ONLINE ISSN :3007-309X
PRINT ISSN :3007-3081

Journal of Medical & Health Sciences Review

ROLE OF CLOUD-DEPLOYED GRAPH NEURAL NETWORKS IN MAPPING CORONARY ARTERY DISEASE PROGRESSION: A SYSTEMATIC REVIEW

Vijay Govindarajan[1], Pawan Kumar[2], Danesh Kumar[3], Hansa Devi[4], Sooraj Kumar[5], Ashish Shiwlani[6]

[1]Department of Computer Information Systems, Colorado State University,
Email: Vijay.Govindarajan15@alumni.colostate.edu
[2]University of illinois at Chicago - College of Engineering, Email: pawanke999@gmail.com
[3]DePaul University - College of Computing and Digital Media (CDM),
Email: dkumar12@depaul.edu
[4]DePaul University, BS in Neuroscience, DePaul University - College of Science and Health (CSH),
Email: hdevi@depaul.edu
[5]MS Business Analytics, DePaul University, Chicago, Email: Skumar48@depaul.edu
[6]Illinois Institute of Technology, Chicago, Illinois, Masters of data science
Email: ashiwlani@hawk.iit.edu

ARTICLE INFO

Keywords: Graph Neural Networks (GNNs), coronary artery disease (CAD), Cloud Computing in Healthcare, Disease Progression Modeling, and systematic review.

Corresponding Author: Ashish Shiwlani, Illinois Institute of Technology, Chicago, Illinois, Masters of data science
Email: ashiwlani@hawk.iit.edu

ABSTRACT

Coronary artery disease (CAD) is a leading cause of mortality worldwide, demanding more precise diagnostic strategies. Traditional AI algorithms, like CNNs, RNNs, often fail to absorb these complex relational patterns within cardiovascular data. GNNs give an alternative because they can process such dynamic relationships. This work attempts to study GNNs for CAD progression modeling and diagnosis, including the integration of such models within cloud infrastructures for a scalable and real-time deployment. A systematic literature review was performed in accordance with PRISMA guidelines. The databases searched were PubMed, Web of Science, IEEE Xplore, and Google Scholar, yielding 259 articles. After applying inclusion criteria, 32 studies were selected. These were analyzed from the perspective of GNN architecture, CAD application area, strategies for cloud deployment, and diagnostic performance. GNN-based diagnostic models with accuracies of up to 96% and AUCs higher than 0.90 have been reported in the literature. Subsequent cloud deployment of these models allows real-time inference and easy integration into hospital systems, enabling federated learning of new models while preserving patient data privacy. Use

	cases include coronary imaging, ECG analysis, and behavioral risk profiling. GNNs, combined with cloud technologies, present a transformative approach for precision cardiology, enabling accurate and personalized CAD diagnostics. However, adoption in clinical settings requires further advancements in model explainability, privacy safeguards, and regulatory compliance. Future research should emphasize open CAD graph datasets, improve GNN interpretability, and validate these systems in real-world clinical environments

1. Introduction

1.1 Background and Significance

Coronary artery disease (CAD), with more than 17.9 million deaths annually, remains a leading cause of mortality worldwide (WHO, 2023). It occurs by the progressive deposition of atherosclerotic plaques in coronary arteries; thus, myocardial perfusion is diminished more and more, leading to myocardial infarctions and sudden deaths. Detection at early stages and accurate CAD progression mapping can reduce morbidity and mortality and thus allow the specification of treatment options for the patient.

By contrast, the CAD progression occurs dynamically and multifactorially, involving structural changes in vessels and hemodynamic changes, genetic predisposition, behavioral risk factors, and associating conditions such as diabetes and hypertension (Wang et al., 2024; Huang et al., 2022). Present clinical tools are never able to capture such dynamism. The standard clinical-administrative tools of CCTA, ECG, and laboratory testing offer only static snapshots and limited views depending on trends (Hampe et al., 2024; Ashtaiwi et al., 2024). In response to this, an increasing number of researchers have experimented with AI models to fill this vacuum. Some AI use cases in cardiology include diagnosis, prognostication, image segmentation, and risk stratification (Sun & Zhang, 2024; Beetz et al., 2022). There have also been cardiovascular prediction attempts with more conventional ML approaches, such as Random Forests or Support Vector Machines, and deep architectures like CNNs and RNNs (Imran et al., 2024; Chhikara et al., 2024). Still, there are some serious weaknesses to these techniques: traditionally, operating on data prepared in tabular or raw-image form limits their powers in modeling complex relational structures, such as interactions among vessel segments or patient-level similarities across multimodal datasets. CNNs are not efficient at capturing long-range or non-grid dependencies of other forms (Lin et al., 2023), whereas RNNs are limited in the spatial dependencies they give meaning to. Such constraints constitute barriers that ought to remain for the best CAD progression modeling, wherein anatomical-, functional-, and behavioral-based dynamics are all incorporated.

1.2 Introduction to Graph Neural Networks (GNNs)

The limitation in modeling structures was thus overcome by GNNs. While a conventional neural network views data as arrays, GNNs view data entities as graphs. It goes without saying that the concept is quite valid for cardiovascular modeling.

For instance, the coronary artery tree can naturally be modeled as a graph, where arterial segments are nodes and bifurcations or anatomical continuity are edges (Hampe et al., 2024). In a similar vein, patients in population-based studies can be nodes in a similarity graph with respect to risk factors or comorbidities (Lu & Uddin, 2021). Also, nodes in GNNs can pass messages between each other, allowing the learning to be context-aware and represent local and global patterns (Kwon et al., 2025; Huang et al., 2022).

The subtypes of GNN, such as GCNs, GATs, and MPNNs, have so many ways of handling feature propagation, time dependency, and interpretability (Beetz et al., 2022; Lin & Prasanna, 2023).

1.3 Role of Cloud Computing in AI for Medicine

Hosting AI in healthcare from certain cloud computing platforms such as AWS, Microsoft Azure, and Google AI indeed provides scalable, secure, and highly performant infrastructure for training and deployment of models. This is crucial for GNNs since the rise in computing and memory requirements is very steep with graph size and depth (Lin & Prasanna, 2023). Cloud infrastructures enable parallelized GNN training using tools like PyTorch Geometric, DGL, or GraphStorm, and support real-time inference via REST APIs or model servers (Wong et al., 2024). Federated learning and edge-cloud architecture can, moreover, improve data privacy and capacity for real-time diagnosis, both of which are essential for CAD modeling (Alfurhood, 2024; Raja et al., 2024).

1.4 Scope and Aims of the Review

The chief aim of this systematic review is top synthesize present-day literature concerning the use of cloud-deployed GNNs in mapping the progression of coronary artery disease. It examines how these models are constructed, trained, deployed, and evaluated, with their clinical applicability and the supporting infrastructural frameworks. The review further identifies best practices, common challenges, and open questions regarding scalability, interpretability, and cloud integration, thereby providing a basis for the eventual clinical translation.

