A Survey on Graph Construction for Geometric Deep Learning in Medicine: Methods and Recommendations

Anonymous authors
Paper under double-blind review

Abstract

Graph neural networks are powerful tools that enable deep learning on non-Euclidean data structures like graphs, point clouds, and meshes. They leverage the connectivity of data points and can even benefit learning tasks on data, which is not naturally graph-structured, like point clouds. In these cases, the graph structure needs to be determined from the dataset, which adds a significant challenge to the learning process. This opens up a multitude of design choices for creating suitable graph structures, which have a substantial impact on the success of the graph learning task. However, so far no concrete guidance for choosing the most appropriate graph construction is available, not only due to the large variety of methods out there but also because of its strong connection to the dataset at hand. In medicine, for example, a large variety of different data types complicates the selection of graph construction methods even more. We here therefore summarise the current state-of-the-art graph construction methods, especially for medical data. In this paper, we identify two main strands of graph construction: static and adaptive methods, discuss their advantages and disadvantages, and formulate recommendations for choosing a suitable graph construction method. We also discuss how a created graph structure can be assessed and to what degree it supports graph learning. We hope to support medical research with graph deep learning with this work by elucidating the wide variety of graph construction methods[1].

1 Introduction

Graphs can be used to represent several kinds of real-world datasets, such as networks, interactions, connections, or information flows. They hold information encoded in a set of nodes and edges, which connect pairs of nodes, and therefore add a structural component to otherwise independent data points. A wide variety of data can be structured as graphs, like knowledge (Ruan et al., 2021), (3D) structures in space (Wolterink & Suk, 2021), brain signals (Kim et al., 2021), or maps (Yu et al., 2021c). Yet, sometimes the question of how to construct an appropriate graph structure from a given dataset is non-trivial.

Graph learning techniques have been designed to apply deep learning (DL) methods directly to non-Euclidean datasets like graphs or meshes (Bronstein et al., 2017). These methods have since been frequently applied to data that can efficiently be structured by using a graph, like social networks (Fan et al., 2019) or molecules (Moreira-Filho et al., 2022), and have addressed tasks like friendship recommendations or drug discovery. Graph deep learning also naturally benefits many applications on medical datasets, as Graph Neural Networks (GNNs) have proven to be powerful algorithms for downstream tasks like medical diagnosis (Parisot et al., 2017), or image segmentation (Xie et al., 2022). They allow for straightforward integration of multi-modal data, such as image features and clinical data, into one coherent data structure, which has been explored in the context of so-called population graphs, where a medical cohort is represented by a graph structure instead of a tabular database (Kazi et al., 2022; Parisot et al., 2017).

Even in cases where a graph representation is not the default choice for a dataset, it has been shown that imposing such a graph structure by leveraging connections between data points can improve the performance

[1]The authors used ChatGPT for minor writing support.

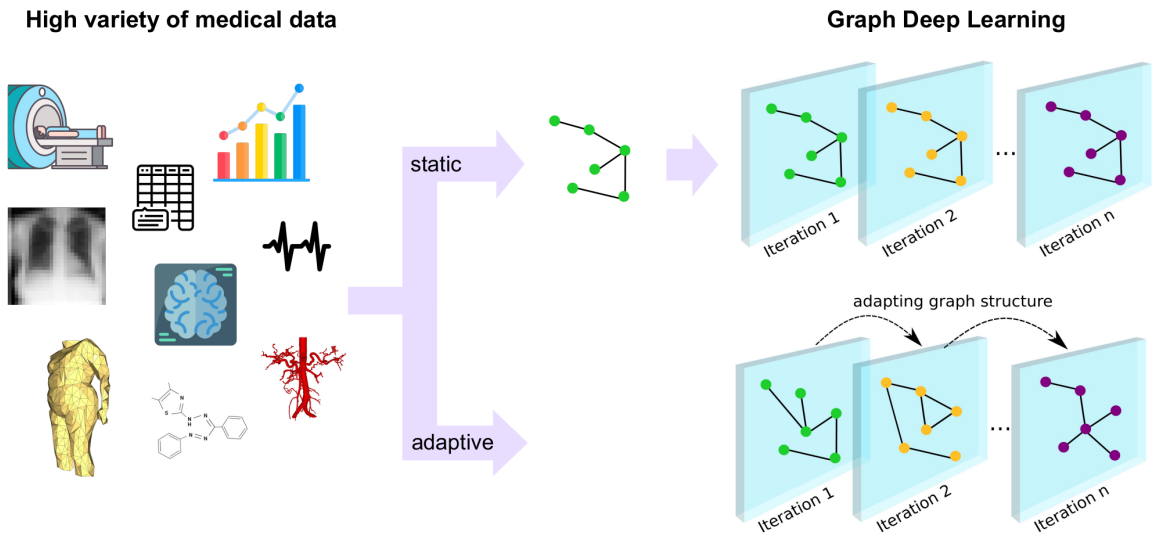


Figure 1: In this work, we summarise the state-of-the-art graph creation methods that allow one to transform a high variety of medical datasets into graph structures to perform graph deep learning. Static graph creation methods (upper row) extract the graph structure prior to learning, while adaptive methods (lower row) change the graph structure during training.

of ML algorithms (Parisot et al., 2017; Ahmedt-Aristizabal et al., 2021; Bessadok et al., 2022; Pellegrini et al., 2022). This way, relations can be utilised or newly discovered, which can be beneficial for the task at hand (Cosmo et al., 2020). The application of GNNs to point cloud datasets, for example, can improve model performance compared to only using individual data points (Wang et al., 2019). This has also been shown in the medical field for tasks like disease prediction (Parisot et al., 2017), vessel segmentation (Paetzold et al., 2021), or the interaction between symptoms, diseases, and medication (Ruan et al., 2021). Here, spatial proximity (Yao et al., 2022; Hansen & Heinrich, 2021), medical knowledge (Mueller et al., 2022a), anatomical structures (Sun et al., 2021), correlations (Kim et al., 2021), or cartographic location (Yu et al., 2021c), have been used to generate a graph structure from a previously disconnected set of data points.

This additional processing step of generating a graph structure introduces new challenges to the overall downstream task (Ahmedt-Aristizabal et al., 2021) and creating or fine-tuning nodes and edges holds a range of crucial design choices. This has turned out to be especially challenging in many medical settings since for medical images or health reports a graph structure is not the default choice of representation. Brain connectivity graphs have, for example, shown to be a suitable representation of the human connectome (Bessadok et al., 2022), which represents a map of neural connections in the brain. However, the construction of the brain graph holds challenges like the temporal component of functional magnetic resonance imaging (fMRI). Tube-like structures like airways (Selvan et al., 2020) and vessels (Paetzold et al., 2021) can be accurately represented by a graph that follows the anatomy of the structure at hand. Still, the concrete extraction of the graph structure requires precisely segmenting the curvilinear structures or transforming branching points into nodes.

In addition to the wide variety of graph construction methods, the strong impact of the graph structure on the success of the learning task makes this especially challenging (Luan et al., 2022). We thus conclude that the construction of a suitable graph structure is crucial to optimally leverage the connectedness inherent to the dataset.

1.1 Why the graph construction matters

Since the introduction of GNNs, many works have shown that GNNs can improve performance on non-Euclidean datasets compared to graph-agnostic DL models (Cosmo et al., 2020; Parisot et al., 2017; Ahmedt-

Aristizabal et al., 2021). However, this assumption does not apply to all settings and datasets and recent works have demonstrated that GNNs outperform graph-agnostic models only under specific circumstances. This can often be attributed to the utilisation of unsuitable graph structures and can even lead to simple graph-agnostic methods outperforming GNNs (Luan et al., 2022; Zhu et al., 2020). One of the reasons for this might be an over-smoothing of node features over unideal neighbourhoods (with, for example, highly diverse labels), which complicates the establishment of suitable feature embeddings required for the downstream task. When node features of neighbours with highly different labels (and therefore different node features) get averaged during message passing, the resulting node embedding might be an over-smoothed representation that merges node features of different labels.

The interaction between the graph structure and the model performance has been investigated in several works. One question of interest here is how and under which circumstances the graph structure hinders or benefits graph deep learning. In this context, several graph metrics that assess the graph structure have been introduced that are strongly correlated with GNN performances. One such metric is *edge homophily* (and its counterpart: heterophily) (Luan et al., 2021; Ma et al., 2022), which quantifies the similarity of neighbouring labels. Originally, GNNs were built on the assumption that connected nodes share similar properties (they are homophilic), and GNNs perform well based on this assumption (McPherson et al., 2001). As a result, several graph DL models underperform on datasets with highly different neighbourhoods (heterophilic graphs). More metrics and their impact on GNN performance are discussed in Section 5.3.

Even though this line of research implies that the graph structure and, therefore, the graph construction method strongly impacts the performance of GNNs, the analysis of the works of GNNs in medicine shows that there is no unique, clearly defined method nor any guidelines for creating the graph structures from the wide variety of medical datasets. In this work, we, therefore, survey recent works that address graph creation methods for graph deep learning tasks with a focus on medical data. The methods summarised in this review are not limited to applications on medical datasets, and we provide links to other non-medical domains in Section 6.

1.2 Contributions and outline

This work provides an overview of graph construction methods in medicine. We performed a literature search on *Google Scholar* based on keywords like "geometric deep learning", "graph neural network", "medicine", "population graph", "disease", "graph construction", and combinations of them. We summarise 76 works and categorise them by their graph construction method. The outline of this work can be summarised as follows:

- We identify four types of graphs that can be distinguished: population level graphs, subject level graphs, knowledge graphs, and real-world structure graphs, which we use to categorise the included works (Section 3);
- In Section 4, we systematise existing works that utilise GNNs in medical application areas by graph creation methods with a focus on static and adaptive graph creation (see Figure 1);
- Based on state-of-the-art works, we suggest guidelines for graph creation methods in Section 5.2;
- We summarise existing graph assessment metrics that allow to evaluate the generated graph structures in Section 5.3;
- We embed our work into the context of related review papers in Section 6;
- We identify open challenges of graph learning in medicine (Section 7) and conclude with promising future directions of research (Section 8).

2 Background

In this section, we give an overview of graphs, GNNS, and homophily, a main graph property linked to the performance of GNNs.

2.1 Formal definition of graphs

Throughout this work, we discuss datasets involving graph structures. A graph $G := (V, E)$ is defined as a collection, including a set of nodes/vertices V and a set of edges E connecting nodes. $n = |V|$ denotes the number of nodes in the graph. An edge $e_{ij} = (v_i, v_j)$ defines the connection from node v_i to node v_j. The Graph G is undirected if and only if $e_{ij} \Rightarrow e_{ji}$, $\forall i, j \in \{1, \ldots, n\}$. All edges E can be represented in the adjacency matrix $\mathbf{A}$ of size $n \times n$, where $\mathbf{A}_{ij} = 1$ if $e_{ij} \in E$ and 0 otherwise. A weighted graph $G_w := (V, E, \mathbf{W})$ additionally requires a weight matrix $\mathbf{W}$ that assigns a weight to every edge in the adjacency matrix. The weight matrix has the same dimensions as the adjacency matrix.

A neighbourhood $\mathcal{N}_v$ of a node $v \in V$ is defined by a set of all nodes that have an incoming edge to node v: $\mathcal{N}_v := \{u \in V | e_{uv} \in E\}$. A node v can be represented by a feature vector $x_v \in \mathbb{R}^m$. The features of all nodes (node features) can be summarised by the feature matrix $\mathbf{X} \in \mathbb{R}^{n \times m}$.

In this work, we summarise, categorise, and investigate different methods to build the node feature matrix $\mathbf{X}$, the adjacency matrix $\mathbf{A}$, and the weight matrix $\mathbf{W}$ of the graphs.

2.2 Graph neural networks

Graph neural networks (GNNs) were first introduced in 2005 by Gori et al. (2005) and further extended by Scarselli et al. (2008). The term summarises a branch of research that expanded DL methods to non-Euclidean datasets, using graph convolutions. Over the last years, several different graph convolutions have been introduced.

GNNs are based on a message-passing scheme, where the information stored in the nodes is propagated among neighbouring nodes, following the graph's edges. We here define the message passing with the example of graph convolutional networks (GCNs) (Kipf & Welling, 2016) but note that the principle is easily transferable to other graph convolutions.

Definition 2.1 (Graph convolutional networks (Kipf & Welling, 2016)) *Let $h_v^{(k)}$ define the feature representation of node v at layer k. For GCNs, the initial node representation for all nodes $v \in V$ is defined as following:*

$$h_v^{(0)} = x_v. \tag{1}$$

The node embedding of node v at step k is then defined as:

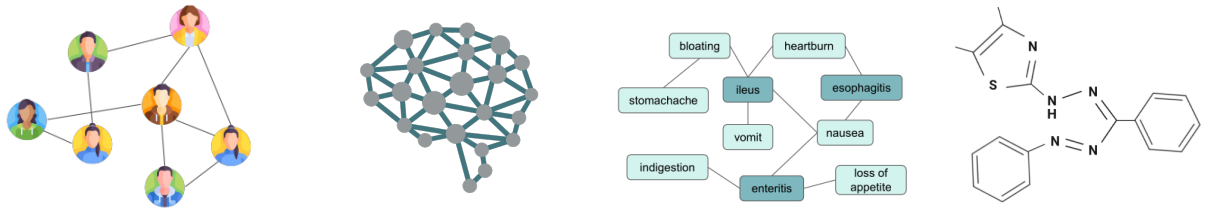
$$h_v^{(k)} = f^{(k)}\Big(W^{(k)} \cdot \frac{\sum_{u \in \mathcal{N}_v} h_u^{k-1}}{|\mathcal{N}_v|} + B^{(k)} * h_v^{(k-1)}\Big), \tag{2}$$

where the function f, the weight matrix W, and the bias B are k-dependent learnable parameters that are shared across all nodes (Daigavane et al., 2021).

We can see that the embedding of node v of the previous step ($h_v^{(k-1)}$), as well as the sum of all neighbouring node embeddings, are combined in the new node feature representation at step k. Different graph convolutions use varied versions of this definition, but they all follow the message-passing scheme. For more information about GNNs, we refer to Wu et al. (2020).

2.3 Graph homophily and heterophily

A key statistical property of graphs that indicates how nodes of different labels are connected throughout the entire graph structure is homophily (Luan et al., 2021). This property has been shown to potentially have a significant impact on the performance of GNNs (Zhu et al., 2020). In general, three types of homophily can be distinguished that all focus on a different nuance of the metric: node, edge, and label homophily (Luan et al., 2021). *Edge* homophily (Zhu et al., 2020; Ma et al., 2022) in graphs is defined as the ratio of edges that connect nodes with the same label vs different labels (see Equation 3). *Node* homophily (Pei et al., 2020) describes the average number of direct neighbours with the same label. *Class* homophily (introduced by Lim et al. (2021) and termed by Luan et al. (2021)) is an extension of edge homophily with the additional



(a) population-level graph (b) subject-level graph (c) knowledge graph (d) real-world structures

Figure 2: Schematic display of the four graph types utilised in medical research with GNNs: (a) a population graph, where each node represents a subject and subjects are connected based on similarity, (b) a subject-level graph with the example of a brain connectivity graph, where each subject is represented by a separate graph structure (c) a knowledge graph, where entities and connections between them form a representation of knowledge and (d) graph based on real-world proximity with the example of a molecule.

consideration for class imbalance. Formal definitions of node and class homophily can be found in Luan et al. (2021). We here define edge homophily since this is the most commonly used metric to assess graph structures.

