of two temporal sequences $x(t)$ and $y(t)$ is measured as

$$S(x, y) = \frac{cov(D_x, D_y)}{\sigma_{D_x}\sigma_{D_y}}$$

where D_x is the degree of the degree sequence of VG of x, $cov(x, y)$ is the covariance between x and y, and σ_x is the standard deviation of x. Here, the range of the synchronization score is $[0, 1]$.

This Visibility Graph Synchronization (VGS) method is more powerful than coherence methods. On the other hand, VGS is not appropriate for time series signals that produce the same degree of sequence, and the same degree of sequence production is possible because VGs are shift- and scale-invariant. To overcome this limitation, the author used the HVDG similarity method. The synchronization is calculated from the degree sequence of the HVDG as

$$S(x, y) = \frac{cov(D_x, D_y)}{\sigma_{D_x}\sigma_{D_y}}$$

where the range is between -1 and 1. A value close to -1 indicates that x and y are desynchronized, and a value of 1 indicates synchronization.

The VGS concept was applied by the author of [44] in a slightly different way. First, the given temporal signal is transformed in state space with multiple concepts [45] with different embedding dimensions. Then, the resulting trajectories are mapped into VGs. Then, the synchronization score is measured as

$$S(x, y) = cov(DS_x, DS_y)/\sigma_{DS_x}\sigma_{DS_y}$$

where DS_x is the degree sequence of trajectory x. The advantage of this approach is that it permits the incorporation of multichannel signal analysis. Sengupta *et al.* [46] used this concept to construct a brain network and observe various stages of sleep-deprived fatigue. It first calculates the synchronization score between the signals from every pair of electrodes to form an adjacency matrix called the VGS matrix. Then, a threshold is applied to decide whether to form an edge between them. The various functional brain stages were observed through the VGS matrix. It has been observed that the greater the number of edges is, the greater the fatigue level and the lower the degree of connection fatigue.

On the other hand, Ahmadi and Pechenizkiy [47] suggested that as VGs are nonlinear transformations of signals into graphs, the notion of state space is not required because of the use of VGs. The author verified that the synchronization score classifies healthy and unhealthy signals.

In 2014, a new concept was introduced by Zhu *et al.* [48] to sparsify the edges of the Visibility Graph (VG) derived from a temporal sequence, called the difference VG. This was used to analyze and classify sleep stages from a single-channel EEG signal.

3) Difference visibility graph

Let $G_{vg}(V, E_1)$ and $G_{hv}\ (V, E_2)$ be the VG and HVG obtained from time series $x = [x_1, x_2, \dots, x_N]$, respectively, where V is the set of vertices $\{1, 2, \dots, \mathrm{N}\}$ and E_1 and E_2 are the sets of edges of the VG and HVG, respectively. Then, the Difference Visibility Graph (DVG) is the graph $G_{dvg}(V, E_3)$, where E_3 is the difference between the set of edges E_1 of the VG and E_2 of the HVG.

Experimentally, Zhu *et al.* [48] verified that DVG extracts more essential features than VG and HVG. The features from the DVG were sent as input to the SVM to classify sleep stages. The mean degree of DVGs in the deep sleep stage was greater than that in the light sleep stage. Wang *et al.* [49] used the mean degree and degree entropy from the VG, HVG and DVG to analyse seizure patterns.

Another way to control sparsifying the edges of VGs was introduced in [50] in terms of a Weighted Horizontal Visibility Graph (WHVG). The author did so based on the fact that if HVGs obtained from time series are consistently monotonic, it is difficult to distinguish them.

4) Weighted horizontal visibility graph

The WHVG of $x = [x_1, x_2, \dots, x_N]$ is a graph $G(V, E, W)$ where V and E behave exactly the same as the HVG and $W = (w_{ij})$ and w_{ij} are the edge weights between nodes i and j given by

$$w_{ij} = \begin{cases} \left|\frac{x_i - x_j}{i - j}\right| + 1 \; if \; e_{ij} \in E \\ 0 \qquad otherwise \end{cases}$$

where, the degree (strength) of node i of WHVG $S_i = \Sigma_j\, w_{ij}$ creates a difference even for monotonic time series. Finally, the features extracted from the WHVG were input to the KNN classifier for epileptic seizure detection corresponding to the EEG signals.

Supriya *et al.* [51] introduced the concept of WVG similar to that of Zhu *et al.* [50] but for VG and with the weight defined differently as

$$w_{ij} = \left|\frac{x(j) - x(i)}{j - i}\right| \quad j > i.$$

This approach of measuring edge weight helps to record sudden changes in EEG signals for the duration of seizure activity. It helps to recognize fluctuations in EEG signals. Mohammadpoory *et al.* [52] also used WVG with the edge weight given by

$$w_{ij} = \begin{cases} \tan^1\left(\frac{x_i - x_j}{i - j}\right), i < j \; if \; e_{ij} \in E \\ 0 \qquad otherwise \end{cases}$$

Cai *et al.* [53] also used the WVG from the theta band for characterizing spontaneous activity in the AD and control groups.

In contrast to sparsifying graph edges, Wang *et al.* [54] proposed the notion of a Limited Penetrable Visibility Graph (LPVG), which adds more edges than VG does. The notion was based on the assumption that VGs cannot detect the spatial location of inverse bifurcation in a chaotic dynamical system.

5) Limited penetrable visibility graph

An LPVG is a generalization of a VG in which two nodes are connected if and only if

$$x_{m+j} < x_n + \left(\frac{n-(m+j)}{n-m}\right)(x_m - x_n)\ \forall\ j \in \mathbb{Z}^+,\quad j < n-m$$

An edge can be formed between two temporal signal time points even if the maximum number of predefined bars blocks their visibility.

Gao *et al.* [55] generalized the LPHVG to a multiplex limited penetrable horizontal visibility graph. It first transforms the time series $x = [x_1, x_2, \ldots, x_N]$ into a coarse time series $\{y_j : j = 1, 2, \ldots, \frac{N}{s}\}$ by

$$y_j^s = \frac{1}{s}\sum_{i=(j-1)s+1}^{js} x_i \quad 1 \le j \le \frac{N}{s}$$

where s is the scale factor.

Gao *et al.* [56] generalizes the concept of VG to time-frequency domain representation for epileptiform classification. First, the EEG signal was transformed into a time-frequency representation by an adaptive optimal kernel, and a VG was formed before extracting features and sending them to the classifier. This time-frequency representation enables the mapping of nonstationary signals to the time-frequency plane with a diagnostic energy distribution. Supriya *et al.* [57] formed WVG from Fourier transformed series to analyse sleep stage classification.

Samanta *et al.* [58] incorporated multichannel EEG signals and created a WVG for each channel. Then, another graph was created by considering each channel as nodes, and edge connectivity was measured based on the correlation between the CCs of each VG.

A similar concept was used in a previous paper [59]. For multichannel EEG signals $\{x_{\alpha,i}\}_{i=1}^{N}$ and $\alpha = 1, 2, \ldots, M$, M is the number of channels, and N is the length of each EEG signal. First, the LPVG is created for each channel separately, which is denoted by $\{A^\alpha\}_{\alpha=1}^{M}$. Then, coherence between a pair of EEG channels was measured by $w = \frac{\sum_i \sum_{j>i} \sum_\alpha a_{i,j}^\alpha}{M \sum_i \sum_{j>i}\left(1-\delta_{0,\sum_\alpha a_{i,j}^\alpha}\right)}$, where δ is Kronecker's delta and a_{ij}^α is the edge between nodes i and j of the LPHVG of channel α. Here, $w \in \left[\frac{1}{M}, 1\right]$, where w close to $\frac{1}{M}$ indicates that the edge is unique in each channel, and $w = 1$ means that M channels are identical.

Based on the degree distribution of the LPHVG for each layer, the interconnection between pairs of channels is calculated as

$$I_{\alpha,\beta} = \sum_{K^\alpha}\sum_{K^\beta} P\left(K^\alpha, K^\beta\right)\log\frac{P\left(K^\alpha, K^\beta\right)}{P(K^\alpha)P(K^\beta)}$$

where $P\left(K^\alpha, K^\beta\right) = \frac{N_{K^\alpha,K^\beta}}{N}$ and N_{K^α,K^β} is the number of nodes with degree K^α in channel α.

This matrix $I_{\alpha,\beta}$ gives the brain network from which graph statistics were computed to classify EEG signals.

B. Spatial Representation

Structurally, the graph representation has been the same for multiple electrode EEG signal analysis. Almost all existing EEG signal paradigms use different brain regions (electrode/channel positions) as the nodes of the graph and the relationships between the signals from these regions as the edges of the graph. However, the number of electrodes from which the signals are extracted is not uniform; some use random numbers of electrodes, and some use random numbers based on prior knowledge.

Incorporating signals from all possible electrodes is a challenging machine learning task because of its high complexity. To reduce the complexity, many authors have used different approaches involving sparsifying techniques. Therefore, in this section, the article will briefly discuss these techniques, which are commonly used.

Lin *et al.* [60] introduced the concept of edge weight based on a two-dimensional vector representation of electrodes given by

$$W_{xy} = \begin{cases} d_1 = \left|C_{x,1} - C_{y,1}\right| \\ d_2 = \left|C_{x,2} - C_{y,2}\right| \end{cases}$$

where $\left(C_{x,1}, C_{x,2}\right)$ and $\left(C_{y,1}, C_{y,2}\right)$ are the coordinates of electrodes x and y, respectively. The concept was generalized by [61–63] 3D coordinates of electrodes, where the Euclidean distance between the electrodes was considered to be the weight of the graph edge. As the human head has a manifold-like structure, Wagh *et al.* [64] and Jia *et al.* [65] incorporated the geodesic distance between two electrodes given by

$$w_{xy} = \cos^{-1}\left(\frac{x_1y_1 + x_2y_2 + z_1z_2}{r^2}\right)$$

where (x_1, x_2, x_3) and (y_1, y_2, y_3) are the Cartesian coordinates of the electrodes and is the edge weight.

Tang *et al.* [66] generalized and computed edge weights by applying a threshold on a Gaussian kernel to the Euclidean distance between two electrodes, i.e.,

$$w_{xy} = \begin{cases} exp\left(-\frac{dist(x,y)^2}{\sigma^2}\right) & if \quad dist(x,y) \le \kappa \\ 0 & otherwise \end{cases}$$

where σ is the standard deviation of the distances and κ is the threshold.

Based on the assumption that the strength of connections between brain regions attenuates as an inverse square function of physical distance, Zhong *et al.* [67] and Cai *et al.* [68] generalized the edge weight of a graph to

$$W_{xy} = min\left\{1, \frac{\delta}{q_{xy}^2}\right\}$$

where $\delta > 0$ denotes the calibration constant and q_{xy} is the physical distance between channels x and y.

