Journal of Computer Science

Research Article

EW-GAT: Edge-Weighted Graph Attention Framework for Patient-Aware Heart Disease Classification Using Feature-to-Feature Similarity Graphs

Ramakrishna Reddy Badveli[1], Nijaguna Gollara Siddappa[2] and Sundeep Kumar Kanipakapatnam[3]

[1]*Department of Computer Science and Engineering, S.E.A College of Engineering and Technology, Bangalore, India*
[2]*Department of Information Science and Engineering, S.E.A College of Engineering and Technology, Bangalore, India*
[3]*Department of Computer Science & Allied Branches, Sri Venkateswara College of Engineering and Technology, Chittoor, India*

Article history
Received: 4-02-2026
Revised: 23-05-2026
Accepted: 17-06-2026

Corresponding Author:
Ramakrishna Reddy Badveli
Department of Computer Science and Engineering, S.E.A College of Engineering and Technology, Bangalore, India
Email: itramkey@gmail.com

Abstract: Heart disease remains an important cause of global death, requiring accurate and clinically interpretable prediction models. However, traditional Machine Learning (ML) methods often fail to capture meaningful dependencies among clinical features and lack robustness across diverse patient populations. To overcome these issues, this research introduces an Edge-Weighted Graph Attention Network (EW-GAT) for patient-aware heart disease classification. The approach first models feature-to-feature correlations to identify clinical dependencies, which are then used as a structural prior to improve patient-to-patient similarity during graph construction. An adaptive edge-weighting mechanism further enhances graph representation by integrating patient-specific variations. These weights are incorporated into the attention mechanism to emphasize clinically relevant relationships during message passing. Finally, a Multi-Layer Perceptron (MLP) layer predicts disease outcomes from refined patient embeddings. From the results, the proposed EW-GAT model attains higher performance, achieving 99.92% accuracy on the Cardiovascular dataset when compared to the existing Denoising Autoencoder (DAE) model. The improvements are supported by strict cross-validation and statistically significant analysis, confirming the model's robustness and generalization capabilities.

Keywords: Adaptive Edge-Weighted Graph, Edge-Weighted Graph Attention Network, Heart Disease Classification, Multi-Layer Perceptron, Patient-Specific Variations

Introduction

In recent years, heart disease has remained a major global health challenge, significantly contributing to mortality and long-term disability (Pandey et al., 2024). Early prediction plays a transformative role in preventing severe complications with timely medical interventions and clinical decision-making (Abbas et al., 2024). Machine Learning (ML) and Deep Learning (DL) models have been increasingly adopted to analyze cardiovascular risk based on clinical characteristics such as chest pain, blood pressure, cholesterol levels, and Electrocardiogram (ECG) results (El-Sofany, 2024; Elshewey et al., 2025). The combined interactions of these attributes produce subtle, interdependent, and complex physiological patterns (Manikandan et al., 2024). Accurate modeling of interactions is essential because traditional clinical assessments might overlook hidden correlations, especially when high-dimensional patient data are involved. Therefore, computational models are now indispensable tools for modern cardiovascular risk assessment (Enad and Mohammed, 2024). Despite advancements in ML and DL, conventional heart disease prediction models still face many challenges. Traditional ML models generally assume feature independence or rely on shallow feature representations. While DL models improve representation learning, they often lack interpretability. Several existing approaches use handcrafted feature engineering and struggle to capture deeper clinical dependencies across attributes (Raoof et al., 2025). Additionally, models often ignore interactions among variables such as cholesterol, heart rate, and exercise-related indicators, under the assumption of feature independence. Particularly, Graph Neural

Networks (GNNs) are recently gaining attention for modeling relational data and enabling structured representation learning. However, graph-based models typically rely on patient-level similarity to construct graphs, which limits their ability to explicitly capture complex medical interdependencies among clinical variables (Natarajan et al., 2024). Furthermore, inconsistent data distributions, limited labeled samples, and sensitivity to noise significantly reduce the generalizability of existing methods (Hangaragi et al, 2025). These limitations highlight the need for more adaptive models that better capture clinical relevance.

The main motivation of this work is to address the recognition of clinical attributes as interacting components rather than isolated variables (Aslam et al., 2024). Diseases such as heart failure, arrhythmia, and ischemic conditions often result from the combined influence of features such as chest pain class, ST-depression, and thallium scan results (Ashtaiwi et al., 2024). Capturing these correlations at the feature level instead of the instance level provides more meaningful medical reasoning (Suresh et al., 2024). Moreover, real-world datasets, namely Cleveland and Statlog, share common features but have different distributions, which creates a need for a scalable and clinically interpretable model (Fang et al., 2025; Devarajan et al., 2025). This motivates the development of an edge-weighted, attention-driven, graph-based framework to improve prediction accuracy and reliability (El-Sofany et al., 2024). Recently, state-of-the-art techniques include autoencoder architectures, optimization-enhanced ML models, deep convolutional networks, and GNNs (Cenitta et al., 2025). Autoencoder-DenseNet (ADN) combinations effectively extract multilevel representations but at a higher computational cost (Narasimhan and Victor, 2025). Optimization-driven models, namely Stochastic Gradient Optimization (SGO)-tuned Random Forest (RF)and eXtreme Gradient Boosting (XGBoost), are used to enhance parameter tuning; however, their capacity for relational modeling is limited. Fuzzy Support Vector Machine (FSVM) methods have significantly improved their handling of outliers but still struggle with the challenges of high-dimensional datasets (Gao et al., 2025). In contrast, deep-dense residual networks have demonstrated excellent performance but require extensive training. Although modern graph-based techniques typically support relational learning only at the instance level, thus ignoring relationships between features that are also important for medical reasoning. Therefore, these deficiencies have driven the development of a new feature-centric Graph Attention Framework designed to have greater clinical awareness. Additionally, a conventional GAT operates on pre-defined instance-level graphs or sample similarities learned dynamically. Explicit modeling enables pre-attention graph formulation for clinical co-dependencies before message passing, which standard GAT formulations do not support. Moreover, the proposed model first estimates feature-to-feature similarities to capture clinical dependencies and uses them as structural priors to refine patient-to-patient relationships during graph construction. An adaptive edge-weighting mechanism introduces patient-specific variations, allowing the model to emphasize clinically relevant connections. These adaptive weights are integrated into the attention mechanism, enabling the model to consider both feature semantics and structural importance during message passing. Unlike traditional weighted GAT approaches that rely solely on node-level feature similarity, the proposed method incorporates a feature-guided similarity mechanism that considers inter-feature correlations during patient graph construction. This approach captures latent clinical relationships that are otherwise overlooked, thereby enhancing representation learning. Consequently, the feature-interaction graph is improved by using patient-aware adaptive edge weighting. The Edge Weighted-Graph Attention Network (EW-GAT) redefines the graph-learning process instead of modifying the existing GAT architectures.

The key contributions of this research include:

- A feature-to-feature similarity measure is introduced to compute clinical dependencies based on correlations, which then serves as a structural prior for refining patient-to-patient similarity
- The patient-aware adaptive weighting function refines graph edges by incorporating patient-specific deviations. This mechanism dynamically strengthens clinically relevant connections, enabling more personalized and robust heart disease reasoning
- An edge-weighted attention module explicitly incorporates adaptive edge strength into the attention coefficient. This module fuses adaptive weights directly into the attention calculations, allowing the model to highlight medically important relations rather than just statistical similarities

Literature Review

Recent advancements in heart disease prediction utilize ML and DL techniques to enhance diagnostic accuracy. These methods focus on learning patterns from clinical data; however, their effectiveness depends on their ability to model complex feature interactions and generalize across datasets.

Alghamdi et al. (2024) presented a DenseNet architecture and autoencoder for heart disease prediction. These models were combined to integrate feature extraction and classification capabilities simultaneously. The limited labeled data was addressed through data augmentation, improving model generalization and robustness. While these models effectively captured

nonlinear patterns and improved the classification performance, they introduced deep hierarchical transformations that increased parameter redundancy and reduced interpretability of learned clinical relationships.

Naik et al. (2025) developed an SGO-enhanced RF and XGBoost classifier for heart disease prediction. This combined approach tackles current healthcare challenges, especially concerning heart disease prediction. The hyperparameter tuning by the SGO contributed to model flexibility by matching exploration and exploitation, which reduced overfitting to specific training data. However, this optimization focused solely on parameter tuning rather than adaptive representation learning, limiting robustness across different data distributions.

Qi and Zhang (2024) implemented an approach to improve the effectiveness of fuzzy SVM. Despite these improvements, the model operated on fixed vector representations and lacked mechanisms to capture nonlinear relational dependencies among patients.