Definition 2.2 (Edge homophily) *Formally, the edge homophily of a graph $G := (V, E)$ and the set of node labels $Y := \{y_u; u \in V\}$ is defined as:*

$$h(G, Y) := \frac{1}{|E|} \sum_{e_{uv} \in E} \mathbb{I}(y_u = y_v), \tag{3}$$

where $\mathbb{I}$ is the indicator function.

For example, in case half of the edges in a graph connect nodes with different labels and the other half connects nodes with the same label, the graph has edge homophily of 0.5. A graph is described as *homophilous* when h is large (typically larger than 0.5) and as *heterophilous* otherwise (Kim & Oh, 2021).

3 Graph structures in medicine

Based on the current literature, we identified four distinct graph types that are used in medical applications: (1) *population graphs*, where typically individuals of a cohort are connected in a large graph, (2) *subject-level graphs*, where each subject is represented by an individual graph –leading to a multi-graph dataset– (3) *knowledge graphs*, where (often heterogeneous) entities are connected based on previous knowledge and databases, and (4) graphs that follow a *real-world structure*, like molecules. Each approach comes with individual challenges and utilises different methods for graph creation. In this section, we give an overview of those four graph types, which are visualised in Figure 2.

3.1 Population graphs

One research area of graph learning in medicine utilises so-called population graphs (Figure 2(a)). They are generated by connecting all subjects in a cohort to a single (usually large) graph. The goal is to improve model performance by using interactions between the subjects/nodes in the graph. The most common structure of a population graph is one where every subject in the dataset is represented as a node and node connectivity is, for example, defined by some distance metric between the subjects. When using population graphs, the learning task of the GNN is usually *node prediction*. Here, a prediction (e.g. classification or regression) is made for every node. This can, e.g. be a disease prediction (Parisot et al., 2017) or age prediction (Kazi et al., 2022). Population graphs are an effective method to integrate multi-modal data and enable the usage of patient data from different data sources and modalities. There are some examples where population graphs are extended with some additional components. In Gao et al. (2021) e.g., the authors create a bipartite graph, where subjects and gene expressions are represented by node entities.

3.2 Subject-level graphs

Another way to represent medical data is in the form of *subject-level* graphs. This term summarises various graphs, where each subject in a dataset is represented by a single graph. The individual graphs are therefore independent of each other and together constitute a multi-graph dataset. One commonly used example for a subject-level graph is the representation of brain images as a brain connectivity graph (see Figure 2 (b)). In general, there are numerous ways to create subject-level graphs, depending on the dataset at hand and the application. They can, for instance, be used to represent structural connectivity in graph representations of arteries (Chen et al., 2020b), brain vessels (Paetzold et al., 2021), or airways (Zhao & Yin, 2021; Selvan et al., 2020), or to model a skeleton of a human in motion (He et al., 2022). When using subject-level graphs, the most common learning task is *graph prediction* (e.g. classification or regression). However, there are also applications where node-level and edge-level predictions (Chen et al., 2020b) are targeted using subject-level graphs.

3.3 Knowledge graphs

Knowledge graphs refer to the encoding of general concepts and knowledge in graph structure, highlighting relations between different entities. This is often used in the context of diseases, symptoms, drugs, or genes to display their correlation or interaction. They are usually not personalised but are "intended to accumulate and convey knowledge of the natural world, whose nodes represent entities of interest and whose edges represent potentially different relations between these entities" (Hogan et al., 2021). An example of a multi-modal medical knowledge graph is PrimeKG (Chandak et al., 2023), which includes information about drugs, diseases, phenotypes, exposures, and genes. Cheng et al. (Cheng et al., 2021a) construct a knowledge graph on stroke data, and Bonner et al. (2022) review different knowledge graphs on biomedical data for drug discovery. A knowledge graph differs from population graphs in the sense that here, no individual patient data is represented as a graph, but general knowledge and connections between entities are modelled in graph form. For a more detailed review of knowledge graphs, we refer to Ye et al. (2022). Knowledge graphs can be used for different applications, either on their own or as an additional source of information for other tasks, like in Pfeifer et al. (2022).

3.4 Real-world structures

The fourth category of graph types is based on *real-world structures* (Figure 2 (d)). This can, for example, be the encoding of cartographical proximity in maps like connecting hospitals in different regions of the country (Jin et al., 2021) or connecting cities based on their local proximity (Yu et al., 2021c). Another example is the representation of molecules, where nodes represent atoms and the edges bindings between them (Bonner et al., 2022). Molecule-based datasets in the medical domain are commonly used for graph-level predictions, for example investigating drug properties (Duvenaud et al., 2015; Kearnes et al., 2016) or potential interactions between different drugs (Xu et al., 2019).

4 Graph construction methods

In the following, we will summarise and categorise the state-of-the-art graph construction methods for medical data, linking them to the different graph types introduced in Section 3. For graph deep learning, graph structures are usually extended to contain node features. This additional knowledge is then propagated along the graph structure during the learning process.

Generally, the graph construction process consists of two aspects: (1) defining the nodes and their features and (2) edge creation. Each node has a feature vector of shape m and all node features of a graph $\mathcal{G}$ with n nodes are summarised by a feature matrix $\mathbf{X} \in \mathbb{R}^{n \times m}$. The edges can be summarised by the adjacency matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$. The definition of the adjacency matrix and node feature matrix are in general intertwined, and both steps are necessary to extract the full graph structure.

In the following sections, we will investigate the definition of nodes and node features as well as the creation of the graph structure (the edges). An overview of the categories for graph construction methods, with a focus on the definition of the graph's edges, is visualised in Figure 3.

4.1 Defining the graph's nodes

The extraction of the graph's nodes is highly dependent on the dataset. When building **population-level graphs**, every node usually represents a subject in the dataset. Node features can, for example, contain tabular data like lab results (Parisot et al., 2017), images (Keicher et al., 2021), image-derived features (Parisot et al., 2017), or combinations of those (Keicher et al., 2021).

For **subject-level graphs**, the node feature extraction can vary greatly. We will here summarise the most prominent node and node feature definition strategies for different subject-level graphs. For the creation of brain connectivity graphs from fMRI data, the most commonly used approach to define node features is to define regions of interest (ROIs). Here, the definition of nodes is often guided by a 3D atlas, which defines the ROIs from the recorded BOLD signal (Wang et al., 2022b). We, therefore, say the definition of the graph nodes relies on prior knowledge (it is prior-driven). There are some examples where slightly different approaches are utilised to define the final nodes of a graph connectivity graph. For example, Zheng et al. (2022a) identify the brain's most informative regions through sub-graph generation. Yao et al. (2021) use several templates with varying ROI parcellation scales to create coarse-to-fine brain connectivity networks for each subject instead of depending on a specific brain parcellation. Yao et al. (2022), for instance, build a graph from image data and use features extracted from a convolutional neural network (CNN) as node features of the graph. To extract a subject-level graph that represents curvilinear structures, like vessels or airways, branching points of the structure can be used to define the nodes of the graph. However, nodes can also represent a section of the tube (Paetzold et al., 2021). Since in most cases, the definition of node features and edges for subject-level graphs are intertwined, we provide more information about the node features of subject-level graphs in Section 4.2.

When building **knowledge graphs**, the nodes typically represent entities of interest, like diseases, symptoms, or medications. Graphs **on real-world structures** follow the given entities of the dataset. For molecules in drug research, for example, the nodes usually represent the molecules' atoms (Zheng et al., 2021; Zhao et al., 2021) and for cartography-based graphs, cities or hospitals can, for instance, be used to encode the graph's nodes (Yu et al., 2021c; Jin et al., 2021).

4.2 Defining the graph's edges

In this section, we introduce different methods for defining the structure of the graph (the edges). We hereby categorise existing works based on the following criteria:

(a) Graph type: population-level, subject-level, knowledge graphs, and real-world structures

(b) Static vs adaptive graph creation

(c) Purely data-driven vs. prior-driven

Graph creation methods can be categorised into *static* and *adaptive* approaches. We consider a graph creation method as *static* if the adjacency matrix is generated prior to training without any adaptions during the learning pipeline and as *adaptive* if the graph structure is adapted during training. The different static and adaptive graph creation methods are summarised in Table 1 and 2, respectively. Both tables also indicate the generated graph type and whether the approach is *data-* or *prior-driven*. We consider methods that only use the dataset at hand as *data-driven* and ones that also include additional prior knowledge as *prior-driven*.

4.3 Static graph construction methods

In this section, we categorise different static graph construction methods by the metric and information utilised to define the edges between the nodes. We mainly discuss graph type, connection mechanism and

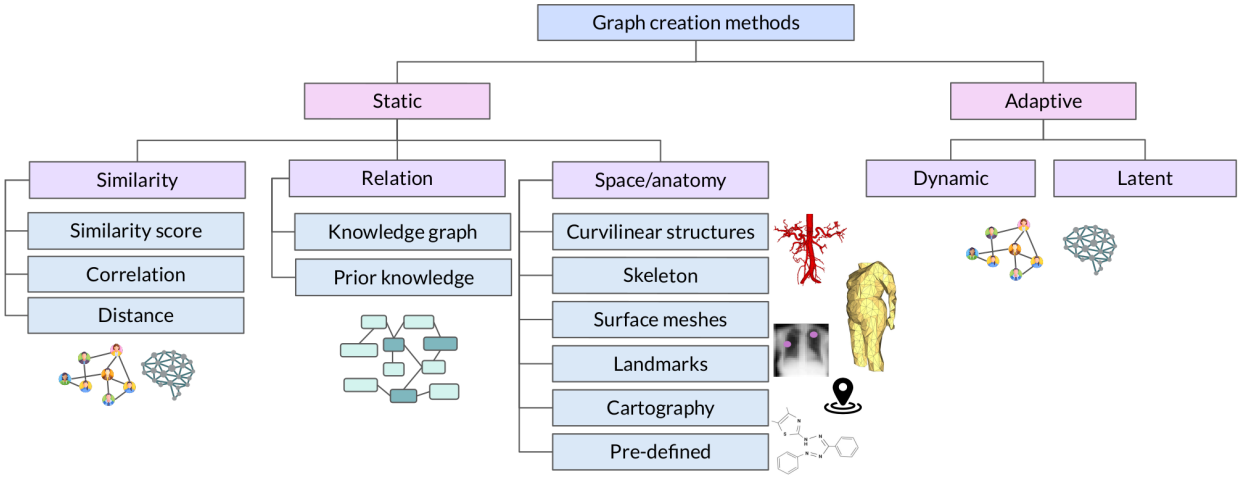


Figure 3: Overview of graph construction methods for GNN training in medicine categorised by static and adaptive graph construction. The icons represent examples of represented data in the respective categories.

the utilised data. The available methods are summarised in Table 1. In column "Category" we categorise each method by the utilised relation for edge construction. "Graph type" indicates the specific type of graph (Section 3) that is generated, "graph construction" refers to the utilised metric or property to decide when an edge is added to the graph, the represented data is listed in the respective column, "P/D" indicates whether the graph is prior- or data-driven, and the listed references give examples of works that utilise the respective methods, including the initial introduction of a specific method.

4.3.1 Graph construction based on similarity or distance

One method that is frequently used to determine whether nodes connect is based on a similarity or distance measure. Simplified speaking, in this approach, two nodes will be connected, if they are "similar" or "similar enough". The similarity/distance between nodes can be defined by different metrics, including a similarity score, the Euclidean distance, cosine similarity, or correlation-based similarity. Furthermore, methods using similarity or distance measures to construct the graph structure can be categorised into ones using the original data at hand, or a (mostly lower-dimensional) embedding of the original data (Lu et al., 2022). When using a similarity or distance measure to determine whether nodes should be connected, usually an additional step of edge selection needs to be performed in order to obtain the final graph (Section 4.3.5).

Similarity score The similarity between nodes in the graph can be defined with the help of a similarity score. This is one of the most commonly used methods for the creation of population graphs in order to determine how similar two subjects in a cohort are. This method was first introduced by Parisot et al. (2017) in 2017 and has since then been used in many variants and applications, e.g. in Ghorbani et al. (2022); Vivar et al. (2021); Pellegrini et al. (2022); Peng et al. (2022) and Lu et al. (2022). One advantage of this method is that the similarity score considers discrete and non-discrete features and can be adapted by adding additional weight factors. All available features are divided into two groups (e.g. imaging and demographic features), then the graph is built based on a similarity score that uses one of these subgroups of features (e.g. demographic features). The second group of features (e.g. imaging features) are used as node features. Parisot et al. (2017) originally used this approach for brain analysis in Alzheimer's disease and autism prediction. They used phenotypic information for the graph creation and imaging features as node features and show an example of multi-modal data integration.

Variations of this method have been used for bone age estimation (Du et al., 2015), brain age prediction (Stankevičiūtė et al., 2020), autism detection (Rakhimberdina et al., 2020), genome inference (Dilthey et al., 2015), disease prediction (Chen et al., 2020a; Kazi et al., 2019; Ghorbani et al., 2021), to only name a few. In each of the here mentioned works, the original approach of creating the population graph by Parisot et al.

Table 1: *Summary of* ***static*** *graph construction methods.* We differentiate between different graph types, construction methods, and source data types. We also indicate if the method is prior-driven (P), data-driven (D) or both (P, D), as well as if the graph is weighted (**W**).

