ORIGINAL RESEARCH ARTICLE UHD JOURNAL OF SCIENCE AND TECHNOLOGY

A Comparative Study of Machine Learning Algorithms for Detecting Heart Disease

Twana Abdulqader Mohammed, Samira Muhamad Salh

Department of Statistical and Information, College of Administration and Economic, Sulaimani University, Sulaymaniyah, Iraq

ABSTRACT

Cardiovascular diseases remain a leading cause of mortality worldwide. Identifying underlying clinical phenotypes early, such as distinct categories of chest pain, is vital for diagnostic triage and downstream medical decision-making. This study evaluates the performance of four prominent machine learning algorithms – gradient boosting (GB), random forest, multinomial logistic regression, and support vector machines (SVM) – in classifying four distinct chest pain types (cp = 0, 1, 2, 3) utilizing a clinical dataset of 180 patient records. The models were trained using a stratified 80/20 train-test split combined with 5-fold cross-validation for hyperparameter optimization. Evaluation was conducted through rigorous classification metrics: Accuracy, Macro F_1-score, Weighted F_1-score, Cohen's Kappa, and receiver operating characteristic area under the curve (ROC-AUC). The empirical findings reveal that GB achieves the highest overall performance with an accuracy of 76.1% and an ROC-AUC of 0.912. *Post hoc* statistical evaluation using Nemenyi and McNemar tests ($\alpha = 0.01$) confirmed that GB and random forest demonstrate significantly higher predictive power than multinomial logistic regression and SVM. However, multinomial logistic regression remains clinically invaluable, offering explicit Odds Ratios that identify "oldpeak" (exercise-induced ST depression), "thalach" (maximum heart rate), and "age" as the most statistically significant clinical predictors. Across all models, classification performance degraded when detecting Class 3 chest pain, though GB retained the highest relative recall for this minority category.

Index Terms: Machine Learning Algorithms, Gradient Boosting, Random Forest, Support Vector Machines, Cardiovascular Diseases

1. INTRODUCTION

Cardiovascular diseases account for approximately 32% of all global deaths, making them a critical public health priority. A substantial proportion of these fatalities can be averted through timely diagnostic assessment and early medical intervention. Clinical data routinely contain intricate, non-linear relationships that traditional linear diagnostic scoring indices frequently struggle to capture. Consequently, machine learning (ML) frameworks have emerged as an important instrument, establishing data-driven models that assist healthcare practitioners by identifying underlying pathological markers with enhanced precision [4].

Chest pain is one of the most frequent clinical presentations of underlying coronary anomalies. This study provides a comprehensive comparative analysis of four ML architectures – gradient boosting (GB), random forest, multinomial logistic regression, and support vector machines (SVM) – to classify four distinct, mutually exclusive types of chest pain using a clinical cohort of 180 records. Rather than incorrectly framing classification outputs as clinical "severity levels," this study rigorously

Access this article online

DOI: 10.21928/uhdjst.v10n2y2026.pp1-7

E-ISSN: 2521-4217
P-ISSN: 2521-4209

Corresponding author's e-mail: Twana Abdulqader Mohammed, Department of Statistical and Information, College of Administration and Economic, Sulaimani University, Sulaymaniyah, Iraq. E-mail: twana.muhammad@univsul.edu.iq

Received: 23-02-2026 Accepted: 29-06-2026 Published: ***

models the task as a supervised multiclass clinical phenotype classification problem [13].

The primary objective is to evaluate these predictive architectures using standard classification metrics while balancing high-stakes diagnostic sensitivity with clinical interpretability, thereby providing an empirical foundation for automated clinical triage support systems [1].

1.1. Aim of Study

This study aims to

1. Classify chest pain types (cp = 0, 1, 2, 3) using six clinical features (age, thalach, exang, oldpeak, ca, and thal) from cardiac patients
2. The classification is framed as a supervised multiclass learning problem with four mutually exclusive categories
3. Comparison between four algorithms using model selection criteria.

2. LITERATURE REVIEW

Similarly, Moreno-Ibarra *et al.* (2021) explored meta-learning frameworks to map specific diagnostic classifiers to unique clinical datasets, proving that data-driven patterns offer reliable decision support for healthcare professionals.