Barfungpa et al. (2023) presented a Deep-DenseAquilaNet for heart disease prediction. Feature selection was performed using the Enhanced Sparrow Search Algorithm (E-SSA) to reduce data dimensionality and select relevant preprocessed data containing optimal features. These features were then input into Deep-DenseAquilaNet, where weights were updated using the Aquila Optimization Algorithm (AOA). This model effectively captured spatial and hierarchical patterns. However, its performance across various datasets was not assessed, limiting its generalizability and increasing the risk of overfitting to training data.

Aggarwal and Kumar (2023) developed a Neighborhood Component Analysis (NCA) using ML for heart disease classification. Classifiers such as RF, Artificial Neural Network (ANN), and K-Nearest Neighbor (KNN) were used for disease classification. Self-Diagnosis Algorithm (SDA) was applied for clustering data based on similar patterns in a high-dimensional input space. While NCA and SDA improved clustering and feature relevance using multiple classifiers, they did not focus on computational efficiency, which affected overall model performance.

Sasikala and Mohanarathinam (2024) introduced ML models for peripheral arterial disease detection and severity classification, along with hyperparameter optimization. These models and the Hyperparameter Optimization (HPO) process were tailored for disease identification; SMOTE-generated samples may not fully reflect population diversity, impacting the robustness of the models.

Research Gap

Existing heart-disease prediction models often struggle with high computational complexity, limited generalizability across diverse populations, and inadequate handling of clinically meaningful feature dependencies. Conventional models, such as GAT-based approaches, typically construct graphs based on sample-level similarities, thereby ignoring feature-level medical correlations and patient-aware variability, thereby reducing interpretability and diagnostic relevance. Moreover, the attention mechanism fails to incorporate edge importance, resulting in weaker modeling of structural relationships. This research gap highlights the need for a patient-adaptive, feature-aware graph learning model to enhance robustness and predictive accuracy.

Problem Statement

This section defines the problem statements for heart disease prediction as follows:

- Heart disease prediction models often face high computational complexity hampers scalability, delays processing, and limits practical applicability
- The absence of a robust validation framework has caused previous models to poorly generalize across various populations and datasets, increasing bias risk and decreasing prediction reliability

Objectives

This section outlines the objectives for heart disease prediction:

- To improve computational efficiency by focusing on advanced computing techniques, enabling analysis of large healthcare datasets
- To develop a robust ML or DL validation framework that ensures the model's generalizability across diverse populations and datasets, reducing bias and boosting reliability

Materials and Methods

This section employs a novel model, EW-GAT, for heart disease classification, incorporating patient-wise node computation and feature-to-feature similarity analysis. First, data were collected from the Cleveland and Statlog datasets, followed by preprocessing with a min-max scaler to ensure consistent feature representation. Patient-to-patient similarity is then computed via correlation analysis to preserve clinically meaningful relationships, with a thresholding mechanism. An adaptive edge-weighting mechanism further refines the graph by incorporating patient-aware relevance into each connection.

This weighted patient graph captures medical dependencies. Next, EW-GAT is applied to emphasize clinically significant patient-to-patient relationships guided by feature correlations, followed by an MLP classification

layer that combines refined patient node embeddings to classify heart disease accurately. The flow illustration of the proposed EW-GAT is presented in Fig. 1.

Data Collection

The dataset is taken from an open-access Kaggle repository for heart disease prediction. The Cleveland dataset (Janosi et al., 1989), Statlog dataset, and Cardiovascular dataset (Hagan et al., 2021) help to assess a person's risk for heart issues based on specific attributes. The Cleveland dataset contains 303 instances, while the Statlog dataset contains 270, as detailed in Table 1.

Data Preprocessing

The input datasets share a common attribute, such as thal, chol, or thalach, while other features vary widely in range. The min-max scaling preprocessing technique is a simple and effective normalization method for preparing data by transforming it into a consistent range, usually [0,1]. This scaling improves convergence, stabilizes gradient-based learning, and reduces the range of large features, as mathematically expressed for a feature x in Eq. 1:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (1)$$

Here, x denotes a real value, $\max(x)$ represents the maximum value of the feature, $and\ x'$ indicates a normalized value within the range ([0,1]). The minimum value of the feature is $\min(x)$. This process preserves the distribution shape while compressing values into a fixed interval, making it suitable for algorithms.

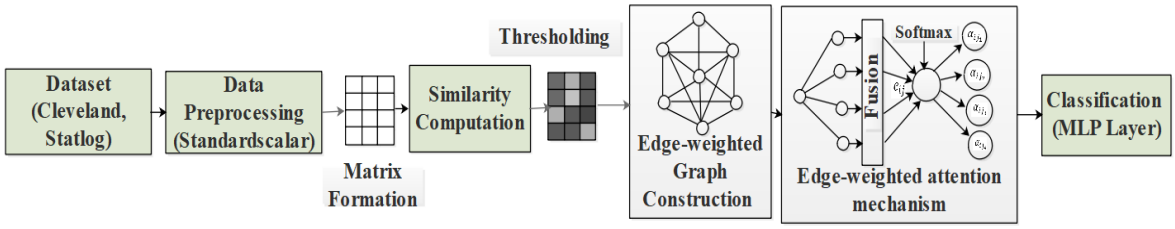

Fig. 1: Architecture of the proposed EW-GAT model for heart disease classification

Table 1: Common features with descriptions and values used in heart disease classification

Feature and Description	Value
Age (in years)	Continuous value
Sex	Male: 1, Female: 0
Chest pain type (Cp)	Four numerical values (1,2,3,4)
Trestbps (Resting blood pressure)	mmHg continuous value
Serum cholesterol	Mg/dL continuous value
Fasting blood sugar (Fbs)	0: < 120 mg/dL, 1: > 120 mg/dL
Restecg: Resting ECG results	0: Normal, 1: ST–T wave abnormality, 2: Probable or definite LVH
Thalach Maximum heart rate achieved	Continuous value
Exercise-induced angina (Exang)	0: No, 1: Yes
Oldpeak (ST depression induced by exercise relative to rest)	Continuous value
Slope (Slope of the peak exercise ST segment)	1: Upsloping, 2: Flat, 3: Downsloping
Ca (Number of major vessels colored by fluoroscopy)	0, 1, 2, 3
Thal (Heart beat)	3: Normal, 6: Fixed defect, 7: Reversible defect
Number of Predicted classes	0, 1

Similarity Computation

Furthermore, these pre-processed features □′ are represented as an □ ∈ ℝ□×□ matrix, where □ specifies the number of patient instances, and □ denotes the number of clinical features. In this framework, graph nodes strictly represent patient instances, while feature-to-feature correlations are used solely to refine patient similarity. The computed feature-to-feature correlation matrix is not used to define graph nodes. Still, it guides the estimation of edge weights between patient nodes by embedding clinically meaningful dependencies into the construction of the adjacency matrix. This approach maintains a patient-centric graph structure, where each node corresponds to an individual patient, and edges reflect adaptive similarity relationships. This design ensures conceptual consistency while preserving medical interpretability without altering the patient-level classification objective. Accordingly, the feature vector used for feature-to-feature similarity computation is defined as x, which captures feature values across all patients. Each feature □□ is represented as a vector □□ ∈ ℝ□ which contains its value across the patients. Consequently, to capture the interdependencies among clinical variables, a feature-to-feature correlation is first

computed to capture clinical dependencies, and it is presented in Eq. 2:

$$S_{fg} = \text{corr}(x_f, x_g) \quad (2)$$

Here, □□□ represents the correlation between features □□ and □□, and patient similarity is computed using Eq. 3:

$$sim(i,j) = \frac{x_i \cdot x_j}{\|x_i\| \, \|x_j\|} \quad (3)$$

Where □□□(□, □) denotes likeness across patients □□ and □□ ; these correlations are then used as a structural basis before refining patient-to-patient similarity. The resulting similarity matrix □ ∈ $\mathbb{R}^{□×□}$ is not used to define graph nodes instead but serves as a structural prior to modulate patient-to-patient similarity during graph construction, as defined in Eq. 4:

$$\tilde{S}_{ij} = \sum_{f=1}^{F} \sum_{g=1}^{F} x_{if} S_{fg} \, x_{jg} \quad (4)$$

Here, x_{if} denotes the value of feature f for patient i, x_{jg} signifies the value of feature g for patient j, and S_{fg} denotes the correlation between feature f and g. The similarity scores capture linear relations among clinical features. A threshold τ is then applied to retain meaningful relations while pruning weaker ones, mathematically defined in Eq. 5:

$$A_{ij} = \begin{cases} \tilde{S}_{ij}, & if \ \tilde{S}_{ij} \geq \tau \\ 0, & otherwise \end{cases} \quad (5)$$

The threshold similarity matrix defines feature-interaction weights, later used in patient similarity computation. Unlike conventional GCNs that model relationships between samples, this framework first captures meaningful dependencies among medical attributes, then constructs a patient-centric graph. Instead of defining graph nodes as features, feature correlation is treated as interpretable before refining patient-to-patient similarity. The resulting graph remains patient-based, integrating medically grounded feature interactions into the edge-weighting process. This strategy separates the estimation of the structural prior from the representation-learning process, thereby enhancing interpretability without altering the patient-level prediction objective.

