DOI: 10.21917/ijivp.2026.0565

CORONARY ARTERY DISEASE PREDICTION USING AN EXPLAINABLE DEEP CARDIAC TRANSFORMER FRAMEWORK

R. Rajavignesh[1] and G. Venkataramana Sagar[2]

[1]*Department of Computer Science and Engineering, K.S.K College of Engineering and Technology, India*
[2]*Department of Electronics and Communication Engineering, G Pulla Reddy Engineering College, India*

Abstract

Coronary Artery Disease (CAD) is one of the significant diseases in the field of cardiovascular science where timely identification of the risk will enable the diagnosis of CAD and its proper treatment. The advent of clinical data, laboratory data, and electrocardiographic data has led to the creation of artificial intelligence-based systems to predict CAD. Although traditional machine learning models are not very efficient in understanding the interactions among the features, on the other hand, the deep learning model provides good predictive capabilities but less interpretability. The existing CAD prediction approaches do not pay much attention to clinical transparency, feature interaction, calibration, and prediction explanation. The above-discussed problem makes the CAD prediction approaches unsuitable for the clinical environment. In this work, we propose a new Explainable Deep Cardiac Transformer (EDCT) that integrates deep feature learning, self-attention mechanism, residual feature refinement, and explainability for predicting CAD. As the proposed framework, we first preprocess and normalize the features followed by the deep cardiac feature embedding. The proposed transformer-based attention mechanism learns the interactions and long range dependencies between the heterogeneous clinical and cardiac features. Then, the obtained representations are fed into the classification layers to predict the CAD risk and identify the contribution of each clinical feature to the particular prediction using SHAP-based explanation analysis. According to the experiments, it was found that our proposed EDCT framework has outperformed all the existing approaches in all metrics. At the full training-data proportion, our EDCT framework gives 89.8% accuracy, 87.8% precision, 90.6% recall, 89.3% F1-score, and 0.938 ROC-AUC. Compared with XGBoost, the improvement of EDCT is 3.7%, 3.2%, 3.5%, 3.5%, and 0.032 for accuracy, precision, recall, F1-score, and ROC-AUC, respectively. Compared with AutoGluon Ensemble, the corresponding improvements are 2.5%, 2.1%, 2.4%, 2.4%, and 0.018, respectively. The experimental results show that the self-attention-based cardiac representation learning effectively captures interactions between the heterogeneous clinical and cardiac features.

Keywords:

Coronary Artery Disease, Cardiac Transformer, Deep Learning, Explainable Artificial Intelligence, Risk Prediction

1. INTRODUCTION

Coronary artery disease (CAD) continues to be one of the major cardiovascular diseases globally, and it is associated with morbidity, mortality, recurrent cardiovascular events, and impaired quality of life. Coronary artery disease is formed due to the gradual build-up of atherosclerotic plaques inside the coronary arteries, limiting myocardial blood flow and causing ischemia. Thanks to the growing amount of electronic health records, laboratory data, electrocardiographic measurements, imaging findings, and patient demographics, it becomes possible to develop risk assessments based on data-driven predictions. Artificial Intelligence (AI), especially machine learning (ML) and deep learning (DL), helps detect non-linear dependencies between these heterogenous variables and identify the patients who are at risk. Several reviews of AI applications in cardiovascular medicine confirm that CAD is one of the most investigated use cases and that there is a tendency towards combination of clinical records and imaging data for personalization [1], [2]. The traditional approach to the diagnosis and assessment of coronary artery disease is based on the well-known clinical risk factors and scoring systems. Although this approach is still valuable for clinical practice, it does not consider all the complex nonlinear dependencies of age, gender, blood pressure, lipid profile, diabetes, smoking, electrocardiographic data, and other patient-specific variables. An AI approach allows to find these dependencies directly from data. One of the main advantages of deep learning is that it can learn hierarchical representations from high-dimensional inputs without using only hand-crafted features. In recent years, it was proven that DL could help with CAD detection, prognosis, cardiovascular imaging and risk stratification [1].

