

Article

Graph Neural Networks in Medical Imaging: Methods, Applications and Future Directions

Ibomoiye Domor Mienye  and Serestina Viriri * 

Computer Science Discipline, School of Agriculture and Science, University of KwaZulu-Natal, Durban 4041, South Africa

* Correspondence: viriris@ukzn.ac.za

Abstract

Graph neural networks (GNNs) extend deep learning to non-Euclidean domains, offering a robust framework for modeling the spatial, structural, and functional relationships inherent in medical imaging. This paper reviews recent progress in GNN architectures, including recurrent, convolutional, attention-based, autoencoding, and spatiotemporal designs, and examines how these models have been applied to core medical imaging tasks, such as segmentation, classification, registration, reconstruction, and multimodal fusion. The review further identifies current challenges and limitations in applying GNNs to medical imaging and discusses emerging trends, including graph-transformer integration, self-supervised graph learning, and federated GNNs. This paper provides a concise and comprehensive reference for advancing reliable and generalizable GNN-based medical imaging systems.

Keywords: graph neural networks; deep learning; machine learning; medical imaging; segmentation; classification; multimodal fusion



Academic Editors: Sandeep Reddy, Mohamed Almorsy Abdelrazek and Varadraj Gurupur

Received: 13 October 2025

Revised: 20 November 2025

Accepted: 24 November 2025

Published: 1 December 2025

Citation: Mienye, I.D.; Viriri, S. Graph Neural Networks in Medical Imaging: Methods, Applications and Future Directions. *Information* **2025**, *16*, 1051. <https://doi.org/10.3390/info16121051>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Medical imaging has transformed modern healthcare by enabling non-invasive visualization of anatomical structures and physiological functions with high precision [1–3]. Modalities such as magnetic resonance imaging (MRI), computed tomography (CT), positron emission tomography (PET), and digital histopathology provide clinicians with essential insights for disease diagnosis, treatment planning, and monitoring. However, medical image analysis presents several computational challenges due to high dimensionality, heterogeneous data sources, and the need to capture complex structural and functional relationships [4–6]. Traditional deep learning approaches, while successful in many domains, typically assume Euclidean grid-like data, which limits their ability to fully exploit the inherent non-Euclidean relationships present in medical images.

Graph representations offer a natural way to capture these relationships by modeling pixels, superpixels, or anatomical regions as nodes, with edges encoding spatial, structural, or functional connections [7]. For example, in neuroimaging, brain regions extracted from MRI can be represented as nodes, with edges denoting their anatomical or functional connectivity, allowing GNNs to capture both local and global relationships among brain regions. Graph neural networks (GNNs), an important class of geometric deep learning models, extend deep learning to such graph-structured data and have demonstrated strong performance across a range of medical imaging tasks [8,9]. These include tumor and organ segmentation, disease classification, image reconstruction, and multimodal fusion. By

leveraging relational structures, GNNs capture complex inter-regional and cross-modal patterns, improving interpretability and robustness over conventional convolutional neural network (CNN) and transformer models.

Several reviews have attempted to consolidate progress in graph-based learning, but their coverage remains either too broad or fragmented for medical imaging. For example, Wu et al. [10] and Zhou et al. [11] provided broad taxonomies of GNN models and theoretical foundations, but with minimal discussion on medical imaging. Li et al. [12] surveyed graph representation learning across biomedicine and healthcare, treating imaging as one modality among many. Ahmedt-Aristizabal et al. [13] offered one of the earliest GNN overviews in healthcare but focused primarily on biosignals and brain connectivity. Bessadok et al. [14] addressed network neuroscience applications, while Li et al. [15] centered on graph signal processing and biological data. More recent works, such as Paul et al. [16], provided broader healthcare perspectives without a structured taxonomy for imaging tasks. Zhang et al. [17] focused on image-guided diagnosis but limited its scope to classification.

However, despite these prior efforts, existing reviews reveal several critical gaps. Most studies are either domain-general, covering diverse biomedical or healthcare applications without emphasizing imaging, or narrowly focused on isolated tasks such as classification or brain connectivity. None provide a unified, imaging-centered survey that jointly examines how graphs are constructed, how different GNN architectures and learning paradigms are employed, and how these map systematically to key medical imaging objectives such as segmentation, registration, reconstruction, and multimodal fusion. This work addresses these gaps through the following contributions:

- Conducting a comparative review of GNN architectures employed in medical imaging, including recurrent, convolutional, attention-based, autoencoding, and spatiotemporal models, and distilling design patterns and performance trends across studies.
- Presenting a unified taxonomy of graph formulations and learning tasks tailored to medical imaging, spanning node-, edge-, subgraph-, and graph-level representations and linking these to core clinical objectives.
- Providing an imaging-centered synthesis of GNN applications organized by clinical task, including segmentation, classification, and multimodal fusion, thereby offering a consolidated view of its robustness in the medical domain.
- Summarizing cross-cutting challenges and outlining future directions grounded in the evidence reviewed in this study.

The remainder of this paper is organized as follows. Section Review Methodology presents the review methodology. Section 2 introduces preliminaries on graphs and graph neural networks, while Section 3 details GNN architectures and learning paradigms tailored to medical imaging. Section 4 presents GNN applications across core imaging tasks, with detailed comparisons. Section 5 discusses open challenges and outlines future research directions. Finally, Section 6 concludes the paper.

Review Methodology

This review followed a structured methodology to ensure comprehensive and unbiased coverage of recent advances in GNNs for medical imaging. The search strategy targeted peer-reviewed journal and conference papers published between 2019 and 2025, retrieved primarily from IEEE Xplore, PubMed, Scopus, SpringerLink, and arXiv. The search terms included combinations of “graph neural network,” “graph convolution,” “geometric deep learning,” “medical imaging,” and task-specific keywords such as “segmentation,” “classification,” “reconstruction,” and “multimodal fusion.”

After removing duplicates and screening titles and abstracts, 312 papers were initially identified. Of these, 182 met the inclusion criteria following full-text assessment. Studies were included if they (i) proposed or evaluated a GNN-based method applied to medical image data, (ii) reported quantitative performance metrics, and (iii) provided sufficient methodological detail to enable reproducibility. Excluded works comprised non-imaging studies (e.g., molecular or social graphs), purely theoretical papers without empirical validation, and short workshop abstracts.

Each selected paper was categorized according to (a) architectural family (recurrent, convolutional, attention-based, autoencoding, or spatiotemporal), (b) learning paradigm (supervised, semi-supervised, unsupervised, or self-supervised), and (c) imaging application domain (segmentation, classification, registration, reconstruction, or multimodal fusion). Quantitative comparisons and qualitative trends were synthesized to identify consistent design patterns, emerging challenges, and future research opportunities. The taxonomy and analysis developed from this process form the basis for the organizational structure adopted in subsequent sections.

2. Preliminaries

Graphs provide a flexible and powerful mathematical representation for data exhibiting relational or topological dependencies, where entities are modeled as nodes and their interactions as edges. Formally, a graph is defined as $G = (V, E, X)$, where V represents the set of $N = |V|$ nodes, E the set of edges, and $X \in \mathbb{R}^{N \times d}$ the node feature matrix of dimension d [18]. The adjacency matrix $A \in \mathbb{R}^{N \times N}$ encodes connectivity, with $A_{ij} = 1$ if an edge exists between nodes i and j , and 0 otherwise. This representation is suited to medical imaging, where nodes can represent pixels, superpixels, patches, or anatomical regions, while edges model spatial, structural, or functional dependencies [19].

2.1. Overview of Graph Neural Networks

GNNs extend deep learning to non-Euclidean domains by propagating and aggregating information along graph edges. The core mechanism, known as message passing, allows each node to update its representation based on information from its neighbors, capturing both local and global dependencies [20,21]. A typical GNN layer can be expressed as:

$$h_i^{(k)} = \sigma\left(W^{(k)} \cdot \text{AGGREGATE}\left(\{h_j^{(k-1)} : j \in \mathcal{N}(i)\}\right)\right), \quad (1)$$

where $h_i^{(k)}$ denotes the embedding of node i at layer k , $\mathcal{N}(i)$ its neighborhood, $W^{(k)}$ a learnable weight matrix, and $\sigma(\cdot)$ a nonlinear activation function. Aggregation functions may take various forms such as mean, sum, max pooling, or attention-based weighting. Through multiple layers, node representations are refined to encode higher-order relational features [22]. Figure 1 illustrates a representative graph-based medical imaging pipeline for brain imaging: functional MRI (fMRI) time series are summarized as blood-oxygen-level dependent (BOLD) signals for each region of interest (ROI); an ROI graph is constructed; spatial graph operations perform neighborhood aggregation, optionally followed by temporal modeling to capture dynamics across time; node features are projected and pooled; and a readout head produces subject-level predictions [23]. The same pipeline structure generalizes to other modalities (e.g., CT, OCT, PET) once image content is encoded as nodes and edges.

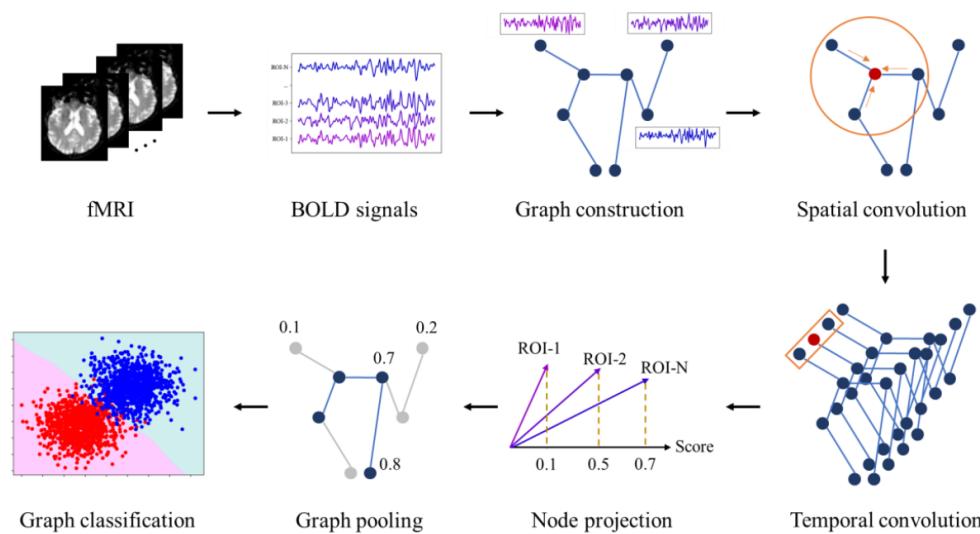


Figure 1. Graph-based medical imaging pipeline. fMRI volumes are converted to BOLD time series per ROI, an ROI graph is constructed, spatial graph operations and temporal modeling extract relational dynamics, node features are projected and pooled, and a readout head yields subject-level predictions.

2.2. Graph Variants

Graphs used in medical imaging can be categorized into several variant pairs that capture different structural, temporal, and semantic properties of the data. Each variant determines how relationships between entities are represented and how information is propagated during graph learning. Figure 2 illustrates the main variants discussed below.

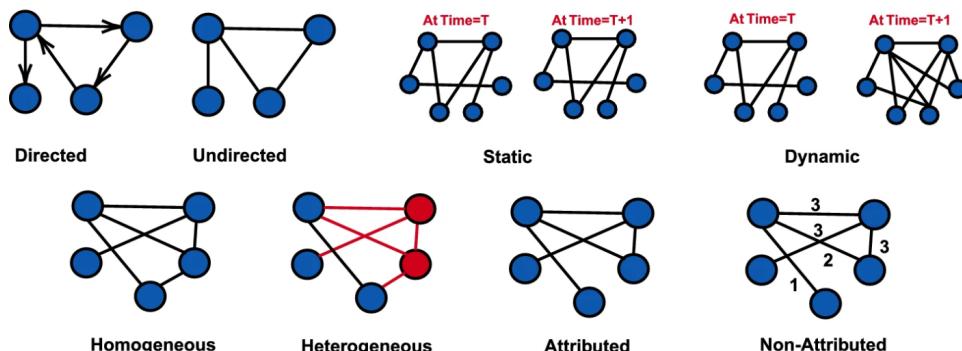


Figure 2. Common graph variants used in medical imaging, including directed vs. undirected, static vs. dynamic, homogeneous vs. heterogeneous, and attributed vs. non-attributed.

2.2.1. Directed and Undirected Graphs

Directed graphs encode asymmetric relationships, where the direction of an edge conveys flow or causality between nodes [24]. They are particularly useful for modeling processes such as blood flow in angiography, disease progression across time, or signal transmission in brain connectivity studies. Meanwhile, undirected graphs, in contrast, assume bidirectional and symmetric dependencies between nodes, making them suitable for structural MRI and histopathology where mutual adjacency or correlation (e.g., between neighboring tissues or regions) is key [25].

2.2.2. Static and Dynamic Graphs

Static graphs are constructed once and remain fixed during training, typically representing stable anatomical or morphological relationships [13]. They are common in structural connectomics and organ segmentation tasks where geometry does not change

over time. Dynamic graphs evolve either in topology or feature space, capturing temporal variations in node states or edge weights [26]. These are widely applied in functional MRI (fMRI), dynamic PET, and cardiac cine-MRI, where relationships between regions change as a function of time or physiological activity.