Zhang *et al.* [69] generalized the distance-based graph in the following three ways:

- N-Graph:

First, it considers the natural neighboring electrodes from brain regions (up, down, left, right, upleft, upright, downleft and downright) as the set of local connectivity E_v of the node v. Then, the edge weight is given by

$$w_{xy} = \begin{cases} 1 & if\ x, y \in E \\ 0 & else \end{cases}$$

This equation gives the connection between nodes x and y.

- D-Graph:

In this case, the graph topology is exactly the same as that of the N-graph topology but with weighted edge connectivity

$$w_{xy} = \begin{cases} \frac{1}{d_{xy}} & if\ d_{xy} < E(L) \\ 0 & if\ d_{xy} \geq E(L) \\ \frac{1}{E(\{d_{xq}: d_{xq} < E(L), q \in [1,2, \ldots n]\})} & if\ x = y \end{cases}$$

- S-Graph:

It is obtained by slight modification of the D-Graph edge weight between nodes x and y as follows:

$$w_{xy} = \begin{cases} \frac{1}{d_{xy}} & if\ d_{xy} < E(L) \\ 0 & if\ d_{xy} \geq E(L) \\ \frac{1}{min\{d_{xq}: q \in [1,2, \ldots n]\}} & if\ x = y \end{cases}$$

C. *Functional Representation*

The functional representation of a graph is based on the functional connectivity of the brain. The functional connectivity in EEG signals indicates several measures of how electrical activity in one brain region differs from electrical activity in another brain region. Some familiar functional graph representations are discussed below.

1) *Correlation-based graph*

For the signals $x(t)$ and y(t) from two electrodes x and y at time t, the edge weight of the graph is given by

$$w_{xy} = \left| \frac{\sum(x(t) - \bar{x}(t))(y(t) - \bar{y}(t))}{\sqrt{\sum(x(t) - \bar{x}(t))^2}\sqrt{\sum(y(t) - \bar{y}(t))^2}} \right|$$

The absolute value is taken here to make the graph undirected so that the weights from x to y and y to x. Zhong *et al.* [68] also used correlation coefficients as edge weights. Lun *et al.* [70] generalized correlation-based edge to

$$w_{xy} = \begin{cases} |r_{xy}| & if\ \ x \neq y \\ |r_{xy}| - 1 & if\ x = y \end{cases}$$

where r_{xy} is the correlation coefficient between x and y.

2) *Phase locking value-based graph*

The proposed approach is sensitive to artifacts. To minimize the effect of artefacts, a PLV-based approach was introduced [71–74].

The edge weight of the graph based on the PLV is given by

$$w_{xy} = \left| \frac{1}{T} \sum_{t=1}^{T} e^{-(\phi_x(t) - \phi_y(t))} \right|$$

where $\phi_x(t)$ is the instantaneous phase of signal $x(t)$ obtained from the Hilbert transform. It quantifies the consistency of phase differences between signals.

Klepl *et al.* [71] also mentioned the weighted PLV

$$w_{xy} = \left| \frac{1}{T} \sum_{t=1}^{T} \frac{\left|\sin\left(\phi_x(t) - \phi_y(t)\right)\right|}{\sin\left(\phi_x(t) - \phi_y(t)\right)} \right|$$

which removes the effect of amplitude and volume conduction by maximally weighting the ± 90 degree phase differences and helps to remove uniformly driven differences. Additionally, Samanta *et al.* [58] mentioned that the relative PLV given by

$$w_{xy} = \left| \frac{1}{T} \sum_{t=1}^{T} e^{-\phi_{xy}(j\Delta t)} \right|$$

where $\phi_{xy}(t)$ is the relative phase between EEG signals from electrodes x and y, N is the number of sample points and $\frac{1}{\Delta t}$ is the sampling frequency.

3) *Phase lag index-based graph*

The phase and amplitude of an EEG signal at time "*t*" can be calculated from the analytic representation

$$z(t) = x(t) + \tilde{x}(t)$$

where $\tilde{x}(t)$ is the Hilbert transform of $x(t)$. Then, the phase and amplitude can be defined as

$$\phi(t) = \tan^{-1}\left(\frac{\tilde{x}(t)}{x(t)}\right)$$

and

$$amp(t) = \sqrt{\left(x(t) + \tilde{x}(t)\right)^2}$$

Then, the edge weight based on the phase lag index (PLI) is given by

$$w_{xy} = \left| \frac{1}{T} \sum_{t=1}^{T} sign \sin\left(\phi_x(t) - \phi_y(t)\right) \right|$$

where $\phi_x(t)$ is the phase of the signal from electrode x.

The concept of a PLI-based graph was introduced [57] using the cross-spectrum density is given by

$$w_{xy} = \left| \frac{\sum_{t=1}^{T} \left| Img\left(S_{xy,t}\right)\right| sign\ \left(Img\ \left(S_{xy,t}\right)\right)}{\sum_{t=1}^{T} \left| Img\left(S_{xy,t}\right)\right|} \right|$$

where $Img\left(S_{xy,t}\right)$ is the imaginary component of the cross-spectrum $S_{xy,t}$ of two signals from two electrodes x and y at time t. Liu *et al.* [75] discussed the weighted phase lag index between the cross-spectra of two signals.

Hasanzadeh *et al.* [76] used the concept of phase transfer entropy to quantify edge weights in brain network graphs.

4) Coherance-based graph

The weight edge for the signals from two electrodes x and y is given by

$$w_{xy} = \frac{\left|S_{xy}(f)\right|^2}{S_{xx}(f) S_{yy}(f)}$$

where $S_{xy}(f)$ is the cross-spectral density and $S_{xx}(f)$ is the power spectral density at frequency f. The coherence within the frequency band is calculated as the mean.

5) Imaginary coherence-based graph

The weight matrix is

$$w_{xy} = \frac{Img\left(S_{xy}(f)\right)}{\sqrt{S_{xx}(f)}\sqrt{S_{yy}(f)}}$$

where Img is the imaginary component. Similar to the previous approach, the frequency within the band is the mean.

It measures phase consistency exactly as the coherence-based graph and describes the volume conduction effect. This is the effect of recording electrical activities at a distance from the source generator.

6) Partial directed coherence-based graph

It is a frequency domain-based analysis method for multichannel EEG signals, and it is suitable for information transformation between multichannel EEG signals [77]. The p-th autoregression

$$x_i(n) = \sum_{i=1}^{p} a_{id} x_i(n-i) + e_i(n)$$

where i and d represent channels, $a_{id}\}$ is the coefficient from the i-th to the j-th channel, e is the deviation, and $x_i(n)$ is the state value of the variable at time n and is used to represent each EEG signal. Then,

$$A_{ij}(f) = \begin{cases} 1 - \sum_{i=1}^{p} a_{ij}(r) e^{-i2\pi f} & if\ i = j \\ -\sum_{i=1}^{p} a_{ij}(r) e^{-i2\pi f} & if\ i \neq j \end{cases}$$

where r is the order, f is the frequency, and the partial directed coherence is defined as

$$P_{ij}(f) = \frac{A_{ij}(f)}{\sqrt{a_i^H(f) a_j(f)}}$$

H is the conjugate transpose.

7) Amplitude envelope correlation-based graph

The graph based on the amplitude envelope correlation has an edge weight

$$w_{xy} = \left| \frac{\sum\left(\phi_x(t) - \bar{\phi}_x(t)\right)\left(\phi_y(t) - \bar{\phi}_y(t)\right)}{\sqrt{\sum\left(\phi_x(t) - \bar{\phi}_x(t)\right)^2}\sqrt{\sum\left(\phi_y(t) - \bar{\phi}_y(t)\right)^2}} \right|$$

This quantifies the coupling based on the amplitude of signals.

8) Mutual information-based graph

The mutual information-based [71] graph has an edge weight

$$w_{xy} = \Sigma_{x_i, x_j} P_{XY} \log\left(\frac{P_{XY}\left(x_i, y_j\right)}{P_X(x_i) P_Y\left(y_j\right)}\right)$$

where P_{XY} is a joint probability distribution and P_X is a marginal probability distribution. It quantifies the amount of known information about a second signal by observing the first signal.

9) Lagged linear coherence-based graph

The lagged linear coherence-based graph [78] has an edge weight related to the value given by

$$w_{xy} = \frac{[Img\ Cov(x,y)]^2}{Var(x) Var(y) - [Re\ Cov(x,y)]^2}$$

It provides a measure of true physiological connectivity not affected by volume conduction and low spatial resolution.

10) Attention-based graph

The learnable attention-based edge weight of the graph [79] is given by

$$w_{xy} = \frac{exp\ \left(ReLU\left(\omega^T \left|F_x - F_y\right|\right)\right.}{\sum_{y=1}^{N} exp\ \left(ReLU\left(\omega^T \left|F_x - F_y\right|\right)\right.}$$

where F_x is the input feature of a node and ω is a learnable vector that is updated by the following loss function $\sum_{x,y=1}^{N} \left|\left|F_x - F_y\right|\right|_2^2 w_{xy} + \lambda \left|\left|W\right|\right|_{Frob}^2$. Here, a larger $\left|\left|F_x - F_y\right|\right|_2$ implies a smaller w_{xy}.

The idea of generating such a graph is due to limited understanding of the brain. λ in the above expression controls the sparsity of the brain network graph.

Li *et al.* [80] used the same learnable attention edge weight but with different activation functions $\tan h$.

11) Neural network-based graph

The edge weight of the GNN-based graph [81] is defined as

$$w_{xy} = \exp\left(\frac{exp\left(\epsilon * P_{xy}\right) - 1}{\sigma}\right)$$

where ϵ, σ are hyperparameters used to control the strength and P_{xy} is the connection probability between channels x and y given by the GNN.

Fig. 2 comprises three images that depict different aspects of network graphs. The first image shows the brain locations where electrodes are systematically placed to collect brain signals. The second image displays a spatio-functional network graph with nodes representing a subsample of the electrodes and edges representing physical distances and/or functional relationships between them. The colors of the edges indicate the strength of connectivity among the nodes. The third image represents the spatiotemporal graph of brain networks, which captures how the network evolves over time. It is challenging to incorporate brain signals from all channels in EEG data, so researchers select a sample of the best representative channels for a specific task and create multiple graphs. However, single graph creation based on labels is a laborious task.