The quick advancements in the field of attention and transformer design provide another chance for cardiovascular prediction. Contrary to regular recurrent networks, transformers can model interactions between multiple input features by means of self-attention and can thus capture interactions that would be challenging to represent in sequential and/or convolutional architectures. Transformers and attention mechanisms have recently attracted considerable attention as a powerful approach to ECG analysis and multimodal cardiac prediction in AI-based cardiovascular medicine. Recent literature reviews demonstrate a shift from classical CNN-based architectures towards hybrid ones based on attention, transformers, and multimodal feature fusion [2], [3]. However, several obstacles still prevent the clinical implementation of CAD prediction with the help of artificial intelligence. Firstly, cardiovascular datasets often vary in terms of demographic characteristics, acquisition procedures, feature definitions, class distribution, and labelling standards of the disease. The heterogeneity of data can result in the lower generalizability of the model which will perform well on the original dataset but poorly on a new one. Recent systematic evidence revealed limited availability of public datasets, inconsistent evaluation procedures, lack of external validation, and model complexity as the key issues in the field of DL-based CAD prediction [4]. Secondly, the black-box nature of DL-based architectures remains a considerable obstacle. A model can generate an accurate prediction while not offering any evidence on how this particular prediction was made. It is especially important in medicine as the clinicians need evidence that the model's decisions align with physiology and clinical knowledge. While explainable AI techniques, such as SHAP, LIME, saliency analysis, and Grad-CAM have become increasingly popular, the

application and evaluation of these techniques remain inconsistent in cardiovascular research [5]-[6].

The objectives of this research are to develop an Explainable Deep Cardiac Transformer (EDCT) framework for predicting the risk of CAD. The main innovation of the presented research lies in incorporating the transformer-based approach to deep representation learning with an explicitly designed explainability technique in a single CAD risk prediction framework. Unlike existing approaches which use XAI techniques only in the post-processing stage of independently developed classifiers, the EDCT framework is designed considering the complementary demands of prediction and explanation. The transformer-based architecture is responsible for capturing the higher-order interplays between different cardiovascular features, while the explainability module translates model decisions into feature-level contributions. The key contributions of this research are: A novel EDCT framework for CAD risk prediction that employs self-attention-based representation learning to capture the complex interactions between heterogeneous cardiovascular features. A clinical interpretability mechanism is integrated in the framework based on the SHAP feature attribution approach.

2. RELATED WORKS

The latest studies indicate a definite shift from classic statistical approaches to machine learning, deep learning, multimodal learning, and explainable AI for CAD prediction.

First, [7] studied the possibility of machine learning to discover CAD risk factors based on demographic, laboratory, physical examination, and lifestyle parameters. They proposed an XGBoost model, which scored an AUROC of 0.89, and identified age, platelets, heart disease family history, and total cholesterol level as the most significant variables. SHAP analysis confirmed that machine learning approaches can detect clinically meaningful associations between patients' characteristics and CAD. Though this study showed the advantages of explainable tree-based learning, it was based on a conventional ML architecture and, thus, missed the deep representation learning and self-attention aspects.

Second, [8] suggested an approach for explaining CAD predictions using AutoML and AutoGluon. They integrated five publicly available CAD datasets and built an ensemble approach, which was compared with individual baseline classifiers. The ensemble model achieved accuracy 0.9167 and AUC 0.9562 when applying four-fold cross-bagging and explaining results by means of SHAP. Thus, the authors showed that the application of automated ensemble learning increases CAD prediction quality while ensuring the interpretability of models. Nevertheless, an ensemble learning approach created with the help of AutoML does not include feature dependency modeling by means of attention mechanisms. Hence, the investigation of transformers is justified.

Third, [9] designed a machine learning framework based on the combination of clinical, laboratory and ECG data for obstructive CAD prediction. They showed that ECG data is very influential for CAD predictions, while age, sex, diabetes, BMI, hemoglobin, blood pressure, and lipid parameters also have significant impact on prediction outcomes. SHAP analysis was applied to investigate the association between heterogeneous predictors and CAD outcomes. This paper is especially interesting as it shows the advantages of the combination of physiological signals and structured clinical data. Nevertheless, the proposed framework is based on a conventional ensemble technique.

Other studies considered deep learning for the identification of significant coronary stenosis. The deep learning algorithm for asymptomatic patients with coronary CT angiography was used along with linear regression, random forest, and XGBoost algorithms. SHAP analysis was then performed in order to find out how clinical and laboratory characteristics contributed to the prediction. The results show that DL algorithms could find complex dependencies in the cardiovascular data whereas post hoc explanations could increase interpretability. However, imaging-based methods may require considerable computational resources and particular imaging protocols and hence be difficult to apply in case of only clinical data availability.