2.2.3. Homogeneous and Heterogeneous Graphs

Homogeneous graphs contain a single type of node and edge, representing uniform entities such as image patches, cortical regions, or superpixels. They are computationally simpler and often used in segmentation or disease classification. Heterogeneous graphs, by contrast, integrate multiple node and edge types to capture multimodal interactions [27]. For instance, imaging-derived features may be connected with genomic or clinical nodes, allowing multimodal GNNs to jointly reason over heterogeneous biomedical information.

2.2.4. Attributed and Non-Attributed Graphs

Attributed graphs associate nodes and edges with feature vectors, such as texture, intensity, or morphological descriptors, enriching the representation of local and global image context. These are widely used in pathology and brain imaging where region-level features enhance prediction accuracy. Non-attributed graphs encode only connectivity information without explicit feature embeddings, focusing instead on the relational topology [28]. They are mainly relevant in structural or functional brain network analysis, where the pattern of connections carries the main diagnostic signal.

Beyond these structural forms, another emerging category involves knowledge graphs, which extend GNN representations by encoding semantic relationships between imaging entities, medical concepts, and prior domain knowledge [29]. Nodes represent entities such as diseases, anatomical structures, or biomarkers, while edges denote semantic or causal relations. In medical imaging, knowledge graphs can link radiological findings to ontology-based entities, facilitating explainable and knowledge-informed reasoning within GNN frameworks.

2.3. Graph Learning Tasks and Levels of Representation

GNNs support a hierarchy of learning objectives that align closely with the representational levels at which medical imaging data can be structured, as illustrated in Figure 3. These objectives span local predictions (node and edge levels) to global or compositional reasoning (subgraph and graph levels), collectively enabling end-to-end analysis from regional patterns to patient-level outcomes.

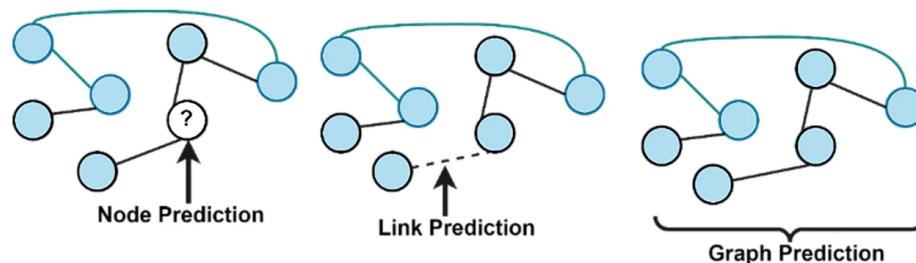


Figure 3. Hierarchies of graph learning tasks relevant to medical imaging. Node-level prediction (**left**) assigns labels to individual nodes such as cortical regions or image patches; edge-level prediction (**middle**) infers missing or abnormal links for connectivity or registration; subgraph- and graph-level prediction (**right**) classify regional structures or entire graphs, enabling subject-level diagnosis and prognosis.

2.3.1. Node-Level Learning

Node-level learning focuses on assigning labels or regression targets to individual nodes [30]. In medical imaging, each node can represent a brain region in MRI, a superpixel patch in histopathology, or a vascular segment in angiography. Tasks include tissue classification, lesion identification, and local atrophy assessment in neurodegenerative diseases [31].

2.3.2. Edge-Level Learning

Edge-level learning, known as link prediction, aims to infer the presence, strength, or type of relationships between nodes [30]. Applications include reconstructing functional connectivity in brain networks, estimating missing edges in structural graphs, and refining correspondences in deformable registration.

2.3.3. Graph-Level Learning

This treats the entire graph as a single sample to be classified or scored [32]. This is prevalent in tasks such as disease diagnosis from subject-level brain graphs, subtype prediction in whole-slide pathology, or multimodal prognosis from patient-level heterogeneous graphs that integrate imaging, genomic, and clinical data.

2.3.4. Subgraph-Level Learning

Subgraph-level learning identifies localized patterns or substructures within the overall graph. Examples include discovering tumor subregions in MRI, identifying abnormal vessel trees, or segmenting cortical modules in population-level studies [32]. Subgraph reasoning often provides interpretability, as it highlights localized biomarkers relevant for diagnosis or prognosis.

These learning levels provide a unified hierarchical framework for medical image analysis: node and edge tasks capture local structural relationships, subgraph tasks reveal mesoscopic organization, and graph-level tasks infer global outcomes. This multiscale design allows GNNs to bridge fine-grained image information with patient-level decision-making.

3. GNN Architectures and Learning Paradigms

Building on the concepts introduced in Section 2, this section details major GNN architectures and learning paradigms, focusing on how they aggregate, propagate, and learn from relational information in medical imaging.

3.1. Architectural Categories of GNNs

Different families of GNN architectures have been developed to capture spatial, spectral, temporal, and structural dependencies within graph-structured data. The following represent the most common formulations:

3.1.1. Recurrent GNNs

Recurrent GNNs (RecGNNs) are the earliest GNN formulations, where node states are iteratively updated until convergence using recurrent message passing [33]. Each node aggregates information from its neighbors and updates its embedding recursively, making RecGNNs conceptually similar to recurrent neural networks but applied on graphs. Although computationally expensive, RecGNNs laid the groundwork for later convolutional formulations.

3.1.2. Convolutional GNNs

Convolutional GNNs (ConvGNNs) generalize convolutional operations to non-Euclidean domains by defining local filtering in the spectral or spatial domain. The Graph

Convolutional Network (GCN) proposed by Kipf and Welling [34] is the most widely adopted variant and can be expressed as:

$$H^{(k+1)} = \sigma\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H^{(k)}W^{(k)}\right), \quad (2)$$

where $\tilde{A} = A + I$ denotes the adjacency matrix with self-loops, \tilde{D} is its degree matrix, and $\sigma(\cdot)$ is a nonlinear activation [34]. ConvGNNs have been widely applied in brain network analysis, tumor segmentation, and cardiac motion modeling due to their efficiency and ability to capture local topology [35,36]. Figure 4 provides a schematic of the GCN workflow. Starting from (V, E, X) , features are optionally standardized and embedded, neighbor information is aggregated using the normalized adjacency, the result is linearly transformed and passed through an activation (with optional dropout/normalization), and a task head performs node-, edge-, or graph-level prediction.

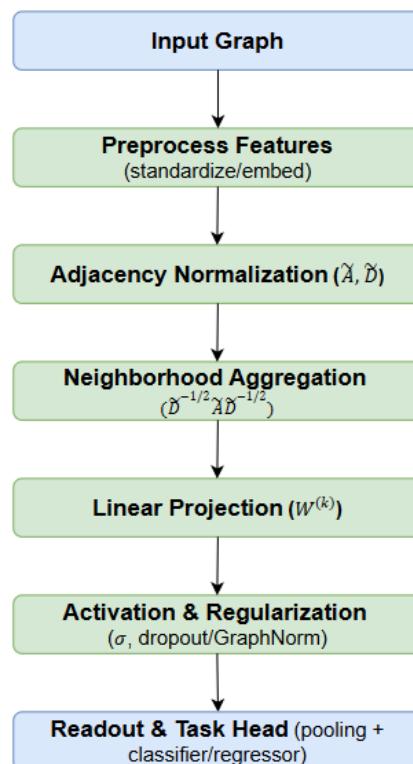


Figure 4. GCN workflow from input graph to prediction. The pipeline ingests a graph (V, E, X) , optionally preprocesses features, aggregates normalized neighbor messages using $\tilde{D}^{-1/2}\tilde{A}\tilde{D}^{-1/2}$, applies a linear projection $W^{(k)}$ with nonlinearity, and performs task-specific readout.

3.1.3. Graph Attention Networks

Graph Attention Networks (GATs) extend graph convolutions by learning attention coefficients that weight the contribution of each neighbor during aggregation [20]. Given node features $h_i^{(k)}$ at layer k , a shared linear projection W first maps them to a new space, and a shared attention vector a scores each pair (i, j) with $j \in \mathcal{N}(i)$:

$$e_{ij} = \text{LeakyReLU}\left(a^\top [Wh_i^{(k)} \| Wh_j^{(k)}]\right), \quad (3)$$

$$\alpha_{ij} = \text{softmax}_{j \in \mathcal{N}(i)}(e_{ij}), \quad (4)$$

where \parallel denotes concatenation. The normalized α_{ij} express the relative importance of neighbor j to node i . Node updates then apply attention-weighted aggregation followed by a nonlinearity $\sigma(\cdot)$:

$$h_i^{(k+1)} = \sigma \left(\sum_{j \in \mathcal{N}(i)} \alpha_{ij} W h_j^{(k)} \right). \quad (5)$$

Multi-head attention averages or concatenates outputs from K heads to stabilize training and enrich features:

$$h_i^{(k+1)} = \parallel_{m=1}^K \sigma \left(\sum_{j \in \mathcal{N}(i)} \alpha_{ij}^{(m)} W^{(m)} h_j^{(k)} \right). \quad (6)$$

Here, W and $W^{(m)}$ are learnable projections, a is a learnable attention vector, $\mathcal{N}(i)$ the neighbor set of i , and $\sigma(\cdot)$ a pointwise activation. Figure 5 illustrates the workflow: starting from an Input Graph, node features are preprocessed (standardization/embedding) and linearly projected. GAT then computes attention scores for neighbors and converts them to coefficients via a softmax over $\mathcal{N}(i)$. These coefficients drive attention-weighted aggregation of neighbor representations, after which outputs from multiple heads are combined (concatenate/average), followed by activation and regularization (e.g., σ , dropout, GraphNorm). Finally, the readout & task head block performs pooling and classification/regression.

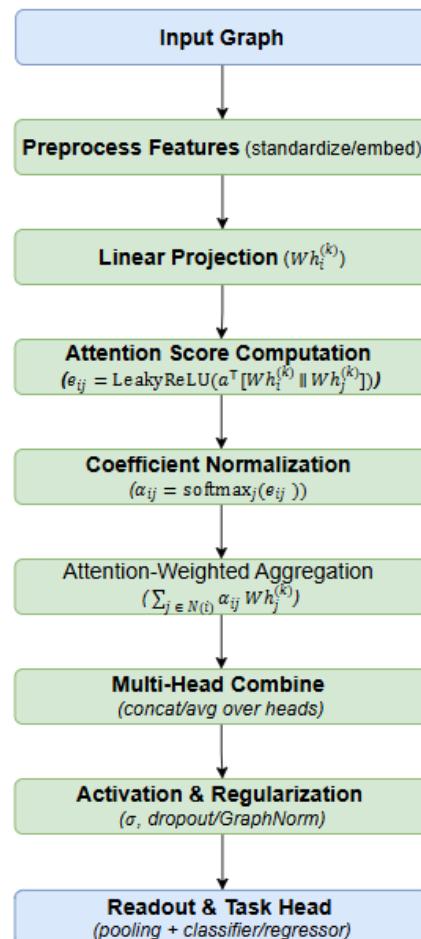


Figure 5. Workflow of a Graph Attention Network: preprocess features, compute attention scores and normalized coefficients over neighbors, aggregate with attention weights, combine multiple heads, then apply activation/regularization and task readout.

GATs have been particularly impactful in medical imaging where contextual relationships vary across modalities or tissues. They have been applied to tumor boundary refinement, disease subtyping, and multimodal feature fusion, exploiting adaptive attention to emphasize clinically informative regions or connections [20,37].

3.1.4. GraphSAGE

Graph Sample and Aggregate (GraphSAGE) [38] addresses the scalability limits of traditional GCNs by enabling inductive learning on unseen nodes or graphs. Instead of full-graph convolutions, GraphSAGE samples a fixed-size neighborhood and aggregates information using functions such as mean, max-pooling, or Long Short-Term Memory (LSTMs):

$$h_i^{(k+1)} = \sigma\left(W^{(k)} \cdot \text{AGGREGATE}\left(\{h_j^{(k)}, \forall j \in \mathcal{N}(i)\}\right) + h_i^{(k)}\right), \quad (7)$$

where $h_i^{(k)}$ denotes the embedding of node i at layer k , $\mathcal{N}(i)$ represents the set of neighboring nodes of i , $W^{(k)}$ is a learnable weight matrix, $\sigma(\cdot)$ is a nonlinear activation function, and $\text{AGGREGATE}(\cdot)$ is a neighborhood aggregation function (e.g., mean, max-pooling, or LSTM). The additive term $h_i^{(k)}$ preserves the node's self-representation, promoting stable feature propagation across layers.

This inductive formulation allows GraphSAGE to perform efficient mini-batch training and generalize to unseen nodes or graphs, making it suitable for large-scale medical imaging applications such as population-level brain connectivity or multi-center cohort analysis [39,40]. However, its simplicity can limit its ability to capture fine-grained structural details compared to attention-based or spectral approaches.