Category	Graph type	Graph construction	Represented data	P/D	W	References
Similarity and distance	Population level	Similarity score	Patient cohort	D	✗	Parisot et al. (2017); Ghorbani et al. (2022); Vivar et al. (2021) Pellegrini et al. (2022); Peng et al. (2022); Pan et al. (2021); Kazi et al. (2019); Ghorbani et al. (2021)
	Population level	Similarity score	Image features, clinical data	D	✓	Lin et al. (2023); Qiu et al. (2021)
	Population level	Mutual information	Images, clinical data	D	✗	Keicher et al. (2021)
	Population level	Euclidean distance	Image features, clinical data	D	✗	Lu et al. (2022); Yu et al. (2021b)
	Subject level	Euclidean distance	EEG signal	D	✗	Demir et al. (2021)
	Subject level	Euclidean distance	Images	D	✗	Sun et al. (2021)
	Subject level	Euclidean distance	Airways	D	✗	Tan et al. (2021)
	Subject level	Cosine similarity	Images	D	✗	Mahapatra et al. (2022)
	Subject level	Morphological similarity	fMRI data	D	✗	Mahjoub et al. (2018)
	Subject level	Pearson correlation	Brain connectivity	D	✗	Kim et al. (2021)
	Subject level	Partial correlation	Brain connectivity	D	✓	Li et al. (2021c)
	Knowledge graph	Conditional probability	Lesion types	D	✗	Cheng et al. (2021b)
Relation	Population level	Medical assessments	Patient cohort	D	✓	Mao et al. (2022)
	Knowledge graph	Known interactions	Disease/symptom/medication	P	✓	Ruan et al. (2021)
	Knowledge graph	Interactions	Ontologies	P	✓	Hao et al. (2021)
	Knowledge graph	Synergism/antagonism	Drugs	P	✓	Zheng et al. (2021)
	Knowledge graph	Drug–protein pairs	Drugs	P	✓	Zhao et al. (2021)
	knowledge graph	Protein interactions	Proteins	P	✗	Schulte-Sasse et al. (2021)
	Knowledge graph	Co-occurrences	(Clinical) abnormalities	P	✓	Liu et al. (2021); Zhou et al. (2021)
	Knowledge graph	Co-occurrences	Medical labels	D	✓	Hou et al. (2021)
	Subject level	Medical importance	ECG leads	P	✗	Mueller et al. (2022a)
	Subject level	Protein interactions	Proteins	P	✓	Pfeifer et al. (2022)
Space, structure, or anatomy	Subject level	Local proximity	Image landmarks	D	✗	Yao et al. (2022); Hansen & Heinrich (2021)
	Subject level	Tree structure	Curvilinear structures	D	✓	Yu et al. (2022a)
	Subject level	Tree structure	Curvilinear structures	D	✗	Paetzold et al. (2021); Wittmann et al. (2023); Shin et al. (2019) Chen et al. (2020b); Wolterink et al. (2019); Yu et al. (2022b)
	Subject level	Local proximity	Curvilinear structures	D	✗	Xu et al. (2022); Xie et al. (2022)
	Subject level	Local proximity	Curvilinear structures	D	✓	Li et al. (2021b)
	Subject level	Anatomy	Variety of patient data	D	✗	Barbiero et al. (2021)
	Subject level	Anatomy	3D point clouds	D	✗	Yu et al. (2021a)
	Subject level	Anatomy	Skeleton	P,D	✗	He et al. (2022)
	Subject level	Anatomy	Curvilinear structures	D	✗	Selvan et al. (2020)
	Subject level	Mesh generation	Cortical surface meshes	D	✗	Azcona et al. (2020); Gopinath et al. (2019a); Wu et al. (2019); Gopinath et al. (2019b)
	Real-world structure	Cartography	Maps	D	✗	Yu et al. (2021c); Jin et al. (2021)
	Real-world structure	Chemical structure	Molecules	D	✗	Bonner et al. (2022); Kearnes et al. (2016); Duvenaud et al. (2015)
Combinations	Knowledge & structure	Text-derived	Disease/symptom/medication	P,D	✓	Vretinaris et al. (2021)
	Knowledge & structure	Chemical structure	Disease/symptom/medication	P,D	✓	Zhang et al. (2022b)

(2017) has been modified to best fit the data and the application. Some adaptations of this method have also used both imaging and non-imaging features for graph creation as well as node feature representation (Lin et al., 2023) to maximise the information about the similarity between individual subjects.

Correlation-based similarity Another option to create the graph structure is by evaluating the correlation between node features and connecting those nodes that show high correlation. Similar to the similarity-based creation, the correlation is thresholded and only nodes that are "correlated enough" will be connected by edges (Section 4.3.5). This method is used mainly for the creation of brain connectivity graphs from functional magnetic resonance imaging (fMRI) data. Here the correlation between the blood oxygen level dependency (BOLD) signal, indicating alterations in blood oxygen levels over time of all previously defined regions of interest (ROIs) –which are represented by the nodes in the graph– is first averaged within each brain ROI. Then, a correlation metric is calculated between pairs of regions, and the nodes that show a high correlation are connected leading to functional connectivity that illustrates the communication between different brain regions.

Pearson's correlation coefficient, which measures the linear correlation between nodes, has been commonly used to evaluate functional connectivity. Values range from -1 to 1 where the former shows perfect negative correlation and the latter perfect positive correlation. A value of 0 indicates no correlation. Alternatively, other correlation methods such as partial correlation, Ledoit-Wolf (LDW) regularised shrinkage estimator (Noman et al., 2021) and Spearman's rank correlation (Yu et al., 2023) are utilised to construct functional connectivity. Partial correlation is an extension of the Pearson correlation coefficient, which measures the linear relationship between two variables. By removing the effects of the controlling variables, partial correlation helps to identify the unique relationship between the two variables of interest, providing a clearer understanding of the associations between them. LDW regularised shrinkage estimator is a statistical method designed to improve the estimation of covariance matrices when the number of observations –number of scans– is small compared to the number of variables –number of ROIs. Noman et al. (2021) use this method to obtain well-conditioned functional connectivity. Spearman's rank correlation is a non-parametric rank correlation method between two nodes that can detect monotonic nonlinear relationships. Yu et al. (2023) propose a multi-graph attention network to use both Pearson's and Spearman's rank correlation measures. Li et al. (2021c) utilise ROI-aware GNN assuming more closely connected ROIs exert a greater effect on each other and apply node (ROI) pooling layer (R-pool) to retain the most representative ROIs while eliminating noisy nodes. Gharsallaoui et al. (2021) predict an affinity matrix of the target view from the source view. Klepl et al. (2022) compare different functional connectivity measures for Alzheimer's disease prediction from EEG data as well as different edge selection methods.

One potential shortcoming of the creation of brain connectivity graphs from the BOLD signal is that the signal and the correlation between ROIs change over time. Kim et al. (2021); Kong et al. (2021); Wang et al. (2022b) target this by creating several brain connectivity graphs for different time steps to allow a more dynamic graph structure over time. Kong et al. (2022), e.g., propose a multi-stage learning module to fully utilise features at different stages and use a deep auto-encoder to extract the graph structure.

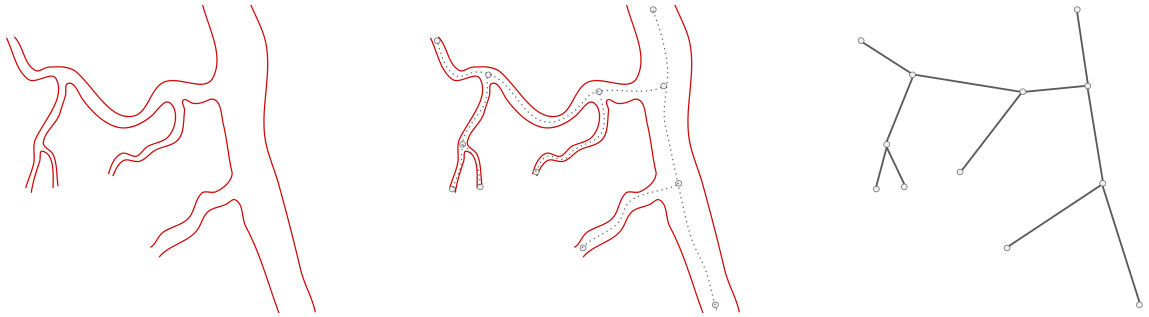
Distance measures and mutual information In line with the usage of a similarity score, the *Euclidean distance* or the *cosine similarity* has also been used to determine the "distance" between nodes (Lu et al., 2022; Yu et al., 2021b). These distance measures can also be interpreted as a similarity between nodes and used as a basis for edge selection (Section 4.3.5), where only the least distant nodes are connected. Keicher et al. (2021) use images, extracted image features, and clinical data as node features and build the graph based on *mutual information*.

4.3.2 Graph construction based on relations

Apart from similarity- and correlation-based edge definitions, the graph structure can also be derived from known relations between the entities. This can e.g. be symptoms that are related to certain diseases, medications that are used to treat diseases, or lab results that are associated with specific doctor's visits (Mao et al., 2022).

Knowledge graphs Knowledge graphs are typically constructed based on known relations between entities that are usually extracted from large knowledge resources or clinical studies. Medical knowledge graphs can for example contain diseases, symptoms and medications, connecting co-occurring symptoms and diagnosed diseases with prescribed medication and reported symptoms. Even though the graph is constructed based on general knowledge, several design choices still remain. Zhang et al. (2020) create a knowledge graph based on clinical studies where abnormalities are connected with organs and body parts. Hou et al. (2021) follow the same graph creation approach, with an additional step of post-processing where missing links are included in a data-driven manner. Protein-protein interactions can be encoded in knowledge graphs and have, for example, been used to identify cancer genes (Schulte-Sasse et al., 2021) and drug–target interactions (Zhao et al., 2021).

The knowledge graph PrimeKG (Chandak et al., 2023) was for example built on large resources like *DrugBank* (Assempour et al., 2018), *Drug central* (Avram et al., 2021), and *Entrez gene* (Maglott et al., 2010) - to mention only a few. The authors highlight that PrimeKG contains edges between drugs and diseases



(a) Vessel tree (b) Centerline and branching points (c) Abstract graph

Figure 4: Example of an abstract graph representation of a vessel tree, with nodes encoding branching points and edges vessel segments.

that specify "indications", "contradictions", and "off-label use", which are often missing in other medical knowledge graphs (Chandak et al., 2023).

These constructed knowledge graphs can then be used as a basis for decision-making and graph learning. The authors in (Vretinaris et al., 2021) for example use a knowledge graph to combine generally applicable information regarding medications and symptoms with text-derived patient-specific information about their individual symptoms. Zhang et al. (2022b) utilise a similar technique and utilise an attribute graph, containing chemical structures of drugs as an additional graph to the knowledge graph.

Prior knowledge The graph structure can also be created or influenced by the inclusion of prior knowledge about the data and the importance of specific features in other areas apart from knowledge graphs. This was for example applied in a graph classification task using electrocardiogram (ECG) data to determine left bundle branch blocks (Mueller et al., 2022a), where prior medical knowledge about the importance of seeds guided the graph construction method.

4.3.3 Graph construction based on space, structure, or anatomy

Another graph construction technique is building spatio-structural or spatio-semantic graphs, where the nodes have associated positions in a metric space. In this case, the relations between the nodes usually correspond to distances in the respective metric space. The starting point can e.g. be a 2D or 3D image or a map, and a common choice for the metric space is the Euclidean space.

Curvilinear structures One prominent application of spatio-structural graphs in medical imaging is represented by curvilinear structures. These are long, thin, structurally constrained tubular structures, following a tree or network configuration. Examples of such structures are: vessel trees, airways, or neurons. As en exemplification of their structural constrains, vessel trees follow blood flow constraints, their branching pattern is dictated by the perfusion requirements of nearby tissue, and the cross-sectional diameter of constituent vessels progressively narrows in the distal direction. The inherent graph structure of these objects makes them ideal candidates for geometric deep-learning: either by achieving a more compact, abstract representation for branch classification or exploiting geometric information for an auxiliary task. The graph representation is typically only a secondary representation derived from an initial imaging representation, making the graph learning highly dependent on the quality of the graph extraction process. Figure 4 shows an example of a graph representation derived from a vessel tree.

The most straightforward representation of curvilinear trees as graphs follows the structure of the trees, where branching points form the node set and individual segments are edges. Such an approach is suitable for multi-class classification tasks, such as airway or vessel labelling. Chen et al. (2020b) construct a graph of the intracranial arteries by choosing nodes as centerline points with more than two neighbours and linking them based on skeleton connectivity. Their approach to multi-class vessel labelling is formulated as simultaneous node and edge classification, where message passing occurs through both node and edge

representations. Many other works opt to formulate the branch labelling problem as node classification instead, since node classification is more widely spread than edge classification. Paetzold et al. (2021) introduced the VesselGraph dataset of graphs derived from mouse brains for the tasks of link prediction and vessel classification. The graph structure is extracted by first generating a segmentation of a 3D volume, then processing it via the graph extraction tool Voreen (Drees et al., 2021). Then the initial graphs with nodes as branching points and vessels as edges are subsequently converted via the line graph approach, such that nodes of the derived graph correspond to the edges of the initial graph. Xie et al. (2022) performed anatomical labelling of airways by directly constructing a graph where nodes correspond to branches and assigning edges based on the connectivity in the image segmentation map. Yu et al. (2022b) used a similar graph construction approach for airway labelling while adding additional hyperedges between the children of a parent, thereby using two information passing pathways.

Another prominent use of curvilinear structures in graph-based deep learning is for the purpose of improving image segmentation. The construction of the graph is less abstract in such scenarios, instead following the voxel structure of the data representation. Nodes are no longer well-defined anatomical structures such as branches or branching points, but tend to be pixels or superpixels, while edges follow their connectivity in the image. Xu et al. (2022) first partition the input image into superpixles, which become the nodes of the vessel graph. Edges are then only defined between nodes corresponding to neighbouring superpixels in the original image. The superpixel-derived graph is used for vessel segmentation refinement, by predicting superpixel vesselness as a binary node classification problem. Similarly, Shin et al. (2019) proposed to train a node classifier for segmentation refinement also using vesselness prediction. However, in their approach nodes are represented by points sampled equidistantly along the centerlines rather than superpixels, whereas edges are based on node connectivity or geodesic distances.

Some works (Yu et al., 2022a; Li et al., 2021b; Tan et al., 2021) only use GNNs during training for improved CNN feature representation for image segmentation, rather than as a standalone task. Yu et al. (2022a) partition the image into super-voxels and sample a node per super-voxel, relying on travel distance between nodes in the ground truth for determining edges, while Li et al. (2021b) rely on corner sampling following skeletonization for node creation and a similar approach for determining edges. Tan et al. (2021) perform airway segmentation via CNNs with concatenated GNN features by combining nodes around landmarks and nodes sampled at random lung locations. The adjacency matrix is created by connecting the k nearest neighbours, in addition to connecting another node with the same predicted semantic class for each node in order to improve homogeneity.

Wolterink et al. (2019) use an entirely different approach for vessel segmentation. Instead of working in the pixel space, they turn to the polar space centred around the vessel centerline and formulate the vessel segmentation problem as regression. For each cross-section of the centerline, a node for the centerline and 24 additional nodes around the centerline are created. Thus, the GNN is trained to regress the distance to the centerline. Each non-centerline node is connected to its corresponding centerline node, the corresponding nodes in adjacent cross-sections, as well as 2 neighbouring nodes. Additional connections are defined in order to establish a triangle mesh, whose vertex coordinates are determined by the regressed distances in the polar space.

Skeleton representation Another application that follows anatomical structures for graph creation, is the representation of the human skeleton in graph structure. This has been used for the detection of Parkinson's disease, where skeletons are extracted from video data of patients (He et al., 2022). The authors use two different methods for generating the graph structure: (1) connecting the joints of the skeleton locally, following the underlying skeleton (prior knowledge), or (2) globally, where every joint of the skeleton is connected to one node at the position of the neck. This shows that even when the data contains natural connectivity, there are multiple ways to extract a suitable graph structure for graph deep learning.