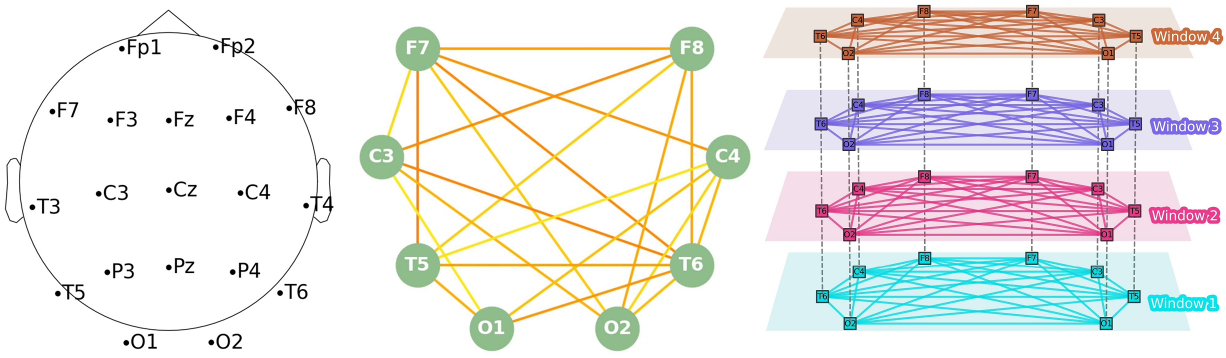

Fig. 2. (Left): Electrode placement on the head based on the international 10-20 system. (Middle): Functional graph considering only eight electrodes as the nodes of the graph. (Right): spatiotemporal graph as the stack of multiple spatiotemporal graphs.

Based on the papers discussed in this section, Fig. 3 gives the detailed classification of graph representation methods for EEG signal analysis.

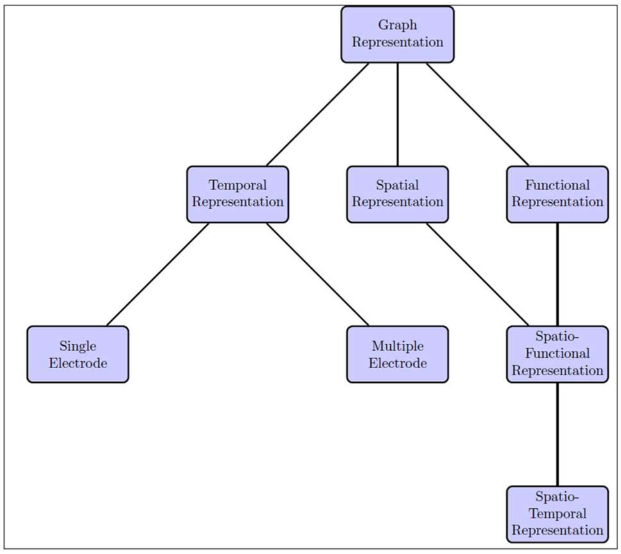

Fig. 3. Graph representation approaches for EEG signals.

From close observation of the different types of graphs, graph representation techniques are summarized in Fig. 3, to analyse and visualize brain connectivity networks.

In this section, we provide an overview of the different approaches that can be used to represent EGG signals in network graphs. These methods may vary depending on the problem being addressed, the metadata available in the data, and the hardware dependencies. In the next section, we discuss several potential techniques for processing graph representations of EEG signals using neural networks. We have conducted a comparative analysis of various existing methods, highlighting their advantages and limitations. To assist newcomers to this field, we have included the methods, datasets, and code materials used in some existing studies.

IV. Graph Neural Networks

Graph Neural Networks (GNNs) are a category of neural network that can learn representations of nodes and edges in a graph and can use this information to perform various tasks on the graph, such as node classification, link prediction, and graph classification. Basically, GNNs contain two major steps, namely, aggregation and updating, which assemble to form a message passing step in graphs. Graph Convolutional Networks (GCNs) are a specific type of GNN that are designed for graphs with a regular structure, such as grids or lattices. GCNs use convolutional operations, similar to those used in image processing, to perform message passing between nodes in a graph. This allows GCNs to learn hierarchical representations of the graph structure, which can be used for various tasks, such as node classification and link prediction. There are abundant GNN variants because of their widespread applicability, and here, we included several methods that are commonly used for EEG signal analysis.

Mathematically, for a hidden representation in the k^{th} layer, h_u^k for each node $u \in V$, a new embedding h_u^{k+1} in the $(k+1)^{th}$ layer can be obtained as

$$\begin{aligned} h_u^{k+1} &= f_{update}^k\left(h_u^k, f_{aggregate}^k\left(\{h_v^k, \forall v \in \mathcal{N}(u)\}\right)\right) \\ &= f_{update}^k\left(h_u^k, m_{\mathcal{N}(u)}^k\right) \end{aligned}$$

where f_{update} and $f_{aggregate}$ are arbitrary differentiable functions and $m_{\mathcal{N}(u)}^k = f_{aggregate}^k\left(\{h_v^k, \forall v \in \mathcal{N}(u)\}\right)$.

The choice of f_{update} and $f\_aggregate$ provides the different variants of GNNs. The basic message passing through the GNN at the node level is

$$m^k_{\mathcal{N}(u)} = f^k_{aggregate}\ (\{h^k_v, \forall\ v \in\ \mathcal{N}(u)\}) = \sum_{v \in \mathcal{N}(u)} h^k_v$$

and

$$h^{k+1}_u = f^k_{update}\left(h^k_u, m^k_{\mathcal{N}(u)}\right) = \sigma(W^k_u\ h^k_u + W^k_{\mathcal{N}(u)} m^k_{\mathcal{N}(u)})$$

where σ is nonlinear, such as ReLU, and W^k_u and $W_{m^k_{\mathcal{N}(u)}}$ are trainable parameters.

Additionally, graph-level message passing can be written as

$$H^{k+1} = \sigma(AH^k W^k_1 + H^k W^k_0)$$

where $H^k \in \mathbb{R}^{|V| \times d}$ is the node representation at layer k of the entire graph, A is the adjacency matrix and W^k_0, W^k_1 are the trainable parameters.

Another variant is

$$m^k_{\mathcal{N}(u)} = f^k_{aggregate}\ (\{h^k_v, \forall\ v \in\ \mathcal{N}(u)\}) \cup \{u\}\}\ )$$

which excludes the explicit update step, and the matrix-level expression is

$$H_{k+1} = \sigma(\ (A + I)H^k W^k)$$

There are different variants of GNNs based on the aggregation function $f_{aggregation}$, some of which are as follows:

- GraphSAGE [59] uses the aggregation function that normalizes the aggregation function based on the connectivity of nodes.

$$m^k_{\mathcal{N}(u)} = \sum_{v \in \mathcal{N}(u)} \frac{h^k_v}{|\mathcal{N}(u)|}$$

- The graph convolutional network (GCN)[60] is a symmetric normalization function; hence, the updated node feature on the $(k+1)^{th}$ layer is

$$h^{k+1}_u = \sigma\left(W^k_u \sum_{v \in \mathcal{N}(u)} \frac{h^k_v}{\sqrt{|\mathcal{N}(u)||\mathcal{N}(v)|}}\right)$$

- The aggregation function of the graph attention network [82, 83]

$$m^k_{\mathcal{N}(u)} = \sigma\left(W^k_u \sum_{v \in \mathcal{N}(u)} \alpha_{u,v} h^k_v\right)$$

where $\alpha_{u,v} = \frac{\exp\left(a^T\left(W h^k_u \| W h^k_v\right)\right)}{\sum_{v' \in \mathcal{N}(u)} \exp\left(a^T\left(W h^k_u \| W h^k_{v'}\right)\right)}$, a^T is the trainable attention vector and W is the trainable parameter.

- A graph isomorphism network has used the following aggregation function [83, 84]:

$$h^k_u = MLP\left((1 + \epsilon^k).\, h^{k-1}_u + \sum_{v \in \mathcal{N}(u)} h^{k-1}_v\right)$$

where MLP is a multilayer perceptron and ϵ^k is a hyperparameter.

The methods discussed above are all static graphs, but the EEG signal is a temporal signal. Thus, to incorporate the temporal dimension, some authors have attempted dynamic graph convolution [85, 86] and temporal graph convolutions [87].

V. EEG Signal Analysis Methods Based on Graphs

The fundamental aspect of machine learning and deep learning pipelines is the representation of data as vectors or tensors. Real numbers are typically used for this purpose, and classical machine learning and deep learning approaches rely on Euclidean geometry for data processing. This Euclidean approach includes frameworks such as ANNs, CNNs, and RNNs. However, due to the irregular shape of the brain and the nonuniform distribution of data generated by EEG electrodes, non-Euclidean methods are considered better for EEG signal analysis. As a result, this survey paper focuses only on non-Euclidean methods for EEG signal processing using neural networks. We broadly categorized various EEG signal processing techniques and illustrated them in Fig. 4. In this section, we described some data-driven approaches proposed in the past and the datasets they used for specific purposes. The details of the datasets are described in Section VI.

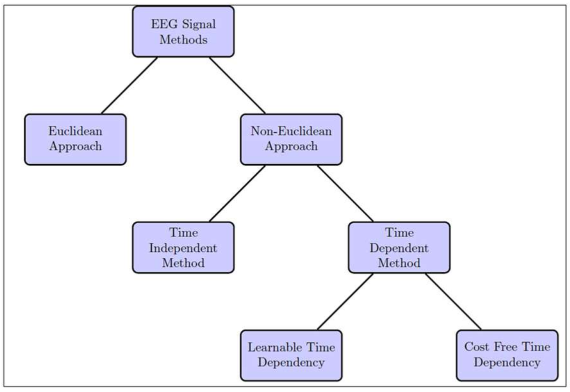

Fig. 4. Structure of the EEG signal analysis methods.

Over time, various handcrafted and automated methods have been proposed for EEG signal analysis. This paper focuses on data-driven automated methods that utilize neural network architectures for specific tasks. In previous work, machine learning methods such as linear SVM, LDA, logistic regression, and KNN have been utilized for EEG signal classification, as demonstrated in previous studies [88, 89]. However, with the rise of deep learning, several deep neural architectures have been introduced for EEG signal analysis for diverse applications. For instance, in Refs. [77, 88, 90], deep graph convolutional networks (GCNs) were utilized for fatigue driving experiments, where the authors proposed a driving fatigue detection dataset and used a Partial Directed Coherence Graph Convolutional Neural Network (PDC-GCNN) to study mental fatigue. In another study [66], a GCN was used to decode EEG Motor Imagery (MI) signals. To evaluate the effectiveness and robustness of the proposed GCN, two benchmark datasets, namely, the PhysioNet dataset (available at https://physionet.org/data/) and the High

Gamma dataset (available at https://www.bbci.de/competition/iv/), are utilized. Furthermore, in Ref. [91], a GCN was employed for EEG signal classification on two public datasets, Error-related Potentials (ErrP) (available at https://www.kaggle.com/c/inria-bci-challenge/) and Rapid Serial Visual Presentation (RSVP) (available at http://bci.med.tsinghua.edu.cn/download.htm). For abnormality detection in EEG signals, Lin *et al.* [60] proposed a GCN and validated their system on the Temple University Hospital EEG (TUH EEG). A GCN was used for Alzheimer's disease classification on their own Alzheimer's disease dataset. Using the CHB-MIT (available at https://physionet.org/content/chbmit/1.0.0/) dataset, Raeisi *et al.* [61] proposed a GCN with three GCN convolution layers, one global average pooling layer, and three Fully Connected (FC) layers for Band Energy and Hjorth. All of the discussed GNNs have similar structures, i.e., convolutions followed by flatten or global average pooling, and fully connected layers at the head of the neural network for EEG signal classification.