3.1.5. Graph Autoencoders

Graph Autoencoders (GAEs) and Variational GAEs (VGAEs) learn unsupervised node or graph embeddings by reconstructing adjacency structures or node attributes. A typical GAE comprises an encoder $f_\theta : (X, A) \mapsto Z$ and a decoder that reconstructs \hat{A} ; VGAЕ further imposes a variational posterior $q_\phi(Z | X, A)$ with a KL regularizer toward a simple prior, improving smoothness and out-of-distribution behavior. GAEs are especially relevant in medical imaging where labels are scarce, supporting anomaly detection, clustering, and latent representation learning for complex anatomical patterns [41].

3.1.6. Spatial–Temporal GNNs

Designed for dynamic medical data, Spatial–Temporal GNNs (ST-GNNs) integrate temporal evolution with spatial graph reasoning [42]. For imaging contexts with trajectories (cine-MRI, fMRI, DCE-MRI, ultrasound), edges can be static (anatomical priors) or dynamic (recomputed by attention/similarity per time step), allowing the model to capture evolving physiology while respecting structure [43]. ST-GNNs have shown strong performance in modeling disease progression (e.g., Alzheimer's trajectory analysis), cardiac motion prediction, and dynamic fMRI connectivity, where both structural and temporal dependencies are key.

Table 1 summarizes the major GNN families in terms of their empirical performance, interpretability, scalability, and clinical applicability. It highlights the trade-offs between efficiency, expressiveness, and explainability that guide architectural selection for specific medical imaging tasks.

Table 1. Comparative Summary of Major GNN Architectures in Medical Imaging.

Architecture	Strengths	Weaknesses	Performance Trends
Recurrent GNN [33]	Captures long-range dependencies via iterative message passing.	Computationally expensive; prone to vanishing gradients.	Moderate accuracy; strong temporal modeling but low scalability.
GCN [34]	Simple and stable; effectively captures local topology.	Limited receptive field; weak on highly irregular graphs.	High accuracy on structured data; efficient with low cost.
GAT [20]	Attention improves contextual reasoning and boundary precision.	Heavy on dense graphs; prone to noise and overfitting.	High interpretability; moderate scalability.
GraphSAGE [38]	Highly scalable and inductive; supports large patient graphs.	Simplified aggregation may lose fine spatial details.	Excellent scalability; slightly reduced interpretability.
GAE/VGAE [44,45]	Useful for unsupervised embedding and anomaly detection.	Limited interpretability; weak spatial reasoning.	Moderate reconstruction accuracy; strong feature learning.
ST-GNN [42]	Models spatial-temporal dynamics in disease progression.	High computational demand; sensitive to missing timepoints.	Very high temporal accuracy; moderate scalability.

3.2. Learning Paradigms for GNNs

The learning paradigm determines how supervision is leveraged to train GNNs. In medical imaging, where labeled data is often limited, selecting the appropriate paradigm is critical for performance and generalization. The major learning settings include:

- Supervised Graph Learning: This paradigm relies on fully labeled data, where ground-truth labels are provided for nodes, edges, or entire graphs [46]. It is commonly used in segmentation, tissue classification, and disease diagnosis tasks. Although accurate, its dependency on expert annotations limits scalability in medical contexts.
- Semi-supervised Graph Learning: Here, only a subset of nodes is labeled, and information propagates through graph connectivity to infer labels for unlabeled nodes [47]. This paradigm has been effective in medical domains such as lesion segmentation and brain network analysis, where graph topology aids label propagation across similar regions [48,49].
- Unsupervised Graph Learning: Unsupervised methods focus on learning embeddings or latent representations without explicit labels, often optimizing reconstruction or contrastive objectives [50]. In medical imaging, they are useful for clustering patients, identifying abnormal subregions, and pretraining models for downstream tasks.
- Self-supervised Graph Learning: Self-supervision exploits auxiliary “pretext” tasks to learn transferable representations from unlabeled data. Examples include predicting masked node features, reconstructing subgraphs, or contrasting augmented graph views. Self-supervised strategies are increasingly employed to pretrain GNNs on large medical imaging datasets before fine-tuning on smaller labeled subsets [51].
- Few-shot and Meta Graph Learning: These paradigms address extreme label scarcity by learning to generalize from very few samples or across related tasks [52]. Meta-learning-based GNNs can adapt quickly to new cohorts or modalities, which is advantageous for rare disease imaging or small multi-center datasets.

4. Applications of GNNs in Medical Imaging

4.1. Segmentation Tasks

GNNs have been deployed across diverse segmentation problems, spanning brain tumors, chest radiographs, retinal imaging, abdominal organs, and polyps. Their key contribution lies in modeling non-Euclidean relations such as inter-region continuity and boundary topology, which are difficult for conventional CNNs to encode.

Neuro-oncology has been a strong focus for graph-based segmentation. Arshad Choudhry et al. [36] designed an interpretable graph convolutional neural network (GCNN) for multimodal brain tumor segmentation, integrating adaptive class activation maps and contrastive backpropagation to enhance transparency. On the Brain Tumor Segmentation (BraTS) benchmark, the model achieved Dice coefficient (Dice) scores up to 0.96 across tumor subregions, outperforming standard CNNs. Mohammadi and Allali [53] proposed a spectral–spatial GCNN that exploited both spectral signatures and non-Euclidean spatial structure, reaching Dice scores of 86.7% (whole tumor), 82.4% (tumor core), and 79.8% (enhancing tumor), improving on a U-Net baseline by 3–5 points. Complementing these designs, Zhou [35] introduced M2GCNet, which employs spatial- and channel-wise graph modules across T1, T1c, T2, and fluid-attenuated inversion recovery (FLAIR) MRI. On BraTS 2018, M2GCNet achieved Dice of 84.3% (tumor core) and 78.3% (enhancing tumor), with reduced Hausdorff distance (HD) of 5.1 mm, highlighting its strength in boundary delineation.

Extending to thoracic imaging, Gaggion et al. [54] presented HybridGNet for chest X-ray segmentation, which links CNN encoders and GCNN decoders through localized image-to-graph skip connections. On Japanese Society of Radiological Technology (JSRT), Shenzhen, and PadChest, the method improved cardiothoracic ratio (CTR) estimation correlations from 0.80 to 0.88 (normal) and 0.70 to 0.85 (abnormal), demonstrating better plausibility under domain shift and occlusion.

In retinal analysis, graph attention and graph–transformer hybrids have shown clear benefits. Shen et al. [55] proposed GAU-Net, a graph attention U-Net for retinal optical coherence tomography (OCT), which achieved mean absolute error (MAE) of 1.82–2.30 μm for retinal layer detection and Dice of 0.917 for choroidal neovascularization (CNV) segmentation, surpassing U-Net and SegNet. Xu and Wu [56] developed G2ViT, which integrates GNN priors with a vision transformer backbone. On Digital Retinal Images for Vessel Extraction (DRIVE), Structured Analysis of the Retina (STARE), and Child Heart and Health Study in England database (CHASE_DB1) vessel datasets, G2ViT delivered Dice of 0.844–0.856, outperforming CNN- and transformer-only variants by 2–4 points. Joshi and Sharma [57] targeted optic disc (OD) and optic cup (OC) segmentation in glaucoma screening with a graph deep network. On DRISHTI-GS, RIM-ONE, DRIONS-DB, and HRF, Dice scores reached 0.97 (OD, DRISHTI-GS) and 0.93 (OC, DRISHTI-GS), consistently surpassing U-Net, GAN, and CE-Net baselines. Collectively, these approaches demonstrate how graph reasoning enhances fine-grained structural segmentation critical to ophthalmic diagnostics.

Several works have expanded GNNs to multi-organ and multi-modality segmentation. Meng et al. [58] proposed RBA-Net, a framework that fuses region and boundary information through graph reasoning. On fundus datasets, it achieved Dice of 97.7% (OD) and 89.4% (OC), outperforming U-Net++ by up to 3.4 points. For polyp segmentation, RBA-Net reported Dice 75.7% and Boundary Intersection over Union (BIoU) 69.3%, exceeding ACSNet and PraNet. Li et al. [59] introduced MedSegViG, combining a hierarchical vision graph encoder with a CNN decoder. Evaluated on seven datasets, it achieved Dice of 0.899 (Kvasir) and 0.908 (CVC-ClinicDB) for polyp segmentation, 0.825 on ISIC2017 for skin lesions, and 0.739 on DRIVE for retinal vessels, consistently outperforming CNN (U-Net, ResUNet++) and transformer (TransFuse, PraNet) baselines.

Abdominal and prostate segmentation further illustrate the advantages of surface-aware graph designs. Yang and Wang [60] presented GRFE, which augments multi-organ CT segmentation with a regional feature-enhancing graph module. On the Synapse dataset, GRFE achieved 83.4% average Dice, improving 2.7 points over a 3D U-Net, with notable gains in pancreas and gallbladder segmentation. Tian et al. [61] proposed Surface-GCN, which incorporates mini-batch adaptive matching to simulate radiologist-like corrections for 3D surface alignment. On PROMISE12 prostate MRI, it achieved a Dice of 91.6% and reduced HD by 1.8 mm compared to U-Net. When extended to multi-organ CT, it showed improved consistency in organ boundary preservation.

Overall, segmentation studies show that GNNs not only raise Dice and Intersection over Union (IoU) scores across tasks but also improve anatomical plausibility, boundary sharpness, and robustness to data heterogeneity. Table 2 summarizes representative GNN-based segmentation methods, their applications, and reported outcomes.

Table 2. GNN-based methods for segmentation tasks in medical imaging. Abbreviations defined in text.

Author(s)	Year	Method(s)	Application	Reported Results
Arshad Choudhry et al. [36]	2025	Interpretable GCNN with class-activation guidance	Multimodal brain tumor segmentation	Dice up to 0.96; surpasses CNNs
Mohammadi and Allali [53]	2024	Spectral–spatial GCNN	Brain tumor segmentation	Dice up to 86.7%; surpasses U-Net
Zhou [35]	2024	M2GCNet	Brain tumor segmentation	Dice up to 84.3%; reduced HD (5.1 mm)
Gaggion et al. [54]	2022	HybridGNet (CNN encoder + GCNN decoder)	Chest X-ray anatomy	CTR correlation: 0.80→0.88 (normal), 0.70→0.85 (abnormal)
Shen et al. [55]	2023	Graph attention U-Net	Retinal OCT (layers/CNV)	MAE 1.82–2.30 μ m; Dice 0.917; surpasses U-Net/SegNet
Xu and Wu [56]	2024	GNN-guided ViT	Retinal vessels	Dice 0.844–0.856 across DRIVE/STARE/CHASE_DB1
Joshi and Sharma [57]	2022	Graph deep network	OD/OC (glaucoma screening)	Dice 0.97 (OD) and 0.93 (OC); IoU up to 0.96
Meng et al. [58]	2021	RBA-Net	OD/OC and polyps	OD 97.7%, OC 89.4%; polyp Dice 75.7%, BIoU 69.3%
Li et al. [59]	2024	MedSegViG	Polyps, skin lesions, vessels	Polyp Dice 0.899–0.908; skin 0.825; vessels 0.739
Yang and Wang [60]	2022	Regional feature-enhancing GCN	Abdominal multi-organ (CT, Synapse)	Avg Dice 83.4%; +2.7 vs. 3D U-Net
Tian et al. [61]	2023	Surface-GCN with adaptive matching	Prostate MRI; abdominal CT	Dice 91.6%; HD –1.8 mm; smoother boundaries

4.2. Classification Tasks

GNNs are increasingly used for classification at the patient, scan, or region level across cancer grading, Alzheimer's disease (AD), cardiac arrhythmia, and retinal disease.

In cancer image analysis, Li et al. [62] proposed AFFC-Net, an adaptive CNN-GNN fusion with attention. On Breast Cancer Semantic Segmentation (BRACS) slides, it achieved a weighted F1-score (F1) of 67.23%; on the LC25000 lung/colon dataset it reached 99.84% across standard metrics, and on BreakHis histopathology attained 96.97% accuracy. Lu et al. [63] introduced SlideGraph+, a graph-based whole-slide modeling approach for human epidermal growth factor receptor 2 (HER2) status; on the Nottingham cohort it achieved an area under the receiver characteristic curve (AUC) of 0.83, surpassing whole-slide imaging (WSI) baselines. Chang et al. [64] combined attention-driven causal reasoning with graph structure learning for hematoxylin and eosin (H&E) slides, reaching 86.36% accuracy for invasive vs. non-invasive breast cancer.

For neuroimaging, Zhang et al. [65] built a population-based GCN with nodes as subjects from the Alzheimer's Disease Neuroimaging Initiative (ADNI) and edges from phenotypic similarity using structural MRI (sMRI) and PET. It achieved 88.95% accuracy for AD vs. normal controls (NC) and 75.64% for stable vs. progressive mild cognitive impairment (sMCI vs. pMCI). Huynh et al. [66] merged generative adversarial networks (GANs) with a GCN, improving robustness under limited data across cohorts. Kim [37] used correlation-based GCNs with GNNExplainer for early AD detection on 506 ADNI subjects, reporting mean AUCs of 0.8851 (cognitively normal, CN), 0.8741 (mild cognitive impairment, MCI), and 0.8632 (combined), while identifying biomarkers such as APOE $\epsilon 4$ and cortical thickness patterns.