Graph Construction

After estimating the threshold feature-interaction matrix, Edges $e_{ij} \in E$ indicate similarity between patients, computed using feature representations and refined with feature-level correlation priors. The graph nodes strictly represent patient instances; clinical features are not treated as graph nodes at any stage. The main goal is to model relationships among patients while incorporating clinically meaningful feature dependencies into edge weights. Once valid edges are identified, the method refines each edge weight using adaptive weighting. For each connected patient pair (i, j), the adaptive weight w_{ij} is computed as follows Eq. 6:

$$w_{ij} = \propto \cdot Sim(i,j) + (1 - \propto) \cdot \tilde{S}_{ij} \quad (6)$$

Where $\propto \in [0,1]$ balances raw similarity and ensures the weight lies between $0 < w_{ij} < 1$. The adaptive weight w_{ij} defines the connection strength between patient nodes i and j, conditioned on feature-level deviations. This thresholding approach differs from traditional linear-transformation-based adjacency learning by explicitly integrating clinically grounded feature relationships into the representation learning process. Although the adaptive weighting function employs a sigmoid-based design, it plays a structural role during graph construction, unlike conventional reweighting schemes that modify important features or attention scores. Moreover, unlike conventional post-attention reweighting, the proposed adaptive weighting operates during graph construction, refining edge strengths before message passing. Simultaneously, this approach acts as a patient-aware structural regularizer, amplifying or suppressing feature interactions based on deviation patterns. This patient-adaptive edge-weighted graph ensures those with medically relevant similarities are strongly connected, avoiding pairs that are mathematically similar but clinically irrelevant. The final weighted adjacency matrix is defined as $\tilde{A}_{ij} = A_{ij} \cdot w_{ij}$, which refines the graph edges using weights informed by feature deviations, encoding patient-aware clinical relevance. Additionally, $A = [a_{ij}] \in \mathbb{R}^{N \times N}$ denotes the adaptive edge-weight matrix derived from the similarity refinement. The adoption of the attention layer is explained in the subsequent section.

Edge-Weighted Graph Attention Layer

The proposed model enhances the traditional graph attention network by integrating adaptive edge weights into the attention mechanism. Message passing occurs between patient nodes, where each node aggregates information from patients with similar clinical profiles. This process is guided by edge weights derived from feature-informed similarity, enabling the model to dynamically emphasize clinically relevant relationships based on feature-level dependencies, making it particularly suitable for heart-disease classification. Each node's embedding is transformed into a learnable representation space where the similarity, attention, and aggregation mechanisms are explained clearly in Eq. 7:

$$h_i = W x_i, W \in \mathbb{R}^{d' \times d} \quad (7)$$

DOI: 10.3844/jcssp.2026.2220.2235

Where h_i is the feature after linear transformation, and W is the learnable weight matrix. This notation differentiates raw features from learnable representations in the attention process.

Edge-Weighted Attention Coefficient

The main goal of the EW-GAT network is to incorporate both features and edge strengths into the attention computation. This ensures message passing reflects not only the feature similarity but also the structural importance encoded in edge weights. After the linear transformation, the core operation is to evaluate how important node v_j is to node v_i by combining their representations and edge weight. A learnable attention vector a $\in \mathbb{R}^{2d'}$ is used to assess compatibility between connected nodes, as shown in Eq. 8:

$$e_{ij} = LeakyReLU\left(a^{\top}\left[h_i \| h_j\right]\right) \tag{8}$$

Where || denotes concatenation, e_{ij} is the important edge between connected nodes, and h_i and hj specify embeddings of patient node i and j. Unlike the traditional GATs, which normalize raw attention scores e_{ij} directly, this model fuses the structural weight w_{ij} into the attention score, combining semantic correlation and graph structure. This fusion is multiplicative, shown in Eq. 9:

$$\grave{e}_{ij} = w_{ij} \cdot e_{ij} \tag{9}$$

Here, $\grave{e}_{ij}$ denotes the edge-aware attention, and this fusion mainly concentrates on modulating the similarity weights, which encode clinically meaningful relationships. However, similar to existing weighted GAT variants, edge weights are typically static and predefined from graph topology; this leads to serving auxiliary scaling factors while passing the message. The normalized attention coefficient is calculated via SoftMax over the neighbors $\mathcal{N}(i)$ of node v_i, as in Eq. 10:

$$\alpha_{ij} = \frac{exp(\grave{e}_{ij})}{\sum_{k \subset \mathcal{N}(i)} exp(\grave{e}_{ik})} \tag{10}$$

This attention score α_{ij} refers to the normalized attention coefficient, which indicates how strongly node v_i connects to its neighbor node v_j, accounting for both semantic similarity and graph structure. The node update is performed using weighted aggregation, shown in Eq. 11:

$$h_i' = \sigma\left(\sum_{j \subset \mathcal{N}(i)} \alpha_{ij} h_j\right) \tag{11}$$

Where $\sigma(\cdot)$ specifies a non-linear activation function and h_i' signifies the updated embedding of patient node i. This weighted aggregation embeds edge weights as core factors into the attention process, allowing the model to focus on medically relevant relationships rather than solely on feature projections. A flow diagram of this edge-aware attention is presented in Fig. 2:

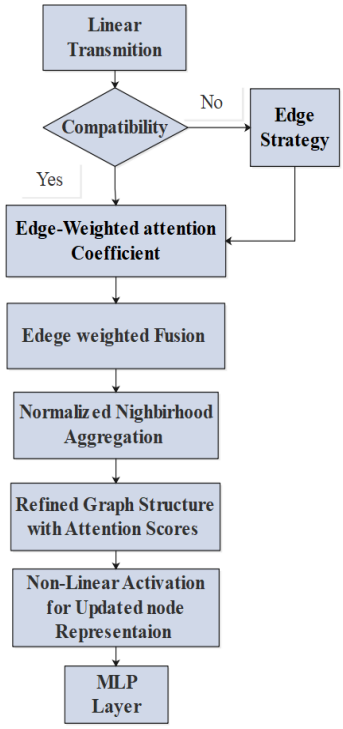

Fig. 2: Flow diagram of proposed EW-GAT model for heart disease classification

Multi-Layer Perceptron as Classification Layer

The EW-GAT model generates rich patient node embeddings, which are further transformed into final class predictions through an MLP classification layer. Let H $\in \mathbb{R}^{N \times d}$ denote the matrix of patient node embeddings obtained from the attention mechanism, where N denotes the number of nodes, each representing a patient instance, and d denotes the embedding dimension. For each node i, an embedding vector is defined as $h_i \in \mathbb{R}^d$. The MLP processes each h_i through fully connected layers; the first transformation of the patient node embeddings is given in Eq. 12:

$$z_i^{(1)} = \sigma\left(W^{(1)} h_i + b^{(1)}\right) \tag{12}$$

Here, the weight matrix is presented as $W^{(1)} \in \mathbb{R}^{d_1 \times d}$, the bias vector is denoted as $b^{(1)} \in \mathbb{R}^{d_1}$, *and* d_1 denotes the layer dimension. This initial transformation projects each patient node embedding into a new feature space, making decision boundaries easier to distinguish. The first hidden layer $\mathbf{z}_i^{(1)}$ is then passed to subsequent layers, as shown in Eq. 13:

$$z_i^{(l)} = \sigma\left(W^{(l)} z_i^{(l-1)} + b^{(l)}\right) \tag{13}$$

Where l denotes the index of the layer number, $\mathrm{W}^{(l)} \in \mathbb{R}^{d_l \times d_{l-1}}$ and $\mathrm{b}^{(l)} \in \mathbb{R}^{d_l}$ specify the learnable weights and bias vectors. These multiple transformations enable the MLP layer to learn hierarchical abstractions from the graph-refined representations. The lower layers in the MLP capture simple combinations of GAT features, while the deeper layers capture higher-level medical patterns. To classify these patient embeddings into C classes, the final layer is described in Eq. 14:

$$o_i = softmax\left(\mathrm{W}^{(L)}\mathrm{z}_i^{(L-1)} + \mathrm{b}^{(L)}\right) \quad (14)$$

Here, $\mathrm{W}^{(L)} \in \mathbb{R}^{C \times d_{L-1}}$ and $\mathrm{b}^{(L)} \in \mathbb{R}^{C}$ represent the learnable weights and bias vectors for L layers. The predicted class for node i is given in Eq. 15:

$$\grave{y}_i = \text{softmax}\,(W_o h_i') \quad (15)$$

Here, $\grave{y}_i$ denotes the predicted label using softmax activation, and cross-entropy loss is trained by the MLP as shown in Eq. 16:

$$\mathcal{L} = -\sum_{i=1}^{N}\sum_{c=1}^{C} y_{i,c}\log\left(o_{i,c}\right) \quad (16)$$

Here, $y_{i,c}$ presents the ground-truth one-hot indicator, and the cross-entropy loss represents $\mathcal{L}$. Unlike traditional MLPs, the proposed model applies an MLP layer after an edge-weighted GAT, feeding in embeddings whose dimensions already reflect the medically meaningful relational patterns learned from the graph. Since the edge-weighted attention effectively incorporates patient-specific and feature-specific importance, and the MLP layer processes representations that encode inter-patient variability, severity differences, and feature correlations, this layer serves as a powerful decision module. It classifies heart disease by assigning 0 to normal and 1 to heart disease. The pseudocode for the proposed EW-GAT is provided below.