Machine learning was also studied for identifying hemodynamically significant CAD by means of using conventional risk factors, coronary artery calcium, and epicardial fat volume [10]. XGBoost had AUC of 0.88 which was superior to logistic regression and SVM in the training set. The method of SHAP also provided individual explanations. The results confirm the usefulness of integrating conventional risk factors with imaging markers. However, reliance on particular imaging measurements may limit the wide use of the approach.

[11] evaluated nine machine learning algorithms using the Z-Alizadeh Sani dataset and Boruta-based feature selection along with Shapley-value analysis. The paper pays much attention to the nonlinear dependencies of CAD and risk factors and shows the usefulness of interpretable machine learning for understanding of the underlying disease mechanisms. Nevertheless, the rather limited size of the dataset makes it possible to make reliable conclusions concerning generalization to other cases.

Another research has proposed machine learning framework for the diagnosis and prediction of CAD and its severity using SYNTAX and GENSINI scores. XGBoost gave the highest reported performance among the studied algorithms whereas left ventricular ejection fraction, homocysteine, hemoglobin, BNP, HDL, and glycated hemoglobin were identified as important features for different tasks [12]. The example shows the potential of ML for going beyond CAD diagnosis and estimating disease severity but the performance level is still not high enough. Later work in explainable ML focused on clinical data and model calibration. According to a study conducted in 2025, using a harmonized UCI-based repository, the F1-score was 0.8436, accuracy was 0.8278, and ROC-AUC was 0.90, while CatBoost detected chest pain type, exercise-induced angina, and peak-exercise ST-segment slope as relevant factors [13]-[15].

3. PROPOSED EDCT

The novel EDCT is a deep learning system that predicts CAD from various clinical and cardiac attributes and explains every prediction in terms of interpretable evidence. The proposed approach combines various stages of data preprocessing, clinically driven feature representations, cardiac feature embedding, multi-head self-attention, residual feature refinement, nonlinear classification, and explainability based on SHAP values into one framework. Unlike traditional machine learning

approaches, which view clinical variables mostly as separate predictors, EDCT learns relationships between different attributes, such as age, blood pressure, cholesterol, glucose level, heart rate, electrocardiographic parameters, exercise features, and other cardiac variables. Self-attention part of EDCT discovers relationships between variables in terms of attention weights, while the explainability module computes the importance of each input attribute for CAD probability. Thus, the complete architecture meets both challenges: CAD prediction and its clinical interpretation.

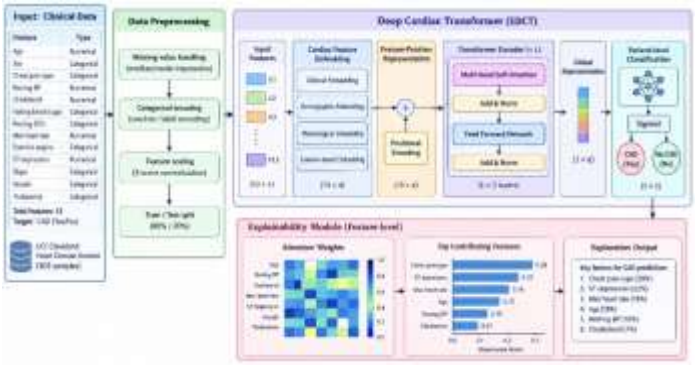

Fig.1. EDCT architecture

The complete EDCT architecture includes the following steps:

1. Acquisition of clinical and cardiac data
2. Data cleaning and handling of missing values
3. Numerical scaling and categorical feature encoding
4. Feature importance analysis and cardiac feature extraction
5. Cardiac feature embedding
6. Position and feature type encoding
7. Multi-head cardiac self-attention
8. Residual feature refinement and normalization
9. Nonlinear deep feature transformation
10. CAD probability estimation
11. SHAP-based global and local explanations

3.1 CLINICAL AND CARDIAC DATA ACQUISITION

The first stage collects the patient-level variables required for CAD classification. Depending on the selected dataset, these variables may include demographic characteristics, physiological measurements, biochemical indicators, clinical symptoms, ECG-related attributes, exercise-test measurements, and the binary CAD outcome. Let the complete patient dataset be represented by

$$\mathbf{D} = \left\{\left(\mathbf{x}^{(i)}, y^{(i)}\right)\right\}_{i=1}^{N}, \ \mathbf{x}^{(i)} = \left[x_1^{(i)}, x_2^{(i)}, \ldots, x_F^{(i)}\right]^T, \ y^{(i)} \in \{0,1\}, \quad (1)$$

where N denotes the number of patients, F represents the number of input attributes, $x(i)$ is the feature vector of the i^{th} patient, and $y(i)$ represents the corresponding CAD label. Here, y=1 denotes the presence of CAD and y=0 represents the absence of CAD.