GCNs are a popular choice for geometric behavior analysis of EEG signals. The hybrid form of GCNs with Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) has also been used. To effectively capture both the spatial and temporal behavior of EEG signals, some studies have proposed multistream neural network architectures. For instance, a two-stream neural network was introduced [92], where one branch employs a graph convolutional neural network for spatial feature extraction and the other branch uses a channelwise convolutional neural network for temporal feature understanding. The proposed model was evaluated on two public datasets, namely, the Stanford University dataset (SU DB) (available at https://purl.stanford.edu/bq914sc3730) and the Max Planck Institute dataset (MPI DB) (available at https://www.mpib-berlin.mpg.de/research-data/data-sets). In addition, Liu *et al.* [75] proposed the use of multiview spatial-temporal graph convolutional networks to extract spatial-temporal features on the ISRUC-S3 (available at https://sleeptight.isr.uc.pt/) and MASS-SS3 (available at http://massdb.herokuapp.com/en/) datasets. Furthermore, a dynamical graph convolutional neural network was proposed by Li *et al.* [93] to perform emotion recognition using multichannel EEG signals on the SJTU emotion EEG dataset (SEED) (available at https://bcmi.sjtu.edu.cn/home/seed/) and DREAMER (available at https://zenodo.org/record/546113#.ZEtoPHZBxD8) dataset. The hierarchical features of neural networks were leveraged by the GCN proposed in Awais *et al.* [94], while the study by Liu *et al.* [95] used EEG signals as a graph based on within-frequency and cross-frequency data. Additionally, regularized graph neural networks [58] and self-organized graph neural networks [76] have been used for EEG classification on the SEED and SEED-IV datasets.

In addition to the previously mentioned studies, a recent study [81] explored the use of unsupervised graph convolutional networks for EEG signal analysis. The authors utilized an entropy-based dynamic graph embedding model to cluster the graphs. Raeisi *et al.* [62] proposed a self-supervised GCN for seizure detection and used the public Temple University Hospital EEG Seizure Corpus (TUSZ) v1.5.2 dataset (available at https://isip.piconepress.com/projects/tuh_eeg/) in their experiments. Neural attention is another widely used technique in deep learning that enables a neural network to selectively focus on specific parts of its input while making decisions or predictions. In the context of EEG signal classification, a previous Jia *et al.* [65] introduced a graph-based convolutional recurrent attention model that explores EEG features across different subjects for motor imagery classification. This GCN network is designed to gather positioning information from EEG nodes, while the convolutional recurrent attention model learns EEG features from both spatial and temporal dimensions. Zhang *et al.* [96] proposed a graph attention network that incorporates an attention-based layer into a graph neural network to identify the brain regions (associated with individual electrodes) that are most critical for accurate classification. This study conducted the experiment on their own dataset.

To conduct a comparative analysis of various approaches used for EEG analysis, we excluded methods based on a VG that rely on statistical features of the signal and use classical Euclidean-based techniques. Instead, we only considered the graph learning approach in Table II, along with its reported advantages and disadvantages from multiple sources.

The comparative analysis performed in Table II clearly indicates that (A) there is a lack of guidelines for large datasets, (B) most of the recent studies carried out fluctuate over their accuracy results, and (C) they lack cross-subject EEG variations.

For the empirical study, the performance of the VG approach is excluded because it is based only on single-channel EEG signals. From the comparative study, we observe that GNNs for EEG signal analysis have used different EEG datasets according to various application areas. This clearly indicates that the performances of graph-based EEG signal analysis models are not consistent. The deviation is very high depending on the architectures and datasets. In addition, the proportions of different graphs in these works can be seen in Fig. 5.

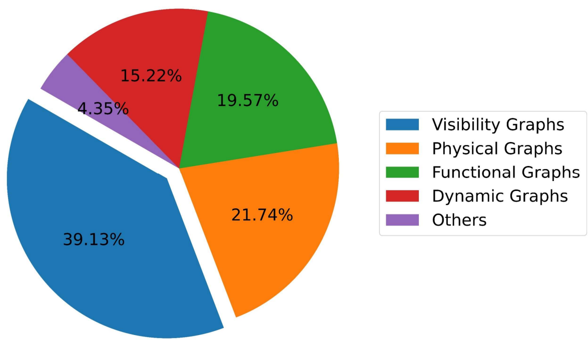

Fig. 5. Graphs involved in EEG signal analysis.

TABLE II. COMPARATIVE STUDY

Proposed Methods	Pros/Cons
Method: Machine learning methods (Linear SVM, LDA, Logistic Regression, and KNN) [99] **Dataset**: EEG Dataset (n =18) **Accuracy**: 97.2% **AUC Score**: 0.90	**Pros**: Proposed graph-to-signal transform method to quantify the strength of relationship with weighted functional connectivity of brain network. It preserves the Euclidean distance between the channels. **Cons**: Tested on only small-scale dataset. No specific guidelines to implement on large-scale dataset.
Method: Directed differential connectivity graphs (dDCGs) using KNN [99] **Dataset**: Own dataset (n = 26+23) **Accuracy**: 92%	**Pros**: Proposed a directed functional network for the classification of major depressive disorder patients. Found these patients have more random brain network in compared to normal person. **Cons**: Tested on only small-scale dataset. No specific guidelines to generalize it on large scale dataset.
Method: GCNN [99] **Dataset**: Fatigue driving recognition dataset (n=600) **Accuracy**: 97.38%	**Pros**: Combine classical deep network and GNN for fatigue driving detection. Created functional brain network based on singular value entropy and fractal dimension. GNN based on B-spline curve spatial convolution. **Cons**: Small dataset. No specific guidelines to generalize it on large scale dataset. Use of conventional Euclidean for feature extraction.
Method: Partial directed coherence graph convolutional neural network (PDC-GCNN) [99] **Dataset**: Own dataset (n=600) **Accuracy**: 96.01%	**Pros**: Implemented PDC for relationship between electrodes. Implemented GCNN for local electrode area information and connection information. **Cons**: Rare chance of model reproduction because of generalization ability.: Implemented PDC for relationship between electrodes. Implemented GCNN for local electrode area information and connection information. Cons: Rare chance of model reproduction because of generalization ability.
Method: Correlation based GCN [99] **Dataset**: The PhysioNet Dataset (n=64 and **accuracy**: 93.06%) and The High Gamma Dataset (n=44) **Accuracy**: 96.24%	**Pros**: Proposed a graph convolutional neural network model on Pearson's correlation based graph for task classification. **Cons**: Tested on moderate datasets and model performance is decreasing as the sample size increasing.
Method: GNN [99] **Dataset**: Error-related Potentials (ErrP) and Rapid Serial Visual Presentation (RSVP) **Accuracies**: 76.73% and 93.49%	**Pros**: Implemented GNN based model for EEG signal classification. Helpful for electrodes selection to reduce computational cost. Implemented different strategies for edge formation in EEG graph. **Cons**: Models are tested on very small datasets, so lack generalization ability. Model performance decreases very fast as the sample size decreases.
Method: EEG-GCNN [99] **Dataset**: Temple University Hospital EEG (TUH EEG) (n= 1385) and Max Planck Institute Leipzig MindBrainBody (MPI LEMON) (n= 208) **Accuracy**: 0.85% https://github.com/ neerajwagh/eeg-gcnn	**Pros**: GCN based on spatial and functional graph for neurological disease classification model. Incorporated large scale datasets **Cons**: Although it has incorporated local temporal dependencies through graph, it assumed signals are temporally independent. Implemented on two different dataset for normal and abnormal patients.
Method: GCN [99] **Dataset**: Alzheimer's disease dataset (own) (n=20+20) **Accuracy**: 91.996%	**Pros**: Implemented GNN based model for EEG signal classification for Alzheimer's Disease. Discussed 8 different functional connectivity based graph. **Cons**: Small dataset Generalization ability Physical topology of brain regions
Method: GCN [61] **Dataset**: CHB-MIT (n=18) **Accuracy**: 96.51%	**Pros**: GCN for seizure classification. Features are combined to reduce the computational cost. **Cons**: Small dataset No specific guidelines to generalize it on large scale dataset. Details on computational cost
Method: GCN+CNN [57] **Dataset**: Stanford University dataset (SU DB) (n=10,) and Max-Planck Institute dataset (MPI DB) (n=24) **Accuracies**: 54.28 and 84.40%	**Pros**: GCN+CNN with spatiotemporal graph and channelwise convolution to classification visual stimuli from EEG signals. **Cons**: Small datasets Use of CNN for feature extraction. Use of wPLI measure the functional connectivity between channel.
Method: Multiview spatialtemporal GCN (MSTGCN) [75] **Dataset**: ISRUC-S3 (n=10) and MASS-SS3 (n=62) **Accuracies**: 82.10 and 89.5% https://github.com/ziyujia/ MSTGCN	**Pros**: GCN for Spatial feature extraction and temporal convolution for transition rule in different sleep stages and apply these for sleep stage classification. **Cons**: Adequate detail Moderate sample sizes Different graphs variants are discussed but physical interpretation is not sufficient.
Method: Dynamic graph convolutional neural network [82] **Dataset**: SJTU emotion EEG dataset (SEED) (n=15) and DREAMER (n=23) **Accuracies**: 90.4 and 84.54%	**Pros**: GCN to adaptively learn the intrinsic relationship among the EEG channels by training neural network and use it to learn more discriminative features for emotion recognition. **Cons**: The proposed method is implemented on comparatively small scale dataset.
Method: GCN with hierarchical feature [95] **Dataset**: SJTU emotion EEG dataset (SEED) (n=15) and DREAMER (n=23) **Accuracies**: 94.24 and 89.32%	**Pros**: Proposed a graph convolutional broad network for deeper information of graph structured EEG-data for emotion recognition. **Cons**: The model is not generalizable as it is tested on very small dataset.
Method: GCN for diverse frequency of EEG Signal for person identification [70] **Dataset**: Multiple datasets and comparisons **Accuracy**: varies	**Pros**: GCN for biometric identification using EEG collected from diverse set of task Address the issues on using univariate signal **Cons**: Usually, person identification requires multisessions of EEG signals but performance is tested on single session Model is unsure about large-scale dataset.
Method: Regularized graph neural network [63] **Dataset**: SEED (n=15) and SEED-IV (n=15) **Accuracies**: 94.24% and 79.37% https://github.com/ zhongpeixiang/RGNN	**Pros**: GCN for emotion recognition. For cross subject EEG variations it incorporated novel concept of regularizes: nodewise domain adversarial training (NodeDAT) and emotion-aware distribution learning (EmotionDL) **Cons**: Tested on only small scale dataset No specific guidelines to implement on large-scale dataset.
Method: Unsupervised GCN [83] **Dataset**: CHB-MIT scalp EEG database (n=23) **ROC-AUC Score**: 97.1 F1 Score: 45.0	**Pros**: Proposed a temporal graph convolutional network to incorporate temporal and structural information with only few parameters **Cons**: Method is implemented only on small temporal window (96 seconds) but real life data might be much lengthier.
Method: Self-supervised graph neural networks [62] **Dataset**: Temple University Hospital EEG Seizure Corpus (TUSZ) v1.5.2 (n= 5,499) **Accuracy**: 74%	**Pros**: GCN for seizure detection. Self-supervised pretraining predicts signals for the next time period for model performance improvement. Implemented on large dataset. **Cons**: Model perform low accuracy. Self-supervised model for labelled data. Physical interpretation of the model is limited.
Method: Graph-based Convolutional Recurrent Attention Model (G-CRAM) [65] **Dataset**: BCI Competition https://www.bbci.de/ competition/ and PhysioNet EEG https://physionet.org/content/?topic=eeg **Accuracies**: 74.71% and 60.11%	**Pros**: GCN spatiotemporal graph attention for motor imagery classification New concept of recurrent attention is applied to value the most distinguishable time period. **Cons**: Tested on moderate datasets Model performance is average. Gave theoretical foundation to robustness to artifact but real implementation not provided.
Method: Graph attention network [67] **Dataset**: Own dataset (n=8) **ROC-AUC**: 79.56	**Pros**: Implemented GNN with attention mechanism for unsupervised seizure localization. Use graph for automatic detection of the brain region on localizing seizure onsets. **Cons**: Lacks generalization ability as the model is tested on very small sample. Provided the experimental comparison on threshold for edge scarification but did not provide theoretical framework.