Algorithm 1: Pseudocode of proposed EW-GAT model for heart disease classification

Input: Clinical dataset $D = \{(x_i, y_i)\}_{i=1}^{N}$, where each $x_i \in \mathbb{R}^F$ represents the clinical feature vector of patient i, and y_i denotes the corresponding class label.
Output: Predicted class labels for patient nodes.

Normalize clinical features and handle missing values using training-fold statistics.
Compute feature vectors $x_f \in \mathbb{R}^N$ for all $i \in \{1, \ldots, d\}$

For each pair of features (f, g):
    Compute feature correlation S_{fg} using a correlation-based metric.
End for
For each pair of patients (i, j)
    Compute patient similarity S_{ij} using a correlation-based metric.
    If $S_{ij} \geq \tau$ then set $A_{ij} = S_{ij}$
    Else $A_{ij} = 0$
End for
Construct the initial patient adjacency matrix A.
For each patient pair (i, j) where $A_{ij} > 0$
    Compute adaptive edge weight w_{ij} using Eq. 4.
End for
Form the final adaptive edge-weighted adjacency matrix Ã.
Initialize a learnable transformation matrix W.
For each epoch from 1 to $t_{\max}$:
    For each patient node i:
        Compute transformed feature $h_i = W \cdot x_i$
        For each neighboring patient j:
            Compute the raw attention score using Eq. 7.
            Fuse structural importance using Eq. 8.
        End for
        Normalize over the neighborhood using Eq. 9.
        Update patient node embeddings with Eq. 11.
        End for
    End for
End for
Feed refined patient embeddings H' into the MLP classifier:
return trained EW-GAT model $H(F)$

Results and Discussion

All experiments are conducted on a Windows 10 system with an Intel Core i7 processor and 16GB RAM, providing sufficient computational capacity for graph-based learning and repeated evaluations. The workflow is implemented using Anaconda Navigator 3.5.2.0 and Python 3.7.6 to ensure a stable, reproducible environment. Visualization libraries, including NumPy, Pandas, Seaborn, and Matplotlib, are installed to support data preprocessing, statistical analysis, and graph visualization. This setup ensures reliable execution of the proposed EW-GAT model throughout all experimental phases. Evaluation metrics such as accuracy, precision, recall, and F1-score are considered, with mathematical formulas provided in Eqs. 17-21:

$$Accuracy = \frac{TP}{TP+FN+TN+FP} \quad (17)$$

$$Precision = \frac{TP}{TP+FP} \quad (18)$$

$$Recall = \frac{TP}{TP+FN} \quad (19)$$

$$F1-score = \frac{2 \times Precision \times Recall}{Precision \times Recall} \quad (20)$$

$$AUC = \int_a^b f(x)dx \quad (21)$$

DOI: 10.3844/jcssp.2026.2220.2235

Here, TP denotes true positives, and TN denotes true negatives; FP and FN denote false positives and false negatives, respectively. The symbols a and b specify the limits of the functions, and $f(\ )$ (x) specifies the function. The hyperparameters and their values are listed in Table 2.

The weight decay is explicitly included in the hyperparameter settings and applied consistently during training to prevent overfitting. All optimization parameters are clearly defined to ensure reproducibility. This includes the selection of τ within the training fold, patient-graph construction, model training, and evaluation of the held-out fold. To isolate the contribution of adaptive edge-weighting effects related to graph density, an ablation study is conducted where all variants have identical edge counts. The comparison includes the standard GAT without adaptive weighting, GAT with fixed edge weights, and the proposed model with adaptive edge weighting. To reduce the risk of overfitting and data leakage, especially for the Cleveland (N = 303) and Statlog (N = 270) datasets, a strict 5-fold cross-validation protocol is used for all evaluations. To ensure reproducibility and statistical robustness, all the experiments were conducted using 5 repeated runs with 5 predefined random seeds. In each run, stratified 5-fold cross-validation was performed under identical experimental settings for both the proposed EW-GAT model and all baseline models. All the operations, including preprocessing, graph construction, threshold selection, and adaptive edge-weight estimation, were performed strictly within the training fold. Then, it is applied to the corresponding test fold to avoid information leakage. Unless otherwise specified, the reported comparative classification results are presented as mean±standard deviation obtained from 5 repeated runs using stratified 5-fold cross-validation.

This controlled setup confirms that performance results stem from adaptive weighting rather than increased graph connectivity. Additionally, the hyperparameters for baseline models are detailed in Table 3. All hyperparameters are selected based on validation within the training folds, and experimental settings matched those of the proposed model across all datasets to ensure fairness and reproducibility.

Performance Evaluation

The evaluation of the proposed EW-GAT model is assessed using various performance metrics and baseline classifier models. The considered baseline models include the Knowledge Graph Convolutional Network (KGCN), GCN, GNN, and GAT models. All models, including the proposed EW-GAT and baseline models, are re-implemented within the same experimental setup, such as hyperparameter settings, preprocessing techniques, and datasets, to ensure a fair comparison, with results reported as mean ± standard deviation. The analysis of these results is presented in Table 4. From Table 2, baseline models, namely GCN, GAT, and KGCN, struggle to learn discriminative features for heart disease classification. The results reported reflect mean performance across 5 repeated runs, along with the standard deviation, indicating model stability. The low standard deviation observed for the proposed EW-GAT demonstrates consistent performance across different data splits and random initializations. Additionally, Tables 2 and 3 are carefully reformatted to ensure consistent alignment of datasets and evaluation metrics across the Cleveland, Statlog, and cardiovascular datasets. The proposed EW-GAT improved discriminative feature learning by combining feature-level similarity and adaptive edge-weighting refinement through an edge-aware attention mechanism.

Table 2: Hyperparameters and their values for the proposed EW-GAT model

Parameter	Value
Layers	2
Hidden dimension	64
Attention heads	4
random seeds	{42,2023,3407,7777,9999}
Train: test	80:20
Dropout	0.5
Optimizer	Adam
Learning rate	0.001
Correlation threshold τ	Selected from {0.3, 0.4, 0.5, 0.6} using training-fold validation
τ selection criterion	Maximum validation F1-score
Weight decay	5×10^{-4}
Batch size	Full batch
Epochs	300
Early stopping	20 patience
Tau	Top-k or fixed threshold 0.6
Alpha	0.7

Table 3: Experimental configurations for baseline models

Hyperparameter	Value
LR – regularization	L2
Max iterations	1000
SVM- Kernel	RBF
Max depth	None
Min sample split	2
XGBoost – No. of trees	100
Max depth	6
Subsample	0.8
Colsample by tree	0.8
LightGBM- No. of trees	31
Max depth	-1
Feature fraction	0.8
MLP baseline – Hidden layers	2
Hidden units	64,32
Activation	ReLU
Batch size	32
Epochs	100

Table 4: Performance of the proposed EW-GAT model for heart disease classification across different baseline classifiers

Methods	Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
KGCN	Cleveland	98.63±0.04	98.42±0.05	98.41±0.05	98.41±0.05
	Statlog	98.41± 0.05	98.66±0.04	98.65±0.04	98.65±0.04
	Cardiovascular	97.86±0.06	98.10±0.05	98.09±0.05	98.09±0.05
GCN	Cleveland	98.36±0.05	98.52±0.04	98.51±0.04	98.51±0.04
	Statlog	97.55±0.06	97.73±0.05	97.72±0.05	97.72±0.05
	Cardiovascular	98.23±0.05	98.53±0.04	98.54±0.04	98.53±0.04
GNN	Cleveland	98.41±0.05	98.28±0.05	98.27±0.05	98.27±0.05
	Statlog	97.85±0.06	98.16±0.05	98.15±0.05	98.15±0.05
	Cardiovascular	97.63±0.06	97.83±0.05	97.82±0.05	97.82±0.05
GAT	Cleveland	98.31±0.05	98.25±0.05	98.24±0.05	98.24±0.05
	Statlog	98.33±0.05	98.55±0.04	98.54±0.04	98.54±0.04
	Cardiovascular	98.14±0.05	98.24±0.05	98.23±0.05	98.23±0.05
EWAGAT	Cleveland	99.81±0.02	99.84±0.02	99.76±0.03	99.79±0.02
	Statlog	99.83±0.02	99.74±0.03	99.91±0.01	99.82±0.02
	Cardiovascular	99.92±0.01	99.92±0.01	99.92±0.01	99.92±0.01

This mechanism enables the model to leverage clinically aligned information, resulting in superior performance across all three datasets. Furthermore, heart disease classification is effectively achieved through a novel attention mechanism incorporated into the EW-GAT model. Evaluating this attention mechanism against different baseline attention models confirms patient-aware relevance. All attention mechanisms are tested under identical conditions, including hyperparameter settings and preprocessing, across the three datasets, with results shown in Table 5.