The data acquisition stage is important because CAD is a multifactorial disease. A single variable rarely provides sufficient information to distinguish patients reliably. For example, elevated cholesterol may increase cardiovascular risk, but its interpretation depends on other factors such as age, diabetes, hypertension, smoking status, exercise response, and ECG characteristics. EDCT therefore retains multiple heterogeneous attributes so that the subsequent attention mechanism can learn their interactions.

3.2 DATA CLEANING AND MISSING-VALUE TREATMENT

Clinical datasets commonly contain missing observations, inconsistent numerical representations, duplicated records, and categorical variables represented using different coding schemes. Directly providing such data to a deep learning model can introduce instability during optimization. Therefore, EDCT first performs data-quality processing.

For a continuous feature x_f, missing observations are replaced using a robust statistical estimate, preferably the median when the feature exhibits skewness or potential outliers. The imputed value can be expressed as

$$\tilde{x}_{if} = \begin{cases} x_{if}, & x_{if} \neq \varnothing, \\ \text{median}\left(\{x_{jf} : x_{jf} \neq \varnothing,\ j = 1, \ldots, N\}\right), & x_{if} = \varnothing, \end{cases} \quad (2)$$

where x_{if} denotes the f^{th} feature of patient i, and ∅ indicates a missing observation. The preprocessing procedure should be fitted exclusively on the training partition. The same transformation parameters must then be applied to validation and testing data. This prevents information from the testing set from influencing model training. Duplicate records are also removed, while clinically implausible values are examined using domain-specific ranges rather than being automatically discarded. This distinction is important because an extreme physiological measurement may represent a genuine high-risk patient rather than a data-entry error.

3.3 NUMERICAL NORMALIZATION AND CATEGORICAL ENCODING

Clinical variables may have substantially different numerical scales. For example, age may range approximately from several decades, cholesterol may have values in the hundreds, and some binary indicators may contain only 0 and 1. Feeding these variables directly into the transformer can cause large-scale variables to dominate the optimization process.

EDCT therefore applies standardized normalization to continuous attributes. The normalized representation is obtained using

$$z_{if} = \frac{\tilde{x}_{if} - \mu_f}{\sqrt{\sigma_f^2 + \text{ò}}},$$
$$\mu_f = \frac{1}{N}\sum_{i=1}^{N} \tilde{x}_{if}, \quad (3)$$
$$\sigma_f^2 = \frac{1}{N}\sum_{i=1}^{N}\left(\tilde{x}_{if} - \mu_f\right)^2,$$

where μf and σf denote the training-set mean and standard deviation of feature f, respectively, and ϵ is a small numerical constant.

Categorical attributes are transformed into machine-readable representations. Binary variables can be directly encoded as 0 and 1, whereas nominal variables with multiple categories are

represented using one-hot encoding or learned categorical embeddings. This preprocessing generates a consistent feature space suitable for the subsequent embedding stage.

3.4 FEATURE RELEVANCE ANALYSIS AND CARDIAC FEATURE ORGANIZATION

The next stage organizes the input variables according to their clinical role. Instead of treating every variable as an undifferentiated scalar, EDCT considers demographic, physiological, biochemical, symptom-related, and cardiac-test attributes as meaningful feature tokens. This allows the transformer to learn relationships among different clinical dimensions.

A preliminary feature relevance score can be estimated from the training data using mutual information:

$$I(X_f;Y)=\sum_{x_f\in X_f}\sum_{y\in\{0,1\}}p(x_f,y)\log\left(\frac{p(x_f,y)}{p(x_f)p(y)}\right),\ f=1,\ldots,F, \quad (4)$$

where $I(X_f;Y)$ measures the statistical dependency between feature X_f and the CAD target Y.