Cardiovascular applications show the flexibility of spatiotemporal graphs. Banus et al. [67] proposed a spatiotemporal graph neural process with neural ordinary differential equations (ODEs) and neural processes (NPs), achieving 99% accuracy on Automated Cardiac Diagnosis Challenge (ACDC) data and 67% on UK Biobank atrial fibrillation trajectories. Lin et al. [68] designed a hybrid GCN-LSTM model with trainable weighted ϵ -neighborhood graphs for electrocardiogram (ECG) segments, reaching 85.87% sensitivity and surpassing a ResNet34 baseline. Andayeshgar et al. [69] modeled inter-lead communication with a GCN and improved arrhythmia classification on the MIT-BIH dataset over traditional classifiers.

In ophthalmology, Zedadra et al. [70] presented DRDiag, fusing DenseNet121 image features with patient metadata via a two-layer GNN. On Messidor-2 and APTOS2019 for diabetic retinopathy (DR), it achieved 97.6–98.0% accuracy with Cohen's kappa of 0.957–0.960, outperforming CNN/transformer baselines. Pandugula et al. [71] combined CNNs with graph attention networks (GATs) for DR grading on APTOS2019; with an InceptionV3 backbone it attained 84.90% accuracy and quadratic kappa of 89.98, exceeding CNN-only models by 3–5 points.

The reviewed studies demonstrate the diversity of GNN applications in classification—from histopathology and neuroimaging to cardiology and ophthalmology. Graph structures enable more effective feature aggregation, enhance generalization, and in some cases deliver interpretable predictions linked to clinically meaningful biomarkers. Table 3 provides a structured overview of the reviewed classification studies, including their methodological designs, applications, and reported outcomes.

Table 3. GNN-based methods for classification tasks in medical imaging.

Author(s)	Year	Method(s)	Application	Reported Results
Li et al. [62]	2025	CNN + GNN with adaptive fusion	Breast, lung, colon pathology classification	BRACS: F1 67.23%; LC25000: 99.84%; BreakHis: 96.97%
Lu et al. [63]	2022	SlideGraph and graph WSI modeling	HER2 status (breast)	AUC 0.83; beats WSI baselines
Chang et al. [64]	2025	Causal attention GNN	Invasive vs. non-invasive (H&E slides)	Accuracy 86.36%; outperforms CNN classifiers
Zhang et al. [65]	2023	Population GCN (phenotypic edges)	AD diagnosis	Accuracy of 88.95%
Huynh et al. [66]	2024	GAN + GCN	AD across cohorts	Robust under limited data; Outperforms CNN/GNN baselines
Kim [37]	2023	Correlation GCN + GNNExplainer	Early AD (ADNI)	AUCs: 0.8851 (CN), 0.8741 (MCI), 0.8632 (all); interpretable
Banus et al. [67]	2025	Spatiotemporal GN process (ODE+NP+GNN)	Cardiac trajectory classification	ACDC: 99% acc.; UKB AF: 67% acc.
Lin et al. [68]	2024	Hybrid GCN-LSTM (ϵ -graphs)	Ventricular arrhythmia (ECG)	Sensitivity 85.87%; > ResNet34
Andayeshgar et al. [69]	2024	Inter-lead GCN	Arrhythmia (MIT-BIH)	Higher accuracy than traditional classifiers
Zedadra et al. [70]	2025	DRDiag (CNN + two-layer GNN)	Diabetic retinopathy (Messidor-2, APTOS2019)	Acc. 97.6–98.0%; kappa 0.957–0.960; Outperforms CNN/transformers
Pandugula et al. [71]	2025	CNN-GAT fusion	DR grading (APTO2019)	Acc. 84.90%; quadratic kappa 89.98; +3–5% vs. CNNs

4.3. Image Retrieval and Reconstruction

Reconstruction and retrieval tasks sit at the interface between low-level enhancement and high-level downstream analysis in medical imaging. GNNs are attractive here because they can encode non-local self-similarity and long-range dependencies that are ubiquitous in anatomical imagery, from repeating textures and symmetries to cross-slice coherence.

MRI reconstruction has been a prime testing ground for graph-based approaches. Ma et al. [72] proposed the Graph Convolutional Enhanced Self-Similarity (GCESS) network, which leverages non-local patch similarities modeled through graph convolution in tandem with spatial convolution. On knee and brain MRI under one-dimensional (1D) Cartesian undersampling with acceleration factor (AF) = 4, GCESS achieved 34.19 dB peak signal-to-noise ratio (PSNR) and 0.899 structural similarity (SSIM), surpassing CNN reconstructions by 1.05 dB PSNR and 0.02 absolute SSIM. Meanwhile, Ahmed et al. [73] introduced FedGraphMRI-net, a federated GNN framework designed for heterogeneous client data. On distributed fastMRI brain datasets, FedGraphMRI-net consistently improved PSNR by 1.8–2.3 dB relative to centralized CNNs while preserving privacy. Complementing these efforts, Wang et al. [74] proposed MD-GraphFormer, a model-driven graph transformer that integrates physics priors for fast multi-contrast MRI. By embedding acquisition models into graph-based attention, it achieved superior reconstruction quality with markedly lower runtime, outperforming state-of-the-art alternatives.

In CT reconstruction, Xia et al. [75] developed MAGIC, a manifold and graph integrative convolutional network that unrolls iterative optimization and incorporates both spatial and graph convolutions. Evaluated on the NIH–AAPM–Mayo low-dose CT (LDCT) dataset, MAGIC reported 36.26 dB PSNR and 0.9696 SSIM at 10% dose, outperforming RED-CNN, LEARN, and learned primal–dual (LPD), and highlighting the strength of graph-based regularization under extreme noise and dose reduction.

Beyond reconstruction, GNNs have been applied to secure image retrieval. Tang et al. [76] introduced MRCG, a retrieval framework combining a VGG16-based triplet CNN for similarity

embedding with a graph-based representation of inter-gallery relationships. A GNN with skip connections and focal loss then refined similarity predictions, yielding mean average precision (MAP) of 88.64% on contrast-enhanced MRI (CE-MRI) and 86.59% on Kaggle MRI, outperforming Siamese CNNs and graph baselines such as Chebyshev networks (ChebNet), GraphSAGE, and GAT. These results suggest that retrieval pipelines can benefit from graph-enhanced feature spaces that capture relational structure in query-gallery pairs.

Graph-based learning has also been explored in super-resolution. Isallari and Rekik [77] designed an adversarial GNN for brain graph super-resolution to enhance functional connectivity analysis. Applied to the ADNI, the framework achieved 87.3% accuracy in distinguishing mild cognitive impairment and Alzheimer's disease, outperforming interpolation and non-adversarial GNN baselines. Extending to multi-image setups, Tarasiewicz and Kawulok [78] proposed MagNet and MagNet++, which aggregate multiple low-resolution images into a unified graph representation via spline-based graph convolutions and recursive fusion. On DIV2K and PROBA-V, these models delivered competitive PSNR and SSIM compared to CNN and transformer-based counterparts, while offering flexibility in handling varying numbers of inputs.

Finally, GNNs have been used for refinement of imperfect reconstructions. Ye et al. [79] introduced vessel-like structure rehabilitation network (VSR-Net) that employs graph clustering to repair ruptures in vascular segmentations. Across DRIVE, OCTA-500, and ImageCAS, VSR-Net improved Dice by 0.67–2.08% over topology-preserving methods and reduced expected calibration error (ECE) from 0.0337 to 0.0281. Liang et al. [80] proposed a complementary refinement framework using dynamic graph convolution to correct mis-segmented boundaries. On multiple biomedical segmentation datasets, their method consistently raised Dice and Jaccard indices compared to CNN baselines, demonstrating the value of explicit region-to-boundary graph modeling.

These studies illustrate the versatility of GNNs in reconstruction and retrieval tasks. Table 4 provides a structured overview of the reviewed methods, applications, and their reported outcomes.

Table 4. GNN-based methods for image retrieval and reconstruction in medical imaging.

Author(s)	Year	Method(s)	Application	Reported Results
Ma et al. [72]	2024	GCESS	Accelerated MRI reconstruction (knee and brain)	PSNR 34.19 dB; SSIM 0.899; +1.05 dB and +2% SSIM vs. CNNs
Ahmed et al. [73]	2025	federated GNN framework	Robust MRI reconstruction across institutions	+1.8–2.3 dB PSNR vs. centralized CNNs; improved cross-client generalization
Xia et al. [75]	2021	MAGIC (manifold + graph integrative convolutional network)	Low-dose CT reconstruction	PSNR 36.26 dB; SSIM 0.9696 at 10% dose; > RED-CNN/LEARN/LPD
Tang et al. [76]	2021	MRCG	MRI retrieval with privacy/security focus	MAP 88.64% (CE-MRI), 86.59% (Kaggle); Outperformed Siamese CNN, ChebNet, GraphSAGE, and GAT
Wang et al. [74]	2023	MD-GraphFormer	Fast multi-contrast MRI reconstruction	Higher accuracy and reduced runtime vs. state-of-the-art reconstructions
Isallari and Rekik [77]	2021	Adversarial GNN for graph super-resolution	Functional brain connectivity reconstruction	87.3% accuracy; outperforming non-adversarial GNNs
Tarasiewicz and Kawulok [78]	2024	spline-based GNN + recursive fusion	Multi-image super-resolution (DIV2K, PROBA-V)	Competitive PSNR/SSIM vs. CNNs/transfomers; flexible input handling
Ye et al. [79]	2025	VSR-Net	Repairing ruptures in vessel-like segmentation	Dice +0.67–2.08% over topology-preserving methods; ECE reduced (0.0337→0.0281)
Liang et al. [80]	2023	Dynamic graph convolution for refinement	Correction of mis-segmented regions in biomedical images	Consistent Dice and Jaccard gains vs. CNN baselines across datasets

4.4. Registration and Alignment

Deformable registration aligns a moving image to a fixed reference by estimating a dense displacement or diffeomorphic flow field. GNNs are well-suited here because they encode non-local anatomical relationships and topology, complementing local appearance cues that dominate standard CNN pipelines.

For volumetric brain MRI, two hybrid designs exemplify how graphs mitigate the locality limits of convolutions while taming the spurious global interactions sometimes induced by pure Transformers. Yang et al. [81] proposed GraformerDIR, which couples Chebyshev graph convolution with Transformer layers to capture long-range correspondences under unsupervised losses. On the Open Access Series of Imaging Studies (OASIS) and MGH10 brain MRI, GraformerDIR improved Dice by 4.6% and 4.1% over VoxelMorph and lowered average symmetric surface distance (ASD) by 0.055 mm and 0.084 mm, respectively, while reducing non-positive Jacobians by up to $60\times$. In a similar spirit but with an emphasis on efficiency, Zhou and Cao [82] introduced H-SGANet, a lightweight hybrid that blends convolution, Vision GNN components, and Transformer blocks via sparse graph attention (SGA) and separable self-attention (SSAFormer); on OASIS and the LONI Probabilistic Brain Atlas 40 (LPBA40) it achieved Dice gains of 3.5% and 1.5% over VoxelMorph with lower model complexity.

Beyond hybridizing with Transformers, graph structure can be leveraged to exchange information across moving and fixed images or to regularize displacement propagation from sparse cues. Wang et al. [83] presented DIEGraph, a dual-branch information-exchange GCN that extracts features from the two images separately and links them with a k -nearest neighbors (k NN) correspondence graph; on OASIS and LPBA40, it reported Dice of 0.806 and 0.686, up to +3.8% versus TransMorph, with very low folding (0.20% and 0.08%). In thoracic CT, Hansen and Heinrich [84] tackled the large respiratory motion by introducing GraphRegNet, which propagates displacements from sparse keypoints through a deep graph-regularization network; mean landmark errors reached ≈ 1.00 mm (COPDGene) and 1.04 mm (DIR-Lab 4DCT), improving 20–30% over B-spline and deep baselines while reducing compute.

On cortical surfaces, graph attention and manifold-aware convolutions have enabled both accuracy and speed at scale. Ren et al. [85] proposed SUGAR, a spherical ultrafast graph attention framework with a U-Net-like backbone and distortion-preserving losses; on large cohorts (e.g., UK Biobank) it delivered sub-second registrations with improved accuracy and lower distortion than conventional surface methods. Complementing this, Suliman et al. [86] introduced GeoMorph, an unsupervised multimodal surface pipeline that marries GCN feature extractors with conditional random fields (CRFs) to yield smoother, biologically consistent deformations while retaining competitive correspondence accuracy.

Earlier graph-centric formulations on cortical manifolds further clarify where graphs help most. Cheng et al. [87] learned intrinsic correspondences directly on cortical meshes without handcrafted features or spherical unwrapping, surpassing iterative closest point (ICP) and spectral-descriptor baselines. Zhang et al. [88] then embedded spherical-harmonic descriptors within graph-enhanced filters (GESH-Net), reducing geodesic error and improving robustness to folding variability over traditional spherical-harmonic approaches. Similarly, Tan et al. [89] demonstrated that an improved GCN (ADGCN) with squeeze-and-excitation (SE) blocks can support downstream cortical parcellation on Mindboggle-101, achieving Dice 88.53% and accuracy 90.27%, outperforming FreeSurfer and matching strong spherical CNNs.