Table 5 compares the proposed edge-aware attention mechanism with various baseline attention models across three datasets. As indicated in Table 5, the proposed edge-aware attention achieved higher performance compared to baseline attention mechanisms, with results reported as mean±standard deviation. Unlike existing attention mechanisms that depend solely on feature similarity, the proposed edge-aware attention effectively incorporates structural edge importance with feature semantics. This mechanism adaptively highlights clinically meaningful dependencies, thereby increasing accuracy and robustness.

Ablation Study

The individual components in the proposed EW-GAT model contribute to effective heart disease classification. The findings for each component are validated through an ablation study across three datasets, and the performance metrics are shown in Table 6.

Table 6 shows the ablation study across all datasets. The baseline GAT model without adaptive edge weighting performs well but is consistently outperformed by the proposed model, underscoring the importance of adaptive edge information. Similarly, the GCN model with fixed edge-aware attention shows moderate improvement over the baseline but cannot capture patient-specific variations because its edge weights are static. The GCN+MLP setup achieves lower performance, indicating that increasing model depth alone does not guarantee better results, and emphasizing the importance of structured relational learning. Overall, the proposed model outperforms the baseline across all datasets, achieving a 1.00% improvement. This consistent gain demonstrates that combining feature-informed graph construction with adaptive edge weighting greatly enhances discriminative ability.

Table 5: Evaluation of the attention mechanism's performance against baseline attention models for heart disease classification

Methods	Dataset	Accuracy %	Precision%	Recall%	F1-score%	AUC-ROC%
Additive Attention GAT	Cleveland	97.51±0.06	97.34±0.05	97.33±0.05	97.33±0.05	97.60±0.05
	Statlog	98.63±0.05	98.73±0.04	98.72±0.04	98.72±0.04	98.70±0.04
	Cardiovascular	98.71±0.04	98.83±0.04	98.82±0.04	98.82±0.04	98.80±0.03
Residual GAT	Cleveland	98.67±0.05	98.53±0.04	98.52±0.04	98.52±0.04	98.75±0.04
	Statlog	98.21±0.05	98.43±0.05	98.42±0.05	98.42±0.05	98.30±0.04
	Cardiovascular	98.85±0.04	98.41±0.05	98.40±0.05	98.40±0.05	98.90±0.03
Pyramid Attention Network	Clevland	97.52±0.06	97.63±0.05	97.32±0.06	97.47±0.05	97.60±0.05
	Statlog	96.98±0.07	97.06±0.06	97.05±0.06	97.05±0.06	97.10±0.05
	Cardiovascular	97.81±0.06	97.62±0.05	97.61±0.05	97.61±0.05	97.90±0.05
Lightweight Attention GAT	Cleveland	98.13±0.05	98.37±0.04	98.36±0.04	98.36±0.04	98.20±0.04
	Statlog	98.67±0.05	98.82±0.04	98.81±0.04	98.81±0.04	98.75±0.04
	Cardiovascular	98.54±0.05	98.72±0.04	98.71±0.04	98.71±0.04	98.60±0.04
EWAGAT	Cleveland	99.81±0.02	99.84±0.02	99.76±0.03	99.79±0.02	99.88±0.01
	Statlog	99.83±0.02	99.74±0.03	99.91±0.01	99.82±0.02	99.89±0.01
	Cardiovascular	99.92±0.01	99.92±0.01	99.92±0.01	99.92±0.01	99.95±0.01

Table 6: Evaluating the contribution of each component in the proposed EW-GAT via an ablation study

Methods	Dataset	Accuracy %	Precision%	Recall%	F1-score%
GAT (without adaptive weighting)	Cleveland	98.31	98.53	98.52	98.52
	Statlog	97.63	97.73	97.72	97.72
	Cardiovascular	98.63	98.85	98.84	98.84
GCN+Edge-aware attention (fixed edge weights)	Cleveland	98.55	98.62	98.61	98.61
	Statlog	98.67	98.72	98.71	98.71
	Cardiovascular	98.79	98.41	98.40	98.40
GCN+MLP layer	Cleveland	96.31	96.42	96.40	96.40
	Statlog	97.62	97.82	97.81	97.81
	Cardiovascular	97.58	97.63	97.62	97.62
Full proposed EW-GAT (including adaptive edge weighting and preprocessing)	Cleveland	99.81	99.84	99.76	99.79
	Statlog	99.83	99.74	99.91	99.82
	Cardiovascular	99.92	99.92	99.92	99.92

Table 7: Analysis of the statistical and computational efficiency of the proposed EW-GAT model

Dataset	Method classifier name	Training Time (s)	P-Value	Inference time (s)	Memory usage in MB	FLOPS (M)	Parameters
HeartCleavaland	KGCN	125.25	0.035	0.452147	3547.14	12.25	2,471
	GCN	107.25	0.041	0.852369	2542.48	11.74	1,741
	GNN	136.74	0.044	1.014785	2987.14	12.45	2,136
	GAT	102.47	0.034	1.102588	2369.14	10.58	1,966
	EWAGAT	92.47	0.022	0.206010	1911.01	9.4	1,342
Statlog	KGCN	102.25	0.023	0.214864	1254.25	11.25	2,741
	GCN	98.41	0.033	0.098742	987.41	9.7	2,147
	GNN	110.57	0.034	0.325663	1258.25	14.25	1,741
	GAT	107.52	0.024	0.098742	1452.25	12.47	3,451
	EWAGAT	91.22	0.014	0.022867	833.04	8.7	1,342
Cardiovascular	KGCN	102.71	0.042	0.485165	2510.25	11.6	1,748
	GCN	123.74	0.032	0.515064	5236.14	14.7	2,741
	GNN	117.58	0.038	0.345715	2365.14	10.8	5,741
	GAT	136.85	0.045	0.485763	4125.36	13.33	3,852
	EWAGAT	93.96	0.011	0.285065	1565.32	9.6	1,342

Statistical and Computational Analysis

To validate the reliability of the proposed EW-GAT model, a statistical analysis is carried out by quantitatively assessing whether performance improvements are statistically significant. To determine the significance of these improvements, a two-tailed t-test is conducted. Each p-value compares the proposed model with the corresponding baseline model. Additionally, the computational analysis evaluates the proposed model's efficiency relative to baseline models using evaluation metrics. FLOPs (in Millions, M) and the number of parameters (in M) are reported. The FLOPs are calculated for a single forward pass with the same input size and network configuration across all methods. To further confirm the reliability of the EW-GAT framework, a two-tailed t-test is performed using fold-wise performance scores. The statistical results and computational analyses are presented in Table 5.

From Table 7, the analysis is performed under a unified experimental environment setup, using identical hardware, memory limits, and execution settings for all models to ensure a fair comparison. This evaluation is based on fold-wise performance scores obtained from stratified k = 5 cross-validation. The p-values ($p < 0.05$) indicate that the performance enhancements of EW-GAT are statistically important and unlikely to be due to random variation. The p-values for the Cleveland (0.022), Statlog (0.014), and Cardiovascular (0.011) datasets are below the 0.05 significance threshold. This confirms that the performance gains achieved by the proposed model over baseline methods are statistically substantial and unlikely to be due to random variation, though $the\ p$-value is higher than that of existing models.

K-Fold Cross-Validation

The proposed EW-GAT model is evaluated across three datasets: Cleveland, Statlog, and Cardiovascular datasets. To validate the model's generalization ability and robustness, k-fold cross-validation is additionally performed during the experimental analysis. However, the main evaluation follows a strict stratified 5-fold cross-validation protocol to ensure accurate and unbiased performance estimation. The additional k-fold validation is conducted further to assess the model's consistency across different data splits and fold configurations. In this process. Results from all the folds are averaged to demonstrate model stability and generalizability across diverse data distributions. Additionally, all preprocessing

steps, including imputation, normalization, and feature correlation computation, are done solely on the training fold and then applied to the validation fold to prevent data leakage.