Understanding and implementing an EEG signal analysis system is a complex task for new researchers because the programmer first needs a strong foundation for the graph and its connectivity, a lack of labelled datasets and less availability of code materials. Second, appropriate features and evaluation metrics are needed. The field of EEG signal analysis encompasses a diverse array of performance metrics, each tailored to specific applications such as regression, classification, detection, and segmentation. In this regard, Botchkarev's [97] work serves as a comprehensive reference, providing valuable insights applicable across various domains.

Moreover, the realm of feature selection in EEG signal analysis remains a fertile ground for research, characterized by a lack of universal guidelines or established rules of thumb. While efforts have been made in this area, researchers are encouraged to explore and develop their own approaches to feature engineering, as in the work of Naser *et al.* [98]. This presents an exciting opportunity for innovative contributions and advancements in the field.

Therefore, in this paper, we tried to provide a list of publicly available codes for EEG signal analysis based on GNNs. Table II lists some articles, the datasets they have experimented with and GitHub repository links for reproduction.

VI. Datasets

In this section, we provide a brief summary of commonly used EEG datasets generated by GNNs. The reference indicates the paper that was implemented on these datasets.

- **MPI LEMON EEG Dataset [99]**

The MPI LEMON dataset is a collection of EEG recordings from 216 healthy participants from Leipzig, Germany, consisting of two age groups: young adults (aged 20–35) and older adults (aged 59–77). The EEG recordings were taken using 62 electrodes in the 10-10 sensor configuration with a sampling rate of 2,500 Hz. Each participant completed 16 trials, 8 with their eyes closed and 8 with their eyes open, each lasting 60 seconds

- **THE Datasets [60]**

The TUH EEG dataset is a large collection of more than 30,000 EEG recordings collected at Temple University Hospital (TUH) since 2002. These recordings vary in terms of patient ages, diagnoses, medications, channel configurations, and sampling frequencies. However, a subset of the recordings in TUH EEG have been expertly annotated as either "normal" or "abnormal" and have been released as a derived dataset called the TUH EEG Abnormal Corpus (TUAB).

- **MASS-SS3 [75]**

The Montreal Archive of Sleep Studies (MASS) is a publicly available database that includes sleep data from more than 2,000 participants. SS3, a subset of the MASS dataset, comprises PSG recordings from 62 healthy participants (28 males and 34 females), each of which includes 20 EEG channels, 2 EOG channels, 3 EMG channels, and 1 ECG channel, all of which are standard PSG signals. The dataset also provides sleep stage classifications for each recording, based on the AASM standard, which defines specific criteria for each of the five sleep stages (W, N1, N2, N3, and REM) based on observed brain waves, eye movements, and muscle tone. This standardization allows for consistent comparisons and analysis of PSG recordings. The dataset can be used by researchers and clinicians to create and test automated sleep stage classification algorithms, examine the physiological characteristics of sleep stages, and investigate the relationship between sleep and various health outcomes.

- **CHB-MIT Dataset [100]**

The Children's Hospital Boston (CHB) and the Massachusetts Institute of Technology (MIT) created the CHB-MIT Scalp EEG Database, which is a compilation of EEG recordings from 23 individuals experiencing uncontrollable seizures. The EEG recordings were acquired using scalp electrodes based on the international 10–20 system, a standardized approach for electrode placement in EEG recordings. The dataset consists of extended EEG recordings sampled at 256 Hz, which are marked with the beginning and end times of seizure activity. The dataset is publicly accessible and widely utilized for scientific investigations involving the prediction and identification of seizures.

- **The PhysioNet Dataset [70]**

The EEG Motor Movement/Imagery Dataset is a collection of more than 1,500 EEG records from 109 subjects. The dataset includes 64 electrodes based on the international 10–10 system, and each subject performed 84 trials, including 3 runs, 7 trials per run, and 4 tasks per trial related to motor imagery. The tasks are identified as L, R, B, and F, which involve imagining the movement of the left fist, right fist, both fists, and both feet, respectively. The EEG signals were sampled at a rate of 160 Hz, and each signal had a duration of 4 s, which translates to 640 time points per trial. This dataset can be useful for researchers studying motor imagery and its impact on brain activity and for developing algorithms and techniques for analysing EEG signals.

- **The PhysioNet Dataset [70]**

The EEG Motor Movement/Imagery Dataset is a collection of more than 1,500 EEG records from 109 subjects. The dataset includes 64 electrodes based on the international 10–10 system, and each subject performed 84 trials, including 3 runs, 7 trials per run, and 4 tasks per trial related to motor imagery. The tasks are identified as L, R, B, and F, which involve imagining the movement of the left fist, right fist, both fists, and both feet, respectively. The EEG signals were sampled at a rate of 160 Hz, and each signal had a duration of 4 s, which translates to 640 time points per trial. This dataset can be useful for researchers studying motor imagery and its impact on brain activity and for developing algorithms and techniques for analysing EEG signals.

- **The High Gamma Dataset [101]**

The High Gamma Dataset, gathered from 14 participants, consists of data from four EEG tasks that pertain to movement and rest, specifically involving the left hand, right hand, both feet, and rest. The EEG signals were captured using 44 electrodes and had a frequency range between 0–125 Hz.

- **TUSZ v1.5.2 [66]**

The Temple University Hospital EEG Seizure Corpus (TUSZ) v1.5.2 is the largest publicly available database of EEG seizures, containing 5,612 EEG recordings, including 3,050 clinically annotated seizures representing eight different seizure types. The corpus includes recordings from both pediatric and adult patients with epilepsy, and the EEG activity was recorded using various EEG systems and sampling rates. The TUSZ corpus is publicly available for download from the International Epilepsy Electrophysiology Portal (IEEG Portal) and has been used in numerous studies aimed at developing new methods for seizure detection and classification and investigating the neural mechanisms underlying seizures.

- **RSVP dataset [102]**

The Rapid Serial Visual Presentation (RSVP) dataset is a collection of EEG recordings from 10 healthy subjects that were collected to develop a Brain-Computer Interface (BCI) system based on the Rapid Serial Visual Presentation (RSVP) paradigm. The dataset contains a total of 41,400 trials of 16-channel EEG data, which were recorded using a g.USBamp biosignal amplifier with active electrodes. The trials were recorded during RSVP keyboard operations and were associated with one of four labels: emotion elicitation, resting-state, or motor imagery/execution tasks. Each trial has 128 time samples, and the dataset is publicly available for download through a repository.

- **ErrP dataset [103]**

The ErrP dataset is a collection of EEG recordings from 16 healthy subjects that are aimed at detecting error-related potentials (ErrP) to improve the accuracy of P300-based Brain-Computer Interfaces (BCIs). The dataset was recorded during an offline P300 spelling task, where each subject performed 340 trials using a fast mode (each item was flashed 4 times) or a slow mode (each item was flashed 8 times). The goal was to detect the ErrP when there was an inconsistency between the subject's intention and the BCI system. The EEG data were recorded from 56 channels and downsampled to a rate of 200 Hz. The dataset is available for download through a publicly accessible repository and has been used to develop and evaluate various methods for detecting ErrP, which can be used to improve the accuracy of P300-based BCI spellers.

- **The Stanford University EEG dataset [104]**

The Stanford University EEG dataset includes a total of 12 EEG recordings from 12 patients with epilepsy that were recorded during a 10-minute period while the patient was experiencing a seizure. Each recording in the dataset consists of 23 channels of EEG data sampled at 500 Hz. The dataset is publicly available for research purposes and includes information on the type and location of the seizure. It has been widely used in studies related to the diagnosis and treatment of epilepsy, as well as for developing algorithms for automatic seizure detection and exploring the neural mechanisms underlying seizures.

VII. Challenges and Future Research Scopes

EEG signal analysis using GNNs has many constraints and has much open space. Some of the most prominent challenges and future research directions are discussed here.

EEG signals are temporal and depend on the position of the electrode that captures brain activity. Typically, 64 or more electrodes are used to capture these signals [105]. However, in practical applications, only signals from a few electrodes are utilized in experiments for simplification. It is a challenging task to identify the most influential EEG signal for accurate analysis.