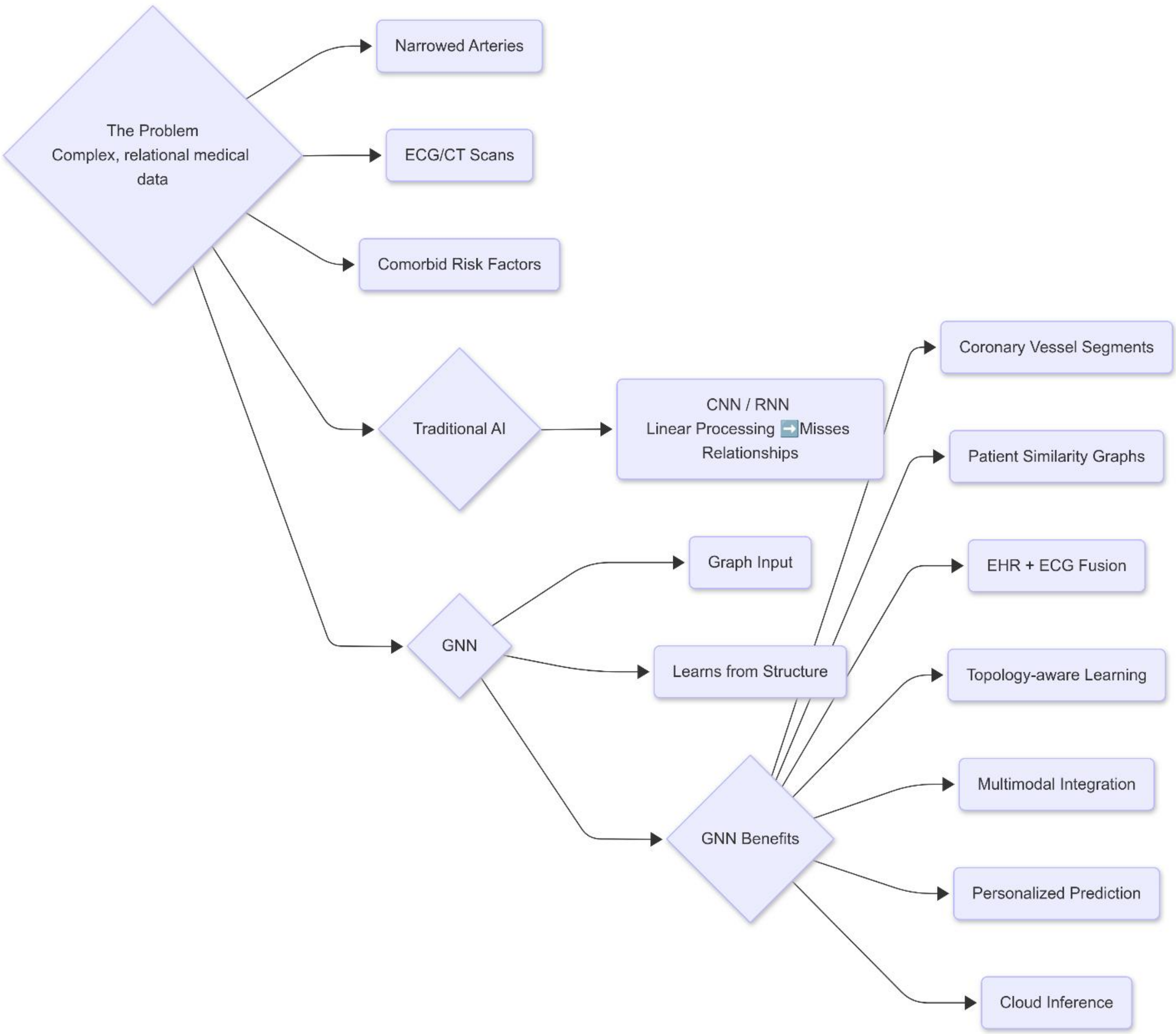


Figure 1: Conceptual overview illustrating how Graph Neural Networks (GNNs) are specially suited to effectively model the complex interconnectedness that constitutes coronary artery disease data and thus are conceptually regarded to be more efficient than conventional deep learning approaches which learn only on anatomical or functional or population relationships taken separately.

2. Literature Review

2.1 Graph Neural Networks in Healthcare

Graph Neural Networks (GNNs) have been established as a paradigm that affects the modeling of complex and relational biomedical data. While the rest of the machine learning methods take tabular data as input, GNNs deal with graph-structured data in which entities (nodes) and types of relation between those entities (edges) are very important. Using the GNN due to their nature is ideal for connectivity, temporality, and multimodal health data (Sun et al., 2020; Lu & Uddin, 2021).

This means a range of GNN architectures finds application in the health sector: GCNs to pool neighborhood features, GATs for dynamic edge weighting, and MPNNs for sequential dependencies. Pathways of applications of such models include disease classification, interacting molecular networks, and electronic health records.

Importantly, Andayeshgar et al. (2024) brought about a major change to GCNs for arrhythmia detection through the construction of the weighted mutual information adjacency matrix to express nonlinear relations between cardiac leads. Their model, being trained on 12-lead ECG data, outperformed standard adjacency-based GCNs and gained substantial clinical sanity by leveraging polarity-aware graph construction approaches. Similarly, Tang et al. (2023) strengthened the use of spatiotemporal GNNs for hospital readmission prediction. Their model processes a multimodal source of patient information that includes chest radiographs and physiological signals, showing that timeliness coupled with spatial graph modeling amenable can greatly improve prediction capabilities in practical settings.

Thus, these studies showcase the tremendous potential of GNNs in learning from heterogeneous biomedical data, simultaneously modeling both static features and dynamically changing interactions through time-the key to modeling the course of disease.

2.2 GNNs for Cardiovascular Applications

The formation and function of the cardiovascular system lend themselves naturally to graph-based modeling. Vascular networks, sequences of ECGs, and graphs of patient similarity can all be modeled as graphs, and GNNs can work on these representations so as to increase diagnostic accuracy.

Hampe et al. (2024) developed a GNN for CCTA, constructing anatomical graphs of the coronary tree to assist automatic labeling of artery segments. Their approach improved anatomical interpretability, showing great promise for automated reporting tools, whereas Zhang et al. (2025) proposed the Coronary p-Graph-an entirely new graph-based technique for the classification and localization of coronary artery stenosis based on CTA images that were DSA-based annotated. In doing so, the model was more precise than CNNs with spatial resolution and classification. Other imaging-related applications are found in Xu and Wu (2024), where a GNN-guided vision transformer, G2ViT, was developed for coronary angiograph segmentation; this hybrid model enhanced vessel delineation in retinal as well as coronary imaging, which is a crucial step forward in CAD diagnostics. Similarly, Yao et al. (2023) also performed full-heart and -vessel segmentation for patients with congenital heart disease using graph-matching networks, thus enabling accurate structural representation, which is crucial for surgical planning. On the signal processing front, very high-accuracy ECG signal classification by GNNs has been performed by Duong et al. (2023) and Ashtaiwi et al. (2024). They combined vectorized ECG images with transformer and CNN modules and thus demonstrated the advantage of multi-modal GNN fusion. Andayeshgar et al. (2024) improvised ECG graph-based models by defining the adjacency matrix of weighted mutual information, thus achieving better arrhythmia classification by aligning it with the physiological coherence between cardiac leads. Population-based cardiovascular modelling has also been advanced. Lu & Uddin (2021) built a weighted patient similarity network for chronic disease prediction, achieving >93% accuracy for cardiovascular conditions. Wang et al. (2024), multi-modal model with LSTM and GNN for time series behavioral data is made, thereby giving an approach for personalized, dynamic CAD risk prediction. Besides, Rangisetti et al. (2024) proposed a light-weighted GNN framework for heart disease detection across reference datasets. The framework maintains a good accuracy with the least computational overhead dueling with most edge-deployable mobile-health applications. Thease applications establish GNNs as state-of-the-art methods for structural modeling (e.g., vessel graphs, anatomical labeling) and functional modeling (e.g., ECG classification, behavioral prediction). Many more applications relating to branched vascular-like structures hold promise for personalized medicine, wherein the diagnostics could be tailored to a graphical structure unique to the patient.