Finally, graph models have also been leveraged for generic shape correspondence under occlusion and cross-domain shifts. Arias-García et al. [90] formulated correspondence as a global linear assignment over a dynamic structural-spatial graph (DGA), solved

with Kuhn–Munkres optimization; on the FAUST human-mesh benchmark they reported a 33.5% reduction in mean geodesic error versus spectral GCNs, with robustness on TOSCA and SHREC-20 benchmarks. Taken together, these volumetric and surface studies indicate two recurring benefits of graphs in registration: stronger non-local reasoning for correspondence (often yielding Dice/ASD gains and fewer folds) and principled geometric regularization on meshes for speed and plausibility at scale.

Table 5 summarizes the reviewed methods, datasets, and reported results.

Table 5. GNN-based methods for registration and alignment in medical imaging.

Author(s)	Year	Method(s)	Application	Reported Results
Yang et al. [81]	2022	GraformerDIR	Brain MRI deformable registration (with cardiac MRI validation)	Dice +4.6 (OASIS), +4.1 (MGH10) vs. VoxelMorph; ASD –0.055/–0.084 mm; folds reduced by up to 60×
Zhou and Cao [82]	2025	H-SGANet	Brain MRI deformable registration	Dice +3.5 (OASIS), +1.5 (LPBA40) vs. VoxelMorph; lower model complexity
Wang et al. [83]	2023	Dual-branch k NN correspondence GCN	Brain MRI deformable registration	Dice 0.806 (OASIS), 0.686 (LPBA40); up to +3.8 vs. TransMorph; folding $\leq 0.20\%$
Hansen and Heinrich [84]	2021	GraphRegNet	3D lung CT respiratory motion registration	Mean landmark error $\approx 1.00\text{--}1.04$ mm; 20–30% lower than B-spline/deep baselines
Ren et al. [85]	2024	Spherical graph attention U-Net	Cortical surface registration at population scale	Sub-second runtime; higher accuracy and lower distortion than conventional methods
Suliman et al. [86]	2023	GeoMorph	Multimodal cortical surface registration	Smoother, biologically consistent deformations; competitive correspondence accuracy
Cheng et al. [87]	2020	Unsupervised cortical manifold correspondence (graph-based)	Cortical surface registration without spherical unwrapping	Higher accuracy than ICP and spectral-descriptor baselines
Zhang et al. [88]	2024	Graph-enhanced spherical-harmonic convolutions	Cortical surface registration under folding variability	Lower geodesic error vs. spherical-harmonic approaches
Tan et al. [89]	2025	U-shaped improved GCN with SE	Cortical surface parcellation	Dice 88.53%, accuracy 90.27%; exceeds FreeSurfer
Arias-García et al. [90]	2025	Dynamic structural–spatial graph + Kuhn–Munkres	Brain shape correspondence under occlusion/domain shift	Reduction in mean geodesic error by 33.5% vs. spectral GCNs; robust on TOSCA/SHREC-20

4.5. Multimodal Fusion

Fusing complementary signals from multiple sources (e.g., MRI, PET/CT, histopathology, genomics, and electronic health records (EHRs) has become central to improving diagnostic accuracy and prognostic modeling in medicine. A major challenge in multimodal learning is the presence of missing modalities or incomplete feature overlap, which is common in real-world clinical datasets where patients may lack one or more imaging or omic modalities. Heterogeneous and multiplex GNNs mitigate this limitation by learning adaptive message-passing paths across available modalities. When one modality is absent, information is propagated through shared nodes or cross-modality attention layers to infer missing features. Approaches such as MPlex-GNN [91], HGL [92], and DMGN [93] explicitly model intra- and inter-modality relationships through distinct graph planes or typed edges, enabling selective information transfer from complete to incomplete subgraphs.

Furthermore, some frameworks perform graph-level imputation using completion or diffusion mechanisms to reconstruct missing modalities [94,95]. For instance, message passing can approximate absent node features by aggregating neighboring modality embeddings, while variational or autoencoding modules infer latent cross-modal representations.

These strategies allow GNNs to maintain predictive performance even when multimodal inputs are only partially observed, a key advantage for clinical deployment where data incompleteness is unavoidable.

In neurodegenerative research, multimodal graph fusion has demonstrated clear benefits. Zhang et al. [96] integrated sMRI and PET scans via a multimodal GNN for early Alzheimer's disease (AD) detection. On the ADNI dataset, their model achieved 96.68% accuracy, 99.19% sensitivity, and 94.49% specificity for AD vs. NC, while reporting 78.0% accuracy and 89.37% specificity for sMCI vs. pMCI prediction, outperforming CNN and GCN baselines. Similarly, Cai et al. [97] proposed MM-GTUNets, which constructs modality-specific graphs and fuses them with graph transformer U-Nets (GTU-Nets) for brain disorder prediction. On ADNI, MM-GTUNets achieved 88.5% accuracy for AD vs. NC and 74.2% for sMCI vs. pMCI, with additional validation on ADHD-200 (73.8% accuracy), consistently surpassing CNN-based multimodal models and early-fusion GCNs. These studies highlight the advantage of graph-aware fusion in capturing complementary structural and functional patterns for neuropsychiatric disorders.

Oncology applications have also benefited from graph-based multimodal fusion. Zheng et al. [98] designed a graph attention-based framework that fuses pathology images with gene expression profiles for non-small cell lung cancer (NSCLC) survival prediction. Their method adaptively weights modalities to highlight cross-modal biomarkers, outperforming CNN, transformer, and GNN baselines on large-scale survival datasets. Wu et al. [99] focused on tumor heterogeneity, proposing ProtoSurv, a heterogeneous graph representation framework for whole-slide images (WSIs). By capturing localized variations in tumor morphology, ProtoSurv consistently outperformed state-of-the-art WSI pipelines across five TCGA cancer cohorts (BRCA, LGG, LUAD, COAD, PAAD), demonstrating the prognostic value of prototype-guided heterogeneous graphs. Complementing these, Fu et al. [93] introduced DMGN, a deep multimodal GNN that fuses imaging mass cytometry (IMC) with patient variables. On METABRIC and Basel cohorts, DMGN achieved concordance indices of 0.7484 and 0.7479, respectively, surpassing DeepHit (0.691) and AttentionSurv (0.708), confirming its superiority for breast cancer survival analysis.

Other works have emphasized structural generalization across heterogeneous modalities. D'Souza et al. [91] proposed MPlex-GNN, which models intra- and inter-modality relationships through multiplex graph planes derived from autoencoder-reduced features. On NIH-TB outcome prediction (3051 patients, five modalities) and ABIDE autism diagnosis (871 subjects), MPlex-GNN delivered significant AUC gains over early, intermediate, and late fusion baselines, achieving 0.754 AUC on ABIDE with improved robustness to missing modalities. Peng et al. [100] addressed multimodal prompt learning, introducing MMGPL, which leverages GPT-4-derived disease concepts to reweight imaging patches and propagate them via concept-guided graphs. On ADNI, MMGPL achieved 82.3% accuracy and 0.851 AUC in AD vs. NC vs. MCI, outperforming transformer-based prompt learners such as MaPLe by 4.5% accuracy; on ABIDE, it reached 72.4% accuracy (AUC 0.754). These contributions demonstrate that concept-aware and multiplexed GNNs can improve fusion robustness, interpretability, and generalization across diverse modalities.

Finally, heterogeneous graph learning strategies explicitly model modality differences. Kim et al. [92] proposed HGL, which represents patients, features, and modalities as distinct node types connected by typed edges. Validated on ADNI, ABIDE, and PPMI datasets, HGL improved accuracy by up to 6.8% over multimodal transformers and boosted AUC by 8.3% compared to late-fusion neural networks. This underscores the strength of heterogeneous graph designs in capturing cross-modal semantics and ensuring robustness when modalities vary across patients. In summary, these designs emphasize that handling incomplete or partially overlapping modalities is critical for clinical viability. GNN-based

fusion models, particularly those employing multiplex or heterogeneous structures, offer principled ways to learn consistent representations under missing data conditions, which remains a major step toward real-world translation.

Table 6 summarizes the reviewed studies.

Table 6. GNN-based multimodal fusion methods for diagnosis and prognosis in medical imaging.

Author(s)	Year	Method	Application	Reported Results
Zhang et al. [96]	2023	Multimodal GNN (sMRI + PET)	AD detection (ADNI)	96.68% acc., 99.19% sens., 94.49% spec. (AD vs. NC); 78.0% acc. (sMCI vs. pMCI)
Zheng et al. [98]	2024	Graph attention fusion (pathology + gene expression)	NSCLC survival prediction	Outperformed multimodal baselines; improved survival prediction accuracy
Wu et al. [99]	2024	Heterogeneous graph WSI	Cancer survival (TCGA cohorts)	Consistently higher C-index than WSI baselines; improved prognostic signatures
Fu et al. [93]	2023	DMGN (IMC + patient variables)	Breast cancer survival	C-index 0.7484/0.7479 vs. DeepHit (0.691), AttentionSurv (0.708)
D'Souza et al. [91]	2024	MPlex-GNN	NIH-TB outcome prediction; autism diagnosis (ABIDE)	AUC gains over fusion baselines; 0.754 AUC on ABIDE
Peng et al. [100]	2024	MMGPL (graph prompt learning)	AD diagnosis (ADNI); autism (ABIDE)	82.3% acc., 0.851 AUC (ADNI); 72.4% acc., 0.754 AUC (ABIDE)
Cai et al. [97]	2025	MM-GTUNets	Brain disorder prediction (ADNI, ADHD-200)	88.5% acc. (AD vs. NC), 74.2% (sMCI vs. pMCI), 73.8% (ADHD-200)
Kim et al. [92]	2023	Heterogeneous graph learning	AD (ADNI), autism (ABIDE), Parkinson's (PPMI)	+6.8% acc., +8.3% AUC over multimodal baselines

5. Challenges, Limitations, and Future Research Directions

Despite notable advances in adapting deep learning to graph-structured data, GNNs continue to face methodological and practical challenges in medical imaging. These challenges, coupled with opportunities for innovation, define the trajectory of current research. This section consolidates both aspects by presenting the key limitations of GNN-based medical imaging and outlining emerging research directions that seek to overcome them. The aim is to provide an integrated perspective linking existing obstacles with potential solutions for improving clinical relevance and deployment readiness.

5.1. Challenges and Limitations

5.1.1. Data Scarcity and Imbalance

The effectiveness of GNNs in medical imaging depends critically on the availability of annotated data for constructing reliable graphs [101]. However, annotated datasets such as BraTS for brain tumors or ADNI for Alzheimer's disease remain limited in scale compared to natural image corpora. Annotation requires costly manual delineation by experts, leading to small sample sizes and class imbalance, especially for rare diseases. While data augmentation and semi-supervised approaches partially mitigate these issues, their success in graph-based pipelines has been modest. Ahmedt-Aristizabal et al. [13] emphasized that scarcity restricts the robustness of learned representations, while Paul et al. [16] noted that imbalance skews GNN performance toward majority classes, undercutting sensitivity to clinically important but less frequent categories. These challenges are compounded by the fact that graph construction itself is data-dependent; errors or noise in small datasets can propagate through the adjacency structure, leading to unstable learning.

5.1.2. Scalability and Computational Overhead

Scaling GNNs to large 3D medical images introduces significant computational burdens. Unlike CNNs, which exploit grid regularity for efficient convolutions, GNNs must operate over irregular topologies that grow rapidly with image resolution [8]. For example, constructing voxel-level graphs from MRI or CT volumes yields millions of nodes and edges, rendering full-batch training infeasible. Downsampling, patching, or pooling can alleviate memory and runtime issues but at the cost of discarding fine-grained anatomical detail. Zhang et al. [17] observed that computational overhead remains a major barrier to applying GNNs at scale, while Li et al. [12] highlighted the lack of principled frameworks for balancing efficiency with representational richness in biomedical applications. Current acceleration strategies such as sparse sampling or hierarchical pooling offer partial relief, but their clinical readiness remains limited.

5.1.3. Interpretability and Clinical Trust

Interpretability remains a critical barrier to clinical deployment. While CNN filters and attention heatmaps provide intuitive cues for clinicians, the node- and edge-level operations of GNNs are less transparent. Existing interpretability techniques such as GNNExplainer or attention visualization provide some insight into influential nodes or subgraphs, yet their reliability in high-stakes medical contexts is questionable [9,15]. Inconsistent reporting of uncertainty and calibration, already noted as a challenge in deep learning surveys, further complicates clinical trust [102,103]. As a result, current GNN-based imaging systems risk being perceived as black boxes, limiting their adoption in clinical practice where transparency is paramount.