Graph learning is a widely interdisciplinary research field with numerous applications, such as Recommender System [106, 107], Neutrino Detection [108], Large Hadron Collider (LHC) [109], Fake News Detection [110], Drug Repurposing [111], Particle Physics and Chemistry [108], Computer Vision and Graphics [107, 112], Robotics and Autonomous Driving [113], Medicine [60], and Drug Design [114]. However, each application requires specific domain knowledge outside the realm of graph learning, and interpreting the graph data structure for these diverse fields of study can be challenging. As a result, there may be a gap between programmers and domain experts when developing a system for EEG signal analysis.

The original EEG signal cannot be processed directly using ML technology because of its high dimensionality. Therefore, several feature extraction methods, such as Time Frequency Distribution (TFD), Fourier Transform (FT), the Autoregressive Method (ARM), Eigenvector Methods (EMs), and Wavelet Transform (WT), have been proposed [115]. However, selecting the optimal feature extraction method for a specific problem can be a challenge.

Various forms of Visibility Graphs (VGs) have been used for analysing short temporal signals in EEGs. VGs can map time series data to scale-free graphs, enabling the capture of long-term dependencies and chaotic properties [38] However, applying VGs to multichannel EEG signals even for an average time signal can result in billions of nodes and edges for a single subject. Learning such colossal graphs poses significant challenges, such as bottlenecks and oversmoothing [116]. Therefore, reducing the size of VGs is an additional challenge to consider.

Currently, dynamic graph classification remains an open problem due to scalability issues with large temporal graph datasets. While dynamic graph convolutional neural networks have been used for temporal data, most applications involve link prediction tasks involving graph snapshots [117, 118], dynamic link prediction [28, 119], and node classification [120, 121] in dynamic graphs. However, we are not aware of any successful applications of dynamic graph classification at this time.

To extract features for deep learning, conventional methods involve reshaping data into vectors or matrices and performing principal component analysis, independent component analysis, and other similar techniques [122]. However, these methods do not preserve the original structure of the data. Currently, selecting attributes and classifying signals are intricate tasks, particularly when dealing with EEG signals that are more complex than other medical signals and images. In fact, even expert clinicians can correctly identify only 50% of abnormal signals as abnormal [123]. Thus, generating EEG signals at the local level is exceptionally challenging. On the other hand, deep learning models are data-hungry and require large-scale datasets [124].

Interpreting EEG signals is a challenging task, and using graph ML and DL techniques requires significant effort to achieve successful results. Therefore, it is crucial to have references to guide the analysis of EEG signals using graph learning methods. Table II provides a list of publicly available repositories with codes for reproduction, but for new researchers entering the field of EEG signal analysis, this can still be a daunting challenge.

VIII. CONCLUSION

This paper explores the mathematical approaches of graph representation for EEG signal analysis. We conduct a comparative and empirical analysis of GNN methods based on EEG signals. This survey of graph representation in EEG signal analysis from 2012 revealed that most research was conducted on small datasets, and there were no specific guidelines for generalizing for larger datasets to adopt a data-driven approach.

The survey also highlighted several challenges in the field, including the selection of the most important EEG signal for analysis and the extraction of appropriate features for graph learning. We concluded that domain knowledge and appropriate graph interpretation are essential for this research domain. Research on EEG signal analysis is in the early stage; hence, we identified several opportunities, such as the use of visibility graphs and dynamic graphs and the development of large labelled datasets for deep learning approaches. Visibility graphs offer an excellent way to represent temporal data in a scale-free manner, but our survey shows that they have only been used for single-channel EEG signals. The results using existing datasets can be improved in two different approaches: first, in graph representation, and second, in graph model training. Multichannel EEG involves the incorporation of a variety of information for EEG signal analysis. This approach is lacking for existing visibility graphs and dynamic graph representations. However, dealing with dynamic graphs is not cost efficient, so multichannel EEGs on visibility graphs may be one possible way to leverage more information.

As the existing methods either address small datasets or large datasets with only a few distinct datasets, the latter can be improved by including EEG signals from multiple subjects. Therefore, incorporating multichannel EEG signals into VGs could be an exciting area for future research.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest to report regarding the present study.

AUTHOR CONTRIBUTIONS

Conceptualization, HCB, and SA; methodology, SJ,and YRP; software, YRP and KJ; validation, SA, HCB, and SJ; writing–original draft preparation, HCB, SA and SJ; writing–review and editing, SJ, KJ; visualization, YRP and KJ ; supervision, SA and SJ.; project administration, SJ, SA; funding, HCB, YRP, KJ, SJ. All authors had approved the final version.

FUNDING

This study is supported via funding from Prince Sattam Bin Abdulaziz University project number (PSAU/2024/R/1445).

ACKNOWLEDGMENT

We thank the Prince Sattam Bin Abdulaziz University, Alkharj, Saudi Arabia for help and support.

REFERENCES

[1] World Health Organization, *Neurological Disorders: Public Health Challenges*, 2006.

[2] V. L. Feigin, "The evolution of neuroepidemiology: Marking the 40-year anniversary of publishing studies on epidemiology of neurological disorders," *Neuroepidemiology*, vol. 56, no. 1, pp. 2–3, 2022.

[3] P. L. Nunez and R. Srinivasan, *Electric Fields of the Brain: The Neurophysics of EEG*, Oxford University Press, USA, 2006.

[4] C. Yen, C.-L. Lin, and M.-C. Chiang, "Exploring the frontiers of neuroimaging: A review of recent advances in understanding brain functioning and disorders," *Life*, vol. 13, no. 7, 1472, 2023.

[5] K. Natarajan, R. Acharya U, F. Alias, T. Tiboleng, and S. K. Puthusserypady, "Nonlinear analysis of EEG signals at different mental states," *Biomedical Engineering Online*, vol. 3, pp. 1–11, 2004.

[6] F. Wendling, P. Chauvel, A. Biraben, and F. Bartolomei, "From intracerebral EEG signals to brain connectivity: Identification of epileptogenic networks in partial epilepsy," *Frontiers in Systems Neuroscience*, vol. 4, 154, 2010.

[7] J. Zhou, G. Cui, S. Hu *et al.*, "Graph neural networks: A review of methods and applications," *AI Open*, vol. 1, pp. 57–81, 2020.

[8] C. Gao, Y. Zheng, N. Li *et al.*, "A survey of graph neural networks for recommender systems: Challenges, methods, and directions," *ACM Transactions on Recommender Systems*, vol. 1, no. 1, pp. 1–51, 2023.

[9] D. Ahmedt-Aristizabal, M. A. Armin, S. Denman, C. Fookes, and L. Petersson, "Graph-based deep learning for medical diagnosis and analysis: Past, present and future," *Sensors*, vol. 21, no. 14, 4758, 2021.

[10] I. Stancin, M. Cifrek, and A. Jovic, "A review of EEG signal features and their application in driver drowsiness detection systems," *Sensors*, vol. 21, no. 11, 3786, 2021.

[11] S. Barua and S. Begum, "A review on machine learning algorithms in handling EEG artifacts," in *Proc. The Swedish AI Society (SAIS) Workshop SAIS*, Stockholm, Sweden, May 2014.

[12] A. R. Mane, S. Biradar, and R. Shastri, "Review paper on feature extraction methods for EEG signal analysis," *Int. J. Emerg. Trend Eng. Basic Sci*, vol. 2, no. 1, pp. 545–552, 2015.

[13] A. Shoka, M. Dessouky, A. El-Sherbeny, and A. El-Sayed, "Literature review on EEG preprocessing, feature extraction, and classifications techniques," *Menoufia J. Electron. Eng. Res*, vol. 28, no. 1, pp. 292–299, 2019.

[14] J. Kaur and A. Kaur, "A review on analysis of EEG signals," in *Proc. the 2015 International Conference on Advances in Computer Engineering and Applications*, 2015, pp. 957–960.

[15] J. S. Kumar and P. Bhuvaneswari, "Analysis of Electroencephalography (EEG) signals and its categorization–a study," *Procedia Engineering,* vol. 38, pp. 2525–2536, 2012.

[16] T. Negi, "Analysis and processing of EEG signal: A review," in *Proc. the International Conference on Smart Data Intelligence (ICSMDI 2021)*, 2021.

[17] J. Reaves, T. Flavin, B. Mitra, K. Mahantesh, and V. Nagaraju, "Assessment and application of EEG: A literature review," *J. Appl. Bioinforma. Comput. Biol.*, vol. 7, no. 2, 2021.

[18] T. Alotaiby, F. E. A. El-Samie, S. A. Alshebeili, and I. Ahmad, "A review of channel selection algorithms for EEG signal processing," *EURASIP Journal on Advances in Signal Processing*, vol. 2015, pp. 1–21, 2015.

[19] B. Li, T. Cheng, and Z. Guo, "A review of EEG acquisition, processing and application," *Journal of Physics: Conference Series*, 012045, IOP Publishing, 2021.

[20] F. Lotte *et al*., "A review of classification algorithms for EEG-based brain–computer interfaces: A 10 year update," *Journal of Neural Engineering*, vol. 15, no. 3, 031005, 2018.

[21] M.-P. Hosseini, A. Hosseini, and K. Ahi, "A review on machine learning for EEG signal processing in bioengineering," *IEEE Reviews In Biomedical Engineering*, vol. 14, pp. 204–218, 2020.

[22] S. Dhivya and A. Nithya, "A review on machine learning algorithm for EEG signal analysis," in *Proc. the 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 2018, pp. 54–57.

[23] D. M. Praveena, D. A. Sarah, and S. T. George, "Deep learning techniques for EEG signal applications—A review," *IETE Journal of Research*, vol. 68, no. 4, pp. 3030–3037, 2022.

[24] K. Rasheed, A. Qayyum, J. Qadir *et al*., "Machine learning for predicting epileptic seizures using EEG signals: A review," *IEEE Reviews in Biomedical Engineering*, vol. 14, pp. 139–155, 2020.

[25] G. Li, C. H. Lee, J. J. Jung, Y. C. Youn, and D. Camacho, "Deep learning for EEG data analytics: A survey," *Concurrency and Computation: Practice and Experience*, vol. 32, no. 18, e5199, 2020.

[26] J. Vandana and N. Nirali, "A review of EEG signal analysis for diagnosis of neurological disorders using machine learning," *Journal of Biomedical Photonics & Engineering*, vol. 7, no. 4, 40201, 2021.

[27] T. N. Kipf and M. Welling, "Semisupervised classification with graph convolutional networks," arXiv preprint, arXiv:1609.02907, 2016.