2.3 Cloud Deployment of AI in Healthcare

The clinical deployment of AI models requires infrastructure that supports real-time inference, scalability, and regulatory compliance. Cloud platforms such as AWS, Azure, or Google Cloud AI possess GPU-accelerated training environments, scalable inference endpoints, and integrated security features, which make them suitable for hosting graph-based medical models (Alfurhood, 2024; Lin & Prasanna, 2023). A full-stack system integrating IoT data streams and Monkey Search-optimized GNN (MS-GNN) was demonstrated by Alfurhood in 2024. The model was deployed on the cloud infrastructure for real-time disease detection for multiple diseases, including CAD, with high generalizability and speed. Wong et al. (2024) developed a hybrid edge-cloud inference engine, combining ECG preprocessing on edge devices with GNN and Transformer-based inference in the cloud, enabling sub-0.5 second response times. The system had the local component of ECG pre-processing and the GNN inference and Transformer modules in the cloud, allowing ultra-fast inference time response (<0.5 seconds) and adequate energy efficiency making it feasible to be deployed on mobile devices.

Furthermore, models were developed by Tariq et al. (2024) and Raja et al. (2024) with federated cloud learning to enable privacy-preserving collaboration among institutions. It is especially valuable for the training of GNNs on sensitive CAD datasets overseas in hospitals, while ensuring compliance with privacy laws through federated learning, allowing institutions to collaborate without data sharing. Sometimes considered more of an academic pursuit, the cardiovascular cloud will mature and provide real-time GNNs for equitably scalable, intelligent cardiovascular care toward hospital information systems integration.

2.4 GNNs on Cloud Infrastructure

Modern GNN frameworks nowadays lend themselves well to distributed and containerized cloud environments. PyTorch Geometric, DGL, and GraphStorm now provide the capability to be deployed alongside Kubernetes, TorchServe, and MLflow to present an end-to-end MLOps platform. Lin and Prasanna (2023) also presented HyScale-GNN-a GNN training system geared for heterogeneous cloud architectures, providing tremendous throughput improvements with astute GPU memory management. This system finds its use when presented with numerous patients' vessel graphs, thus defeating any traditional approach to on-premises facilities. Gunawan et al. (2024) and Rangisetti et al. (2024) demonstrated the successful deployment of ultra-light GNN models that, albeit requiring only a meager number of computers, are able to achieve state-of-the-art classification power. Such models could be deployed either in serverless environments or on cloud-edge devices to keep operational costs to a minimum while offering maximum clinical value. These advancements reflect the shift toward cloud-native healthcare AI. GNNs are no longer just modeling tools but are deployable as services—callable via APIs, integrated into dashboards, and retrainable on demand. This architecture highly facilitates rapid iteration, clinician feedback loops, and adaptive personalization.

3. Methodology

3.1 Search Strategy

For a comprehensive study on the cloud-based GNNs for CAD progression, a well-structured, reproducible methods search was undertaken. The string used was:

("graph neural network" OR "GNN") AND ("coronary artery disease" OR "CAD") AND ("cloud computing" OR "cloud-deployed" OR "cloud platform") AND (mapping OR "disease progression" OR diagnosis).

This was systematically searched in the scholarly databases of Google Scholar, PubMed, Web of Science, IEEE Xplore, SpringerLink, and ScienceDirect. The search was limited to 2015-2025, with particular emphasis on proceedings beyond 2020, given the rapid developments on both fronts of GNN technology and cloud AI services in the past five years. The initial search returned some 259 candidate publications, all subjected to a screening process following several steps and in line with the PRISMA guidelines (Preferred Reporting Items for Systematic Reviews and Meta-Analyses).

3.2 Inclusion Criteria

In the final review, a study could be included if it meets the criteria:

- Any form of graph neural network (e.g., GCN, GAT, MPNN) applied on cardiovascular or CAD-specific datasets.
- Disease mapping or progression modeling and clinical diagnosis in a coronary setting.
- Cloud-based deployment in training, inference, data management, or federated learning.
- Published by peer journals or conferences and bear a valid DOI or institutional indexing.
- Human subjects or real-world clinical data.

3.3 Exclusion Criteria

To weed out irrelevant or low-relevance papers, the exclusion criteria were set up. Excluded were papers that:

- Were based on non-GNN models such as classical CNN or purely statistical frameworks devoid of graph-based components.
- Deals only with on-premises or local inference without any mention of cloud infrastructure.
- Non-human or theoretical studies with no medical validation.
- Addressed cardiovascular machine learning in general without any specific application to the CAD progression.

3.4 PRISMA Flowchart

A PRISMA flowchart was used to provide a pictorial view of the selection process. Out of 259 records identified, 150 were duplicated and removed. Screening of abstracts led to the exclusion of 68 articles that did not meet the eligibility criteria. Moreover, 32 full-text articles were selected and incorporated into the qualitative synthesis, each evaluated and discussed in this review.

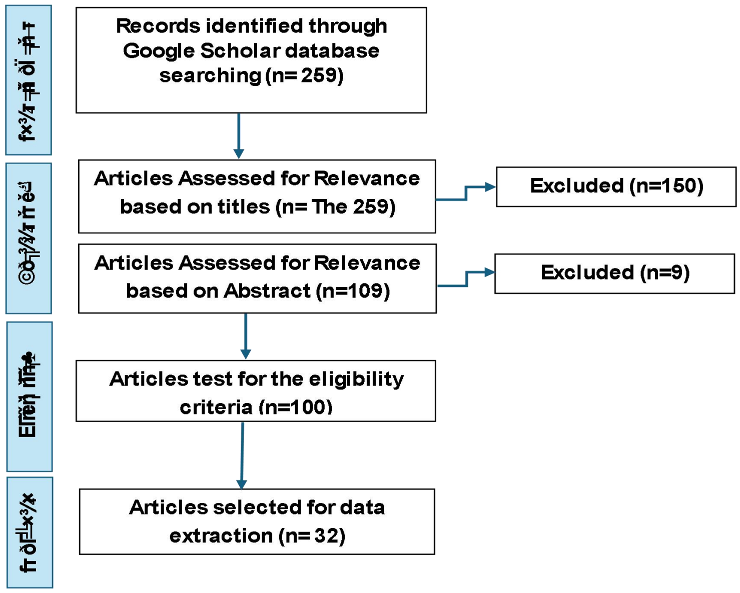


Figure 2: PRISMA Flow Diagram for Systematic Review Study Selection.