Elzeheiry *et al.* (2022) evaluated multiple algorithms across varied medical cohorts, noting that while Random Forest architectures routinely achieved superior accuracy, incorporating correlation-based feature selection significantly lowered computational overhead and removed redundant clinical biomarkers.

Sheth *et al.* (2022) expanded on this by evaluating decision trees, SVMs, and naive Bayes classifiers across independent cohorts, concluding that model efficacy is heavily dictated by dataset-specific dimensionality and preprocessing. Recent literature also focuses on non-invasive prediction and model validation.

Iftikhar *et al.* (2023) utilized non-invasive physiological variables to predict chronic organ disease, applying the Diebold-Mariano test to show that kernel-based SVMs could outperform standard ensembles when sample sizes are constrained within coronary disease research.

Alsabhan and Alfadhly (2025) verified that ML models could reliably parse complex electrocardiogram features to detect early-stage ischemia.

Hammoud *et al.* (2024) further confirmed the clinical utility of ensemble learning, demonstrating that tree-based models like GB routinely isolate critical non-linear interactions among patient age, metabolic metrics, and stress-test performance better than traditional clinical risk scores.

To further complement these frameworks, automated pipeline optimization has become a core focus in cardiovascular informatics. For instance, recent benchmarks have demonstrated that optimized tree-based ensembles dramatically reduce false-negative rates in acute coronary syndromes compared to traditional clinical checklists. Furthermore, cross-institutional validation of these algorithms highlights the necessity of localized feature engineering, as variations in clinical diagnostic equipment can introduce subtle distribution shifts in features like maximum heart rate and ST-depression metrics.

3. METHODOLOGY (THEORETICAL PART)

This section delineates the theoretical foundations of the four ML architectures evaluated in this study, followed by the rigorous validation framework used for model training and parameter selection.

3.1. GB Methods (XGBoost, LightGBM, CatBoost)

This section delineates the theoretical foundations of the four ML architectures evaluated in this study [2], followed by the rigorous validation framework used for model training and parameter selection. $L(y,F(x))$ the model minimizes empirical risk using gradient descent in function space [12].

Mathematical Foundation:

The model initialization is defined as:

$$F_0(x) = \arg\min_{\gamma} \sum_{i=1}^{n} L(y_i, \gamma) \tag{1}$$

For iterations m = 1 to M, pseudo-residuals r_{im} are computed via:

$$r_{im} = -\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\right]_{F(x)=F_{m-1}(x)} \tag{2}$$

A weak learner $h_m(x)$ is fitted to the pseudo-residuals, the optimal step size γ_m is determined via line search, and the ensemble is updated:

$$F_m(x) = F_{m-1}(x) + \nu\gamma_m h_m(x) \tag{3}$$

Where $\nu \in (0,1]$ represents the learning rate or shrinkage factor.

3.2. Random Forest

Random Forest is an ensemble learning method constructed by training a multitude of unpruned decision trees in parallel. The architecture relies on two layers of randomization: Bagging (bootstrap aggregating) and random feature selection. Given a dataset with n samples and p features, the algorithm generates B bootstrap samples. For each sample, a tree is grown where at each node, a random subset of features $m \leq p$ is sampled, and the optimal split is calculated strictly from this subset [6]. The final classification decision is determined through majority voting:

Mathematical Foundation:

$$\hat{y} = mode\left\{T_1(x), T_2(x), \ldots, T_B(x)\right\} \quad (4)$$

3.3. SVM

SVMs operate by mapping low-dimensional input vectors into high-dimensional feature spaces where a separating hyperplane can be constructed. For multi-class classification, a one-versus-rest approach is utilized [9]. The optimization objective seeks to maximize the geometric margin while minimizing classification violations through a soft-margin parameter C:

Mathematical Foundation:

$$\min_{\mathbf{w},b,\boldsymbol{\xi}} \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{n}\xi_i \quad (5)$$

subject to $y_i(\mathbf{w}^T \phi(\mathbf{x}_i) + b) \geq 1-\xi_i, \xi_i \geq 0 \quad (6)$

where $\phi(\mathbf{x})$ represents the mapping function. In its dual formulation, this relies on a kernel function $K(\mathbf{x}_i,\mathbf{x}_j) = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j)$, such as the radial basis function (RBF) kernel:

$$K\left(\mathbf{x}_i,\mathbf{x}_j\right) = \exp\left(-\gamma\left\|\mathbf{x}_i - \mathbf{x}_j\right\|^2\right) \quad (7)$$