[28] S. M. Kazemi, R. Goel, K. Jain *et al.*, "Representation learning for dynamic graphs: A survey," *The Journal of Machine Learning Research*, vol. 21, no. 1, pp. 2648–2720, 2020.

[29] E. Rossi, B. Chamberlain, F. Frasca, D. Eynard, F. Monti, and M. Bronstein, "Temporal graph networks for deep learning on dynamic graphs," arXiv preprint, arXiv:2006.10637, 2020.

[30] P. Mercado, F. Tudisco, and M. Hein, "Spectral clustering of signed graphs via matrix power means," in *Proc. the International Conference on Machine Learning*, 2019, pp. 4526–4536.

[31] D. Zhou, J. Huang, and B. Schölkopf, "Learning with hypergraphs: Clustering, classification, and embedding," *Advances in Neural Information Processing Systems*, vol. 19, 2006.

[32] M. Hein, S. Setzer, L. Jost, and S. S. Rangapuram, "The total variation on hypergraphs-learning on hypergraphs revisited," *Advances in Neural Information Processing Systems*, vol. 26, 2013.

[33] Z. Liu, C. Chen, X. Yang, J. Zhou, X. Li, and L. Song, "Heterogeneous graph neural networks for malicious account detection," in *Proc. the 27th ACM International Conference On Information And Knowledge Management*, 2018, pp. 2077–2085.

[34] X. Wang, D. Bo, C. Shi, S. Fan, Y. Ye, and S. Y. Philip, "A survey on heterogeneous graph embedding: Methods, techniques, applications and sources," *IEEE Transactions on Big Data*, vol. 9, no. 2, pp. 415–436, 2022.

[35] X. Z. Fern and C. E. Brodley, "Solving cluster ensemble problems by bipartite graph partitioning," in *Proc. the Twenty-First International Conference on Machine Learning*, 2004.

[36] X. He, M. Gao, M.-Y. Kan, and D. Wang, "Birank: Towards ranking on bipartite graphs," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 1, pp. 57–71, 2016.

[37] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nuno, "From time series to complex networks: The visibility graph," in *Proc. the National Academy of Sciences*, 2008, vol. 105, no. 13, pp. 4972–4975.

[38] J. Wu, K. Xu, X. Chen, S. Li, and J. Zhao, "Price graphs: Utilizing the structural information of financial time series for stock prediction," *Information Sciences*, vol. 588, pp. 405–424, 2022.

[39] M. Ahmadlou, H. Adeli, and A. Adeli, "Improved visibility graph fractality with application for the diagnosis of autism spectrum disorder," *Physica A: Statistical Mechanics and its Applications*, vol. 391, no. 20, pp. 4720–4726, 2012.

[40] H. Ji, T. Xu, W. Wu, and J. Wang, "Visibility graph analysis on EEG signal," in *Proc. the 2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, 2016, pp. 1557–1561.

[41] C. Hao, Z. Chen, and Z. Zhao, "Analysis and prediction of epilepsy based on visibility graph," in *Proc. the 2016 3rd International Conference on Information Science and Control Engineering (ICISCE)*, 2016, pp. 1271–1274.

[42] B. Luque, L. Lacasa, F. Ballesteros, and J. Luque, "Horizontal visibility graphs: Exact results for random time series," *Physical Review E*, vol. 80, no. 4, 046103, 2009.

[43] G. Zhu, Y. Li, and P. P. Wen, "An efficient visibility graph similarity algorithm and its application on sleep stages classification," in *Proc. the International Conference on Brain Informatics, BI 2012*, Macau, China, 2012, pp. 185–195.

[44] M. Ahmadlou and H. Adeli, "Visibility graph similarity: A new measure of generalized synchronization in coupled dynamic systems," *Physica D: Nonlinear Phenomena*, vol. 241, no. 4, pp. 326–332, 2012.

[45] E. Pereda, R. Rial, A. Gamundi, and J. Gonzalez, "Assessment of changing interdependencies between human electroencephalograms using nonlinear methods," *Physica D: Nonlinear Phenomena*, vol. 148, no. 1–2, pp. 147–158, 2001.

[46] A. Sengupta, A. Routray, and S. Kar, "Complex brain networks using visibility graph synchronization," in *Proc. the 2013 Annual IEEE India Conference (INDICON)*, 2013, pp. 1–4.

[47] N. Ahmadi and M. Pechenizkiy, "Application of horizontal visibility graph as a robust measure of neurophysiological signals synchrony," in *Proc. the 2016 IEEE 29th International Symposium on Computer-Based Medical Systems (CBMS)*, 2016, pp. 273–278.

[48] G. Zhu, Y. Li, and P. Wen, "Analysis and classification of sleep stages based on difference visibility graphs from a single-channel EEG signal," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 6, pp. 1813–1821, 2014.

[49] L. Wang, X. Long, J. B. Arends, and R. M. Aarts, "EEG analysis of seizure patterns using visibility graphs for detection of generalized seizures," *Journal of Neuroscience Methods*, vol. 290, pp. 85–94, 2017.

[50] G. Zhu, Y. Li, and P. P. Wen, "Epileptic seizure detection in EEGs signals using a fast weighted horizontal visibility algorithm," *Computer Methods and Programs in Biomedicine*, vol. 115, no. 2, pp. 64–75, 2014.

[51] S. Supriya, S. Siuly, and Y. Zhang, "Automatic epilepsy detection from EEG introducing a new edge weight method in the complex network," *Electronics Letters*, vol. 52, no. 17, pp. 1430–1432, 2016.

[52] Z. Mohammadpoory, M. Nasrolahzadeh, and J. Haddadnia, "Epileptic seizure detection in EEGs signals based on the weighted visibility graph entropy," *Seizure*, vol. 50, pp. 202–208, 2017.

[53] L. Cai, B. Deng, X. Wei, R. Wang, and J. Wang, "Analysis of spontaneous EEG activity in Alzheimer's disease using weighted visibility graph," in *Proc. the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018, pp. 3100–3103.

[54] J. Wang, C. Yang, R. Wang, H. Yu, Y. Cao, and J. Liu, "Functional brain networks in Alzheimer's disease: EEG analysis based on limited penetrable visibility graph and phase space method," *Physica A: Statistical Mechanics and its Applications*, vol. 460, pp. 174–187, 2016.

[55] Z.-K. Gao, Q. Cai, Y.-X. Yang, W.-D. Dang, and S.-S. Zhang, "Multiscale limited penetrable horizontal visibility graph for analysing nonlinear time series," *Scientific Reports*, vol. 6, no. 1, 35622, 2016.

[56] Z.-K. Gao, Q. Cai, Y.-X. Yang, N. Dong, and S.-S. Zhang, "Visibility graph from adaptive optimal kernel time-frequency representation for classification of epileptiform EEG," *International Journal of Neural Systems*, vol. 27, no. 04, 1750005, 2017.

[57] S. Supriya, S. Siuly, H. Wang, and Y. Zhang, "EEG sleep stages analysis and classification based on weighed complex network features," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 5, no. 2, pp. 236–246, 2018.