To address this, several explainability frameworks have been developed for GNNs. For example, GNNExplainer [104] identifies the most influential subgraphs and node features that drive model predictions, while PGExplainer [105] learns probabilistic edge masks through a separate explanation network. GraphSHAP [106] adapts the SHapley additive explanation principle to estimate the marginal contribution of each node or edge, and GraphLIME [107] employs local surrogate models to approximate feature importance around a specific prediction. These frameworks enable visualization of salient nodes, regional connections, or biomarkers responsible for diagnostic outcomes.

However, their application in clinical imaging remains constrained by multiple factors. Firstly, explanations generated by these methods often vary across runs, raising reproducibility concerns when decisions affect patient care. Secondly, most frameworks focus on structural attribution but rarely quantify confidence or causal influence, making it difficult for clinicians to judge reliability. Thirdly, interpretability outputs are not yet aligned with clinical reasoning, for example, highlighted nodes may not correspond to recognizable anatomical regions.

5.1.4. Generalization Across Institutions and Populations

Generalization across diverse imaging centers and populations is another unresolved issue. Variability in acquisition protocols, scanner vendors, and demographic distributions introduces domain shifts that undermine the robustness of graph-based models. Bessadok et al. [14] highlighted how graph-based neuroscience pipelines often fail to generalize beyond the datasets they were trained on, a challenge mirrored in broader medical imaging studies. Wu et al. [10] noted that graph learning is highly sensitive to the definition of edges and node features, meaning that preprocessing choices in one dataset may not transfer to another. Without robust domain adaptation strategies or standardized graph construction pipelines, achieving reproducibility across institutions remains elusive.

5.1.5. Regulatory and Deployment Barriers

Beyond technical limitations, regulatory and practical considerations complicate real-world deployment. Unlike conventional imaging pipelines, graph-based models require explicit preprocessing and graph construction steps that may vary across institutions, making standardization difficult. Li et al. [12] emphasized that reproducibility is a prerequisite for regulatory approval, yet current graph learning pipelines often lack transparent reporting of graph definitions, hyperparameters, and evaluation protocols. Ethical concerns also persist, particularly regarding privacy when integrating imaging with sensitive clinical or genomic data. Moreover, as Liu et al. [108] argued in the context of epidemic modeling, graph-based systems require careful governance to ensure fairness and mitigate unintended biases when applied to diverse populations. These factors highlight the gap between research progress and regulatory readiness for clinical adoption.

5.2. Future Research Directions

Several emerging directions have the potential to enhance the reliability, interpretability, and scalability of GNNs in medical imaging. These directions bridge fundamental research and applied development, focusing on methods that improve both model performance and translational value.

5.2.1. Graph–Transformer Hybrids

Recent work suggests that combining GNNs with transformers can leverage the strengths of both paradigms: transformers excel at modeling global dependencies, while GNNs preserve structured relational context. Hybrid models such as graph transformers and vision–graph integration have already shown promise in brain connectivity analysis and histopathology [16,17]. Future designs may adopt lightweight adapters or hierarchical schemes where GNN layers enforce local biological plausibility while transformer blocks capture long-range context. Such approaches could help overcome the scalability challenges noted in Section 5.1, particularly for high-resolution 3D imaging.

5.2.2. Self-Supervised and Foundation Models

Label scarcity continues to limit GNN performance in medical imaging. Self-supervised learning, already popular in CNN and transformer pipelines, has only recently been adapted for graphs [10]. Techniques such as contrastive learning on patient graphs, masked node prediction, and graph autoencoding can reduce reliance on expensive annotations. Moreover, foundation models pre-trained on large multimodal biomedical datasets [109] present opportunities for GNNs to serve as relational adapters, distilling broad knowledge into downstream tasks. Research in this direction could alleviate data scarcity and improve generalization across populations.

5.2.3. Federated and Privacy-Preserving GNNs

Privacy remains a central barrier to multi-institutional adoption. While federated learning has been explored for CNNs in medical imaging, federated graph learning is still nascent. Paul et al. [16] noted that cross-institution reproducibility is fragile in graph pipelines, where node and edge definitions vary. Federated GNNs with secure aggregation, personalization layers, and harmonized graph construction protocols offer a path forward. This would allow GNNs to scale across diverse sites without centralized data pooling, addressing regulatory and ethical concerns outlined earlier.

5.2.4. Multimodal and Longitudinal Integration

Medical imaging rarely exists in isolation; clinical practice involves multimodal data (imaging, omics, and EHRs) and longitudinal follow-ups. Heterogeneous and multiplexed

GNNs already demonstrate the ability to integrate modalities [93,98], but scalability and interpretability remain issues. Future research should design principled multimodal fusion strategies, potentially combining graph prompts or attention-guided message passing to capture complementary signals across modalities. Longitudinal graph construction, where temporal edges track disease trajectories, also remains underexplored. Advances in spatiotemporal GNNs could enable clinically meaningful predictions of progression, relapse, or treatment response.

5.2.5. Interpretability and Clinical Trust

Even as predictive performance improves, clinician adoption hinges on interpretability and uncertainty estimation. Current explanation methods such as GNNExplainer or attention maps remain insufficient for high-stakes contexts [9,15]. Future research must focus on case-level rationale, calibrated uncertainty reporting, and clinician-facing artifacts that can be validated in reader studies. Interpretable graph modules—for example, attention highlighting of cortical regions in Alzheimer's studies or subgraph rationales in tumor survival prediction—represent a promising step toward clinical trust and regulatory approval.

To reinforce the link between current challenges and prospective research efforts, Table 7 summarizes these aspects side by side. This synthesis provides a clear map connecting technical barriers with active or emerging strategies aimed at resolving them, ensuring logical continuity between past limitations and future progress.

Table 7. Summary of major challenges and corresponding research directions.

Challenge	Description	Emerging Research Directions
Data scarcity and imbalance	Limited annotated datasets restrict model robustness, especially for rare conditions.	Self-supervised and foundation models to exploit unlabeled data; efficient augmentation and transfer learning.
Scalability and computational overhead	High graph dimensionality and irregular topologies increase computational cost.	Lightweight graph-transformer hybrids, sparse attention, hierarchical pooling, and model compression.
Interpretability and clinical trust	Graph reasoning remains opaque, reducing clinician confidence and accountability.	Explainability via GNNExplainer or GraphSHAP; uncertainty calibration and human-in-the-loop validation.
Generalization across institutions and populations	Domain shifts from scanner and demographic variability limit reproducibility.	Domain adaptation, cross-cohort pretraining, and federated graph learning with standardized graph design.
Regulatory and deployment barriers	Inconsistent graph construction hinders reproducibility and regulatory approval.	Privacy-preserving federated GNNs, transparent reporting, and alignment with FAIR and ethical AI standards.

6. Conclusions

This survey has presented a comprehensive review of GNNs in medical imaging, detailing their architectures, learning paradigms, and diverse applications across segmentation, classification, registration, reconstruction, and multimodal fusion. By integrating recent advances in convolutional, attention-based, and spatiotemporal GNNs, the paper highlights how graph-based representations capture structural and functional dependencies that conventional deep learning models often overlook. The taxonomy and comparative analysis developed in this paper provide a consolidated reference for understanding how these architectures translate relational information into clinically relevant insights.

Despite these advances, the clinical translation of GNNs remains limited by several practical barriers. Firstly, integrating graph-based models into radiology workflows requires standardized data pipelines and computational efficiency suited for real-time analysis. Secondly, regulatory approval further depends on transparent model reporting, reproducible graph construction, and independent external validation to meet clinical

safety standards. Moreover, interpretability and uncertainty quantification remain essential for clinician trust, demanding explainable reasoning that aligns with medical expertise. Addressing these challenges will determine whether GNNs evolve from promising research tools to deployable systems capable of supporting reliable, ethical, and regulatory-compliant medical imaging practice.

Author Contributions: Conceptualization, I.D.M. and S.V.; methodology, I.D.M. and S.V.; validation, I.D.M. and S.V.; formal analysis, I.D.M.; investigation, I.D.M.; resources, S.V.; writing—original draft preparation, I.D.M.; writing—review and editing, S.V.; visualization, I.D.M.; supervision, S.V.; project administration, I.D.M. and S.V.; funding acquisition, S.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AD	Alzheimer's Disease
ADGCN	Attention-Driven Graph Convolutional Network
AUC	Area Under the Curve
BraTS	Brain Tumor Segmentation Challenge
CNN	Convolutional Neural Network
CT	Computed Tomography
CTR	Cardiothoracic Ratio
DL	Deep Learning
DMGN	Deep Multimodal Graph Network
ECG	Electrocardiogram
fMRI	Functional Magnetic Resonance Imaging
GCN	Graph Convolutional Network
GCNN	Graph Convolutional Neural Network
GAT	Graph Attention Network
GAE	Graph Autoencoder
GNN	Graph Neural Network
GNNEExplainer	Graph Neural Network Explainer
GRFE	Graph Regional Feature Enhancer
GTU-Net	Graph Transformer U-Net
HGL	Heterogeneous Graph Learning
MAP	Mean Average Precision
MCI	Mild Cognitive Impairment
MRI	Magnetic Resonance Imaging
NSCLC	Non-Small Cell Lung Cancer
PET	Positron Emission Tomography
PSNR	Peak Signal-to-Noise Ratio
RBA-Net	Region-Boundary Aggregation Network
ROI	Region of Interest
SSIM	Structural Similarity Index Measure
VGAE	Variational Graph Autoencoder

References

1. Anil, S.; Vikas, B.; Thomas, N.G.; Sweety, V.K. Biomedical imaging: Scope for future studies and applications. In *Multimodal Biomedical Imaging Techniques*; Springer: Berlin/Heidelberg, Germany, 2025; pp. 319–338.
2. Nazir, A.; Hussain, A.; Singh, M.; Assad, A. Deep learning in medicine: Advancing healthcare with intelligent solutions and the future of holography imaging in early diagnosis. *Multimed. Tools Appl.* **2025**, *84*, 17677–17740. [[CrossRef](#)]
3. Fan, X.; Liu, X.; Xia, Q.; Chen, G.; Cheng, J.; Shi, Z.; Fang, Y.; Khadaroo, P.A.; Qian, J.; Lin, H. Advanced Image-Guidance and Surgical-Navigation Techniques for Real-Time Visualized Surgery. *Adv. Sci.* **2025**, *12*, e09294. [[CrossRef](#)]
4. Li, M.; Jiang, Y.; Zhang, Y.; Zhu, H. Medical image analysis using deep learning algorithms. *Front. Public Health* **2023**, *11*, 1273253. [[CrossRef](#)] [[PubMed](#)]
5. Zhang, S.; Metaxas, D. On the challenges and perspectives of foundation models for medical image analysis. *Med. Image Anal.* **2024**, *91*, 102996. [[CrossRef](#)]
6. Ijebu, F.F.; Liu, Y.; Sun, C.; Jere, N.; Mienye, I.D.; Usip, P.U. Ensemble Answer Selection Leveraging Cross-lingual Dealignment for Improved Question Answering with Mixture-of-Experts Setup. *IEEE Open J. Comput. Soc.* **2025**, *6*, 1599–1610. [[CrossRef](#)]
7. Jiao, L.; Chen, J.; Liu, F.; Yang, S.; You, C.; Liu, X.; Li, L.; Hou, B. Graph representation learning meets computer vision: A survey. *IEEE Trans. Artif. Intell.* **2022**, *4*, 2–22. [[CrossRef](#)]
8. Li, D.; Lu, C.; Chen, Z.; Guan, J.; Zhao, J.; Du, J. Graph neural networks in point clouds: A survey. *Remote Sens.* **2024**, *16*, 2518. [[CrossRef](#)]
9. Georgousis, S.; Kenning, M.P.; Xie, X. Graph deep learning: State of the art and challenges. *IEEE Access* **2021**, *9*, 22106–22140. [[CrossRef](#)]
10. Wu, Z.; Pan, S.; Chen, F.; Long, G.; Zhang, C.; Yu, P.S. A comprehensive survey on graph neural networks. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *32*, 4–24. [[CrossRef](#)]
11. Zhou, J.; Cui, G.; Hu, S.; Zhang, Z.; Yang, C.; Liu, Z.; Wang, L.; Li, C.; Sun, M. Graph neural networks: A review of methods and applications. *AI Open* **2020**, *1*, 57–81. [[CrossRef](#)]
12. Li, M.M.; Huang, K.; Zitnik, M. Graph representation learning in biomedicine and healthcare. *Nat. Biomed. Eng.* **2022**, *6*, 1353–1369. [[CrossRef](#)]
13. Ahmedt-Aristizabal, D.; Armin, M.A.; Denman, S.; Fookes, C.; Petersson, L. Graph-based deep learning for medical diagnosis and analysis: Past, present and future. *Sensors* **2021**, *21*, 4758. [[CrossRef](#)]
14. Bessadok, A.; Mahjoub, M.A.; Rekik, I. Graph neural networks in network neuroscience. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 5833–5848. [[CrossRef](#)] [[PubMed](#)]
15. Li, R.; Yuan, X.; Radfar, M.; Marendy, P.; Ni, W.; O'Brien, T.J.; Casillas-Espinosa, P.M. Graph signal processing, graph neural network and graph learning on biological data: A systematic review. *IEEE Rev. Biomed. Eng.* **2021**, *16*, 109–135. [[CrossRef](#)] [[PubMed](#)]
16. Paul, S.G.; Saha, A.; Hasan, M.Z.; Noori, S.R.H.; Moustafa, A. A systematic review of graph neural network in healthcare-based applications: Recent advances, trends, and future directions. *IEEE Access* **2024**, *12*, 15145–15170. [[CrossRef](#)]
17. Zhang, L.; Zhao, Y.; Che, T.; Li, S.; Wang, X. Graph neural networks for image-guided disease diagnosis: A review. *Iradiology* **2023**, *1*, 151–166. [[CrossRef](#)]
18. Khosraftar, S.; An, A. A survey on graph representation learning methods. *ACM Trans. Intell. Syst. Technol.* **2024**, *15*, 1–55. [[CrossRef](#)]
19. Khatun, Z.; Jónsson, H., Jr.; Tsirilaki, M.; Maffulli, N.; Oliva, F.; Daval, P.; Tortorella, F.; Gargiulo, P. Beyond pixel: Superpixel-based MRI segmentation through traditional machine learning and graph convolutional network. *Comput. Methods Programs Biomed.* **2024**, *256*, 108398. [[CrossRef](#)]
20. Vrahatis, A.G.; Lazaros, K.; Kotsiantis, S. Graph attention networks: A comprehensive review of methods and applications. *Future Internet* **2024**, *16*, 318. [[CrossRef](#)]
21. Wang, Z.; Zhang, Z.; Ma, T.; Chawla, N.V.; Zhang, C.; Ye, Y. Beyond Message Passing: Neural Graph Pattern Machine. *arXiv* **2025**, arXiv:2501.18739.
22. Mohammadi, H.; Karwowski, W. Graph neural networks in brain connectivity studies: Methods, challenges, and future directions. *Brain Sci.* **2024**, *15*, 17. [[CrossRef](#)]
23. Zhang, S.; Yang, J.; Zhang, Y.; Zhong, J.; Hu, W.; Li, C.; Jiang, J. The combination of a graph neural network technique and brain imaging to diagnose neurological disorders: A review and outlook. *Brain Sci.* **2023**, *13*, 1462. [[CrossRef](#)]
24. Beaini, D.; Passaro, S.; Létourneau, V.; Hamilton, W.; Corso, G.; Liò, P. Directional graph networks. In Proceedings of the International Conference on Machine Learning, Virtual, 18–24 July 2021; pp. 748–758.
25. Tanglay, O.; Dadario, N.B.; Chong, E.H.; Tang, S.J.; Young, I.M.; Sughrue, M.E. Graph theory measures and their application to neurosurgical eloquence. *Cancers* **2023**, *15*, 556. [[CrossRef](#)] [[PubMed](#)]
26. Barros, C.D.; Mendonça, M.R.; Vieira, A.B.; Ziviani, A. A survey on embedding dynamic graphs. *ACM Comput. Surv. (CSUR)* **2021**, *55*, 1–37. [[CrossRef](#)]