Surface Meshes Meshes are a special type of graph structure, that can benefit from graph deep learning Bronstein et al. (2017). A mesh usually represents a 3D structure in space, where nodes are aligned around the surface of an object. Edges usually build triangular faces between neighbouring nodes, which is why they are often also referred to as *triangulated* meshes. Meshes have also been used for medical data representation,

(a) Local connection module: here the graph structure follows the underlying skeleton of a person.

(b) Global connection module: all relevant joints are connected to one global node at the neck of a person.

Figure 5: Schematic visualisation of two approaches to create a graph structure for skeleton representations that are introduced in He et al. (2022)

such as cortical structures (Azcona et al., 2020; Gopinath et al., 2019a; Wu et al., 2019; Gopinath et al., 2019b) in neuroscience or organ surfaces (Mueller et al., 2022b). Meshes can, for example, be extracted from segmentations of 3D images, using methods like marching cubes (Lorensen & Cline, 1998). Other works utilise finite element meshes (Salehi & Giannacopoulos, 2022), for example, using methods like TetGen Hang (2015). Surface meshes are frequently used in neuroscience. The authors in Salehi & Giannacopoulos (2022) use cortical meshes for soft tissue prediction in image-guided neurosurgery. Salehi & Giannacopoulos (2022), e.g., use meshes for predicting soft tissue deformation in image–guided neurosurgery. Azcona et al. (2020) use surface meshes for brain morphology estimation for Alzheimer's disease classification. They analyse cortical and subcortical irregularities that have been shown to be correlated with Alzheimer's disease. A review by Zhao et al. (2023) summarises methods for cortical surface-based neuroimage analyses, including methodologies for cortical surface extraction and parcellation. Popular software tools like FreeSurfer Fischl (2012) have integrated methods for cortical surface extraction from MR images.

Image landmarks Graphs can also be extracted from images by identifying landmarks in the image and then connecting them based on local proximity. Hansen & Heinrich (2021) apply this method to extract lung landmarks from 3D computer tomography (CT) images in order to perform graph-based registration. They use the Foerstner interest operator (Förstner & Gülch, 1987) to identify sparse landmarks from the image, and then use a minimum spanning tree to construct the edges of the graph structure (Heinrich et al., 2015). Zhao (2021) use several graph structures in their work. One of them uses simple linear iterative clustering (SLIC) (Achanta et al., 2012) to determine the connection between super-pixels in chest X-ray images.

Cartography-based graphs There are a few works, where the graph structure follows cartographic proximity, like maps or regions. Jin et al. (2021) e.g. use hospitals and regions as nodes in their graphs, connecting geologically close entities with each other. Yu et al. (2021c) suggest representing nodes by cities in the graph structure, and the proximity of geolocation is used to connect them with each other.

Pre-defined structural connections Some datasets come with internal structural information, from which the graph can be extracted. An example of this would be molecules, which have been used for drug discoveries (Bonner et al., 2022) or protein analysis (Pfeifer et al., 2022). Here, the graph creation usually follows the molecular bindings. For more information about these pre-defined structural connections, we refer the interested reader for example to a review on ML approaches for the design of multi-target drugs (Moreira-Filho et al., 2022).

4.3.4 Fully connected graphs

Similar to pre-defined graph structures, there are some applications, where a fully connected graph is used (Chao et al., 2020). In those scenarios, no graph creation method has to be chosen. Instead, all nodes are

connected to all other nodes. Chao et al. (2020), for example, use fully connected graphs for lymph node gross tumour volume detection.

4.3.5 Edge selection

In many of the above-mentioned graph creation methods, a distance/similarity/correlation/proximity between nodes is derived. However, this does not directly lead to a final graph structure. In order to create the final (sparse) graph from the determined distance between the nodes, a set of –sometimes weighted– edges needs to be selected. This is usually either done by (a) thresholding or (b) selecting a fixed number of neighbours (k-nearest neighbours).

Thresholding edges Thresholding techniques can be further split into *absolute* thresholding and *proportional* thresholding. In *absolute* thresholding, a fixed threshold is chosen to remove weak connections which results in varying graph densities for different subjects. In *proportional* thresholding, a portion of the strongest connections is kept resulting in all graphs having the same edge density. The latter method is for instance used by (Noman et al., 2021; ElGazzar et al., 2022). In addition to those methods, Yang et al. (2021) applied the spatial-constrained sparse representation optimisation method to obtain a sparse representation matrix that captures the relationships between brain regions, while considering their spatial proximity.

k-nearest neighbour selection Alternatively, the k-nearest neighbours (k-NN) of each node can be selected based on the similarity/distance/proximity between the nodes. This way, every node is connected to a fixed number (k) of neighbours. K-NN graph creation methods are applied in several application areas, e.g. for population-level graph creation and brain graph creation. In the latter, a *k-NN graph* is derived from each densely connected functional connectivity matrix. It connects each brain region of interest to the k neighbours that show the highest connectivity. This is for example used in Zhang et al. (2021b); Qu et al. (2021); El Ouahidi et al. (2022); Wang et al. (2022a); Yao et al. (2021). For population-level graphs, usually, the k-nearest neighbours are selected based on the features extracted from the dataset (Lu et al., 2022; Yu et al., 2021b). An implementation of a k-NN graph creation is for example provided by PyTorch Geometric (Fey & Lenssen, 2019).

4.4 Adaptive graph construction methods

In contrast to the previously investigated static graph creation methods, we here introduce adaptive methods, where the graph structure is adapted during GNN training. The methods using adaptive graph construction are summarised in Table 2.

Zhu et al. (2021) propose a categorisation of adaptive graph creation methods into (a) metric-based, (b) neural-based, and (c) direct approaches. Metric-based approaches (a) use kernel functions to compute a similarity between node features or embeddings and use those to determine edge weights. Neural approaches (b) use neural networks to predict the edge weights, and direct approaches (c) interpret the adjacency matrix as free variables. They refer to adaptive graph creation methods in general as "graph structure learning" (GSL). However, these terms are not common in the medical domain, where instead the terms *dynamic* and *latent* are used. *Dynamic* approaches adapt the graph structure in-between training steps, but do not specifically learn the graph structure in an end-to-end manner, which *latent* graph construction methods do.

4.4.1 Dynamic graph construction

Dynamic graph creation methods do not construct a fixed/static graph structure before training but change the graph structure depending on the embeddings during GNN training. Wang et al. (2019) propose dynamic graph CNNs (DGCNNs) for the analysis of point clouds. They extract a k-nearest neighbour graph based on the feature space of the applied neural network, which changes in between training steps. In Tran et al. (2018), the authors extend this method with a new definition of graph convolutional filters with the goal of improving the impact of distant neighbours. They use the shortest path connections and also do not optimise the latent space with respect to the graph structure but with respect to the downstream task. Their k-nearest

Table 2: Summary of **adaptive** graph creation methods with their corresponding graph types and examples of original data, from which the graph is extracted. D indicates, that the graph creation method is data-driven and P that it is prior-driven.

Category	Graph type	Method	Represented data	P/D	References
Dynamic	Real-world structure	Euclidean distance	Point clouds (non-medical)	D	Wang et al. (2019)
	Real-world structure	Euclidean distance	Proteins, molecules	D	Tran et al. (2018)
	Population level	Cosine distance	Patient cohort	D	Zheng et al. (2022b)
	Population level	Contrastive learning	Brain connectivity	D	Wang et al. (2022b)
	Subject level	Euclidean distance	Airways	D	Garcia-Uceda Juarez et al. (2019)
Latent	Population-level	End-to-end learned	Patient cohort	P,D/D	Kazi et al. (2022); Cosmo et al. (2020); Mullakaeva et al. (2022); Huang & Chung (2020); de Ocáriz Borde et al. (2023); Song et al. (2021)
	Subject level	End-to-end learned	Brain connectivity	D	Kan et al. (2022); Campbell et al. (2022); Mahmood et al. (2021); El-Gazzar et al. (2021); Zhu et al. (2022)
	Subject level	End-to-end learned	MR images	D	Mo et al. (2021)

neighbour operation is not differentiable. Tran et al. (2018) also apply their method to biological graphs like molecules and proteins, however, they have not been applied to medical datasets specifically.

Zheng et al. (2022b) introduce a framework that utilises a dynamic graph generation method for disease prediction with a focus on the integration of multi-modal data. With the addition of separate loss terms, their method optimises for the downstream task and certain properties of the adjacency matrix of the graph, using attention. Wang et al. (2022b) use contrastive learning to generate a population-level graph from fMRI data using multiple views of the fMRI scans of every subject. They use partial correlations between ROIs to learn the graph structure that represents the fMRI data. Zhao et al. (2022) dynamically generate the graph structure by aggregating calculated k-NN graphs during training. This method retains more comprehensive non-local information diffusion compared to the proximity derived from a fixed input space.

Garcia-Uceda Juarez et al. (2019) use GNNs for improved CNN feature representation, by placing a GNN as the bottleneck layer inside a U-Net (Ronneberger et al., 2015). Features from the last encoder layer corresponding to encoded super-voxels serve as nodes connected either through regular grid connectivity, or based on the nearest neighbours in the feature space.

4.4.2 Latent graph learning

Here, we summarise methods that specifically learn the graph structure in an end-to-end manner. This requires the adjacency matrix of the graph to be differentiable. When using static graph construction methods, the adjacency matrix A is often a binary discrete matrix, where an entry $a_{ij} = 1$ indicates that there is a connection from node i to node j. When using latent graph creation methods, A needs to be continuous, to allow backpropagation through the adjacency matrix (Figure 6). This is usually implemented by a *fully connected graph* with continuous edge weights that are updated during learning.

The first works in this area used the spectral domain in order to learn the latent structure of the graph (Zhan et al., 2018; Li et al., 2018; Huang et al., 2018). These spectral methods require an initial graph which will then be changed over training and are limited to transductive learning settings since the extraction of the graph Laplacian requires the full graph including test and validation sets. Later, latent graph learning methods have also been transformed into the spatial domain, which also allows inductive training.

Cosmo et al. (2020) were the first to apply this approach in medical research for computer-aided diagnosis (CADx) and disease prediction. They could report improved performance when using a learned graph structure on two CADx problems (brain age prediction and Alzheimer's Disease prediction) compared to the approach of extracting a graph structure prior to training the GNN. In terms of graph structure learning, their method is classified as a "metric-based" GSL (Zhu et al., 2021). Mullakaeva et al. (2022) suggest extending the same graph learning method by adding additional loss to control the graph sparsity and applying it to a graph-in-graph learning task. Mo et al. (2021) explore a similar technique for multi-modal data integration of MR images that also utilises an end-to-end trainable adaptive graph learning method. Huang & Chung (2020) also utilise an adaptive graph learning mechanism for disease prediction with the

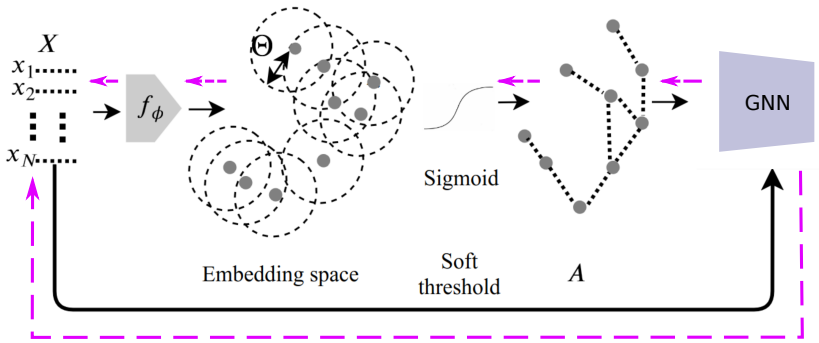


Figure 6: A schematic overview of a latent graph learning approach. Figure adapted from Cosmo et al. (2020); Kazi et al. (2022). The purple arrows indicate the backpropagation tracks through the networks. f_ϕ is an embedding network with learnable weights ϕ.

aid of population graphs. Their method includes a Monte-Carlo edge dropout for uncertainty estimation. This can be used to quantify the uncertainty of the edges learned in the graph creation method.

As an extension of Cosmo et al. (2020), Kazi et al. (2022) introduce the *differentiable graph module* (DGM). They propose two variants of their method with different sampling strategies: continuous and discrete. The latter has the advantage that a sparse graph can be sampled without the need for a fully connected adjacency matrix and requires the utilisation of the Gumbel-top-k trick (Kool et al., 2019) to ensure differentiability. In general, DGM can be applied based on an initial graph structure (prior-driven) or solely on the node features (data-driven). DGM uses Euclidean space to encode the latent features from which the latent graphs are decided. de Ocáriz Borde et al. (2023), extend the DGM to work on more complex embedding spaces by utilising Riemannian geometry.

One aspect in which static methods seem to show shortcomings in representing the full dimensionality of the data is when a temporal component is introduced. Functional brain connectivity networks e.g. fluctuate over time and can therefore benefit from an adaptive creation of the graph structure (Kim et al., 2021). We found five works that use similar approaches to learning the graph structure of functional brain connectivity networks end-to-end (Kan et al., 2022; Campbell et al., 2022; Mahmood et al., 2021; El-Gazzar et al., 2021; Zhu et al., 2022). In the method proposed by El-Gazzar et al. (2021), the adjacency matrix is randomly initialised, and its weights are adaptively learned along with the GNN weights by using gradient descent. Zhu et al. (2022) couple feature learning with dynamic graph learning into the GCN architecture. Campbell et al. (2022) and Mahmood et al. (2021) utilise self-attention in their model architecture. Kan et al. (2022) introduce loss components that enforce sparsity in the graph and maximise the similarity between adjacency matrices of the same group, while minimising similarity between differently labelled subjects.

There are some methods for adaptive graph construction, that have –to the best of our knowledge– so far not been applied to medical data that we consider promising methods from which medical data analysis with GNNs could benefit. Some examples are pointer graph networks (Velickovic et al., 2020), dynamic graph message passing networks (Zhang et al., 2022a), and deep heterophily graph rewiring (Bi et al., 2022).

5 Discussion and recommendations

In this section, we summarise the main advantages of static and adaptive graph creation methods (Table 3) and give recommendations for graph creation approaches. Furthermore, we provide an overview of graph metrics, that can be used to assess the generated graph structure.

5.1 Static vs. adaptive graph construction

Comparing Tables 1 and 2, we can see that fewer works utilise approaches that work with an adaptive graph structure creation, compared to the static approaches. In the following, we will compare both methods and discuss the shortcomings and benefits of both (Table 3).

Table 3: Overview of the main advantages and disadvantages of static and adaptive graph construction methods with respect to different aspects of the methods.