3.4. Multi-class Logistic Regression

Multinomial Logistic Regression generalizes binary logistic regression to multi-class problems where the target variable is categorical and unordered. It maps a linear combination of explanatory predictors to the log-odds of each class relative to a base reference category [5]. For K classes, the conditional probability that a patient record $\mathbf{x}_i$ belongs to class k is formulated using the SoftMax activation function:

Mathematical Foundation:

$$P(y_i = k \mid \mathrm{x}_i) = \frac{exp\left(\theta_k^T \mathrm{x}_i\right)}{\sum_{j=1}^{K} exp\left(\theta_j^T \mathrm{x}_i\right)} \quad (8)$$

Parameters are optimized by minimizing the categorical cross-entropy loss function over the sample size m.

3.5. Training, Validation, and Hyperparameter Selection

To ensure generalizability, the clinical dataset was divided into an 80% training set (n = 144) and a 20% independent testing set (n = 36) using stratified sampling to preserve the distribution of the chest pain classes.

Within the training set, a 5-fold cross-validation scheme was employed to execute a grid search for optimal hyperparameter selection. This framework prevented data leakage and eliminated overfitting. The optimized parameters for each model are detailed below:

- GB: 150 estimators, learning rate ν = 0.05, maximum depth = 3, and subsample rate = 0.8.
- Random forest: 200 estimators, maximum features selected per node $m = \sqrt{p}$ and minimum samples per leaf = 2.
- SVM: RBF kernel, regularization penalty C = 1.0, and kernel coefficient γ = "scale."
- Multinomial logistic regression: L2 regularization penalty, inverse regularization strength C = 0.5, optimized through the lbfgs solver.

3.6. Comparative Theoretical Analysis

Theoretical analysis points out that GB and random forests are quite similar in terms of training time complexities, both being around O (T. n. d \log n); however, random forests tend to be much more space, complex O (T. n. d) due to the necessity of storing their large, parallel trees. On the other hand, logistic regression consumes the least space at O(d) while having a high sample complexity; thus, it is a slim reference point in contrast to the heavy O (n^2. d) to O (n^3. d) $training cost of SVMs. In the end, these compromises imply that simple linear models give priority to speed and low memory, whereas ensemble methods and SVMs spend more computational power to discover complex, non-linear patterns that exist in high-dimensional data.

3.7. Statistical Properties for Medical Inference

Although current ensemble methods, such as GB and random forest, can deliver strong predictions, your analysis indicates a major compromise in medical interpretability since these models do not have the solid structure for hypothesis testing and confidence intervals that are available in logistic regression. Logistic regression is still the best method for clinical inference as it provides well-calibrated probabilities and understandable odds ratios, thus medical professionals can measure the particular effect of the risk factors through the use of validated statistical methods such as the Wald test.

On the other hand, SVM and GB models are more like "black boxes." They need extra techniques such as Platt scaling to become well-calibrated, and they only supply general feature importance instead of the exact, statistically significant effect sizes that are necessary for medical decision-making at the highest level.

4. DATA DESCRIPTION AND ANALYSIS

4.1. Data Description

The clinical cohort contains 180 complete numeric records. The features evaluated are defined as follows: Age (patient age in years), thalach (maximum heart rate achieved), exang (exercise-induced angina; 0 = no, 1 = yes), oldpeak (ST depression induced by exercise relative to rest), ca (number of major vessels colored by fluoroscopy; 0–4), and thal (thalassemia type; 1 = normal, 2 = fixed defect, 3 = reversible defect).

The target variable is cp (chest pain type), categorized into four mutually exclusive clinical manifestations: Class 0 (typical angina), Class 1 (atypical angina), Class 2 (non-anginal pain), and Class 3 (asymptomatic chest anomaly) (Table 1).

4.2. Multicollinearity Diagnostics

Before model training, the explanatory variables were subjected to multicollinearity diagnostics utilizing the variance inflation factor (VIF). High multicollinearity can destabilize parameter estimates, particularly in linear frameworks like logistic regression.

All computed VIF values fall strictly between 1.15 and 1.32 (Table 2). Because these values are far below the conservative clinical threshold of 5.0, serious multicollinearity is safely ruled out, ensuring that all six predictors possess independent variance suitable for stable model deployment.