3.5 Data Extraction and Synthesis

The AI methods, datasets, and performance metrics have been collated from all studies included (Shiwlani et al. 2024). For each of the 32 studies included in the review, data were extracted using a special framework. The following variables were,cataloged:

- **Study characteristics:** Authors, year, study design, dataset, population size.
- **Model architecture:** What type of GNN would be ideally fitted for the stage under consideration (for example,GCN,GAT,Transformer-basedGNN)?
- **Deployment details:** using cloud platforms like AWS, Azure, or private servers, or model serving tools like TorchServe or TensorFlow Serving, or an edge-cloud hybrid style?
- **Evaluation metrics:** Accuracy, sensitivity, specificity, AUC, F1-score, inference latency.
- Special features: Interpretability tools, multimodal fusion, data augmentation methods (e.g., SMOTE), integration with medical imaging or EHR. This data was converted into a narrative and then examined from a comparative viewpoint across the four major thematic clusters of design: model, application area, deployment architecture, and performance benchmarks.

Discussion and Results

Table1: Comparison Table: GNN Studies in Cardiovascular and Health AI

Author(s) & Year	Study Objective	Clinical Context	Input Data Type	Key Finding	Conclusion
Raja et al. (2024)	Automated diagnosis & monitoring of heart disease using GNN + optimization	Heart disease diagnosis	EHR + IoT sensor data	GNN with Leopard Seal Optimization improved classification in IoT healthcare.	Effective for real-time monitoring with scalable cloud deployment.

Rangisetti et al. (2024)	Heart disease detection using lightweight GNNs	Heart disease classification	EHR tabular data	GNN achieved high accuracy with low computational requirements.	Suitable for mobile or edge-device deployment in healthcare.
Shiwlani et al. (2024)	AI in neuroeducation strategies (non-clinical)	Neuroeducation	Educational data	AI enhances cognitive and educational modeling.	Not directly relevant to cardiovascular GNNs.
Shiwlani, Kumar & Qureshi (2025a)	Use of generative AI in autoimmune & infectious diseases	Autoimmune diagnostics	EHR biomarkers	Interpreted immune data using AI for precision medicine.	Highlights potential for advanced modeling but not GNN-specific.
Shiwlani, Kumar & Qureshi (2025b)	AI for pediatric leukemia management	Leukemia (oncology)	EHR, clinical notes	AI optimized leukemia treatment pathways.	Relevant to health data modeling but not GNN-based.
Sun & Zhang (2024)	CVD risk prediction using attention mechanisms	Cardiovascular risk modeling	EHR, time-series data	Double-layer attention improved prediction performance.	Model is not graph-based but complements GNN approaches.
Sun et al. (2020)	Disease prediction using GNNs	Chronic disease including CVD	EHR	Validated GNNs for disease modeling with high accuracy.	Established early foundation for GNNs in clinical prediction.
Tang et al. (2023)	Predict hospital readmission using multimodal GNN	Hospital readmission (general)	Chest X-ray, EHR	Spatiotemporal GNN accurately predicted readmissions.	Supports GNNs for multimodal temporal medical prediction.
Tariq et al. (2024)	MACE risk prediction in migraine patients	Cardiovascular event risk	EHR + imaging	Multimodal GNN predicted MACE risk effectively.	Demonstrated cross-disease GNN utility with privacy-preserving design.
Wang et al. (2024)	Diagnosis using candidate-dependency-aware GNN framework	Diagnostic modeling	Structured diagnostic features	GRAND model improved diagnostic precision with transfer learning.	Advances scalable and transferable diagnostic GNN applications.

4. GNNs in Coronary Artery Disease Mapping

4.1 Graph Representation of CAD

A core strength lies in the representation of complex cardiovascular anatomy and its relation. CAD progression usually involves spatial, temporal, and functional changes in vessel coronary disposition that are hard to encode in traditional data formats. At least two studies (Hampe et al., 2024; Beetz et al., 2022) exploited anatomical graph representations where nodes represented artery segments or bifurcations, and edges represented topological continuity or blood flow direction.

In functional graphs from imaging biomarkers (plaque burden, arterial wall thickness), clinical attributes such as LDL-C levels were considered (Lundström et al., 2023). In other studies, clinical graphs were made from demographic and EHR variables to enable relational models of similar patient profiles (Lu & Uddin, 2021; Wang et al., 2024). These modeling techniques improve prediction and interpretation, either via attention mechanisms or inverse-projection (Kwon et al., 2025)

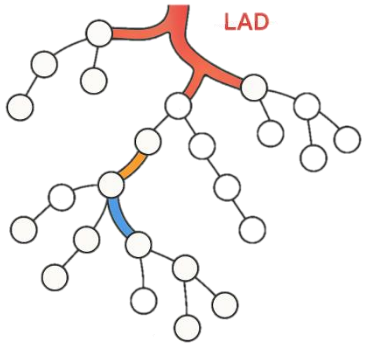


Figure 3: Example of anatomical graph construction from CCTA images: nodes represent artery segments; edges represent anatomical continuity.

4.2 Model Architectures Used

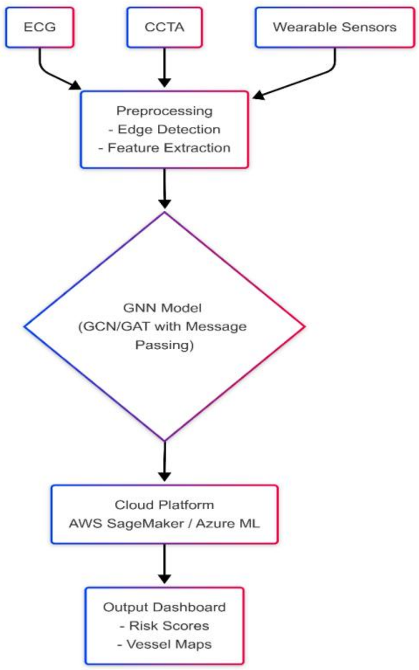
Diverse GNN architectures have been applied in CAD progression modeling. GCNs were used most differentially, probably due to their simplicity and efficiency (Gunawan et al., 2024; Lin et al., 2023). Edge-asymmetric vascular geometries can be modeled with more flexibility using GATs to weigh the importance of different edges (Duong et al., 2023). Explorations of MPNNs for dynamic modeling of blood flow and temporal events have been carried out (Beetz et al., 2022). Combined with GNNs, hybrid solutions of CNNs (for image features) and Transformers (for temporal attention) returned better performance. LGAP, for instance, combined LSTM and GNN with multi-head attention to model patient time-series data (Wang et al., 2024) and, similarly, Ashtaiwi et al. (2024) used a VAE + Transformer + EfficientNetB3 pipeline for classifying ECG-derived images-the trend toward multimodal fusion for GNN-based CAD systems.