Aspect	**Static**	**Adaptive**
Training efficiency	Computationally more efficient during training	No pre-processing w.r.t graph creation required
Complexity of training	Easier to train (fewer (hyper-)parameters)	More difficult to train
Final graph structure	Generally applicable for different problems	Adjacency matrix is fine-tuned to the problem
Further utility of graph structure	Application independent graph structure	Adjacency matrix usable for interpretability
Generalisability of method	Different datasets require different methods	General method, dataset independent
Prior knowledge	Can be easily included in the graph	Some methods allow inclusion of prior knowledge

Static graph creation methods are more computationally efficient during training since they do not require updating the graph structure. They require fewer trainable parameters and are easier to train in general. Backpropagating through the adjacency matrix and fine-tuning the adjacency matrix during training (adaptive approach) adds additional complexity to the learning problem and can make training more difficult. On the other hand, the end-to-end learning of the graph structure eliminates critical design choices, since the adjacency matrix does not need to be defined prior to training.

Another advantage of adaptive graph creation methods is that the resulting adjacency matrix is fine-tuned to the specific task. This way, the most suitable adjacency matrix can be learned. This is not the case for static methods, where the edges are defined prior to learning. Whereas the latter makes the adjacency matrix more general and the same matrix can be used for different learning tasks on the same dataset.

Regarding the utility of the created graph structure beyond the specific downstream task, the adeptly generated adjacency matrix can also be used for interpretability purposes. Learned connections can indicate correlations between node components or highlight more important regions of the graph. Campbell et al. (2022) use their adaptive graph creation functional connectivity networks as a basis to identify sex-discriminate brain regions, which show higher importance within the decision-making progress of the algorithm.

5.2 Recommendations for graph construction

Based on the state-of-the-art literature on graph learning in medicine, we here formulate some recommendations for selecting a graph construction method for different applications.

We consider the choice between population-level, subject-level, knowledge graph, and graphs based on real-world structures to be determined by the application and the problem to solve. If a graph structure can be extracted for each subject individually, *subject-level graphs* are of use, whereas *population graphs* aim to represent each subject in a feature vector or matrix and include additional information by connecting the subjects into a graph. There are also works studying graph-in-graph problems, where subject-level graphs can be additionally connected to form a population-level graph (Mullakaeva et al., 2022). The authors show that nested graph approaches can improve performance. *Knowledge graphs* are usually generated when a large dataset of entities and connections between them is available. They can structure general medical knowledge, which can be used in several applications. They are often text-derived and combined with additional graph types. Creating a graph based on *real-world structures* seems reasonable whenever there is such a defined structure at hand, like when modelling molecules.

The most commonly used method to generate population graphs is based on similarity (Parisot et al., 2017) or distance (e.g. Euclidean) (Lu et al., 2022). In both settings, a subset of the available features can be used to define the edges of the graph (e.g. non-imaging features), while the remaining features (e.g. imaging features) can be used as node features (Parisot et al., 2017). This –in itself– implements an opportunity for multi-modal data integration. Later, it has also been shown that using all available features for both, creating the edges and the node features might be beneficial over splitting the features up (Keicher et al., 2021).

The most frequently used approach for generating brain graphs is based on correlation, and graphs from curvilinear structures are mostly extracted based on the underlying tree structure of the vessels/airways.

The choice between a *static* or *adaptive* graph construction method is more challenging and depends both on the application and its goal. Knowledge graphs are typically created statically, since here, the connections between entities (e.g. symptoms and diseases) are based on literature and probably do not benefit from end-to-end learning. The same holds for graphs constructed from real-world structures like molecules. Here, the graph structure is intrinsic to the data, and it would be unintuitive to change the edges (i.e. chemical bonds) of molecules. On the other hand, population graphs and some subject-level graphs have shown to benefit from an adaptive graph construction (Kazi et al., 2022; Campbell et al., 2022; Kan et al., 2022). They can, for example, be used to incorporate a temporal component into the learning process, like with fMRI data. It can also be a suitable method to altogether avoid the necessity of picking the *right* graph creation method before training since no specific metric or feature selection is required in this case. The main advantages and shortcomings of adaptive and static graph creation methods are summarised in Section 5.1. We believe that choosing an adaptive graph construction method is especially beneficial in settings like population graphs, where neither an initial graph structure nor a ground truth is available. We believe that this is a promising area for future research. Especially, the possibility of using the learned graphs to gain insight into the dataset at hand for interpretability purposes shows high potential for further investigation.

Regardless of the graph construction method of choice, we highly recommend evaluating graph assessment metrics to see whether the graph structure is beneficial to the learning task or might hinder GNN performance. This is especially useful when using static graph construction methods so that the graph structure can be judged prior to training. In the next section, we give an overview of such metrics.

5.3 Graph assessment

We investigated different methods to construct a graph structure from medical data and have evaluated the advantages and disadvantages of different methods. Now we will discuss methods that can be used to assess the constructed graph in terms of similarities in neighbourhoods. This can be done before or during model training, depending on the utilisation of static or dynamic methods. A number of graph assessment metrics have been introduced to evaluate the impact of the graph structure on the learning of GNNs. Table 4 summarises the different metrics as well as a short summary of their targets.

Table 4: Graph assessment metrics that have been shown to be correlated to GNN performance and can be used to evaluate the generated graph structure;

Graph Assessment Metric	**Reference**	**Details**
Edge homophily	Zhu et al. (2020)	Ratio between edges connecting same and differently labelled nodes
Node homophily	Pei et al. (2020)	Average amount of neighbours with the same label
Class homophily	Lim et al. (2021)	Edge homophily with consideration for class imbalance
Normalized total variation (NTV)	Luan et al. (2022)	Variation of the graph signal w.r.t. graph filters
Normalized smoothness value (NSV)		Effect of the edge bias
Neighbourhood entropy	Xie et al. (2020)	Similarity of node embeddings within neighbourhoods
Center-neighbour similarity		(Dis)similarity of the neighbours of a node
Aggregation similarity score	Luan et al. (2021)	Proportion of similarity weights in neighbourhoods after aggregation
Diversification distinguishability		Proportion of nodes that benefit from diversification operations
Cross-class neighbourhood similarity	Ma et al. (2022)	Similarity measure for neighbourhoods

Edge homophily (Zhu et al., 2020) (defined in Section 2.3) is one of the most commonly used metrics known to be correlated with the performance of GNNs. It has been shown that several graph convolutions do not perform well on heterophilic graph structures. However, homophily is not a *necessary* property for successful graph learning and heterophily affects different graph models at different rates (Zhu et al., 2020). Zhu et al. (2020) identify some critical design choices that can improve the performance of GNNs on heterophilous graphs. They show that a separate embedding of neighbourhood node features and node-internal features improves the performance of heterophilous graphs as well as a separate embedding of

higher-order neighbourhoods and introduces a new model architecture that works well on both homophilous and heterophilous graphs.

Additional metrics have been introduced that have shown to impact the performance of GNNs (Luan et al., 2021; Ma et al., 2022; Luan et al., 2022). Luan et al. (2022) discuss in which settings GNNs are beneficial and when GNNs under-perform other ML models and introduce two metrics called normalized total variation (NTV) and normalized smoothness value (NSV) to measure the effect of edge bias. Xie et al. (2020) decide when neighbourhood aggregation may be unnecessary by evaluating the so-called neighbourhood entropy and centre-neighbour similarity in graphs. Luan et al. (2021) assess graph statistics with the aid of an aggregation similarity score and a diversification distinguishability metric, mainly capturing the linear relations of the aggregated node features. Ma et al. (2022) investigate the similarity in neighbourhoods of same-labelled nodes and call their metric cross-class neighbourhood similarity.

We advise using these graph assessment metrics ideally before (and for adaptive methods also during) model training. They can shed light on the "quality" of the constructed graph and can potentially reveal reasons for poor performance.

6 Related work

In this section, we put our work in context with existing research. Table 5 summarises adjacent works in similar fields and more information about GNNs and their application in different areas of medicine.

Table 5: Summary of recent reviews in connected areas with their specific application areas, the year of publication, and the reviewed methods. *Categorisation methodology* lists the main categorisation schemes followed in the respective works. GSP refers to graph signal processing, and AE to auto encoders.

Domain	**Reference**	**Year**	**Reviewed methods**	**Categorisation methodology**
Medical	Ahmedt-Aristizabal et al. (2021)	2021	GNNs	Model architecture, application
Medical imaging	Ding et al. (2022)	2022	GNNs	medical image type
Bioinformatics	Yi et al. (2022)	2021	Graph embedding, GNNs	Algorithm, application levels
Bioinformatics	Zhang et al. (2021a)	2020	GNNs	Application, prediction task
Neuroscience	Bessadok et al. (2022)	2020	GNNs	Application, loss function Architecture, graph creation
Biological data	Li et al. (2021a)	2022	GNNs, GSP	Application, feature extraction
Medical	Ours	2023	GNNs	Graph construction, graph type
Knowledge graphs	Ye et al. (2022)	2022	GNNs	Task, application
Recommender systems	Wu et al. (2022)	2022	GNNs	Type of information used
Fault analysis	Chen et al. (2021)	2021	GNNs	Application, data representation
Finance	Wang et al. (2021)	2021	GNNs	Graph construction, feature extraction
Text classification	Pham et al. (2022)	2022	GNNs, RNNs CNN, AE	Text to graph transformation, Model architectures
Domain invariant	Zhu et al. (2021)	2022	GNNs	Graph structure learning type

Reviews on GNNs on medical data Ahmedt-Aristizabal et al. (Ahmedt-Aristizabal et al., 2021) give a detailed overview of GNN training on medical data, organised by application type and graph architectures. The authors investigate different graph creation methods, however, putting them mostly into context with varying application areas. They specifically name graph creation as one of the open challenges in research in this area. For an overview of works covering GNNs, different graph embedding techniques (homogeneous, heterogeneous, and attributed), and generative graph models, we refer to Yi et al. (2022), who target works in bioinformatics. Zhang et al. (2021a) focus on GNNs in bioinformatics, and Bessadok et al. (2022) give an overview of GNNs in network neuroscience. The latter also categorise existing works by graph creation methods; however only refer to research in neuroscience, and summarise different loss functions for different applications. For a review of graph signal processing and graph learning on biological data, we refer to Li et al. (2021a). The authors also provide a summary of feature extraction techniques in the graph domain. Ding et al. (2022) give an overview of applications of GNN methods to multi-modal medical imaging datasets.

Graph construction in other domains Graph construction can be a challenging aspect of graph deep learning pipelines, regardless of the domain and application. Therefore, we list some works investigating graph creation methods in different domains.

For *recommender systems*, Wu et al. (2022) provide an extensive survey on GNN applications, including an elaboration of the graph creation methods in this area. However, they only focus on the post-processing of the graph structure before training since the initial graph structure is usually always provided by the bipartite user-item graph - which differs from several medical datasets. Han et al. (2022) highlight the difficulties of creating a graph for *retrosynthetic planning* with GNNs and propose a semi-dynamic approach for connecting different molecules with each other. Different graph construction methods for association graphs for *fault analysis* are discussed in Chen et al. (2021). Here, the authors categorise the graph construction methods into (a) using the k-nearest neighbour method, (b) using prior knowledge, and (c) using matrix completion. This categorisation is similar to the one we use in this review. However, we add the additional component of adaptive graph construction. For a review on GNNs in *finance*, including different graph construction methods, we refer to Wang et al. (2021). They distinguish three graph creation methods: data-based, knowledge-based, and similarity-based. This categorisation is similar to the one utilised in our work. However, we provide a more detailed methodological distinction between the graph creation methods. Pham et al. (2022) give an overview of graph creation methods for text analysis, categorising existing works based on text graphs and model architectures. In Zhu et al. (2021), the authors investigate general methods for graph structure learning, limited to the here termed "adaptive" graph creation methods.

7 Open challenges and future directions

In this section, we discuss open challenges that arise when working with GNNs in medicine, linked to graph construction.

7.1 GNNs vs. graph agnostic models

As mentioned in Section 1.1, recent works have shown that GNNs do not consistently outperform graph agnostic models (Luan et al., 2022). In most of the here summarised works, the original data does not come naturally in graph structure. Therefore, one could raise the question of when it is beneficial to construct a graph and perform graph learning at all. Most methodological works show that GNNs lead to improved performances compared to graph-agnostic baselines. However, sometimes GNNs show only little or no improvement over other ML methods. We see a detailed investigation of these connections to be a highly relevant area of future work.

7.2 Bias introduction

Another challenge that arises when choosing a graph creation method is the introduction of bias (Mehrabi et al., 2021). When creating a graph structure statically, every design choice (e.g. which features to select for graph creation and which to use as node features, the incorporation of prior knowledge) introduces (human) bias to the resulting graph structure. Sohn et al. (2015), for example, discuss the impact of the selection of ROIs for brain graph creation from fMRI data and argue that individual ROI selections would be more accurate. When using purely adaptive graph creation methods, no human bias will be introduced into the graph creation method. Here, the bias in the dataset will dominate the graph creation. If the available data is highly imbalanced, questions of fairness might arise. This is something we believe should be considered when selecting a graph creation method, and we believe this to be an interesting area of research.

7.3 Undirected vs. directed graphs

Most approaches discussed in this work have been applied to undirected graphs, where edges are always bidirectional and only a few works discuss methods that work on directed graphs (Keicher et al., 2021). The effects of using directed versus undirected graphs have not been studied in detail. Changing an undirected graph to a directed one could be an extension of some of the here mentioned methods for graph construction.

7.4 Availability of public datasets

One significant challenge in exploring graph construction methods on medical data is the shortage of publicly available datasets. For population graphs, two commonly used public datasets are TADPOLE (Yu et al., 2020), a subset of the ADNI dataset , and the autism brain imaging data exchange (ABIDE) (Di Martino et al., 2014), which address Alzheimer's disease and autism detection, respectively. The private dataset UK Biobank (Sudlow et al., 2015) is also frequently used. The open graph benchmark (Hu et al., 2020) holds a set of benchmark datasets for graph learning. It also contains molecule datasets and a vessel graph dataset from Paetzold et al. (2021). For the construction of brain connectivity graphs, the human connectome project (HCP) (Elam et al., 2021) and the brain connectivity challenge dataset [2] have been used apart from several private datasets. For surface mesh representations, MedShapeNet Scharinger et al. (2023) offers a variety of medical mesh structures.

7.5 Data privacy

Medical research is usually performed on private data such as medical images or patient records, which hold highly sensitive information about patients. A well-established, formal method to allow DL while giving privacy guarantees to individuals is Differential Privacy (DP) (Dwork et al., 2014). It ensures that the output of a randomised algorithm is approximately the same, independent of a single data point being in the dataset or not. It has been shown that guaranteeing DP in GNNs is more challenging than in tabular data since the individual data points (nodes) in a graph are inter-connected (Mueller et al., 2023). Furthermore, graph-structured data is more sensitive to privacy attacks like Membership Inference Attacks (MIAs), which aim to identify whether a certain data point was part of the training dataset. This is due to the fact that additional information lies in the graph structure itself and can be leveraged by adversaries (Olatunji et al., 2021). We see a requirement for the development of high-utility privacy-preserving techniques for graph learning in medicine.