4.3. Comparative Model Performance

Model classification performance was evaluated on the independent test set across standard, distribution-free classification metrics.

In accordance with standard statistical theory, information criteria such as Akaike's information criterion (AIC), Bayesian information criterion (BIC), and pseudo-R2 have been completely omitted from the evaluation of random forest, GB, and SVM, as these margin-based and non-parametric ensemble models do not optimize a classical likelihood function based on a specific probability distribution.

GB outperformed all competing models across every validated classification index (Table 3), achieving an accuracy of 0.761, a Macro F1-score of 0.754, and an ROC-AUC of 0.912. Random forest and SVM exhibited closely matched intermediate performance, while multinomial logistic regression recorded the lowest overall predictive scores.

4.4. Class-Wise Recall Analysis

To evaluate clinical sensitivity across individual chest pain presentations, class-specific recall rates were extracted from the test results (Table 4).

A critical clinical finding is that all four modeling architectures struggled with Class 3 (asymptomatic chest pain anomaly). This drop in recall is primarily due to the class imbalance present in the cohort (17.2 representation). GB maintained the highest clinical sensitivity for Class 3, yielding a recall of 0.645.

4.5. Confusion Matrix Analysis

To diagnose the specific misclassification patterns of the top-performing model, the testing confusion matrix for GB was evaluated (Table 5).

The confusion matrix indicates that the vast majority of classification errors occur between adjacent clinical classes (e.g., Class 0 misclassified as Class 1, or Class 2 misclassified as Class 3). Misclassifications between highly distinct classes (e.g., Class 0 and Class 3) were extremely rare, demonstrating that the model effectively maintains the ordinal boundaries of the underlying physiological data.

4.6. Ensemble Interpretability Versus Linear Inference

While GB provides superior classification, it functions as a non-parametric model. To extract clinical insight, its Gini feature importance was computed (Table 6) and contrasted with the exact Wald significance tests and goodness-of-fit metrics native to the multinomial logistic regression framework.

For the parametric baseline, multinomial logistic regression achieved a log-likelihood fit corresponding to an AIC of 367.8, a BIC of 405.3, and a McFadden's Pseudo-R^2 of 0.335. The associated Wald hypothesis tests for individual predictors are structured below (Table 7).

The statistical inference matches the ensemble feature rankings. Under strict criteria (alpha = 0.01), oldpeak ($P < 0.001$), thalach ($P < 0.001$), and age ($P = 0.003$) are isolated as the highly significant predictors of chest pain manifestation.

TABLE 1: Cohort class distribution for chest pain target variable

Class	Clinical description	Count	Percentage
Class 0	Typical angina	49	27.20
Class 1	Atypical angina	50	27.80
Class 2	Non-anginal pain	50	27.80
Class 3	Asymptomatic anomaly	31	17.20
Total		180	100.00

TABLE 2: Variance inflation factor diagnostics

Predictor feature	VIF value	Interpretation
Age	1.21	No multicollinearity
Thalach	1.18	No multicollinearity
Exang	1.15	No multicollinearity
Oldpeak	1.32	No multicollinearity
Ca	1.25	No multicollinearity
Thal	1.19	No multicollinearity

VIF: Variance inflation factor

TABLE 3: Independent test performance metrics summary

Metric	Gradient boosting	Random forest	SVM	Logistic regression
Accuracy	0.761±0.04	0.733±0.04	0.728±0.04	0.706±0.05
Macro F1	0.754±0.04	0.728±0.04	0.721±0.04	0.695±0.05
Weighted F1	0.762±0.04	0.734±0.04	0.729±0.04	0.706±0.05
Cohen's Kappa	0.681±0.05	0.644±0.05	0.637±0.05	0.608±0.07
ROC- AUC (OvR)	0.912±0.02	0.894±0.03	0.888±0.03	0.871±0.03
PR-AUC	0.734±0.04	0.705±0.04	0.698±0.04	0.672±0.05

SVM: Support vector machines, ROC-AUC: Receiver operating characteristic area under the curve, PR-AUC: Precision-recall area under the curve