27. Bing, R.; Yuan, G.; Zhu, M.; Meng, F.; Ma, H.; Qiao, S. Heterogeneous graph neural networks analysis: A survey of techniques, evaluations and applications. *Artif. Intell. Rev.* **2023**, *56*, 8003–8042. [[CrossRef](#)]
28. Cui, H.; Lu, Z.; Li, P.; Yang, C. On positional and structural node features for graph neural networks on non-attributed graphs. In Proceedings of the 31st ACM International Conference on Information & Knowledge Management, Atlanta, GA, USA, 17–21 October 2022; pp. 3898–3902.
29. Ji, S.; Pan, S.; Cambria, E.; Marttinen, P.; Yu, P.S. A survey on knowledge graphs: Representation, acquisition, and applications. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *33*, 494–514. [[CrossRef](#)]
30. Zhang, L.; Wang, S.; Liu, J.; Chang, X.; Lin, Q.; Wu, Y.; Zheng, Q. MuL-GRN: Multi-level graph relation network for few-shot node classification. *IEEE Trans. Knowl. Data Eng.* **2022**, *35*, 6085–6098. [[CrossRef](#)]
31. Khemani, B.; Patil, S.; Kotecha, K.; Tanwar, S. A review of graph neural networks: Concepts, architectures, techniques, challenges, datasets, applications, and future directions. *J. Big Data* **2024**, *11*, 18. [[CrossRef](#)]
32. Han, Z.; Hu, C.; Li, T.; Qi, Q.; Tang, P.; Guo, S. Subgraph-level federated graph neural network for privacy-preserving recommendation with meta-learning. *Neural Netw.* **2024**, *179*, 106574. [[CrossRef](#)]
33. Pflueger, M.; Cucala, D.T.; Kostylev, E.V. Recurrent graph neural networks and their connections to bisimulation and logic. In Proceedings of the AAAI Conference on Artificial Intelligence, Vancouver, BC, Canada, 26–27 February 2024; Volume 38, pp. 14608–14616.
34. Kipf, T. Semi-supervised classification with graph convolutional networks. *arXiv* **2016**, arXiv:1609.02907.
35. Zhou, T. M2GCNet: Multi-modal graph convolution network for precise brain tumor segmentation across multiple MRI sequences. *IEEE Trans. Image Process.* **2024**, *33*, 4896–4910. [[CrossRef](#)]
36. Arshad Choudhry, I.; Iqbal, S.; Alhussein, M.; Aurangzeb, K.; Qureshi, A.N.; Hussain, A. A novel interpretable graph convolutional neural network for multimodal brain tumor segmentation. *Cogn. Comput.* **2025**, *17*, 24. [[CrossRef](#)]
37. Kim, S.Y. Personalized explanations for early diagnosis of alzheimer’s disease using explainable graph neural networks with population graphs. *Bioengineering* **2023**, *10*, 701. [[CrossRef](#)] [[PubMed](#)]
38. Hamilton, W.; Ying, Z.; Leskovec, J. Inductive representation learning on large graphs. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017; Volume 30. Available online: https://proceedings.neurips.cc/paper_files/paper/2017/file/5dd9db5e033da9c6fb5ba83c7a7ebea9-Paper.pdf (accessed on 12 October 2025).
39. Kumar, S.; Hazarika, S.; Gupta, C.N. SAGEFusionNet: An Auxiliary Supervised Graph Neural Network for Brain Age Prediction as a Neurodegenerative Biomarker. *Brain Sci.* **2025**, *15*, 752. [[CrossRef](#)]
40. Jemima, D.D.; Selvarani, A.G.; Lovenia, J.D.L. A Novel Approach for Recognition of Autism Spectrum Disorder based on GraphSAGE. In Proceedings of the 2024 5th International Conference on Data Intelligence and Cognitive Informatics (ICDICI), Tirunelveli, India, 18–20 November 2024; pp. 1497–1502.
41. Ijebu, F.F.; Liu, Y.; Sun, C.; Jere, N.; Mienye, I.D.; Inyang, U.G. Cross-Encoder-Based Semantic Evaluation of Extractive and Generative Question Answering in Low-Resourced African Languages. *Technologies* **2025**, *13*, 119. [[CrossRef](#)]
42. Bui, K.H.N.; Cho, J.; Yi, H. Spatial-temporal graph neural network for traffic forecasting: An overview and open research issues. *Appl. Intell.* **2022**, *52*, 2763–2774. [[CrossRef](#)]
43. Kim, B.H.; Ye, J.C.; Kim, J.J. Learning dynamic graph representation of brain connectome with spatio-temporal attention. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 4314–4327.
44. Ahn, S.J.; Kim, M. Variational graph normalized autoencoders. In Proceedings of the 30th ACM International Conference on Information & Knowledge Management, Gold Coast, Queensland, Australia, 1–5 November 2021; pp. 2827–2831.
45. Zhang, R.; Zhang, Y.; Lu, C.; Li, X. Unsupervised graph embedding via adaptive graph learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 5329–5336. [[CrossRef](#)]
46. Song, Z.; Yang, X.; Xu, Z.; King, I. Graph-based semi-supervised learning: A comprehensive review. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *34*, 8174–8194. [[CrossRef](#)]
47. Li, Z.; Liu, Y.; Zhang, Z.; Pan, S.; Gao, J.; Bu, J. Cyclic label propagation for graph semi-supervised learning. *World Wide Web* **2022**, *25*, 703–721. [[CrossRef](#)]
48. Shin, H.K.; Uhm, K.H.; Choi, K.; Xu, Z.; Jung, S.W.; Ko, S.J. Graph segmentation-based pseudo-labeling for semi-supervised pathology image classification. *IEEE Access* **2022**, *10*, 93960–93970. [[CrossRef](#)]
49. Zhang, Y.; Li, Y.; Kong, Y.; Wu, J.; Yang, J.; Shu, H.; Coatrieux, G. GSCFN: A graph self-construction and fusion network for semi-supervised brain tissue segmentation in MRI. *Neurocomputing* **2021**, *455*, 23–37. [[CrossRef](#)]
50. Rani, V.; Kumar, M.; Gupta, A.; Sachdeva, M.; Mittal, A.; Kumar, K. Self-supervised learning for medical image analysis: A comprehensive review. *Evol. Syst.* **2024**, *15*, 1607–1633. [[CrossRef](#)]
51. Wen, G.; Cao, P.; Liu, L.; Yang, J.; Zhang, X.; Wang, F.; Zaiane, O.R. Graph self-supervised learning with application to brain networks analysis. *IEEE J. Biomed. Health Inform.* **2023**, *27*, 4154–4165. [[CrossRef](#)]

52. Yu, X.; Fang, Y.; Liu, Z.; Wu, Y.; Wen, Z.; Bo, J.; Zhang, X.; Hoi, S.C. A Survey of Few-Shot Learning on Graphs: From Meta-Learning to Pre-Training and Prompt Learning. *arXiv* **2024**, arXiv:2402.01440.
53. Mohammadi, S.; Allali, M. Advancing brain tumor segmentation with spectral–spatial graph neural networks. *Appl. Sci.* **2024**, *14*, 3424. [[CrossRef](#)]
54. Gaggion, N.; Mansilla, L.; Mosquera, C.; Milone, D.H.; Ferrante, E. Improving anatomical plausibility in medical image segmentation via hybrid graph neural networks: Applications to chest x-ray analysis. *IEEE Trans. Med. Imaging* **2022**, *42*, 546–556. [[CrossRef](#)] [[PubMed](#)]
55. Shen, Y.; Li, J.; Zhu, W.; Yu, K.; Wang, M.; Peng, Y.; Zhou, Y.; Guan, L.; Chen, X. Graph attention u-net for retinal layer surface detection and choroid neovascularization segmentation in oct images. *IEEE Trans. Med. Imaging* **2023**, *42*, 3140–3154. [[CrossRef](#)]
56. Xu, H.; Wu, Y. G2ViT: Graph neural network-guided vision transformer enhanced network for retinal vessel and coronary angiograph segmentation. *Neural Netw.* **2024**, *176*, 106356. [[CrossRef](#)]
57. Joshi, A.; Sharma, K. Graph deep network for optic disc and optic cup segmentation for glaucoma disease using retinal imaging. *Phys. Eng. Sci. Med.* **2022**, *45*, 847–858. [[CrossRef](#)]
58. Meng, Y.; Zhang, H.; Zhao, Y.; Yang, X.; Qiao, Y.; MacCormick, I.J.; Huang, X.; Zheng, Y. Graph-based region and boundary aggregation for biomedical image segmentation. *IEEE Trans. Med. Imaging* **2021**, *41*, 690–701. [[CrossRef](#)]
59. Li, X.; Chen, G.; Wu, Y.; Yang, J.; Zhou, T.; Zhou, Y.; Zhu, W. MedSegViG: Medical Image Segmentation with a Vision Graph Neural Network. In Proceedings of the 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Lisboa, Portugal, 3–6 December 2024; pp. 3408–3411.
60. Yang, Z.; Wang, Y. Graph-based regional feature enhancing for abdominal multi-organ segmentation in CT. In Proceedings of the 2022 IEEE 35th International Symposium on Computer-Based Medical Systems (CBMS), Shenzhen, China, 21–23 July 2022; pp. 125–130.
61. Tian, F.; Tian, Z.; Chen, Z.; Zhang, D.; Du, S. Surface-GCN: Learning interaction experience for organ segmentation in 3D medical images. *Med. Phys.* **2023**, *50*, 5030–5044. [[CrossRef](#)] [[PubMed](#)]
62. Li, L.; Xu, M.; Chen, S.; Mu, B. An adaptive feature fusion framework of CNN and GNN for histopathology images classification. *Comput. Electr. Eng.* **2025**, *123*, 110186. [[CrossRef](#)]
63. Lu, W.; Toss, M.; Dawood, M.; Rakha, E.; Rajpoot, N.; Minhas, F. SlideGraph+: Whole slide image level graphs to predict HER2 status in breast cancer. *Med. Image Anal.* **2022**, *80*, 102486. [[CrossRef](#)] [[PubMed](#)]
64. Chang, X.; Zhang, Z.; Sun, J.; Lin, K.; Song, P. Breast cancer image classification based on H&E staining using a causal attention graph neural network model. *Med. Biol. Eng. Comput.* **2025**, *63*, 1965–1979.
65. Zhang, Y.; Qing, L.; He, X.; Zhang, L.; Liu, Y.; Teng, Q. Population-based GCN method for diagnosis of Alzheimer’s disease using brain metabolic or volumetric features. *Biomed. Signal Process. Control* **2023**, *86*, 105162. [[CrossRef](#)]
66. Huynh, N.; Yan, D.; Ma, Y.; Wu, S.; Long, C.; Sami, M.T.; Almudaifer, A.; Jiang, Z.; Chen, H.; Dretsch, M.N.; et al. The use of generative adversarial network and graph convolution network for neuroimaging-based diagnostic classification. *Brain Sci.* **2024**, *14*, 456. [[CrossRef](#)]
67. Banus, J.; Ogier, A.C.; Hullin, R.; Meyer, P.; van Heeswijk, R.B.; Richiardi, J. Spatiotemporal graph neural process for reconstruction, extrapolation, and classification of cardiac trajectories. *arXiv* **2025**, arXiv:2509.12953. [[CrossRef](#)]
68. Lin, Q.; Ogleć, D.; Lam, H.K.; Curtis, M.J.; Cvetkovic, Z. A Hybrid GCN-LSTM model for ventricular arrhythmia classification based on ECG pattern similarity. In Proceedings of the 2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 15–19 July 2024; pp. 1–4.
69. Andayeshgar, B.; Abdali-Mohammadi, F.; Sepahvand, M.; Almasi, A.; Salari, N. Arrhythmia detection by the graph convolution network and a proposed structure for communication between cardiac leads. *BMC Med. Res. Methodol.* **2024**, *24*, 96. [[CrossRef](#)]
70. Zedadra, A.; Zedadra, O.; Salah-Salah, M.Y.; Guerrieri, A. Graph-Aware Multimodal Deep Learning for Classification of Diabetic Retinopathy Images. *IEEE Access* **2025**, *13*, 74799–74810. [[CrossRef](#)]
71. Pandugula, V.K.; Choudhary, A.; Uyyala, R.; Vurubindi, P. Hybrid CNN-Graph Attention Networks for Diabetic Retinopathy Grading: A Multimodal Feature Fusion Approach. In Proceedings of the 2025 3rd International Conference on Inventive Computing and Informatics (ICICI), Bangalore, India, 4–6 June 2025; pp. 953–958.
72. Ma, Q.; Lai, Z.; Wang, Z.; Qiu, Y.; Zhang, H.; Qu, X. MRI reconstruction with enhanced self-similarity using graph convolutional network. *BMC Med. Imaging* **2024**, *24*, 113. [[CrossRef](#)]
73. Ahmed, S.; Jinchao, F.; Manan, M.A.; Yaqub, M.; Ali, M.U.; Raheem, A. FedGraphMRI-net: A federated graph neural network framework for robust MRI reconstruction across non-IID data. *Biomed. Signal Process. Control* **2025**, *102*, 107360. [[CrossRef](#)]
74. Wang, J.; Yang, Y.; Yang, H.; Lian, C.; Xu, Z.; Sun, J. MD-GraphFormer: A model-driven graph transformer for fast multi-contrast MR imaging. *IEEE Trans. Comput. Imaging* **2023**, *9*, 1018–1030. [[CrossRef](#)]
75. Xia, W.; Lu, Z.; Huang, Y.; Shi, Z.; Liu, Y.; Chen, H.; Chen, Y.; Zhou, J.; Zhang, Y. MAGIC: Manifold and graph integrative convolutional network for low-dose CT reconstruction. *IEEE Trans. Med. Imaging* **2021**, *40*, 3459–3472. [[CrossRef](#)]