8 Conclusion

This in-depth review of state-of-the-art works on graph creation methods for medical data shows that graph construction is challenging and requires various design choices as well as a careful consideration of the dataset and task at hand. There are numerous ways to construct a suitable graph structure from a dataset and the best method needs to be selected with caution. We categorise graph construction methods by *static* or *adaptive* approaches. Static graph construction methods generate a graph structure prior to learning, whereas adaptive methods change the graph over the course of GNN training. We analyse advantages and disadvantages of both approaches and formulate recommendations about how to pick a suitable graph construction method.

So far, adaptive methods have only been applied to a small subset of graph-learning tasks in medicine. We believe that this will be explored further in the coming years. Especially, the post-hoc interpretability of the learned adjacency matrix can hold valuable information about the dataset and task, which might not be possible to be extracted from a statically generated graph.

Given the findings about the strong impact of the graph structure on GNN performance (Section 1.1 and 5.3), we want to raise awareness that in cases where the graph structure is not clearly defined by the dataset, it is important to consider different graph creation methods as well as suitable graph convolutions. One characteristic of static graph creation methods is that the graph structure is defined prior to training. This opens up the possibility of evaluating the graph structure based on the metrics summarised in Section 5.3. We believe this to be a valuable step in evaluating the generated graph structure. This way, for example, the homophily of the graph can be analysed and the potential need for graph convolutions that can handle heterophilous graph structures can be evaluated (Du et al., 2022; Zhu et al., 2020). This could lead to graph learning methods that are more suitably linked to the constructed/available graph structures. These graph metrics could also potentially be incorporated into the end-to-end learning of adaptive graph structures,

[2] miccai.brainconnectivity.net

where they could be included in the loss function to push the generated graph in the direction of a specific metric.

Given the various options to construct graphs from medical data, we think that explicit explanations of the applied graph creation method in each work would be beneficial to the research community. Even though the current state-of-the-art results summarised in this work indicate several recommendations for graph-creation methods on medical data, we believe that there are still a lot of open questions regarding which graph-creation method works best for which application. Furthermore, we consider a more thorough investigation of why a certain method might be beneficial over others to be necessary to learn more about the fundamental methodologies of GNNs and how they can be applied in the most favourable ways. We hope to contribute to further research in this area with this review.

References

Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Süsstrunk. Slic superpixels compared to state-of-the-art superpixel methods. *IEEE transactions on pattern analysis and machine intelligence*, 34(11):2274–2282, 2012.

David Ahmedt-Aristizabal, Mohammad Ali Armin, Simon Denman, Clinton Fookes, and Lars Petersson. Graph-based deep learning for medical diagnosis and analysis: past, present and future. *Sensors*, 21(14): 4758, 2021.

N Assempour, I Iynkkaran, YF Liu, A Maciejewski, N Gale, A Wilson, L Chin, R Cummings, D Le, A Pon, et al. 5.0: a major update to the drugbank database for 2018. *Nucleic Acids Res*, 46:D1074–D1082, 2018.

Sorin Avram, Cristian G Bologa, Jayme Holmes, Giovanni Bocci, Thomas B Wilson, Dac-Trung Nguyen, Ramona Curpan, Liliana Halip, Alina Bora, Jeremy J Yang, et al. Drugcentral 2021 supports drug discovery and repositioning. *Nucleic acids research*, 49(D1):D1160–D1169, 2021.

Emanuel A Azcona, Pierre Besson, Yunan Wu, Arjun Punjabi, Adam Martersteck, Amil Dravid, Todd B Parrish, S Kathleen Bandt, and Aggelos K Katsaggelos. Interpretation of brain morphology in association to alzheimer's disease dementia classification using graph convolutional networks on triangulated meshes. In *Shape in Medical Imaging: International Workshop, ShapeMI 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, Proceedings*, pp. 95–107. Springer, 2020.

Pietro Barbiero, Ramon Vinas Torne, and Pietro Lió. Graph representation forecasting of patient's medical conditions: Toward a digital twin. *Frontiers in genetics*, 12:652907, 2021.

Alaa Bessadok, Mohamed Ali Mahjoub, and Islem Rekik. Graph neural networks in network neuroscience. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2022.

Wendong Bi, Lun Du, Qiang Fu, Yanlin Wang, Shi Han, and Dongmei Zhang. Make heterophily graphs better fit gnn: A graph rewiring approach. *arXiv preprint arXiv:2209.08264*, 2022.

Stephen Bonner, Ian P Barrett, Cheng Ye, Rowan Swiers, Ola Engkvist, Andreas Bender, Charles Tapley Hoyt, and William L Hamilton. A review of biomedical datasets relating to drug discovery: a knowledge graph perspective. *Briefings in Bioinformatics*, 23(6), 2022.

Michael M Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4):18–42, 2017.

Alexander Campbell, Antonio Giuliano Zippo, Luca Passamonti, Nicola Toschi, and Pietro Lio. DBGSL: Dynamic brain graph structure learning. *arXiv:2209.13513*, 2022.

Payal Chandak, Kexin Huang, and Marinka Zitnik. Building a knowledge graph to enable precision medicine. *Scientific Data*, 10(1):67, 2023.

Chun-Hung Chao, Zhuotun Zhu, Dazhou Guo, Ke Yan, Tsung-Ying Ho, Jinzheng Cai, Adam P Harrison, Xianghua Ye, Jing Xiao, Alan Yuille, et al. Lymph node gross tumor volume detection in oncology imaging via relationship learning using graph neural network. In *MICCAI*, pp. 772–782, 2020.

Lang Chen, Yangmin Huang, Bin Liao, Kun Nie, Shoubin Dong, and Jinlong Hu. Graph learning approaches for graph with noise: application to disease prediction in population graph. In *BIBM*, pp. 2724–2729. IEEE, 2020a.

Li Chen, Thomas Hatsukami, Jenq-Neng Hwang, and Chun Yuan. Automated intracranial artery labeling using a graph neural network and hierarchical refinement. In *MICCAI*, pp. 76–85, 2020b.

Zhiwen Chen, Jiamin Xu, Cesare Alippi, Steven X Ding, Yuri Shardt, Tao Peng, and Chunhua Yang. Graph neural network-based fault diagnosis: a review. *arXiv:2111.08185*, 2021.

Binjie Cheng, Jin Zhang, Hong Liu, Meiling Cai, and Ying Wang. Research on medical knowledge graph for stroke. *Journal of Healthcare Engineering*, 2021, 2021a.

Yinlin Cheng, Mengnan Ma, Xingyu Li, and Yi Zhou. Multi-label classification of fundus images based on graph convolutional network. *BMC Medical Informatics and Decision Making*, 21(2):1–9, 2021b.

Luca Cosmo, Anees Kazi, Seyed-Ahmad Ahmadi, Nassir Navab, and Michael Bronstein. Latent-graph learning for disease prediction. In *MICCAI*, pp. 643–653, 2020.

Ameya Daigavane, Balaraman Ravindran, and Gaurav Aggarwal. Understanding convolutions on graphs. *Distill*, 2021. doi: 10.23915/distill.00032.

Haitz Sáez de Ocáriz Borde, Anees Kazi, Federico Barbero, and Pietro Lio. Latent graph inference using product manifolds. In *ICLR*, 2023.

Andac Demir, Toshiaki Koike-Akino, Ye Wang, Masaki Haruna, and Deniz Erdogmus. EEG-GNN: Graph neural networks for classification of electroencephalogram (eeg) signals. In *EMBC*, pp. 1061–1067, 2021.

Adriana Di Martino, Chao-Gan Yan, Qingyang Li, Erin Denio, Francisco X Castellanos, Kaat Alaerts, Jeffrey S Anderson, Michal Assaf, Susan Y Bookheimer, Mirella Dapretto, et al. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular psychiatry*, 19(6):659–667, 2014.

Alexander Dilthey, Charles Cox, Zamin Iqbal, Matthew R Nelson, and Gil McVean. Improved genome inference in the mhc using a population reference graph. *Nature genetics*, 47(6):682–688, 2015.

Kexin Ding, Mu Zhou, Zichen Wang, Qiao Liu, Corey W Arnold, Shaoting Zhang, and Dimitri N Metaxas. Graph convolutional networks for multi-modality medical imaging: Methods, architectures, and clinical applications. *arXiv:2202.08916*, 2022.

Dominik Drees, Aaron Scherzinger, René Hägerling, Friedemann Kiefer, and Xiaoyi Jiang. Scalable robust graph and feature extraction for arbitrary vessel networks in large volumetric datasets. *BMC bioinformatics*, 22(1):1–28, 2021.

Lun Du, Xiaozhou Shi, Qiang Fu, Xiaojun Ma, Hengyu Liu, Shi Han, and Dongmei Zhang. GBK-GNN: Gated bi-kernel graph neural networks for modeling both homophily and heterophily. In *WWW*, pp. 1550–1558, 2022.

Shaoyi Du, Yanrong Guo, Gerard Sanroma, Dong Ni, Guorong Wu, and Dinggang Shen. Building dynamic population graph for accurate correspondence detection. *Medical image analysis*, 26(1):256–267, 2015.

David K Duvenaud, Dougal Maclaurin, Jorge Iparraguirre, Rafael Bombarell, Timothy Hirzel, Alán Aspuru-Guzik, and Ryan P Adams. Convolutional networks on graphs for learning molecular fingerprints. *Advances in neural information processing systems*, 28, 2015.

Cynthia Dwork, Aaron Roth, et al. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3-4):211–407, 2014.

Ahmed El-Gazzar, Rajat Mani Thomas, and Guido van Wingen. Dynamic adaptive spatio-temporal graph convolution for fMRI modelling. In *MLCN@MICCAI*, pp. 125–134, 2021.

Yassine El Ouahidi, Hugo Tessier, Giulia Lioi, Nicolas Farrugia, Bastien Pasdeloup, and Vincent Gripon. Pruning graph convolutional networks to select meaningful graph frequencies for fMRI decoding. In *EUSIPCO*, pp. 937–941. IEEE, 2022.

Jennifer Stine Elam, Matthew F Glasser, Michael P Harms, Stamatios N Sotiropoulos, Jesper LR Andersson, Gregory C Burgess, Sandra W Curtiss, Robert Oostenveld, Linda J Larson-Prior, Jan-Mathijs Schoffelen, et al. The human connectome project: a retrospective. *NeuroImage*, 244:118543, 2021.

Ahmed ElGazzar, Rajat Thomas, and Guido Van Wingen. Benchmarking graph neural networks for fmri analysis. *arXiv:2211.08927*, 2022.

Wenqi Fan, Yao Ma, Qing Li, Yuan He, Eric Zhao, Jiliang Tang, and Dawei Yin. Graph neural networks for social recommendation. In *WWW*, pp. 417–426, 2019.

Matthias Fey and Jan E. Lenssen. Fast graph representation learning with PyTorch Geometric. In *ICLR Workshop on Representation Learning on Graphs and Manifolds*, 2019.

Bruce Fischl. Freesurfer. *NeuroImage*, 62(2):774–781, 2012. ISSN 1053-8119. doi: https://doi.org/10.1016/j.neuroimage.2012.01.021. URL https://www.sciencedirect.com/science/article/pii/S1053811912000389. 20 YEARS OF fMRI.

Wolfgang Förstner and Eberhard Gülch. A fast operator for detection and precise location of distinct points, corners and centres of circular features. In *Proc. ISPRS intercommission conference on fast processing of photogrammetric data*, volume 6, pp. 281–305. Interlaken, 1987.

Jianliang Gao, Tengfei Lyu, Fan Xiong, Jianxin Wang, Weimao Ke, and Zhao Li. Predicting the survival of cancer patients with multimodal graph neural network. *IEEE/ACM Trans. on Computational Biology and Bioinformatics*, 19(2):699–709, 2021.

Antonio Garcia-Uceda Juarez, Raghavendra Selvan, Zaigham Saghir, and Marleen de Bruijne. A joint 3d unet-graph neural network-based method for airway segmentation from chest cts. In *MLMI@MICCAI*, pp. 583–591, 2019.

Mohammed Amine Gharsallaoui, Furkan Tornaci, and Islem Rekik. Investigating and quantifying the reproducibility of graph neural networks in predictive medicine. In *PRIME@MICCAI*, pp. 104–116, 2021.

Mahsa Ghorbani, Mojtaba Bahrami, Anees Kazi, Mahdieh Soleymani Baghshah, Hamid R Rabiee, and Nassir Navab. Gkd: Semi-supervised graph knowledge distillation for graph-independent inference. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part V 24*, pp. 709–718. Springer, 2021.

Mahsa Ghorbani, Anees Kazi, Mahdieh Soleymani Baghshah, Hamid R Rabiee, and Nassir Navab. RA-GCN: Graph convolutional network for disease prediction problems with imbalanced data. *Medical Image Analysis*, 75:102272, 2022.

Karthik Gopinath, Christian Desrosiers, and Herve Lombaert. Graph convolutions on spectral embeddings for cortical surface parcellation. *Medical Image Analysis*, 54:297–305, 2019a. ISSN 1361-8415. doi: https://doi.org/10.1016/j.media.2019.03.012. URL https://www.sciencedirect.com/science/article/pii/S1361841518305243.

Karthik Gopinath, Christian Desrosiers, and Herve Lombaert. Adaptive graph convolution pooling for brain surface analysis. In Albert C. S. Chung, James C. Gee, Paul A. Yushkevich, and Siqi Bao (eds.), *Information Processing in Medical Imaging*, pp. 86–98, Cham, 2019b. Springer International Publishing.

Marco Gori, Gabriele Monfardini, and Franco Scarselli. A new model for learning in graph domains. In *Proc. 2005 IEEE Int. joint conf. on neural networks*, volume 2(2005), pp. 729–734, 2005.

Peng Han, Peilin Zhao, Chan Lu, Junzhou Huang, Jiaxiang Wu, Shuo Shang, Bin Yao, and Xiangliang Zhang. Gnn-retro: Retrosynthetic planning with graph neural networks. In *AAAI*, volume 36(4), pp. 4014–4021, 2022.

Si Hang. Tetgen, a delaunay-based quality tetrahedral mesh generator. *ACM Trans. Math. Softw*, 41(2):11, 2015.

Lasse Hansen and Mattias P Heinrich. GraphRegNet: Deep graph regularisation networks on sparse keypoints for dense registration of 3D lung CTs. *IEEE Trans. on Medical Imaging*, 40(9):2246–2257, 2021.

Junheng Hao, Chuan Lei, Vasilis Efthymiou, Abdul Quamar, Fatma Özcan, Yizhou Sun, and Wei Wang. Medto: Medical data to ontology matching using hybrid graph neural networks. In *Proc. of the 27th ACM SIGKDD Conf. on Knowledge Discovery & Data Mining*, pp. 2946–2954, 2021.

Yefei He, Tao Yang, Cheng Yang, and Hong Zhou. Integrated equipment for parkinson's disease early detection using graph convolution network. *Electronics*, 11(7):1154, 2022.