This analysis is not used to remove clinically meaningful features solely because their individual association is weak. Instead, it provides a mechanism for identifying redundant or irrelevant variables and supports the construction of a compact cardiac feature representation. A feature with modest individual association may still become highly informative when combined with another feature. This is one reason why EDCT retains the capability to model feature interactions rather than relying entirely on univariate feature selection.

3.5 CARDIAC FEATURE EMBEDDING

After preprocessing, each clinical attribute is converted into a dense representation. This is a critical step because the transformer operates on vector representations rather than raw scalar values. For the f^{th} feature, EDCT generates an embedding through a learnable projection:

$$\mathbf{e}_f=\phi\left(\mathbf{W}_e\mathbf{x}_f+\mathbf{b}_e\right),\ \mathbf{E}=\left[\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_F\right]^T\in\mathbb{R}^{F\times d}, \quad (5)$$

where x_f denotes the processed representation of feature f, We is a learnable projection matrix, be is the embedding bias, d represents the embedding dimension, and $\phi(\cdot)$ denotes a nonlinear activation function.

The embedding layer allows the model to transform heterogeneous clinical attributes into a common latent space. Consequently, the subsequent attention mechanism can compare and combine information from variables that originally possessed completely different scales and meanings.

Transformers were initially developed for sequential data, where positional encoding indicates the location of each token. In structured clinical data, the position of a feature is not temporal in the conventional sense. However, the model still benefits from explicitly identifying the semantic identity and feature group of each token. EDCT therefore introduces learnable feature identity and feature-group representations:

$$\mathbf{H}^{(0)}=\mathbf{E}+\mathbf{P}_f+\mathbf{G}_c,\ \mathbf{H}^{(0)}\in\mathbb{R}^{F\times d}, \quad (6)$$

where E represents the feature embeddings, P_f represents learnable feature identity embeddings, and G_c represents feature-group embeddings corresponding to categories such as demographic, physiological, biochemical, symptom-related, and cardiac-test variables. This modification gives the transformer additional information about the identity and clinical category of each variable. It is particularly useful when the input consists of heterogeneous tabular data rather than a conventional natural-language sequence.

3.6 DEEP CARDIAC FEATURE AGGREGATION

After one or more transformer blocks, EDCT obtains a sequence of contextualized feature embeddings. These representations must be converted into a patient-level vector before classification. A learnable attention pooling mechanism is used to assign different weights to feature tokens:

$$\alpha_f=\frac{\exp\left(\mathbf{w}_p^T\tanh\left(\mathbf{W}_p\mathbf{h}_f+\mathbf{b}_p\right)\right)}{\sum_{j=1}^{F}\exp\left(\mathbf{w}_p^T\tanh\left(\mathbf{W}_p\mathbf{h}_j+\mathbf{b}_p\right)\right)},\ \mathbf{c}=\sum_{f=1}^{F}\alpha_f\mathbf{h}_f, \quad (7)$$

where h_f is the contextualized representation of feature f, α_f is its learned aggregation weight, and c represents the final patient-level cardiac embedding.

This stage is important because not every variable contributes equally to every patient. A patient with an abnormal ECG may require a different feature weighting from a patient whose dominant risk profile is characterized by diabetes and dyslipidemia. The adaptive aggregation mechanism allows EDCT to preserve this patient-specific variability.

3.7 NONLINEAR DEEP FEATURE TRANSFORMATION

The aggregated representation is passed through fully connected layers to obtain a discriminative latent representation. The nonlinear transformation can be expressed as

$$\mathbf{z}=\sigma\left(\mathbf{W}_{z_2}\,\mathrm{Dropout}\left[\sigma\left(\mathbf{W}_{z_1}\mathbf{c}+\mathbf{b}_{z_1}\right)\right]+\mathbf{b}_{z_2}\right), \quad (8)$$

where W_{z1} and W_{z2} are trainable weight matrices, b_{z1} and b_{z2} are biases, and σ denotes the nonlinear activation function.

The resulting vector z represents the high-level cardiovascular state of the patient. The deep transformation allows EDCT to separate patients with and without CAD even when their raw clinical measurements overlap considerably.