[58] K. Samanta, S. Chatterjee, and R. Bose, "Cross-subject motor imagery tasks EEG signal classification employing multiplex weighted visibility graph and deep feature extraction," *IEEE Sensors Letters*, vol. 4, no. 1, pp. 1–4, 2019.
[59] Z.-K. Gao, W. Guo, Q. Cai, C. Ma, Y.-B. Zhang, and J. Kurths, "Characterization of SSMVEP-based EEG signals using multiplex limited penetrable horizontal visibility graph," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 29, no. 7, 2019.
[60] Z. Lin, T. Qiu, P. Liu, L. Zhang, S. Zhang, and Z. Mu, "Fatigue driving recognition based on deep learning and graph neural network," *Biomedical Signal Processing and Control*, vol. 68, 102598, 2021.
[61] J. Kalafatovich, M. Lee, and S.-W. Lee, "Learning spatiotemporal graph representations for visual perception using EEG signals," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 97–108, 2022.
[62] K. Raeisi, 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," *Computer Methods and Programs in Biomedicine*, vol. 222, 106950, 2022.
[63] 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. the 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2021, pp. 1061–1067.
[64] N. Wagh and Y. Varatharajah, "EEG-GCNN: Augmenting electroencephalogram-based neurological disease diagnosis using a domain-guided graph convolutional neural network," in *Proc. the Machine Learning for Health*, 2020, pp. 367–378.
[65] M. Jia, W. Liu, J. Duan *et al.*, "Efficient graph convolutional networks for seizure prediction using scalp EEG," *Frontiers in Neuroscience*, vol. 16, 967116, 2022.
[66] S. Tang et al., "Self-supervised graph neural networks for improved electroencephalographic seizure analysis," arXiv preprint, arXiv:2104.08336, 2021.
[67] P. Zhong, D. Wang, and C. Miao, "EEG-based emotion recognition using regularized graph neural networks," *IEEE Transactions on Affective Computing*, vol. 13, no. 3, pp. 1290–1301, 2020.
[68] S. Cai, H. Li, Q. Wu, J. Liu, and Y. Zhang, "Motor imagery decoding in the presence of distraction using graph sequence neural networks," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 1716–1726, 2022.
[69] D. Zhang, K. Chen, D. Jian, and L. Yao, "Motor imagery classification via temporal attention cues of graph embedded EEG signals," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 9, pp. 2570–2579, 2020.
[70] Y. Hou, S. Jia, X. Lun *et al.*, "GCNs-Net: A graph convolutional neural network approach for decoding time-resolved eeg motor imagery signals," *IEEE Transactions on Neural Networks and Learning Systems*, 2022.
[71] D. Grattarola, L. Livi, C. Alippi, R. Wennberg, and T. A. Valiante, "Seizure localization with attention-based graph neural networks," *Expert Systems with Applications*, vol. 203, 117330, 2022.
[72] 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 Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 2651–2660, 2022.
[73] T. T. K. Munia and S. Aviyente, "Graph-to-signal transformation based classification of functional connectivity brain networks," *Plos One*, vol. 14, no. 8, e0212470, 2019.
[74] M. Wang, H. El-Fiqi, J. Hu, and H. A. Abbass, "Convolutional neural networks using dynamic functional connectivity for EEG-based person identification in diverse human states," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 12, pp. 3259–3272, 2019.
[75] H. Liu, J. Zhang, Q. Liu, and J. Cao, "Minimum spanning tree based graph neural network for emotion classification using EEG," *Neural Networks*, vol. 145, pp. 308–318, 2022.
[76] F. Hasanzadeh, M. Mohebbi, and R. Rostami, "Graph theory analysis of directed functional brain networks in major depressive disorder based on EEG signal," *Journal of Neural Engineering*, vol. 17, no. 2, 026010, 2020.
[77] W. Zhang, F. Wang, S. Wu, Z. Xu, J. Ping, and Y. Jiang, "Partial directed coherence based graph convolutional neural networks for driving fatigue detection," *Review of Scientific Instruments*, vol. 91, no. 7, 2020.
[78] F. Vecchio, F. Miraglia, P. Bramanti, and P. M. Rossini, "Human brain networks in physiological aging: A graph theoretical analysis of cortical connectivity from EEG data," *Journal of Alzheimer's Disease*, vol. 41, no. 4, pp. 1239–1249, 2014.
[79] Z. Jia, Y. Lin, J. Wang *et al.*, "Multiview spatial-temporal graph convolutional networks with domain generalization for sleep stage classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 1977–1986, 2021.
[80] J. Li, S. Li, J. Pan, and F. Wang, "Cross-subject EEG emotion recognition with self-organized graph neural network," *Frontiers in Neuroscience*, vol. 15, 611653, 2021.
[81] Z. Li, K. Hwang, K. Li, J. Wu, and T. Ji, "Graph-generative neural network for EEG-based epileptic seizure detection via discovery of dynamic brain functional connectivity," *Scientific Reports*, vol. 12, no. 1, 18998, 2022.
[82] W. Hamilton, Z. Ying, and J. Leskovec, "Inductive representation learning on large graphs," *Advances in Neural Information Processing Systems*, vol. 30, 2017.
[83] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph attention networks," arXiv preprint, arXiv:1710.10903, 2017.
[84] K. Xu, W. Hu, J. Leskovec, and S. Jegelka, "How powerful are graph neural networks?," arXiv preprint, arXiv:1810.00826, 2018.
[85] G. Li and J. J. Jung, "Seizure detection from multichannel EEG using entropy-based dynamic graph embedding," *Artificial Intelligence in Medicine*, vol. 122, 102201, 2021.
[86] T. Song, W. Zheng, P. Song, and Z. Cui, "EEG emotion recognition using dynamical graph convolutional neural networks," *IEEE Transactions on Affective Computing*, vol. 11, no. 3, pp. 532–541, 2018.
[87] I. C. Covert, B. Krishnan, I. Najm *et al.*, "Temporal graph convolutional networks for automatic seizure detection," in *Proc. the Machine Learning for Healthcare Conference*, 2019, pp. 160–180.
[88] B. Hosseinifard, M. H. Moradi, and R. Rostami, "Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal," *Computer Methods and Programs in Biomedicine*, vol. 109, no. 3, pp. 339–345, 2013.
[89] V. Doma and M. Pirouz, "A comparative analysis of machine learning methods for emotion recognition using EEG and peripheral physiological signals," *Journal of Big Data*, vol. 7, no. 1, pp. 1–21, 2020.
[90] Z. Mu, L. Jin, J. Yin, and Q. Wang, "Application of graph neural network in driving fatigue detection based on EEG signals," *Computational Intelligence and Neuroscience*, vol. 2022, 2022.
[91] J. Han, X. Wei, and A. A. Faisal, "EEG decoding for datasets with heterogenous electrode configurations using transfer learning graph neural networks," arXiv preprint, arXiv:2306.13109, 2023.
[92] Z. Ding, Y. Zhao, A. Li, and Z. Zheng, "Spatial–temporal attention two-stream convolution neural network for smoke region detection," *Fire*, vol. 4, no. 4, 66, 2021.
[93] G. Li, D. Ouyang, Y. Yuan *et al.*, "An EEG data processing approach for emotion recognition," *IEEE Sensors Journal*, vol. 22, no. 11, pp. 10751–10763, 2022.
[94] M. Awais, S. B. Belhaouari, and K. Kassoul, "Graph neural networks for neurological diseases detection in EEG signals," *Research Square Preprint*, 2023. https://doi.org/10.21203/rs.3.rs-3086079/v1
[95] S. Liu, C. M. Wong, X. Liu *et al.*, "Driving fatigue effects on cross-frequency phase synchrony embedding in multilayer brain network," *IEEE Transactions on Instrumentation and Measurement*, 2023.
[96] D. Zhang, L. Yao, K. Chen, S. Wang, P. D. Haghighi, and C. Sullivan, "A graph-based hierarchical attention model for movement intention detection from EEG signals," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 11, pp. 2247–2253, 2019.
[97] A. Botchkarev, "A new typology design of performance metrics to measure errors in machine learning regression algorithms," *Interdisciplinary Journal of Information, Knowledge, and Management*, vol. 14, pp. 045–076, 2019.
[98] M. Naser, "An engineer's guide to explainable artificial intelligence and interpretable machine learning: Navigating causality, forced

goodness, and the false perception of inference," *Automation in Construction*, vol. 129, 103821, 2021.

[99] A. Babayan, M. Erbey, D. Kumral *et al.*, "A mind-brain-body dataset of MRI, EEG, cognition, emotion, and peripheral physiology in young and old adults," *Scientific Data*, vol. 6, no. 1, pp. 1–21, 2019.

[100] A. H. Shoeb, "Application of machine learning to epileptic seizure onset detection and treatment," Doctoral theses, MIT Division of Health Sciences and Technology, Harvard University, 2009.

[101] B. Blankertz, K. R. Muller, D. J. Krusienski *et al.*, "The BCI competition III: Validating alternative approaches to actual BCI problems," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 153–159, 2006.

[102] S. Zhang, Y. Wang, L. Zhang, and X. Gao, "A benchmark dataset for RSVP-based brain–computer interfaces," *Frontiers in Neuroscience*, vol. 14, 568000, 2020.

[103] P. Margaux, M. Emmanuel, D. Sébastien, B. Olivier, and M. Jérémie, "Objective and subjective evaluation of online error correction during P300-based spelling," *Advances in Human-Computer Interaction*, vol. 2012, 2012.

[104] B. Kaneshiro, M. P. Guimaraes, H.-S. Kim, A. M. Norcia, and P. Suppes, "A representational similarity analysis of the dynamics of object processing using single-trial EEG classification," *Plos One*, vol. 10, no. 8, e0135697, 2015.

[105] J. Montoya-Martínez, J. Vanthornhout, A. Bertrand, and T. Francart, "Effect of number and placement of EEG electrodes on measurement of neural tracking of speech," *Plos One*, vol. 16, no. 2, e0246769, 2021.

[106] Z. Yu, C. Wang, J. Bu, X. Wang, Y. Wu, and C. Chen, "Friend recommendation with content spread enhancement in social networks," *Information Sciences*, vol. 309, pp. 102–118, 2015.

[107] F. Monti, D. Boscaini, J. Masci, E. Rodola, J. Svoboda, and M. M. Bronstein, "Geometric deep learning on graphs and manifolds using mixture model cnns," in *Proc. the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 5115–5124.

[108] N. Choma, F. Monti, L. Gerhardt *et al.*, "Graph neural networks for icecube signal classification," in *Proc. the 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2018, pp. 386–391.

[109] J. A. Martínez, O. Cerri, M. Spiropulu, J. Vlimant, and M. Pierini, "Pileup mitigation at the Large Hadron Collider with graph neural networks," *The European Physical Journal Plus*, vol. 134, no. 7, 2019.

[110] F. Monti, F. Frasca, D. Eynard, D. Mannion, and M. M. Bronstein, "Fake news detection on social media using geometric deep learning," arXiv preprint, arXiv:1902.06673, 2019.

[111] Y. Zhou, F. Wang, J. Tang, R. Nussinov, and F. Cheng, "Artificial intelligence in COVID-19 drug repurposing," *The Lancet Digital Health*, vol. 2, no. 12, pp. e667–e676, 2020.

[112] M. M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, and P. Vandergheynst, "Geometric deep learning: Going beyond euclidean data," *IEEE Signal Processing Magazine*, vol. 34, no. 4, pp. 18–42, 2017.

[113] F. P. Such *et al.*, "Robust spatial filtering with graph convolutional neural networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 6, pp. 884–896, 2017.

[114] X. Xia, J. Hu, Y. Wang, L. Zhang, and Z. Liu, "Graph-based generative models for de Novo drug design," *Drug Discovery Today: Technologies*, vol. 32, pp. 45–53, 2019.

[115] M.Z.Naser, "An engineer's guide to eXplainable artificial intelligence and interpretable machine learning: Navigating causality, forced goodness, and the false perception of inference," *Automation in Construction*, 129, 103821, 2021.

[116] U. Alon and E. Yahav, "On the bottleneck of graph neural networks and its practical implications," arXiv preprint, arXiv:2006.05205, 2020.

[117] R. Hisano, "Semisupervised graph embedding approach to dynamic link prediction," in *Proc. the 9th Conference on Complex Networks, CompleNet 2018*, 2018, pp. 109–121.

[118] N. M. A. Ibrahim and L. Chen, "Link prediction in dynamic social networks by integrating different types of information," *Applied Intelligence*, vol. 42, pp. 738–750, 2015.

[119] H. Kasat, S. Markan, M. Gupta, and V. Pudi, "Temporal link prediction in dynamic networks," in *Proc. the Mining and Learning on Graphs workshop*, 2019.

[120] C. Zheng, B. Zong, W. Cheng *et al.*, "Node classification in temporal graphs through stochastic sparsification and temporal structural convolution," in *Proc. the Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2020*, Ghent, Belgium, 2021, pp. 330–346.

[121] D. Xu, W. Cheng, D. Luo, X. Liu, and X. Zhang, "Spatio-temporal attentive RNN for node classification in temporal attributed graphs," in *Proc. the Twenty-Eighth International Joint Conference on Artificial Intelligence*, 2019, pp. 3947–3953.

[122] A. Botchkarev, "A new typology design of performance metrics to measure errors in machine learning regression algorithms," *Interdisciplinary Journal of Information, Knowledge, and Management*, vol. 14, pp. 45–76, 2019. https://doi.org/10.28945/4184

[123] M. K. I. Rahmani, S. Ahmad, M. R. Hussain, A. K. Ameen, A. Ali, F. Shaman, A. Alshehri, M. S. Dildar, R. R. Irshad, and A. Islam, "Enhanced nanoelectronic detection and classification of motor imagery electroencephalogram signal using a hybrid framework," *Journal of Nanoelectronics and Optoelectronics*, vol. 18, no. 10, pp. 1254–1263, 2023.

[124] J. R. A. Solares, F. E. D. Raimondi, Y. Zhu *et al.*, "Deep learning for electronic health records: A comparative review of multiple deep neural architectures," *Journal of Biomedical Informatics*, vol. 101, 103337, 2020.