Mattias P Heinrich, Heinz Handels, and Ivor JA Simpson. Estimating large lung motion in COPD patients by symmetric regularised correspondence fields. In *MICCAI*, pp. 338–345, 2015.

Aidan Hogan, Eva Blomqvist, Michael Cochez, Claudia d'Amato, Gerard de Melo, Claudio Gutierrez, Sabrina Kirrane, José Emilio Labra Gayo, Roberto Navigli, Sebastian Neumaier, et al. Knowledge graphs. *ACM Computing Surveys (CSUR)*, 54(4):1–37, 2021.

Daibing Hou, Zijian Zhao, and Sanyuan Hu. Multi-label learning with visual-semantic embedded knowledge graph for diagnosis of radiology imaging. *IEEE Access*, 9:15720–15730, 2021.

Weihua Hu, Matthias Fey, Marinka Zitnik, Yuxiao Dong, Hongyu Ren, Bowen Liu, Michele Catasta, and Jure Leskovec. Open graph benchmark: Datasets for machine learning on graphs. *Advances in neural information processing systems*, 33:22118–22133, 2020.

Wenbing Huang, Tong Zhang, Yu Rong, and Junzhou Huang. Adaptive sampling towards fast graph representation learning. *Advances in neural information processing systems*, 31, 2018.

Yongxiang Huang and Albert CS Chung. Edge-variational graph convolutional networks for uncertainty-aware disease prediction. In *MICCAI*, pp. 562–572, 2020.

Ruidong Jin, Tianqi Xia, Xin Liu, Tsuyoshi Murata, and Kyoung-Sook Kim. Predicting emergency medical service demand with bipartite graph convolutional networks. *Ieee Access*, 9:9903–9915, 2021.

Xuan Kan, Hejie Cui, Joshua Lukemire, Ying Guo, and Carl Yang. FBNETGEN: Task-aware gnn-based fMRI analysis via functional brain network generation. In *MIDL*, 2022.

Anees Kazi, Shayan Shekarforoush, S Arvind Krishna, Hendrik Burwinkel, Gerome Vivar, Karsten Kortüm, Seyed-Ahmad Ahmadi, Shadi Albarqouni, and Nassir Navab. InceptionGCN: receptive field aware graph convolutional network for disease prediction. In *Proc. IPMI 2019, June 2–7*, pp. 73–85. Springer, 2019.

Anees Kazi, Luca Cosmo, Seyed-Ahmad Ahmadi, Nassir Navab, and Michael M Bronstein. Differentiable graph module (DGM) for graph convolutional networks. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 45(2):1606–1617, 2022.

Steven Kearnes, Kevin McCloskey, Marc Berndl, Vijay Pande, and Patrick Riley. Molecular graph convolutions: moving beyond fingerprints. *Journal of computer-aided molecular design*, 30:595–608, 2016.

Matthias Keicher, Hendrik Burwinkel, David Bani-Harouni, Magdalini Paschali, Tobias Czempiel, Egon Burian, Marcus R Makowski, Rickmer Braren, Nassir Navab, and Thomas Wendler. U-gat: Multimodal graph attention network for covid-19 outcome prediction. *arXiv:2108.00860*, 2021.

Byung-Hoon Kim, Jong Chul Ye, and Jae-Jin Kim. Learning dynamic graph representation of brain connectome with spatio-temporal attention. *Advances in Neural Information Processing Systems*, 34:4314–4327, 2021.

Dongkwan Kim and Alice Oh. How to find your friendly neighborhood: Graph attention design with self-supervision. In *ICLR*, 2021.

Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv:1609.02907*, 2016.

Dominik Klepl, Fei He, Min Wu, Daniel J Blackburn, and Ptolemaios Sarrigiannis. EEG-based graph neural network classification of alzheimer's disease: An empirical evaluation of functional connectivity methods. *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, 30:2651–2660, 2022.

Youyong Kong, Shuwen Gao, Yingying Yue, Zhenhua Hou, Huazhong Shu, Chunming Xie, Zhijun Zhang, and Yonggui Yuan. Spatio-temporal graph convolutional network for diagnosis and treatment response prediction of major depressive disorder from functional connectivity. *Human brain mapping*, 42(12):3922–3933, 2021.

Youyong Kong, Shuyi Niu, Heren Gao, Yingying Yue, Huazhong Shu, Chunming Xie, Zhijun Zhang, and Yonggui Yuan. Multi-stage graph fusion networks for major depressive disorder diagnosis. *IEEE Trans. on Affective Computing*, 13(4):1917–1928, 2022.

Wouter Kool, Herke Van Hoof, and Max Welling. Stochastic beams and where to find them: The gumbel-top-k trick for sampling sequences without replacement. In *ICML*, pp. 3499–3508, 2019.

Rui Li, Xin Yuan, Mohsen Radfar, Peter Marendy, Wei Ni, Terence J O'Brien, and Pablo M Casillas-Espinosa. Graph signal processing, graph neural network and graph learning on biological data: a systematic review. *IEEE Reviews in Biomedical Engineering*, 2021a.

Ruikun Li, Yi-Jie Huang, Huai Chen, Xiaoqing Liu, Yizhou Yu, Dahong Qian, and Lisheng Wang. 3d graph-connectivity constrained network for hepatic vessel segmentation. *IEEE Journal of Biomedical and Health Informatics*, 26(3):1251–1262, 2021b.

Ruoyu Li, Sheng Wang, Feiyun Zhu, and Junzhou Huang. Adaptive graph convolutional neural networks. In *Proc. of the AAAI conference on artificial intelligence*, volume 32(1), 2018.

Xiaoxiao Li, Yuan Zhou, Nicha Dvornek, Muhan Zhang, Siyuan Gao, Juntang Zhuang, Dustin Scheinost, Lawrence H Staib, Pamela Ventola, and James S Duncan. Braingnn: Interpretable brain graph neural network for fmri analysis. *Medical Image Analysis*, 74:102233, 2021c.

Derek Lim, Xiuyu Li, Felix Hohne, and Ser-Nam Lim. New benchmarks for learning on non-homophilous graphs. *arXiv:2104.01404*, 2021.

Lan Lin, Min Xiong, Ge Zhang, Wenjie Kang, Shen Sun, Shuicai Wu, and Initiative Alzheimer's Disease Neuroimaging. A convolutional neural network and graph convolutional network based framework for ad classification. *Sensors*, 23(4), 2023. ISSN 1424-8220. doi: 10.3390/s23041914. URL `https://www.mdpi.com/1424-8220/23/4/1914`.

Fenglin Liu, Chenyu You, Xian Wu, Shen Ge, Xu Sun, et al. Auto-encoding knowledge graph for unsupervised medical report generation. *Advances in Neural Information Processing Systems*, 34:16266–16279, 2021.

William E Lorensen and Harvey E Cline. Marching cubes: A high resolution 3d surface construction algorithm. In *Seminal graphics: pioneering efforts that shaped the field*, pp. 347–353. 1998.

Siyuan Lu, Ziquan Zhu, Juan Manuel Gorriz, Shui-Hua Wang, and Yu-Dong Zhang. NAGNN: classification of covid-19 based on neighboring aware representation from deep graph neural network. *Int. Journal of Intelligent Systems*, 37(2):1572–1598, 2022.

Sitao Luan, Chenqing Hua, Qincheng Lu, Jiaqi Zhu, Mingde Zhao, Shuyuan Zhang, Xiao-Wen Chang, and Doina Precup. Is heterophily a real nightmare for graph neural networks to do node classification? *arXiv:2109.05641*, 2021.

Sitao Luan, Chenqing Hua, Qincheng Lu, Jiaqi Zhu, Xiao-Wen Chang, and Doina Precup. When do we need gnn for node classification? *arXiv:2210.16979*, 2022.

Yao Ma, Xiaorui Liu, Neil Shah, and Jiliang Tang. Is homophily a necessity for graph neural networks? In *ICLR*, 2022.

Donna Maglott, Jim Ostell, Kim D Pruitt, and Tatiana Tatusova. Entrez gene: gene-centered information at NCBI. *Nucleic acids research*, 39(suppl_1):D52–D57, 2010.

Dwarikanath Mahapatra, Steven Korevaar, Behzad Bozorgtabar, and Ruwan Tennakoon. Unsupervised domain adaptation using feature disentanglement and gcns for medical image classification. In *ECCV Workshops*, pp. 735–748, 2022.

Ines Mahjoub, Mohamed Ali Mahjoub, and Islem Rekik. Brain multiplexes reveal morphological connectional biomarkers fingerprinting late brain dementia states. *Scientific reports*, 8(1):1–14, 2018.

Usman Mahmood, Zening Fu, Vince D Calhoun, and Sergey Plis. A deep learning model for data-driven discovery of functional connectivity. *Algorithms*, 14(3):75, 2021.

Chengsheng Mao, Liang Yao, and Yuan Luo. MedGCN: Medication recommendation and lab test imputation via graph convolutional networks. *Journal of Biomedical Informatics*, 127:104000, 2022.

Miller McPherson, Lynn Smith-Lovin, and James M Cook. Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1):415–444, 2001. doi: 10.1146/annurev.soc.27.1.415. URL `https://doi.org/10.1146/annurev.soc.27.1.415`.

Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. A survey on bias and fairness in machine learning. *ACM Computing Surveys (CSUR)*, 54(6):1–35, 2021.

Shaocong Mo, Ming Cai, Lanfen Lin, Ruofeng Tong, Qingqing Chen, Fang Wang, Hongjie Hu, Yutaro Iwamoto, Xian-Hua Han, and Yen-Wei Chen. Mutual information-based graph co-attention networks for multimodal prior-guided magnetic resonance imaging segmentation. *IEEE Trans. on Circuits and Systems for Video Technology*, 32(5):2512–2526, 2021.

José Teófilo Moreira-Filho, Meryck Felipe Brito da Silva, Joyce Villa Verde Bastos Borba, Arlindo Rodrigues Galvão Filho, Eugene Muratov, Carolina Horta Andrade, Rodolpho de Campos Braga, and Bruno Junior Neves. Artificial intelligence systems for the design of magic shotgun drugs. *Artificial Intelligence in the Life Sciences*, pp. 100055, 2022.

Tamara T Mueller, Johannes C Paetzold, Chinmay Prabhakar, Dmitrii Usynin, Daniel Rueckert, and Georgios Kaissis. Differentially private graph classification with gnns. *arXiv:2202.02575*, 2022a.

Tamara T. Mueller, Johannes C. Paetzold, Chinmay Prabhakar, Dmitrii Usynin, Daniel Rueckert, and Georgios Kaissis. Differentially private graph neural networks for whole-graph classification. *IEEE Trans. Pattern Anal. Mach. Intell.*, 45(6):7308–7318, dec 2022b. ISSN 0162-8828. doi: 10.1109/TPAMI.2022.3228315. URL `https://doi.org/10.1109/TPAMI.2022.3228315`.

Tamara T Mueller, Dmitrii Usynin, Johannes C Paetzold, Daniel Rueckert, and Georgios Kaissis. Differential privacy guarantees for analytics and machine learning on graphs: A survey of results. *Journal of Privacy and Confidentiality (Accepted)*, 2023.

Kamilia Mullakaeva, Luca Cosmo, Anees Kazi, Seyed-Ahmad Ahmadi, Nassir Navab, and Michael M Bronstein. Graph-in-graph (GiG): Learning interpretable latent graphs in non-euclidean domain for biological and healthcare applications. *arXiv:2204.00323*, 2022.

Fuad Noman, Chee-Ming Ting, Hakmook Kang, Raphael C-W Phan, Brian D Boyd, Warren D Taylor, and Hernando Ombao. Graph autoencoders for embedding learning in brain networks and major depressive disorder identification. *arXiv:2107.12838*, 2021.

Iyiola E Olatunji, Wolfgang Nejdl, and Megha Khosla. Membership inference attack on graph neural networks. In *TPS-ISA*, pp. 11–20, 2021.

Johannes C Paetzold, Julian McGinnis, Suprosanna Shit, Ivan Ezhov, Paul Büschl, Chinmay Prabhakar, Anjany Sekuboyina, Mihail Todorov, Georgios Kaissis, Ali Ertürk, et al. Whole brain vessel graphs: A dataset and benchmark for graph learning and neuroscience. In *NeurIPS Datasets and Benchmarks*, 2021.

Li Pan, Jundong Liu, Mingqin Shi, Chi Wah Wong, and Kei Hang Katie Chan. Identifying autism spectrum disorder based on individual-aware down-sampling and multi-modal learning. *arXiv:2109.09129*, 2021.

Sarah Parisot, Sofia Ira Ktena, Enzo Ferrante, Matthew Lee, Ricardo Guerrerro Moreno, Ben Glocker, and Daniel Rueckert. Spectral graph convolutions for population-based disease prediction. In *MICCAI*, pp. 177–185, 2017.

Hongbin Pei, Bingzhe Wei, Kevin Chen-Chuan Chang, Yu Lei, and Bo Yang. Geom-GCN: Geometric graph convolutional networks. In *ICLR*, 2020.

Chantal Pellegrini, Nassir Navab, and Anees Kazi. Unsupervised pre-training of graph transformers on patient population graphs. *arXiv:2207.10603*, 2022.

Liang Peng, Nan Wang, Nicha Dvornek, Xiaofeng Zhu, and Xiaoxiao Li. Fedni: Federated graph learning with network inpainting for population-based disease prediction. *IEEE Trans. on Medical Imaging*, 2022.

Bastian Pfeifer, Afan Secic, Anna Saranti, and Andreas Holzinger. GNN-SubNet: disease subnetwork detection with explainable graph neural networks. *bioRxiv*, 2022.

Phu Pham, Loan TT Nguyen, Witold Pedrycz, and Bay Vo. Deep learning, graph-based text representation and classification: a survey, perspectives and challenges. *Artificial Intelligence Review*, pp. 1–35, 2022.

Yali Qiu, Shuangzhi Yu, Yanhong Zhou, Dongdong Liu, Xuegang Song, Tianfu Wang, and Baiying Lei. Multi-channel sparse graph transformer network for early alzheimer's disease identification. In *ISBI*, pp. 1794–1797. IEEE, 2021.

Gang Qu, Li Xiao, Wenxing Hu, Junqi Wang, Kun Zhang, Vince D Calhoun, and Yu-Ping Wang. Ensemble manifold regularized multi-modal graph convolutional network for cognitive ability prediction. *IEEE Trans. on Biomedical Engineering*, 68(12):3564–3573, 2021.

Zarina Rakhimberdina, Xin Liu, and Tsuyoshi Murata. Population graph-based multi-model ensemble method for diagnosing autism spectrum disorder. *Sensors*, 20(21):6001, 2020.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional networks for biomedical image segmentation. In *MICCAI*, pp. 234–241, 2015.

Chunyang Ruan, Yingpei Wu, Yun Yang, and Guangsheng Luo. Semantic-aware graph convolutional networks for clinical auxiliary diagnosis and treatment of traditional chinese medicine. *IEEE Access*, 9: 8797–8807, 2021.