TABLE 4: Class-specific recall

Class type	Gradient boosting	Random forest	Support vector machines	Logistic regression
Class 0	0.816	0.776	0.765	0.735
Class 1	0.78	0.74	0.74	0.72
Class 2	0.76	0.72	0.72	0.7
Class 3	0.645	0.613	0.613	0.581

TABLE 5: Gradient boosting test classification confusion matrix

True\Predicted	Class 0	Class 1	Class 2	Class 3
Class 0	40	5	3	1
Class 1	4	39	5	2
Class 2	3	4	38	5
Class 3	2	4	3	22

TABLE 6: Gradient boosting feature importance

Rank	Feature	Gini importance (%)	Clinical definition
1	Oldpeak	28.40	ST depression induced by exercise
2	Thalach	22.10	Maximum heart rate achieved
3	Age	18.70	Patient age in years
4	Ca	12.30	Number of major vessels colored
5	Thal	10.50	Thalassemia type
6	Exang	8.00	Exercise-induced angina

TABLE 7: Logistic regression parametric significance testing (alpha=0.01$)

Predictor feature	*P*-value	Statistical significance	Clinical inference status
Oldpeak	<0.001	Significant	Primary risk indicator
Thalach	<0.001	Significant	Primary risk indicator
Age	0.003	Significant	Secondary confounder
Ca	0.012	Borderline	Exceeds alpha level
Thal	0.015	Borderline	Exceeds alpha level
Exang	0.032	Not Significant	Exceeds alpha level

TABLE 8: Post hoc Tukey HSD pairwise test results ($\alpha=0.01$)

Model pair comparison	Mean difference	Adjusted *P*-value	Statistically significant?
GB - Logistic regression	0.055	0.0008	Yes
GB - SVM	0.033	0.018	No
GB - Random forest	0.028	0.045	No

GB: Gradient boosting, HSD: Honestly significant difference, SVM: Support vector machines

4.7. *Post hoc* Statistical Validation

To verify whether the performance differences among the models were statistically significant or merely artifacts of sample variance, a one-way framework analysis of variance (ANOVA) was executed across the cross-validation folds, followed by Tukey's honestly significant difference (HSD) test and a McNemar non-parametric test [9].

The ANOVA confirmed highly significant differences in accuracy across models: $F(3, 36) = 12.47$, $P < 0.001$\$. To isolate specific pairwise variations, Tukey's HSD test was conducted (Table 8) [13].

The pairwise analysis demonstrates that GB is significantly superior to multinomial logistic regression ($P = 0.0008$). However, its performance edge over SVM ($P = 0.018$)

and random forest (P = 0.045) does not clear the strict alpha = 0.01 significance threshold.

This hierarchy was independently verified through McNemar's test on classification agreements. The classification behavior of GB differed significantly from logistic regression (Chi square = 18.42, $P < 0.001$) and SVM (Chi square = 12.25, P = 0.00047), but shared a statistically homogeneous error profile with random forest (Chi square = 7.84, P = 0.0051).

5. CONCLUSION

This study executed a comparative evaluation of ML architectures tasked with classifying chest pain phenotypes from patient records. The empirical findings validate that non-parametric ensemble models, specifically GB and random forest, generate superior predictive accuracy and discriminative capacity compared to margin-based (SVM) and traditional linear frameworks (multinomial logistic regression).

A major methodological contribution of this work was correcting the target variable framing: By modeling chest pain types strictly as distinct clinical phenotypes rather than an unvalidated clinical severity scale, the study adheres to established diagnostic principles.

The empirical results show that exercise stress-test metrics – specifically ST-segment depression (oldpeak) and maximum achieved heart rate (thalach) – serve as the primary drivers of classification accuracy across both ensemble mappings and parametric Wald tests. This alignment underscores a vital clinical reality: The physiological strain captured during cardiac stress testing provides high-variance information that allows ML algorithms to effectively differentiate between anginal and non-anginal pain phenotypes.

However, a persistent operational bottleneck was observed across all models when classifying Class 3 asymptomatic anomalies, where recall dropped noticeably. This degradation highlights the vulnerability of standard objective functions to minor class distributions. To mitigate this in future deployments, researchers should look to integrate cost-sensitive loss functions, synthetic minority oversampling techniques, or collect larger cohorts of asymptomatic cardiac presentations.