Table 8: 5-fold cross-validation results for the proposed EW-GAT model in heart disease classification

Methods	Fold	Dataset	Accuracy	Precision	Recall	F1-score
KGCN	2	Cleveland	97.21	97.63	97.62	97.62
	3		97.88	97.42	97.41	97.41
	5		98.63	98.42	98.41	98.41
	7		97.52	97.68	97.67	97.67
	2	Statlog	98.12	98.20	98.19	98.19
	3		97.26	97.37	97.36	97.36
	5		98.41	98.66	98.65	98.65
	7		96.86	96.72	98.71	97.70
	2	Cardiovascular	97.53	97.92	97.91	97.91
	3		97.27	98.06	98.05	98.05
	5		97.86	98.10	98.09	98.09
	7		97.33	97.88	97.87	97.87
GCN	2	Cleveland	97.82	98.22	98.21	98.21
	3		97.36	97.86	97.85	97.85
	5		98.36	98.52	98.51	98.51
	7		97.51	98.10	98.09	98.09
	2	Statlog	97.08	96.55	96.54	96.54
	3		96.89	97.12	97.12	97.12
	5		97.55	97.73	97.72	97.72
	7		96.76	96.12	96.11	96.11
	2	Cardiovascular	97.22	97.48	97.47	97.47
	3		97.56	97.69	97.68	97.68
	5		98.23	98.53	98.54	98.53
	7		97.81	97.52	97.51	97.51
GNN	2	Cleveland	97.51	98.10	98.09	98.09
	3		97.08	96.55	96.54	96.54
	5		98.41	98.28	98.27	98.27
	7		97.56	97.69	97.68	97.68
	2	Statlog	97.12	97.20	97.19	97.19
	3		97.26	97.37	97.36	97.36
	5		97.85	98.16	98.15	98.15
	7		97.33	97.88	97.87	97.87
	2	Cardiovascular	97.82	98.22	98.21	98.21
	3		97.36	97.86	97.85	97.85
	5		97.63	97.83	97.82	97.82
	7		97.08	96.55	96.54	96.54
GAT	2	Cleveland	98.12	98.20	98.19	98.19
	3		97.26	97.37	97.36	97.36
	5		98.31	98.25	98.24	98.24
	7		97.52	97.68	97.67	97.67
	2	Statlog	98.12	98.20	98.19	98.19
	3		97.26	97.37	97.36	97.36
	5		98.33	98.55	98.54	98.54
	7		97.52	97.72	97.71	97.71
	2	Cardiovascular	97.51	98.10	98.09	98.09
	3		97.08	96.55	96.54	96.54
	5		98.14	98.24	98.23	98.23
	7		97.21	97.55	97.54	97.54
EWAGAT	2	Cleveland	98.24	98.56	98.55	98.55
	3		97.82	98.76	98.75	98.75
	5		99.81	99.84	99.76	99.79
	7		98.32	98.14	98.13	98.13
	2	Statlog	98.41	98.37	98.35	98.35
	3		98.31	98.82	98.84	98.82
	5		99.83	99.74	99.91	99.82
	7		98.42	98.36	98.35	98.35
	2	Cardiovascular	98.23	98.46	98.45	98.45
	3		98.56	98.67	98.66	98.66
	5		99.92	99.92	99.92	99.92
	7		98.43	98.61	98.60	98.60

The same baseline models are used across the three datasets, each partitioned into folds to estimate how well the proposed model simplifies relative to the baseline models. The results of this K-fold cross-validation are presented in Table 8.

From Table 8, both the baseline models and the proposed model achieved higher results at fold 5 for each dataset. Variations in recall across different folds are mainly due to differences in class distribution and sample composition, especially for smaller datasets with mild class imbalance. The higher performance across folds is due to the proposed model better capturing feature-to-feature relationships, and patient-aware edge weights remaining stable across different train-test splits. The existing models lack structural adaptiveness, leading to larger variance as training data distribution changes. The proposed model addresses this issue by learning meaningful clinical relationships, enabling it to maintain superior accuracy across all folds and datasets.

Class-Wise Results With Proposed EW-GAT

The proposed EW-GAT model is evaluated across three datasets, with class-wise results presented for each. Initially, the model is tested on the Cleveland dataset, which has two binary classes, 0 and 1. Each class-wise result and the average are explained in Table 9.

From Table 9, the class-wise results on the Cleveland dataset show consistent balanced performance across the normal and disease classes, demonstrating strong discrimination. The accuracy for each class is evaluated using balanced splits and average confusion-matrix-derived metrics, including TP and TN rates, across the entire dataset.

The edge-aware attention module highlights medically relevant feature interactions, ensuring high precision and recall across classes. Subsequently, class-wise outcomes for the Statlog dataset are represented in Table 10.

From Table 10, the class-wise results on the Statlog dataset show consistent performance for both classes, mainly by focusing on patient-aware adaptive edge weights that capture subtle differences in clinical patterns. A similar analysis is performed on the cardiovascular dataset, with results shown in Table 11.

From Table 11, the class-wise results on the cardiovascular dataset show near-perfect, balanced performance, demonstrating the robustness of EW-GAT across larger, more diverse patient populations.

The proposed model consistently achieved high scores across all classes by capturing deeper medical dependencies through an edge-weighted attention mechanism.

Visualization Results of Heart Disease Classification

The heart disease classification is performed using the proposed EW-GAT model. To validate this model, three datasets are considered: Cleveland, Statlog, and Cardiovascular. Additionally, the confusion matrices shown in Fig. 3 present the aggregated predictions across all folds of cross-validation, rather than a single train-test split. This collective evaluation provides sensitivity and specificity across the entire dataset and shows that misclassification rates remain low across different validation splits.

The confusion matrices illustrate the proposed EW-GAT model's ability to classify the reported performance metrics. The diagonal values represent correctly classified samples, while the non-diagonal (light-shaded) samples indicate misclassifications. As shown, the number of misclassified samples is very low, indicating the model's strong classification capability.

Table 9: Class-wise outcomes of the proposed EW-GAT model

Class Name	Accuracy%	Precision %	Recall %	F1-score%
0		99.83	99.75	99.78
1		99.85	99.77	99.90
Average	99.81	99.84	99.76	99.79

Table 10: Class-wise outcomes of the proposed EW-GAT for heart disease classification on the Statlog Dataset

Class Name	Accuracy%	Precision %	Recall %	F1 score%
0		99.73	99.90	99.81
1		99.75	99.92	99.83
Average	99.83	99.74	99.91	99.82

Table 11: Class-wise outcomes of the proposed EW-GAT on the Cardiovascular Dataset for heart disease classification

Class Name	Accuracy%	Precision %	Recall %	F1 score%
0		99.91	99.93	99.92
1		99.93	99.91	99.92
Average	99.92	99.92	99.92	99.92

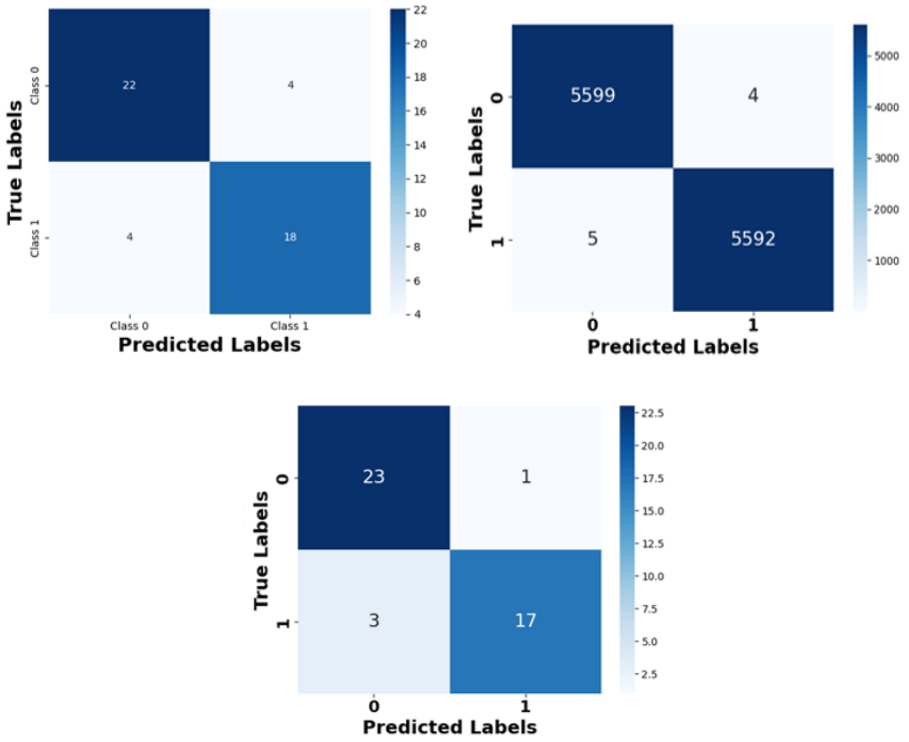

Fig. 3: Confusion matrices of the proposed EW-GAT on Cleveland, Statlog, and Cardiovascular datasets for heart disease classification

Furthermore, for the Cleveland, Statlog, and Cardiovascular datasets, accuracy and loss are evaluated over epochs, along with Area Under Curve (ROC-AUC), as shown in Fig. 4.