Yasmin Salehi and Dennis Giannacopoulos. Physgnn: A physics–driven graph neural network based model for predicting soft tissue deformation in image-guided neurosurgery. *Advances in Neural Information Processing Systems*, 35:37282–37296, 2022.

Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. The graph neural network model. *IEEE trans. on neural networks*, 20(1):61–80, 2008.

Bernhard Scharinger, Antonio Pepe, Yuan Jin, Christina Gsaxner, Jianning Li, and Jan Egger. Multicenter aortic vessel tree extraction using deep learning. In *Medical Imaging 2023: Biomedical Applications in Molecular, Structural, and Functional Imaging*, volume 12468, pp. 341–347. SPIE, 2023.

Roman Schulte-Sasse, Stefan Budach, Denes Hnisz, and Annalisa Marsico. Integration of multiomics data with graph convolutional networks to identify new cancer genes and their associated molecular mechanisms. *Nature Machine Intelligence*, 3(6):513–526, 2021.

Raghavendra Selvan, Thomas Kipf, Max Welling, Antonio Garcia-Uceda Juarez, Jesper H Pedersen, Jens Petersen, and Marleen de Bruijne. Graph refinement based airway extraction using mean-field networks and graph neural networks. *Medical image analysis*, 64:101751, 2020.

Seung Yeon Shin, Soochahn Lee, Il Dong Yun, and Kyoung Mu Lee. Deep vessel segmentation by learning graphical connectivity. *Medical image analysis*, 58:101556, 2019.

William Sohn, Kwangsun Yoo, Young-Beom Lee, Sang Seo, Duk Na, and Yong Jeong. Influence of ROI selection on resting state functional connectivity: an individualized approach for resting state fmri analysis. *Frontiers in Neuroscience*, 9, 2015. doi: 10.3389/fnins.2015.00280. URL `https://www.frontiersin.org/articles/10.3389/fnins.2015.00280`.

Xiaofan Song, Mingyi Mao, and Xiaohua Qian. Auto-metric graph neural network based on a meta-learning strategy for the diagnosis of alzheimer's disease. *IEEE Journal of Biomedical and Health Informatics*, 25 (8):3141–3152, 2021.

Kamilė Stankevičiūtė, Tiago Azevedo, Alexander Campbell, Richard Bethlehem, and Pietro Liò. Population graph gnns for brain age prediction. *bioRxiv*, 2020.

Cathie Sudlow, John Gallacher, Naomi Allen, Valerie Beral, Paul Burton, John Danesh, Paul Downey, Paul Elliott, Jane Green, Martin Landray, et al. Uk biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS medicine*, 12(3):e1001779, 2015.

Li Sun, Ke Yu, and Kayhan Batmanghelich. Context matters: Graph-based self-supervised representation learning for medical images. In *Proc. AAAI*, volume 35(6), pp. 4874–4882, 2021.

Zimeng Tan, Jianjiang Feng, and Jie Zhou. Sgnet: Structure-aware graph-based network for airway semantic segmentation. In *MICCAI*, pp. 153–163, 2021.

Dinh V Tran, Nicolò Navarin, and Alessandro Sperduti. On filter size in graph convolutional networks. In *SSCI*, pp. 1534–1541. IEEE, 2018.

Petar Velickovic, Lars Buesing, Matthew C Overlan, Razvan Pascanu, Oriol Vinyals, and Charles Blundell. Pointer graph networks. *stat*, 1050:11, 2020.

Gerome Vivar, Anees Kazi, Hendrik Burwinkel, Andreas Zwergal, Nassir Navab, Seyed-Ahmad Ahmadi, et al. Simultaneous imputation and classification using multigraph geometric matrix completion (MGMC): Application to neurodegenerative disease classification. *Artificial Intelligence in Medicine*, 117:102097, 2021.

Alina Vretinaris, Chuan Lei, Vasilis Efthymiou, Xiao Qin, and Fatma Özcan. Medical entity disambiguation using graph neural networks. In *SIGMOD Conf.*, pp. 2310–2318, 2021.

Jianian Wang, Sheng Zhang, Yanghua Xiao, and Rui Song. A review on graph neural network methods in financial applications. *arXiv:2111.15367*, 2021.

Qianqian Wang, Long Li, Lishan Qiao, and Mingxia Liu. Adaptive multimodal neuroimage integration for major depression disorder detection. *Frontiers in Neuroinformatics*, 16, 2022a.

Xuesong Wang, Lina Yao, Islem Rekik, and Yu Zhang. Contrastive functional connectivity graph learning for population-based fmri classification. In *MICCAI*, pp. 221–230, 2022b.

Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E Sarma, Michael M Bronstein, and Justin M Solomon. Dynamic graph cnn for learning on point clouds. *Acm Trans. On Graphics (tog)*, 38(5):1–12, 2019.

Bastian Wittmann, Johannes C Paetzold, Chinmay Prabhakar, Daniel Rueckert, and Bjoern Menze. Link prediction for flow-driven spatial networks. *arXiv:2303.14501*, 2023.

Jelmer M Wolterink, Tim Leiner, and Ivana Išgum. Graph convolutional networks for coronary artery segmentation in cardiac ct angiography. In *GLMI@MICCAI*, pp. 62–69, 2019.

Jermer Wolterink and Julian Suk. Geometric deep learning for precision medicine. *Key enabling technology for scientific machine learning*, pp. 60, 2021.

Shiwen Wu, Fei Sun, Wentao Zhang, Xu Xie, and Bin Cui. Graph neural networks in recommender systems: a survey. *ACM Computing Surveys*, 55(5):1–37, 2022.

Zhengwang Wu, Fenqiang Zhao, Jing Xia, Li Wang, Weili Lin, John H. Gilmore, Gang Li, and Dinggang Shen. Intrinsic patch-based cortical anatomical parcellation using graph convolutional neural network on surface manifold. In Dinggang Shen, Tianming Liu, Terry M. Peters, Lawrence H. Staib, Caroline Essert, Sean Zhou, Pew-Thian Yap, and Ali Khan (eds.), *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, pp. 492–500, Cham, 2019. Springer International Publishing.

Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and S Yu Philip. A comprehensive survey on graph neural networks. *IEEE Trans. on neural networks and learning systems*, 32(1):4–24, 2020.

Weiyi Xie, Colin Jacobs, Jean-Paul Charbonnier, and Bram van Ginneken. Structure and position-aware graph neural network for airway labeling. *arXiv:2201.04532*, 2022.

Yiqing Xie, Sha Li, Carl Yang, Raymond Chi-Wing Wong, and Jiawei Han. When do gnns work: Understanding and improving neighborhood aggregation. In *Proc. {IJCAI}*, volume 2020(1), 2020.

Nuo Xu, Pinghui Wang, Long Chen, Jing Tao, and Junzhou Zhao. MR-GNN: multi-resolution and dual graph neural network for predicting structured entity interactions. In *IJCAI*, 2019.

Xiayu Xu, Peiwei Yang, Hualin Wang, Zhanfeng Xiao, Gang Xing, Xiulan Zhang, Wei Wang, Feng Xu, Jiong Zhang, and Jianqin Lei. Av-casnet: Fully automatic arteriole-venule segmentation and differentiation in oct angiography. *IEEE Trans. on Medical Imaging*, 2022.

Chunde Yang, Panyu Wang, Jia Tan, Qingshui Liu, and Xinwei Li. Autism spectrum disorder diagnosis using graph attention network based on spatial-constrained sparse functional brain networks. *Computers in Biology and Medicine*, 139:104963, 2021.

Dongren Yao, Jing Sui, Mingliang Wang, Erkun Yang, Yeerfan Jiaerken, Na Luo, Pew-Thian Yap, Mingxia Liu, and Dinggang Shen. A mutual multi-scale triplet graph convolutional network for classification of brain disorders using functional or structural connectivity. *IEEE transactions on medical imaging*, 40(4): 1279–1289, 2021.

Melissa Min-Szu Yao, Hao Du, Mikael Hartman, Wing P Chan, and Mengling Feng. End-to-end calcification distribution pattern recognition for mammograms: An interpretable approach with GNN. *Diagnostics*, 12 (6):1376, 2022.

Zi Ye, Yogan Jaya Kumar, Goh Ong Sing, Fengyan Song, and Junsong Wang. A comprehensive survey of graph neural networks for knowledge graphs. *IEEE Access*, 10:75729–75741, 2022.

Hai-Cheng Yi, Zhu-Hong You, De-Shuang Huang, and Chee Keong Kwoh. Graph representation learning in bioinformatics: trends, methods and applications. *Briefings in Bioinformatics*, 23(1):bbab340, 2022.

Hao Yu, Jie Zhao, and Li Zhang. Vessel segmentation via link prediction of graph neural networks. In *MMMI@MICCAI*, pp. 34–43, 2022a.

Jianhui Yu, Chaoyi Zhang, Heng Wang, Dingxin Zhang, Yang Song, Tiange Xiang, Dongnan Liu, and Weidong Cai. 3D medical point transformer: Introducing convolution to attention networks for medical point cloud analysis. *arXiv:2112.04863*, 2021a.

Renping Yu, Cong Pan, Xuan Fei, Mingming Chen, and Dinggang Shen. Multi-graph attention networks with bilinear convolution for diagnosis of schizophrenia. *IEEE Journal of Biomedical and Health Informatics*, 2023.

Shuangzhi Yu, Shuqiang Wang, Xiaohua Xiao, Jiuwen Cao, Guanghui Yue, Dongdong Liu, Tianfu Wang, Yanwu Xu, and Baiying Lei. Multi-scale enhanced graph convolutional network for early mild cognitive impairment detection. In *MICCAI*, pp. 228–237, 2020.

Weihao Yu, Hao Zheng, Yun Gu, Fangfang Xie, Jie Yang, Jiayuan Sun, and Guang-Zhong Yang. Tnn: Tree neural network for airway anatomical labeling. *IEEE Trans. on Medical Imaging*, 42(1):103–118, 2022b.

Xiang Yu, Siyuan Lu, Lili Guo, Shui-Hua Wang, and Yu-Dong Zhang. ResGNet-C: A graph convolutional neural network for detection of COVID-19. *Neurocomputing*, 452:592–605, 2021b.

Zehua Yu, Xianwei Zheng, Zhulun Yang, Bowen Lu, Xutao Li, and Maxian Fu. Interaction-temporal GCN: A hybrid deep framework for covid-19 pandemic analysis. *IEEE Open Journal of Engineering in Medicine and Biology*, 2:97–103, 2021c.

Kun Zhan, Xiaojun Chang, Junpeng Guan, Ling Chen, Zhigang Ma, and Yi Yang. Adaptive structure discovery for multimedia analysis using multiple features. *IEEE transactions on cybernetics*, 49(5):1826–1834, 2018.

Li Zhang, Mohan Chen, Anurag Arnab, Xiangyang Xue, and Philip HS Torr. Dynamic graph message passing networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(5):5712–5730, 2022a.

Xiao-Meng Zhang, Li Liang, Lin Liu, and Ming-Jing Tang. Graph neural networks and their current applications in bioinformatics. *Frontiers in genetics*, 12:690049, 2021a.

Yingying Zhang, Xian Wu, Quan Fang, Shengsheng Qian, and Chengsheng Xu. Knowledge-enhanced attributed multi-task learning for medicine recommendation. *ACM Trans. on Information Systems (TOIS)*, 2022b.

Yixiao Zhang, Xiaosong Wang, Ziyue Xu, Qihang Yu, Alan Yuille, and Daguang Xu. When radiology report generation meets knowledge graph. In *Proc. AAAI*, volume 34(07), pp. 12910–12917, 2020.

Yu Zhang, Loïc Tetrel, Bertrand Thirion, and Pierre Bellec. Functional annotation of human cognitive states using deep graph convolution. *NeuroImage*, 231:117847, 2021b.

Fenqiang Zhao, Zhengwang Wu, and Gang Li. Deep learning in cortical surface-based neuroimage analysis: a systematic review. *Intelligent Medicine*, 3(1):46–58, 2023. ISSN 2667-1026. doi: https://doi.org/10.1016/j.imed.2022.06.002. URL `https://www.sciencedirect.com/science/article/pii/S2667102622000493`.

Gangming Zhao. Cross chest graph for disease diagnosis with structural relational reasoning. In *Proc. ACM Multimedia*, pp. 612–620, 2021.

Kanhao Zhao, Boris Duka, Hua Xie, Desmond J Oathes, Vince Calhoun, and Yu Zhang. A dynamic graph convolutional neural network framework reveals new insights into connectome dysfunctions in ADHD. *NeuroImage*, 246:118774, 2022.

Tianyi Zhao and Zhaozheng Yin. Airway anomaly detection by prototype-based graph neural network. In *MICCAI*, pp. 195–204, 2021.

Tianyi Zhao, Yang Hu, Linda R Valsdottir, Tianyi Zang, and Jiajie Peng. Identifying drug–target interactions based on graph convolutional network and deep neural network. *Briefings in bioinformatics*, 22(2):2141–2150, 2021.

Kaizhong Zheng, Shujian Yu, Baojuan Li, Robert Jenssen, and Badong Chen. Brainib: Interpretable brain network-based psychiatric diagnosis with graph information bottleneck. *arXiv:2205.03612*, 2022a.

Shuai Zheng, Zhenfeng Zhu, Zhizhe Liu, Zhenyu Guo, Yang Liu, Yuchen Yang, and Yao Zhao. Multi-modal graph learning for disease prediction. *IEEE Trans. on Medical Imaging*, 41(9):2207–2216, 2022b.

Zhi Zheng, Chao Wang, Tong Xu, Dazhong Shen, Penggang Qin, Baoxing Huai, Tongzhu Liu, and Enhong Chen. Drug package recommendation via interaction-aware graph induction. In *Proc. WWW*, pp. 1284–1295, 2021.

Yi Zhou, Tianfei Zhou, Tao Zhou, Huazhu Fu, Jiacheng Liu, and Ling Shao. Contrast-attentive thoracic disease recognition with dual-weighting graph reasoning. *IEEE Trans. on Medical Imaging*, 40(4):1196–1206, 2021.

Jiong Zhu, Yujun Yan, Lingxiao Zhao, Mark Heimann, Leman Akoglu, and Danai Koutra. Beyond homophily in graph neural networks: Current limitations and effective designs. *Advances in Neural Information Processing Systems*, 33:7793–7804, 2020.

Yanqiao Zhu, Weizhi Xu, Jinghao Zhang, Yuanqi Du, Jieyu Zhang, Qiang Liu, Carl Yang, and Shu Wu. A survey on graph structure learning: Progress and opportunities. *arXiv e-prints*, pp. arXiv–2103, 2021.

Yonghua Zhu, Junbo Ma, Changan Yuan, and Xiaofeng Zhu. Interpretable learning based dynamic graph convolutional networks for alzheimer's disease analysis. *Information Fusion*, 77:53–61, 2022.