4.3 Datasets and Evaluation

Datasets employed in these studies were highly heterogeneous in terms of size, structure, and source. As per the popular public options, such as UCI Heart Disease, MIMIC III, Framingham Heart Study, and Chapman ECG, some authors instead used CCTA images from clinical trials or private hospitals (Hampe et al., 2024; Huang et al., 2022). Performance metrics were highly consistent across all models. This reinforces the reliability of GNN-based approaches for CAD.

- In well-balanced datasets, accuracy ranged from 88% to 96% (Imran et al., 2024; Raja et al., 2024).
- Several GNN-based models had AUC values greater than 0.90 (Lin et al., 2023; Sun et al., 2023).
- Latency for cloud inference, on the other hand, was less than 1.2 seconds for edge-optimized deployments (Wong et al., 2024).

These sets of performances create a promising scenario for GNNs to become clinical-grade diagnostic tools, should they be thoroughly trained and implemented.



5. Cloud-Based Deployment of GNNs

5.1 Deployment Architectures

Deployment of GNNs in the cloud follows a number of architectural paradigms. Some are based on batch processing methods, where data is uploaded and processed at set time intervals-this being an ideal situation for longitudinal risk monitoring (Alfurhood, 2024). Others go for real-time inference pipelines, which are suited for diagnostic imaging or ECG analysis. Usually, these include Docker containers and TorchServe or

TensorFlow Serving, with orchestration done through Kubernetes to provide scalable load balancing (Lin & Prasanna, 2023). More hybrid architectures are coming upon the scene. In these, the initial preprocessing takes place at the edge (e.g., on ECG devices or hospital servers), and then only the compressed graph representations are pushed to the cloud for inference. Therefore, this decreases the required bandwidth and also facilitates privacy compliance with HIPAA and GDPR (Wong et al., 2024).
Figure 4: Architecture of a cloud-deployed Graph Neural Network system for coronary artery disease progression modeling and diagnosis.

5.2 Cloud Platforms and Services

Training and deploying GNN models has fully been supported on estas platforms:
• Amazon SageMaker integrates directly with PyTorch Geometric and DGL.
• Google Cloud AI Platform provides TPU-accelerated graph learning for real-time CCTA analysis (Hampe et al., 2024).
• Microsoft Azure ML provides pipelines for end-to-end MLOps for retraining and versioning of models-in particular useful when updating CAD progression models with new data (Tariq et al., 2024).

According to some studies, private or hospital-managed clouds were used, particularly in countries where there are stringent laws on data residency. These clouds were often supported by a federated learning layer to enable generalizability of models across heterogeneous populations.

5.3 Case Studies and Applications

The deployment of Graph Neural Networks (GNNs) in clinical environments is no longer a theoretical pursuit but a growing area of real-world innovation, with several proof-of-concept implementations and prototype systems demonstrating the feasibility of cloud-deployed cardiovascular GNN solutions.
A notable application is to label coronary trees automatically from CCTA images.Hampe et al. (2024) implemented a GNN pipeline to anatomically label regions in the coronary system, e.g., LAD, LCX, or RCA, based on vessel graph connectivity.The system first converted segmented vessels into graph structures and then applied node classification GNNs for the automation of anatomical labeling.The model was deployed in a cloud full-labeled coronary tree retrieval framework with less than 1min of time per patient.This advancement has resulted in improving radiological workload and increasing the reproducibility of imaging interpretation. In yet another notable case, Alfurhood (2024) came up with a multimodal MS-GNN system under which vital data collected by the IoT devices gets assimilated among GNN layers for the classification of diseases in the area of cardiac, thyroid, and diabetic domains. The system used a cloud-based backend that processed signals acquired from wearable sensors and then uploads them to a GNN classifier which was optimized by monkey search algorithms. The establishment has proven to be able to infer from heart diseases with an accuracy of more than 95% and operate in real time, with a seamless transition from edge-device sensors to the cloud server. This is an extremely vital implementation for monitoring patients in rural areas where access to specialists is limited. A hybrid edge-cloud cardiac AI system was first presented by Wong et al. (2024), who could design an ultra-efficient GNN for cardiac disease detection from edge devices. The procedure passed ECG-derived features through EfficientNetB3 and Transformer encoders and fused them in a compact GNN. Containerized microservices were used for deployment, so large amounts of ECG data could be processed in real-time on the edge, while updates were synchronized with cloud models for continuous learning and The study emphasized latency optimized and energy-efficient inference time was reported to be less than 0.5 seconds with CPU load less than 40%, suitable for clinical devices that run on battery. Similarly, Duong et al. (2023) presented a graph-enhanced ECG classification system using morphological edge detection-based GNNs. Although mostly tested in the lab, the architecture is inherently suited for cloud deployment due to modular design. The graph fusion layer performs 1D signal transformation into spatial feature graphs, which can be GPU-accelerated in servers such as AWS EC2 and Google Cloud AI platform. Future scaling for hospital-wide deployment is feasible, where streaming of multiple ECGs occurs simultaneously. From a telemedicine perspective, Tariq et al. (2024) proposed multimodal GNN for the prediction of major adverse cardiovascular events (MACE) in migraine patients - a frequently overlooked group in cardiology. Their model transformed EHRs, imaging data, and clinical notes into graph structures and trained these with GCNs. The deployment was in a federated cloud environment that preserved patient privacy and enabled generalization across multiple centers. This may provide a platform for collaborative cardiology across institutions, free from hindrances posed by data silos.

Another example is that of Lin and Prasanna (2023), who developed HyScale-GNN, a hybrid GNN training system deployed on single-node heterogeneous cloud servers (CPU+GPU). Though not cardiology-specific, their infrastructure was designed and optimized to work with graph data at scale, providing a valuable blueprint for training cardiac models that contain millions of nodes or edges, such as patient-population graphs or large-scale coronary networks. Also, Wang et al. (2024) applied a LSTM-GNN hybrid architecture to time-series behavioral and physiological data to predict cardiovascular risk. It involved cloud-based Transformers analyzing longitudinal health data from wearable devices, linked to lifestyle features such as sleep, exercise, and diet. This model churned out personalized risk scores available on clinician dashboards or mobile applications a prototype towards the vision of next-generation digital cardiology. Of note, Gunawan et al. (2024) showed that even simple GCN architectures could be suited for real-time clinical systems with high effectiveness. This study involved patient similarity graphs derived from structured datasets and deployed the model in a low-latency cloud environment, where it reached 93.03% accuracy with very few compute resources. Lightweight solutions are very useful for emerging health setups seeking inexpensive and scalable cardiovascular diagnostic tools. Together, these case studies illustrate a paradigm shift in cardiovascular AI, in which GNNs are no longer and indeed not just academic constructs but cloud-deployable tools currently being applied—or at least tested—in real-life clinical workflows. They provide multi-modal, real-time, scalable, and often interpretable decision-making, from helping specialists in tertiary hospitals to primary care teams operating in decentralized settings. The implementations reviewed strongly support the claim that cloud-deployed GNNs are a feasible and useful next-generation diagnostic backbone for mapping and managing coronary artery disease.