From Fig. 4, the training and validation curves illustrate how often the model predicts the correct samples and how the loss measures errors during training; the vertical axis shows the loss value. The curves are plotted over the full training period using a fixed learning rate, ensuring consistent loss scales for both training and validation. The AUC-ROC curve then indicates how well the model separates classes, with higher AUC values (close to 1) indicating perfect separation, and values below 0.5 suggesting random guessing.

Comparative Evaluation

The performance of the proposed EW-GAT model is compared with existing heart-disease classification models, including DAE (Alghamdi et al., 2024), SGO-tuned RF (Naik et al., 2025), and other models, namely Fuzzy ANN (Aggarwal and Kumar, 2023). This comparison establishes a baseline for assessing the proposed EW-GAT, which is outlined in Table 12.

As shown in Table 12, the proposed EW-GAT outperformed all existing models by integrating edge-weighted attention and adaptive feature interaction modeling. Existing methods such as DAE (Alghamdi et al., 2024), SGO-tuned RF (Naik et al., 2025), and Deep-DANet (Barfungpa et al., 2023) relied heavily on static learning, limiting their adaptability. Additionally, these models overlook relational dependencies and clinical variability. In contrast, the proposed model incorporates medically meaningful correlations into graph learning, enabling it to establish clearer decision boundaries. Consequently, EW-GAT achieves a higher accuracy of 99.81% on Cleveland. The comparison outcomes for the Statlog dataset are displayed in Table 13.

From Table 13, most existing models lack relational modeling and patient-aware weighting, making it challenging for them to adapt to the Statlog dataset. The proposed EW-GAT emphasizes adaptive edge-weighted attention to better capture non-linear feature dependencies and patient-specific variations, leading to superior results, achieving 99.83% accuracy compared to existing models.

Collectively, the strong performance results from feature-guided graph construction, adaptive edge weighting, and rigorous fold-wise validation, which together enhance learning robustness and prevent data leakage. The structured nature of the datasets further supports stable and highly accurate predictions. However, it is significant to note that these evaluations are conducted on relatively small, structured benchmark datasets. Such datasets can produce overly optimistic performance estimates due to limited variability.

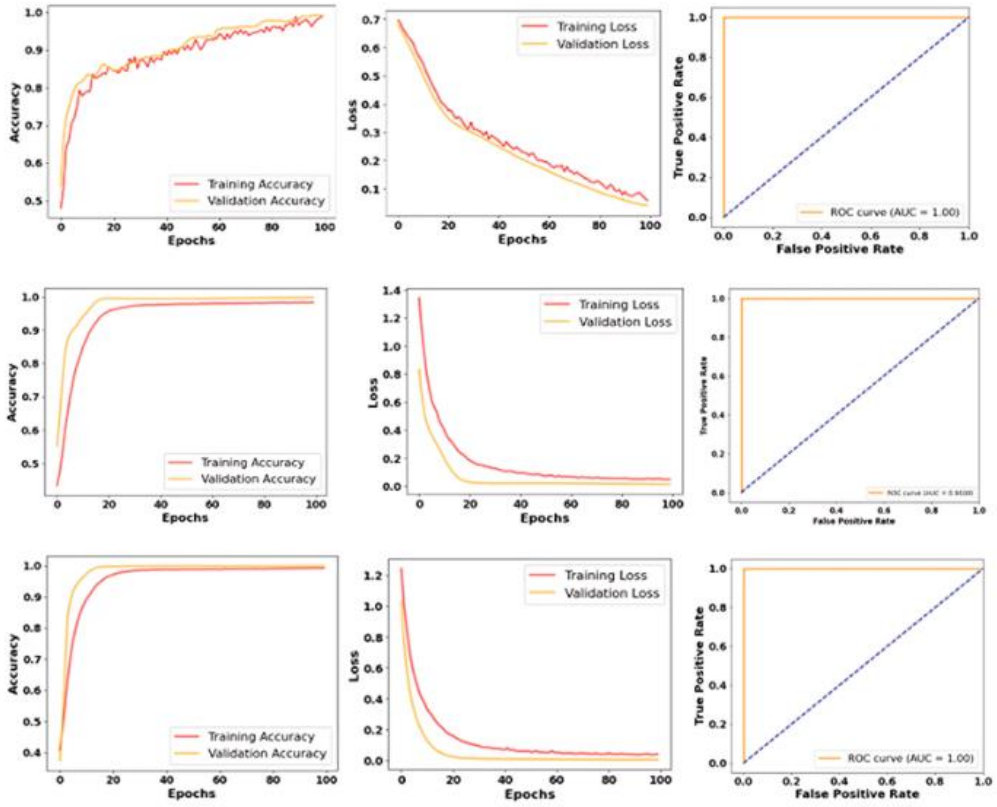

Fig. 4: Training and validation curves of the proposed EW-GAT model on Cleveland, Statlog, and Cardiovascular datasets for heart disease classification

Table 12: Comparative evaluation of the proposed EW-GAT against existing models on the Cleveland dataset

Comparative models	Accuracy (%)
DAE (Alghamdi et al., 2024)	99.60
SGO-tuned RF (Naik et al., 2025)	95.08
Fuzzy SVM (Qi and Zhang, 2024)	91.88
Deep-DANet (Barfungpa et al., 2023)	99.57
Fuzzy ANN (Aggarwal and Kumar, 2023)	98.56
HB with SMOTE and EDT (Sasikala and Mohanarathinam, 2024)	99.20
Proposed EW-GAT	99.81

Table 13: Comparative evaluation of the proposed EW-GAT model against existing models on the Statlog dataset

Comparative models	Accuracy (%)
SGO-tuned RF (Naik et al., 2025)	95.24
Fuzzy SVM (Qi and Zhang, 2024)	90.27
Deep-DANet (Barfungpa et al., 2023)	99.57
Fuzzy ANN (Aggarwal and Kumar, 2023)	98.66
HB with SMOTE and EDT (Sasikala and Mohanarathinam, 2024)	98.52
Proposed EW-GAT	99.83

To reduce overfitting, strict fold-wise validation, training-only preprocessing, and repeated runs are employed.

Discussion

The proposed WE-GAT framework directly addresses the limitations in existing heart-disease prediction models. Models such as DAE (Alghamdi et al., 2024) and Deep-DANet (Barfungpa et al., 2023) suffer from computational complexity, limiting their practicality in real-time settings. As a result, optimization-based ensemble models, such as SGO-tuned RF (Naik et al., 2025), showed strong performance. Still, they lacked robust generalization because they are highly sensitive to data variations and vulnerable to bias. Additionally, models such as fuzzy SVM (Qi and Zhang, 2024) improve outlier handling but struggle with high-dimensional feature interactions and fail to preserve clinically relevant feature dependencies. The proposed EW-GAT model addresses these issues by constructing a feature-to-feature interaction graph that captures medically relevant correlations prior to graph learning, enabling more reliable disease reasoning. The adaptive edge-weighting mechanism incorporates patient-aware contextual

relevance, ensuring that clinically significant dependencies are highlighted over mere numerical similarities. By embedding these relationships within an edge-weighted attention mechanism, EW-GAT overcomes the interpretability limitations of conventional GNNs and provides transparent explanations for diagnostic outcomes. Although the framework demonstrates strong performance on benchmark datasets, validating its clinical use in real-world settings requires testing on larger, diverse datasets. External validation on autonomous clinical datasets has not yet been conducted, which makes the framework susceptible to overfitting due to limited dataset size. Future work will focus on evaluating the model across multi-center clinical data and on enhancing interpretability to support clinical decision-making. Furthermore, understanding the interpretability of learned graph structures in a clinical context requires further research. The lightweight linear transformations and efficient MLP layer for classification contribute to improved scalability and inference speed. Overall, EW-GAT enhances predictive accuracy while establishing a meaningful clinical paradigm for heart-disease classification.