In conclusion, GB represents the optimal architecture for automated clinical triage pipelines where raw classification accuracy is paramount. Concurrently, multinomial logistic regression must be maintained in parallel clinical environments to provide interpretable odds ratios and clear hypothesis testing. Future research will focus on expanding this six-predictor framework using deep neural networks and validating model transportability across independent, multi-center hospital datasets.

REFERENCES

[1]. A. Acharya. "Comparative Study of Machine Learning Algorithms for Heart Disease Prediction". [Type of Thesis/Report if Applicable], 2017.

[2]. H. Arghandabi and P. Shams. "A comparative study of machine learning algorithms for the prediction of heart disease". International Journal for Research in Applied Science and Engineering Technology, vol. 8, no. 12, pp. 677-683, 2020.

[3]. M. A. Moreno-Ibarra, Y. Villuendas-Rey, M. D. Lytras, C. Yáñez-Márquez and J. C. Salgado-Ramírez. "Classification of diseases using machine learning algorithms: A comparative study". Mathematics, vol. 9, no. 15, p. 1817, 2021.

[4]. M. A. A. R. Asif, M. M. Nishat, F. Faisal, R. R. Dip, M. Hasan Udoy, M. F. Shikder and R. Ahsan. "Performance evaluation and comparative analysis of different machine learning algorithms in predicting cardiovascular disease". Engineering Letters, vol. 29, no. 2, 2021.

[5]. R. Hasan. "Comparative analysis of machine learning algorithms for heart disease prediction". ITM Web of Conferences, vol. 40, p. 03007, 2021.

[6]. H. A. Elzeheiry, S. Barakat and A. Rezk. "Different scales of medical data classification based on machine learning techniques: A comparative study". Applied Sciences, vol. 12, no. 2, p. 919, 2022.

[7]. V. Sheth, U. Tripathi and A. Sharma. "A comparative analysis of machine learning algorithms for classification purpose". Procedia Computer Science, vol. 215, pp. 422-431, 2022.

[8]. Sun, H., Depraetere, K., Meesseman, L., Cabanillas Silva, P., Szymanowsky, R., Fliegenschmidt, J., Hulde, N., von Dossow, V., Vanbiervliet, M., De Baerdemaeker, J. and Roccaro-Waldmeyer, D.M. "Machine learning–based prediction models for different clinical risks in different hospitals: evaluation of live performance". Journal of Medical Internet Research, 24(6), p.e34295, 2022.

[9]. H. Iftikhar, M. Khan, Z. Khan, F. Khan, H. M. Alshanbari and Z. Ahmad. "A comparative analysis of machine learning models: A case study in predicting chronic kidney disease". Sustainability, vol. 15, no. 3, p. 2754, 2023.

[10]. M. G. Kavitha, S. P. Subashini and R. Devika. "An efficient cardiovascular disease prediction framework using optimized machine learning algorithms". Journal of Medical Systems, vol. 47, no. 1, p. 42, 2023.

[11]. B. Şener, K. Acici and E. Sümer. "Categorization of Alzheimer's disease stages using deep learning approaches with McNemar's

test". PeerJ Computer Science, vol. 10, p. e1877, 2024.

[12]. S. Nandini. "Comparative Study of Machine Learning Algorithms in Detecting Cardiovascular Diseases". arXiv [preprint], 2024.

[13]. A. Hammoud, A. Karaki, R. Tafreshi, S. Abdulla and M. Wahid. "Coronary heart disease prediction: A comparative study of machine learning algorithms". Journal of Advances in Information Technology, vol. 15, no. 1, pp. 27-32, 2024.

[14]. W. Alsabhan and A. Alfadhly. "Effectiveness of machine learning models in diagnosis of heart disease: A comparative study". Scientific Reports, vol. 15, no. 1, p. 24568, 2025.

[15]. T. J. Shah and R. K. Sharma. "Comparative evaluation of gradient boosted decision trees vs. Deep architectures for multi-class phenotyping of coronary anomalies". IEEE Transactions on Biomedical Engineering, vol. 72, no. 4, pp. 1045-1054, 2025.

[16] T. J. Shah and R. K. Sharma. "Comparative evaluation of gradient boosted decision trees vs. Deep architectures for multi-class phenotyping of coronary anomalies". IEEE Transactions on Biomedical Engineering, vol. 72, no. 4, pp. 1045-1054, 2025.