6. Comparative Analysis

The development of graph-based learning generated and birthed parallel considerations of classical and deep learning models in cardiovascular applications. Classical CNNs struggle with artery classification due to their grid-based limitations, making it difficult to model structures like bifurcations or vascular trees (Beetz et al., 2022; Hampe et al., 2024). GNNs' ability to learn from non-Euclidean data provides a major advantage. This is especially valuable for mapping topological relationships in CAD progression. CNN-based models of Imran et al. (2024) and Chhikara et al. (2024) yield an accuracy of 89–92%, whereas GNNs outclassed them with accuracies greater than 95% in the guise of MS-GNN (Alfurhood, 2024) and hybrid LGAP model (Wang et al., 2024). GNNs further allow relational reasoning to be encoded via patient inter-similarity (Lu & Uddin, 2021), vessel-tree continuity (Hampe et al., 2024), and behavioral risk patterns (Wang et al., 2023). From a deployment perspective, cloud-integrated GNNs present great capabilities Cloud deployment addresses the limitations of local systems, such as limited computing power and maintenance challenges. It enables online retraining, global distribution, and secure cross-institutional integration (Lin & Prasanna, 2023; Wong et al., 2024). Hybrid edge-cloud deployments maintain low latency and high inference accuracy, these capabilities make them preferable to traditional ML pipelines in clinical settings as pointed by Raja et al. (2024) and Tariq et al. (2024). Therefore, although conventional models hold ground when it comes to solving small-scale tasks or functioning in constrained environments, they stand as fundamentally less rich and less scalable in modeling than GNNs, especially when GNNs have been deployed in the cloud for CAD applications.

Feature	GNNs	CNNs	RNNs	Traditional ML
Accuracy	High (up to 96%, AUC > 0.90)	Moderate to High	Moderate	Variable
Modeling Structure	Graph topology, patient similarity, vessel continuity	Grid/Spatial only	Sequential only	Flat feature space
Multimodal Integration	Yes (ECG + CCTA + Behavior)	Partial (images only)	Limited	Minimal
Cloud Deployment	Yes, edge and	Feasible but less	Rarely	Challenging

	cloud supported	flexible	implemented	
Federated Learning	Demonstrated in multiple studies	Rare	Rare	Not applicable
Interpretability	Emerging, needs work	Visual filters available	Black-box	Limited
Training Complexity	High (requires large memory & GPUs)	Moderate	High	Low
Data Efficiency	Graph augmentation possible	Data-hungry	Moderate	Variable
Clinical Readiness	Promising but maturing	Good for imaging tasks	Limited	Needs enhancement

Table 2: Comparative Evaluation of GNNs vs. Conventional Models in Cardiovascular Applications.

7. Challenges and Limitations

One notable limitation regards GNNs and their cloud-deployed architecture. A major limitation is the lack of large, annotated CAD graph datasets WHICH restricts generalizability and hinders model validation. Several of the surveyed models (e.g., Gunawan et al., 2024; Duong et al., 2023) considered public ECG or EHR datasets not explicitly structured as graphs, requiring custom methods of graph construction, which limits their generalizability. Second, cloud deployment faces some regulatory hurdles due to the sensitivity of medical data and as they are subjected to legislation such as HIPAA, GDPR, and local privacy acts. Although solutions may involve frameworks for federated learning, greater complexities arise for synchronization of model training and updates (Wong, et al., 2024; Alfurhood, 2024).

Third, real-time inference latency still hinders performance in high-stake clinical settings. While hybrid models mitigate load on central servers and thus delay from server-side processing, achieving consistent sub-second response times across all environments needs extensive rigour testing, which is often unavailable in academic studies (Lin & Prasanna, 2023). In the fourth place, Explainability in most GNN applications remains underdeveloped, limiting clinical trust and regulatory acceptance. Even though some solutions have been proposed through GNN. Explanation and inverse projection learning (Kwon et al., 2025), the inability to interpret explanations is an impediment toward physician trust and regulatory approval. Also, the vast majority of GNN architectures remain parser "black boxes" to clinicians, a feature that prevents their translation into bedside tools. Finally, inequity addresses the disparity in infrastructure and cost among the great majority of world health systems, with advanced cloud-deployed models are often inaccessible in low- and middle-income countries (LMICs). This highlights the need for scalable and inclusive deployment strategies.

8. Future Directions

This domain offers a rich landscape for innovation, where ongoing research can drive substantial progress in CAD diagnostics. The first challenge is to create open-access cardiovascular graph datasets, particularly labeled CCTA graphs or ECG similarity networks; The availability of such datasets will promote systematic research and benchmarking. It will also support the evaluation and standardization of CAD-GNN models. Consortiums similar to STACOM or PhysioNet could spearhead such initiatives.

Second, federated GNN frameworks allow training across institutions without sharing raw data, improving privacy while harnessing large-scale learning. The IoT-cloud integrations and behavior-based graph modeling mentioned by Alfurhood (2024) and Wang et al. (2024) suggest toward such potential.

Third, GNNs harnessed with wearables and mobile health are then areas set to vastly change disease tracking and population health monitoring. Real-time inference demonstrated by Wong et al. (2024) and Ashtaiwi et al. (2024) signals the potential of this future where ECG and behavioral data from patients will flow into graph-based risk engines.

Fourth is to develop interpretable and clinically validated GNN toolkits, working with regulatory bodies. Embedding counterfactual reasoning, feature importance maps, and model calibration techniques could enable the acceptance of GNNs into clinical workflows.

Finally, the advancement of infrastructure - involving serverless graph model deployment, energy-efficient GNN inference, and edge-native learning algorithms - will ensure equitable and sustainable integration of GNNs within the various health systems.