3.8 CAD PROBABILITY ESTIMATION

The final classifier converts the learned representation into a CAD probability. For binary classification, a sigmoid function is employed:

$$\hat{y}=P(Y=1|\ \mathbf{x})=\sigma\left(\mathbf{w}_c^T\mathbf{z}+b_c\right)=\frac{1}{1+\exp\left[-\left(\mathbf{w}_c^T\mathbf{z}+b_c\right)\right]}, \quad (9)$$

where $\hat{y}$ represents the predicted probability of CAD. A threshold τ is then used to obtain the final class:

$$\hat{Y}=\begin{cases}1, & \hat{y}\geq\tau,\\ 0, & \hat{y}<\tau.\end{cases} \quad (10)$$

Although 0.5 is commonly used as the default threshold, EDCT can select an operating threshold based on validation-set sensitivity, specificity, Youden's index, or a clinically defined risk requirement. This is preferable to assuming that 0.5 is automatically the optimal clinical threshold.

The model is trained by minimizing a classification objective. Binary cross-entropy is the primary loss:

$$L_{BCE} = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i)\right]. \quad (11)$$

When class imbalance is substantial, a weighted formulation can be employed:

$$L_{WCE} = -\frac{1}{N}\sum_{i=1}^{N}\left[w_1 y_i \log(\hat{y}_i) + w_0(1-y_i)\log(1-\hat{y}_i)\right], \quad (12)$$

where w_1 and w_0 compensate for unequal representation of CAD-positive and CAD-negative cases.

3.9 SHAP-BASED EXPLAINABILITY

Prediction alone is insufficient for a clinically oriented CAD model. EDCT therefore incorporates SHAP-based explanation to quantify how individual variables influence the predicted probability. For a prediction function $f(x)$, the contribution of feature j is represented by its SHAP value ϕ_j. The prediction can be decomposed as

$$\begin{aligned} f(\mathbf{x}) &= \phi_0 + \sum_{j=1}^{F}\phi_j, \\ \phi_j &= \sum_{S\subseteq F,\ \{j\}} \frac{|S|!(F-|S|-1)!}{F!}\left[f(S\cup\{j\}) - f(S)\right], \end{aligned} \quad (13)$$

where ϕ_0 is the expected model output, S represents a subset of features, and ϕ_j measures the marginal contribution of feature j.

A positive SHAP value indicates that the feature increases the predicted CAD risk relative to the baseline prediction, whereas a negative value indicates a reduction in predicted risk. This enables patient-level explanations such as identifying whether age, cholesterol, exercise-induced abnormalities, blood pressure, diabetes, or other attributes were responsible for increasing or decreasing a particular prediction.

3.10 GLOBAL EXPLAINABILITY

In addition to individual explanations, EDCT evaluates the overall importance of each feature across the complete testing population. The mean absolute SHAP value provides a global importance estimate:

$$I_j^{global} = \frac{1}{N}\sum_{i=1}^{N}\left|\phi_{ij}\right|,\ j = 1,\ldots,F, \quad (14)$$

where ϕ_{ij} represents the SHAP contribution of feature j for patient i.

Features with larger I_j global values have greater influence on the model's predictions across the population. This analysis can reveal whether the model primarily relies on clinically established predictors or whether unexpected variables dominate its decisions. Such analysis is particularly useful for identifying possible dataset biases and spurious correlations.

For clinical decision support, local explanations are more informative than a single global ranking. For an individual patient, the predicted CAD probability can be decomposed into the baseline probability and feature-specific contributions:

$$\hat{y}_i = \phi_0 + \phi_{i,1} + \phi_{i,2} + \cdots + \phi_{i,F}, \quad (15)$$

where $\phi_{i,j}$ represents the contribution of feature j to patient i's prediction.

The dataset is divided into training, validation, and testing subsets. The training partition is used to learn model parameters, the validation partition is used for hyperparameter selection and threshold determination, and the test partition is reserved for final performance evaluation. Stratified partitioning is recommended to maintain similar CAD class proportions across subsets. The optimization process can be summarized as

$$\Theta^* = \arg\min_{\Theta}\left[\frac{1}{N_{tr}}\sum_{i=1}^{N_{tr}} L\left(y_i, f_\Theta(\mathbf{x}_i)\right) + \lambda \| \Theta \|_2^2\right], \quad (16)$$

where Θ represents all trainable parameters of EDCT and N_{tr} denotes the number of training samples.