Conclusion

This study introduced EW-GAT for heart-disease classification, addressing key limitations of existing models in capturing clinical feature dependencies and patient variability. By integrating feature-to-feature correlations as structural priors alongside adaptive edge weighting and edge-aware attention, the proposed model achieves more discriminative, clinically relevant learning. Experimental results across multiple benchmark datasets demonstrate consistent and superior predictive performance under strict validation settings. Despite these promising results, the model was evaluated on smaller, structured clinical datasets. Future work will aim to adapt the framework to large-scale, diverse, real-world clinical data, including multimodal inputs such as time-series ECG signals, and to incorporate explainable Artificial Intelligence (AI) techniques to improve interpretability and practical application. The model can also be expanded to dynamic or temporal graph structures to track disease progression over time. Lastly, optimizing for real-time deployment and reducing computational demands will further enhance its practical use in clinical environments.

Acknowledgment

The author expresses sincere thanks to SEA College of Engineering and Technology, Bangalore, and the Department of Computer Science and Engineering for providing a supportive environment and the necessary facilities for this research. They also extend gratitude to their colleagues for their valuable guidance and cooperation. Additionally, the author acknowledges the works of the researchers who contributed to this study, and appreciates the constant encouragement and support of their family throughout the research process.

Funding Information

This research received no external funding

Author's Contributions

Ramakrishna Reddy Badveli: Methodology, software, conceptualization, resources, write-up of original draft.

Nijaguna Gollara Siddappa: Project administration, data curation, resources, formal analysis, investigation.

K. Sundeep Kumar: Visualization, Validation, supervision, data curation, write review and edited.

Ethics

This research did not involve any human participants or animals. Therefore, ethical approval and informed consent were not required.

References

Abbas, S., Ojo, S., Al Hejaili, A., Sampedro, G. A., Almadhor, A., Zaidi, M. M., & Kryvinska, N. (2024). Artificial intelligence framework for heart disease classification from audio signals. *Scientific Reports*, *14*(1), 3123. https://doi.org/10.1038/s41598-024-53778-7

Aggarwal, R., & Kumar, S. (2023). Classification model for meticulous presaging of heart disease through NCA using machine learning. *Evolutionary Intelligence*, *16*(5), 1689–1698. https://doi.org/10.1007/s12065-023-00830-6

Alghamdi, N. S., Zakariah, M., Shankar, A., & Viriyasitavat, W. (2024). Heart disease prediction using autoencoder and DenseNet architecture. *Egyptian Informatics Journal*, *28*, 100559. https://doi.org/10.1016/j.eij.2024.100559

Ashtaiwi, 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

Aslam, M. U., Xu, S., Hussain, S., Waqas, M., & Abiodun, N. L. (2024). Machine learning-based classification of valvular heart disease using cardiovascular risk factors. *Scientific Reports*, *14*(1), 24396. https://doi.org/10.1038/s41598-024-67973-z

Barfungpa, S. P., Sarma, H. K. D., & Samantaray, L. (2023). An intelligent heart disease prediction system using hybrid deep dense Aquila network. *Biomedical Signal Processing and Control*, *84*, 104742. https://doi.org/10.1016/j.bspc.2023.104742

Cenitta, D., Arjunan, R. V., Paramasivam, G., Arul, N., Palkar, A., & Chadaga, K. (2025). Ischemic Heart Disease Prognosis: A Hybrid Residual Attention-Enhanced LSTM Model. *IEEE Access*, *13*, 4281–4289. https://doi.org/10.1109/access.2024.3524604

Devarajan, M. V., Yallamelli, A. R. G., Yalla, R. K. M. K., Mamidala, V., Ganesan, T., & Sambas, A. (2025). An Enhanced IOMT anAn Enhanced IOMT and Blockchain-Bonitoring System Using BS-THA and OA-CNN. *Transactions on Emerging Telecommunications Technologies*, *36*(2), e70055. https://doi.org/10.1002/ett.70055

Elshewey, A. M., Abed, A. H., Khafaga, D. S., Alhussan, A. A., Eid, M. M., & El-kenawy, E.-S. M. (2025). Enhancing heart disease classification based on greylag goose optimization algorithm and long short-term memory. *Scientific Reports*, *15*(1), 1277. https://doi.org/10.1038/s41598-024-83592-0

El-Sofany, H. F. (2024). Predicting Heart Diseases Using Machine Learning and Different Data Classification Techniques. *IEEE Access*, *12*, 106146–106160. https://doi.org/10.1109/access.2024.3437181

El-Sofany, Hh., Bouallegue, Belgacem, & El-Latif, Y. M. A. (2024a). A proposed technique for predicting heart disease using machine learning algorithms and an explainable AI method. *Scientific Reports*, *14*(1), 23277. https://doi.org/10.1038/s41598-024-74656-2

Enad, H. G., & Mohammed, M. A. (2024). Cloud computing-based framework for heart disease classification using quantum machine learning approach. *Journal of Intelligent Systems*, *33*(1), 20230261. https://doi.org/10.1515/jisys-2023-0261

Fang, Y., Liu, D., Guo, Z., Leng, H., Liu, X., & Wu, X. (2025). Advanced heart disease classification based on multi-channel heart sound coupling features. *PLOS One*, *20*(5), e0321209. https://doi.org/10.1371/journal.pone.0321209

Gao, Y., Ma, P., Pan, J., Yang, H., Guo, T., & Wang, W. (2025). Non-invasive ML methods for diagnosis of congenital heart disease associated with pulmonary arterial hypertension. *Frontiers in Physiology*, *15*, 1502725. https://doi.org/10.3389/fphys.2024.1502725

Hagan, R., Gillan, C. J., & Mallett, F. (2021). Comparison of machine learning methods for the classification of cardiovascular disease. *Informatics in Medicine Unlocked*, *24*, 100606. https://doi.org/10.1016/j.imu.2021.100606

Hangaragi, S., Neelima, N., Jegdic, K., & Nagarwal, A. (2025). Integrated fusion approach for multi-class heart disease classification through ECG and PCG signals with deep hybrid neural networks. *Scientific Reports*, *15*(1), 8129. https://doi.org/10.1038/s41598-025-92395-w

Janosi, A., Steinbrunn, W., Pfisterer, M., & Detrano, R. (1989). Heart Disease UCI Machine Learning Repository. *UCI Machine Learning Repository*. https://doi.org/10.24432/C52P4X

Manikandan, G., Pragadeesh, B., Manojkumar, V., Karthikeyan, A. L., Manikandan, R., & Gandomi, A. H. (2024). Classification models combined with Boruta feature selection for heart disease prediction. *Informatics in Medicine Unlocked*, *44*, 101442. https://doi.org/10.1016/j.imu.2023.101442

Naik, A., Tejani, G. G., & Mousavirad, S. J. (2025). SGO enhanced random forest and extreme gradient boosting framework for heart disease prediction. *Scientific Reports*, *15*(1), 18145. https://doi.org/10.1038/s41598-025-02525-7

Narasimhan, G., & Victor, A. (2025). Empirical analysis of predicting heart disease using diverse datasets and classification procedures of machine learning. *Ain Shams Engineering Journal*, *16*(8), 103470. https://doi.org/10.1016/j.asej.2025.103470

Natarajan, K., Kumar, V. V., Mahesh, T. R., Abbas, Mohamed, Kathamuthu, Nirmaladevi, Mohan, E., & Annand, J. R. (2024). Efficient heart disease classification through stacked ensemble with optimized firefly feature selection. *International Journal of Computational Intelligence Systems*, *17*(1), 174. https://doi.org/10.1007/s44196-024-00538-0

Pandey, V., Lilhore, U. K., Walia, R., Alroobaea, R., Alsafyani, M., Baqasah, A. M., & Algarni, S. (2024). Enhancing heart disease classification with M2MASC and CNN-BiLSTM integration for improved accuracy. *Scientific Reports*, *14*(1), 24221. https://doi.org/10.1038/s41598-024-74993-2

Qi, Z., & Zhang, Z. (2025). Diagnosis of Heart Disease Using a Novel Membership Computation Method Within a Fuzzy SVM Framework. *International Journal of Fuzzy Systems*, *27*(7), 2077–2093. https://doi.org/10.1007/s40815-024-01888-9

Raoof, M., Mahtab, M., Masood Bhatti, S., Rashid, M., & Jaffar, A. (2025). Heart Disease Classification From Echocardiogram Images Using Deep Learning. *IEEE Access*, *13*, 8011–8022. https://doi.org/10.1109/access.2024.3524732

Sasikala, P., & Mohanarathinam, A. (2024). A powerful peripheral arterial disease detection using machine learning-based severity level classification model and hyper parameter optimization methods. *Biomedical Signal Processing and Control*, *90*, 105842. https://doi.org/10.1016/j.bspc.2023.105842

Suresh, N., Thomas, B., & Joseph, J. (2024). Bibliometric Analysis and Visualization of Scientific Literature on Heart Disease Classification Using a Logistic Regression Model. *Cureus*, *16*(6), e63337. https://doi.org/10.7759/cureus.63337