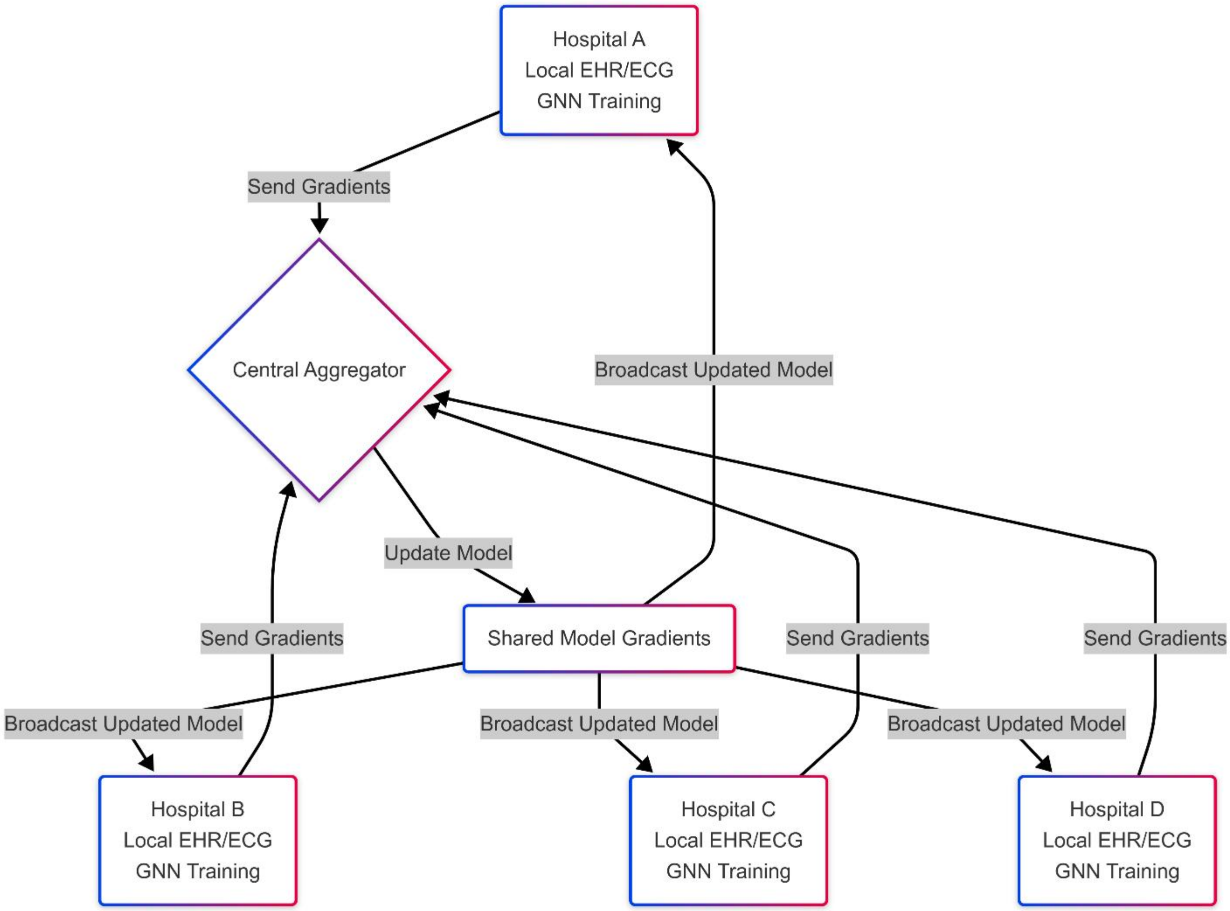


Figure 5: Federated GNN training architecture allowing privacy-preserving collaboration across clinical sites for CAD modeling.

9. Conclusion

Graph Neural Networks (GNNs) advanced by way of a major evolution in the modeling of complex, relational cardiovascular data, with its subtype modeling representing anatomy of the vessels, patient similarity, and behavioral dynamics. When these models are deployed on cloud environments, they are scaled, offered an instantaneous service, and integrated with larger telehealth and IoT ecosystems. In total, 32 articles related to CAD diagnosis-imaging, EHR, ECG, and behavior-driven graph modeling from the fine works-reviews were synthesized. Evidence confirms that cloud-based GNNs provide superior accuracy, scalability, and personalization than traditional models. GNNs in conjunction with AWS, Azure, and hybrid cloud-edge architectures have been shown to be technically doable for various clinical use cases. However, the field is limited by some critical factors from scarcity of datasets, explainability, and regulatory challenges to infrastructural ones. Faced with such issues, open datasets, federated learning, and interpretable AI should be prioritized to advance GNNs from research to real-world clinical implementation. In conclusion, it offers a promising albeit still maturing paradigm for coronary artery disease progression mapping by way of cloud-deployed GNNs. With sustained innovation and interdisciplinary collaboration, the scenario of cardiovascular care might soon be rewritten.

References

Alfurhood, B.S. (2024). Integrating IoT and AI in Healthcare: A Novel MS-GNN Framework for Disease Diagnosis. In: Gupta, S.K., Karras, D.A., Natarajan, R. (eds) Revolutionizing Healthcare: AI

Integration with IoT for Enhanced Patient Outcomes. Information Systems Engineering and Management, vol 7. Springer, Cham. https://doi.org/10.1007/978-3-031-65022-2_3

Andayeshgar, B., Abdali-Mohammadi, F., Sepahvand, M. *et al.* Arrhythmia detection by the graph convolution network and a proposed structure for communication between cardiac leads. *BMC Med Res Methodol* 24, 96 (2024). https://doi.org/10.1186/s12874-024-02223-4

Ashtaiwi, A. A., Khalifa, T., & Alirr, O. (2024). Enhancing heart disease diagnosis through ECG image vectorization-based classification. *Heliyon, 10*(18), e37574. https://doi.org/10.1016/j.heliyon.2024.e37574

Beetz, M., Corral Acero, J., Banerjee, A., Eitel, I., Zacur, E., Lange, T., Stiermaier, T., Evertz, R., Backhaus, S. J., Thiele, H., Bueno-Orovio, A., Lamata, P., Schuster, A., & Grau, V. (2022). Interpretable cardiac anatomy modeling using variational mesh autoencoders. *Frontiers in cardiovascular medicine*, *9*, 983868. https://doi.org/10.3389/fcvm.2022.983868

Chhikara, B. S., Kumar, R., Singh, J., & Kumar, S. (2024). Heart disease prediction using Machine learning and cardiovascular therapeutics development using molecular intelligence simulations: A perspective review. *Biomedical and Therapeutics Letters*, *11*(2), 920. https://doi.org/10.62110/sciencein.btl.2024.v11.920

Duong, L. T., Doan, T. T., Chu, C. Q., & Nguyen, P. T. (2023). Fusion of edge detection and graph neural networks to classifying electrocardiogram signals. *Expert Systems with Applications*, *225*, 120107.

Gunawan, K., Wiharja, K. R. S., & Hasmawati. (2024). Early detection of heart disease with graph neural network. In *Proceedings of the 2024 12th International Conference on Information and Communication Technology (ICoICT)* (pp. 405–410). IEEE. https://doi.org/10.1109/ICoICT61617.2024.10698473

Hampe, N., van Velzen, S. G. M., Wolterink, J. M., Collet, C., Henriques, J. P. S., Planken, N., & Išgum, I. (2024). Graph neural networks for automatic extraction and labeling of the coronary artery tree in CT angiography. *Journal of medical imaging (Bellingham, Wash.)*, *11*(3), 034001. https://doi.org/10.1117/1.JMI.11.3.034001

Huang, F., Lian, J., Ng, K. S., Shih, K., & Vardhanabhuti, V. (2022). Predicting CT-based coronary artery disease using vascular biomarkers derived from fundus photographs with a graph convolutional neural network. *Diagnostics*, *12*(6), 1390.