76. Tang, Z.; Sun, Z.H.; Wu, E.Q.; Wei, C.F.; Ming, D.; Chen, S.D. MRCG: A MRI retrieval framework with convolutional and graph neural networks for secure and private IoMT. *IEEE J. Biomed. Health Inform.* **2021**, *27*, 814–822. [CrossRef] [PubMed]
77. Isallari, M.; Rekik, I. Brain graph super-resolution using adversarial graph neural network with application to functional brain connectivity. *Med. Image Anal.* **2021**, *71*, 102084. [CrossRef] [PubMed]
78. Tarasiewicz, T.; Kawulok, M. Multi-Image Super-Resolution Using Graph Neural Networks. In *Super-Resolution for Remote Sensing*; Springer: Berlin/Heidelberg, Germany, 2024; pp. 93–153.
79. Ye, H.; Zhang, X.; Hu, Y.; Fu, H.; Liu, J. Vsr-net: Vessel-like structure rehabilitation network with graph clustering. *IEEE Trans. Image Process.* **2025**, *34*, 1090–1105. [CrossRef] [PubMed]
80. Liang, H.; Lv, J.; Wang, Z.; Xu, X. Medical image mis-segmentation region refinement framework based on dynamic graph convolution. *Biomed. Signal Process. Control* **2023**, *86*, 105064. [CrossRef]
81. Yang, T.; Bai, X.; Cui, X.; Gong, Y.; Li, L. GraformerDIR: Graph convolution transformer for deformable image registration. *Comput. Biol. Med.* **2022**, *147*, 105799. [CrossRef]
82. Zhou, Y.; Cao, W. H-SGANet: Hybrid sparse graph attention network for deformable medical image registration. *Neurocomputing* **2025**, *633*, 129810. [CrossRef]
83. Wang, L.; Yan, Z.; Cao, W.; Ji, J. Diegraph: Dual-branch information exchange graph convolutional network for deformable medical image registration. *Neural Comput. Appl.* **2023**, *35*, 23631–23647. [CrossRef]
84. Hansen, L.; Heinrich, M.P. GraphRegNet: Deep graph regularisation networks on sparse keypoints for dense registration of 3D lung CTs. *IEEE Trans. Med. Imaging* **2021**, *40*, 2246–2257. [CrossRef]
85. Ren, J.; An, N.; Zhang, Y.; Wang, D.; Sun, Z.; Lin, C.; Cui, W.; Wang, W.; Zhou, Y.; Zhang, W.; et al. SUGAR: Spherical ultrafast graph attention framework for cortical surface registration. *Med. Image Anal.* **2024**, *94*, 103122. [CrossRef]
86. Suliman, M.A.; Williams, L.Z.; Fawaz, A.; Robinson, E.C. Unsupervised multimodal surface registration with geometric deep learning. *arXiv* **2023**, arXiv:2311.13022. [CrossRef]
87. Cheng, J.; Dalca, A.V.; Fischl, B.; Zöllei, L.; The Alzheimer’s Disease Neuroimaging Initiative. Cortical surface registration using unsupervised learning. *NeuroImage* **2020**, *221*, 117161. [CrossRef]
88. Zhang, R.; Wang, L.; Tang, K.; Xu, J.; Wei, H. GESH-Net: Graph-Enhanced Spherical Harmonic Convolutional Networks for Cortical Surface Registration. *arXiv* **2024**, arXiv:2410.14805.
89. Tan, J.; Ren, X.; Chen, Y.; Yuan, X.; Chang, F.; Yang, R.; Ma, C.; Chen, X.; Tian, M.; Chen, W.; et al. Application of improved graph convolutional network for cortical surface parcellation. *Sci. Rep.* **2025**, *15*, 16409. [CrossRef]
90. Arias-García, J.; García, H.F.; Escobar-Mejía, A.; Cárdenas-Peña, D.; Orozco, Á.A. Dynamic Graph Analysis: A Hybrid Structural-Spatial Approach for Brain Shape Correspondence. *Mach. Learn. Knowl. Extr.* **2025**, *7*, 99. [CrossRef]
91. D’Souza, N.S.; Wang, H.; Giovannini, A.; Foncubierta-Rodriguez, A.; Beck, K.L.; Boyko, O.; Syeda-Mahmood, T.F. Fusing modalities by multiplexed graph neural networks for outcome prediction from medical data and beyond. *Med. Image Anal.* **2024**, *93*, 103064. [CrossRef] [PubMed]
92. Kim, S.; Lee, N.; Lee, J.; Hyun, D.; Park, C. Heterogeneous graph learning for multi-modal medical data analysis. In Proceedings of the AAAI Conference on Artificial Intelligence, Washington, DC, USA, 7–14 February 2023; Volume 37, pp. 5141–5150.
93. Fu, X.; Patrick, E.; Yang, J.Y.; Feng, D.D.; Kim, J. Deep multimodal graph-based network for survival prediction from highly multiplexed images and patient variables. *Comput. Biol. Med.* **2023**, *154*, 106576. [CrossRef] [PubMed]
94. Wang, Z.; Liu, Z.; Ma, T.; Li, J.; Zhang, Z.; Fu, X.; Li, Y.; Yuan, Z.; Song, W.; Ma, Y.; et al. Graph Foundation Models: A Comprehensive Survey. *arXiv* **2025**, arXiv:2505.15116. [CrossRef]
95. Tang, H.; Yang, H.; Zhang, W. DAHG: A Dynamic Augmented Heterogeneous Graph Framework for Precipitation Forecasting with Incomplete Data. *Information* **2025**, *16*, 946. [CrossRef]
96. Zhang, Y.; He, X.; Chan, Y.H.; Teng, Q.; Rajapakse, J.C. Multi-modal graph neural network for early diagnosis of Alzheimer’s disease from sMRI and PET scans. *Comput. Biol. Med.* **2023**, *164*, 107328. [CrossRef]
97. Cai, L.; Zeng, W.; Chen, H.; Zhang, H.; Li, Y.; Feng, Y.; Yan, H.; Bian, L.; Siok, W.T.; Wang, N. MM-GTUNets: Unified multi-modal graph deep learning for brain disorders prediction. *IEEE Trans. Med. Imaging* **2025**, *44*, 3705–3716. [CrossRef] [PubMed]
98. Zheng, Y.; Conrad, R.D.; Green, E.J.; Burks, E.J.; Betke, M.; Beane, J.E.; Kolachalam, V.B. Graph attention-based fusion of pathology images and gene expression for prediction of cancer survival. *IEEE Trans. Med. Imaging* **2024**, *43*, 3085–3097. [CrossRef] [PubMed]
99. Wu, J.; Ke, X.; Jiang, X.; Wu, H.; Kong, Y.; Shao, L. Leveraging tumor heterogeneity: Heterogeneous graph representation learning for cancer survival prediction in whole slide images. *Adv. Neural Inf. Process. Syst.* **2024**, *37*, 64312–64337.
100. Peng, L.; Cai, S.; Wu, Z.; Shang, H.; Zhu, X.; Li, X. Mmgpl: Multimodal medical data analysis with graph prompt learning. *Med. Image Anal.* **2024**, *97*, 103225. [CrossRef]
101. Sharma, A.; Sharma, A.; Guo, K. Intelligent Medical Diagnosis Model Based on Graph Neural Networks for Medical Images. *CAAI Trans. Intell. Technol.* **2025**, *10*, 1201–1216. [CrossRef]

102. Gawlikowski, J.; Tassi, C.R.N.; Ali, M.; Lee, J.; Humt, M.; Feng, J.; Kruspe, A.; Triebel, R.; Jung, P.; Roscher, R.; et al. A survey of uncertainty in deep neural networks. *Artif. Intell. Rev.* **2023**, *56*, 1513–1589. [[CrossRef](#)]
103. Mienye, I.D.; Obaido, G.; Emmanuel, I.D.; Ajani, A.A. A survey of bias and fairness in healthcare AI. In Proceedings of the 2024 IEEE 12th International Conference on Healthcare Informatics (ICHI), Orlando, FL, USA, 3–6 June 2024; pp. 642–650.
104. Ying, Z.; Bourgeois, D.; You, J.; Zitnik, M.; Leskovec, J. Gnnexplainer: Generating explanations for graph neural networks. In Proceedings of the 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, BC, Canada, 8–14 December 2019; Volume 32. Available online: https://proceedings.neurips.cc/paper_files/paper/2019/file/d80b7040b773199015de6d3b4293c8ff-Paper.pdf (accessed on 12 October 2025).
105. Luo, D.; Cheng, W.; Xu, D.; Yu, W.; Zong, B.; Chen, H.; Zhang, X. Parameterized explainer for graph neural network. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 19620–19631.
106. Perotti, A.; Bajardi, P.; Bonchi, F.; Panisson, A. GRAPHSHAP: Explaining Identity-Aware Graph Classifiers Through the Language of Motifs. *arXiv* **2022**, arXiv:2202.08815.
107. Huang, Q.; Yamada, M.; Tian, Y.; Singh, D.; Chang, Y. Graphlime: Local interpretable model explanations for graph neural networks. *IEEE Trans. Knowl. Data Eng.* **2022**, *35*, 6968–6972. [[CrossRef](#)]
108. Liu, Z.; Wan, G.; Prakash, B.A.; Lau, M.S.; Jin, W. A review of graph neural networks in epidemic modeling. In Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Barcelona, Spain, 25–29 August 2024; pp. 6577–6587.
109. Mienye, I.D.; Swart, T.G. Ensemble Large Language Models: A Survey. *Information* **2025**, *16*, 688. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.