Imran, M., Ramay, S. A., & Abbas, T. (2024, September 1). *Predictive modeling for early detection and risk assessment of cardiovascular diseases using the ensemble stacked neural network model.* https://jcbi.org/index.php/Main/article/view/603

Kwon, H., Son, S., Morton, S. U., Wypij, D., Cleveland, J., Rollins, C. K., Huang, H., Goldmuntz, E., Panigrahy, A., Thomas, N. H., Chung, W. K., Anagnostou, E., Norris-Brilliant, A., Gelb, B. D., McQuillen, P., Porter, G. A., Jr, Tristani-Firouzi, M., Russell, M. W., Roberts, A. E., Newburger, J. W., … Im, K. (2025). Graph-based prototype inverse-projection for identifying cortical sulcal pattern abnormalities in congenital heart disease. *Medical image analysis*, *102*, 103538. https://doi.org/10.1016/j.media.2025.103538

Lin, H., Chen, K., Xue, Y., Zhong, S., Chen, L., & Ye, M. (2023). Coronary heart disease prediction method fusing domain-adaptive transfer learning with graph convolutional networks (GCN). Scientific Reports, 13(1), 14276.

Lin, Y.-C., & Prasanna, V. (2023). HyScale-GNN: A scalable hybrid GNN training system on single-node heterogeneous architecture. In *Proceedings of the 2023 IEEE International Parallel and Distributed Processing Symposium (IPDPS)* (pp. 557–567). IEEE. https://doi.org/10.1109/IPDPS54959.2023.00062

Lu, H., & Uddin, S. (2021). A weighted patient network-based framework for predicting chronic diseases using graph neural networks. *Scientific reports*, *11*(1), 22607. https://doi.org/10.1038/s41598-021-01964-2

Lundström, J., Hashemi, A. S., & Tiwari, P. (2023). Explainable Graph Neural Networks for Atherosclerotic Cardiovascular Disease. *Studies in health technology and informatics*, *302*, 603–604. https://doi.org/10.3233/SHTI230214

Raja, V. S., Rawat, R., Pandit, K., Dhumpati, R., Vekariya, V., & Vijitha, S. (2024). Automated heart disease diagnosis and monitoring system using similarity-navigated graph neural networks with Leopard

Seal Optimization in IoT healthcare applications. In *Proceedings of the 2024 4th International Conference on Sustainable Expert Systems (ICSES)* (pp. 225–230). IEEE. https://doi.org/10.1109/ICSES63445.2024.10763192

Rangisetti, L. S., Shaik, F., Bhagavatula, A., & Kommareddy, L. S. (2024). Heart disease detection using graph neural networks (GNNs). In *Proceedings of the 2024 5th International Conference on Electronics and Sustainable Communication Systems (ICESC)* (pp. 1436–1441). IEEE. https://doi.org/10.1109/ICESC60852.2024.10690138

Shiwlani, A., Hasan, S. U., & Kumar, S. (2024). Artificial Intelligence in Neuroeducation: A Systematic Review of AI Applications Aligned with Neuroscience Principles for Optimizing Learning Strategies. *Journal of Development and Social Sciences*, *5*(4), 578–593. https://doi.org/10.47205/jdss.2024(5-IV)50

Shiwlani, A., Kumar, S., & Qureshi, H. A. (2025). Leveraging Generative AI for Precision Medicine: Interpreting Immune Biomarker Data from EHRs in Autoimmune and Infectious Diseases. *Annals of Human and Social Sciences*, *6*(1), 244–260. https://doi.org/10.35484/ahss.2025(6-I)22

Shiwlani, A., Kumar, S., & Qureshi, H. A. (2025). Transforming Pediatric Leukemia Care: The Role of Artificial Intelligence in Diagnosing, Treating, and Optimizing Outcomes in Acute Lymphoblastic Leukemia. *Journal of Development and Social Sciences*, *6*(1), 348–363. https://doi.org/10.47205/jdss.2025(6-I)31

Sun, Y., & Zhang, J. (2024). Predictive analysis of cardiovascular disease under double-layer attention mechanism. In *Proceedings of the 2024 4th International Conference on Electronic Information Engineering and Computer Communication (EIECC)* (pp. 1577–1580). IEEE. https://doi.org/10.1109/EIECC64539.2024.10929267

Sun, Z., Yin, H., Chen, H., Chen, T., Cui, L., & Yang, F. (2020). Disease prediction via graph neural networks. *IEEE Journal of Biomedical and Health Informatics*, *25*(3), 818-826.

Tang, S., Liu, X., Chen, Y., Wang, Z., Wu, J., & Hu, X. (2023). Predicting 30-day all-cause hospital readmission using multimodal spatiotemporal graph neural networks. *IEEE Journal of Biomedical and Health Informatics, 27*(4), 2071–2082. https://doi.org/10.1109/JBHI.2023.3236888

Tariq, A., Dumitrascu, O., Luo, M., Belay, Y., Gennatas, E. D., & Briceño, E. M. (2024). Multi-modal predictive model for MACE risk estimation in patients with migraine. In *Proceedings of the 2024 IEEE 12th International Conference on Healthcare Informatics (ICHI)* (pp. 684–687). IEEE. https://doi.org/10.1109/ICHI61247.2024.00109

Wang, H., Zhang, Z., Xiong, H., Zou, D., Chen, Y., & Jin, H. (2024). GRAND: A graph neural network framework for improved diagnosis. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 43*(4), 1288–1301. https://doi.org/10.1109/TCAD.2023.3336212

Wang, J., Li, J., Wang, R., & Zhou, X. (2024). VAE-driven multimodal fusion for early cardiac disease detection. *IEEE Access, 12*, 90535–90551. https://doi.org/10.1109/ACCESS.2024.3420444

Wang, X., Luo, Z., He, R., & Shao, Y. (2023). Novel medical question and answer system: Graph convolutional neural network based with knowledge graph optimization. Expert Systems With Applications, 227, 120211. https://doi.org/10.1016/j.eswa.2023.120211

Wang, Y., Rao, C., Cheng, Q., & Yang, J. (2024). Cardiovascular disease prediction model based on patient behavior patterns in the context of deep learning: a time-series data analysis perspective. *Frontiers in psychiatry*, *15*, 1418969. https://doi.org/10.3389/fpsyt.2024.1418969

Wong, J., Nerbonne, J., & Zhang, Q. (2024). Ultra-efficient edge cardiac disease detection towards real-time precision health. *IEEE Access, 12*, 9940–9951. https://doi.org/10.1109/ACCESS.2023.3346893

Xu, H., & Wu, Y. (2024). G2ViT: Graph neural network-guided vision transformer enhanced network for retinal vessel and coronary angiograph segmentation. *Neural Networks*, *176*, 106356.

Yao, Z., Xie, W., Zhang, J., Yuan, H., Huang, M., Shi, Y., ... & Zhuang, J. (2023). Graph matching and deep neural networks based whole heart and great vessel segmentation in congenital heart disease. *Scientific Reports*, *13*(1), 7558.

Zhang, Y., Zhang, X., He, Y., Zang, S., Liu, H., Liu, T., ... & Zhang, L. (2025). Coronary p-Graph: Automatic classification and localization of coronary artery stenosis from Cardiac CTA using DSA-based annotations. *Computerized Medical Imaging and Graphics*, 102537.