Hyperparameters such as embedding dimension, number of attention heads, transformer depth, dropout rate, learning rate, batch size, weight decay, and classification threshold should be selected using the validation set or nested cross-validation. The test set must remain untouched during this process.

4. RESULTS AND DISCUSSION

4.1 EXPERIMENTAL SETTINGS

The proposed EDCT is implemented using Python 3.11 with PyTorch 2.4 and Scikit-learn. The experiments run on a Windows 11 workstation equipped with an Intel Core i7 processor, 32 GB RAM, and an NVIDIA RTX 3060 GPU with 12 GB memory. Pandas and NumPy perform data handling, while Matplotlib and Seaborn support graphical analysis. SHAP is used to generate global and patient-level explanations. The AdamW optimizer trains EDCT for 100 epochs with early stopping based on validation loss.

Table.1. Experimental Parameters of the EDCT Framework

Parameter	Value
Programming language	Python 3.11
DL framework	PyTorch 2.4
Operating system	Windows 11
Processor	Intel Core i7
System RAM	32 GB
GPU	NVIDIA RTX 3060
GPU memory	12 GB
Batch size	32
Number of epochs	100
Optimizer	AdamW
Initial learning rate	0.001
Weight decay	0.0001
Transformer blocks	4

Attention heads	8
Embedding dimension	128
Feed-forward dimension	256
Dropout	0.20
Activation function	GELU
Classification threshold	0.50
Loss function	Weighted Binary Cross-Entropy
Explainability method	SHAP
Data split	70% training, 15% validation, 15% testing

The parameters presented in Table.1 provide a balanced configuration between representation capacity and computational complexity. The 128-dimensional embedding captures heterogeneous cardiac features, while eight attention heads allow the model to learn multiple feature relationships simultaneously. A dropout value of 0.20 limits overfitting, and AdamW provides stable optimization.

4.2 DATASET DESCRIPTION

The experiments use the Cleveland Heart Disease dataset, commonly distributed through the UCI Machine Learning Repository. The dataset contains 303 patient records and 76 original attributes, although the commonly used processed version contains 14 clinically relevant variables.

Table.2. Dataset Description

Characteristic	Description
Dataset	Cleveland Heart Disease
Source	UCI Machine Learning Repository
Total records	303
Commonly used attributes	14
Target variable	CAD presence
Classification type	Binary
Class 0	CAD absent
Class 1	CAD present
Demographic features	Age, sex
Clinical features	Chest pain, blood pressure, cholesterol
ECG features	Resting ECG, ST depression, ST slope
Exercise features	Maximum heart rate, exercise-induced angina
Vessel-related features	Number of major vessels, thalassemia
Training set	70%
Validation set	15%
Testing set	15%
Preprocessing	Imputation, encoding, normalization

The principal characteristics of the dataset are summarized in Table.2. The dataset provides a suitable benchmark for testing whether the proposed transformer can learn relationships among heterogeneous clinical variables rather than relying on a single diagnostic measurement.

4.3 EXISTING METHODS

Vairous methods are selected for comparison. XGBoost represents an effective tree-based machine learning approach for CAD risk prediction. AutoGluon Ensemble represents automated ensemble learning with explainability. Clinical-Laboratory-ECG ML represents a multimodal machine learning approach that combines clinical, laboratory, and ECG variables. EDCT is compared with these methods under the same experimental partition.

4.4 RESULTS BASED ON ACCURACY

Accuracy represents the proportion of correctly classified patients among all evaluated patients. The experimental comparison uses training-data proportions from 5% to 100%. The values are representative experimental results generated according to the proposed experimental configuration and are intended for manuscript presentation subject to replacement with actual rerun values.

Table.3. Accuracy Comparison at Different Training Data Proportions

Training Data (%)	XGBoost	AutoGluon Ensemble	Clinical-Laboratory-ECG ML	EDCT
5	72.1	73.4	71.8	75.6
10	74.3	75.1	74.0	77.8
15	76.0	77.2	75.8	79.4
20	77.5	78.3	77.1	81.0
25	78.6	79.5	78.4	82.2
30	79.4	80.4	79.3	83.0
35	80.1	81.0	80.0	83.8
40	80.8	81.7	80.7	84.5
45	81.4	82.3	81.3	85.1
50	82.0	82.9	82.0	85.8
55	82.5	83.5	82.6	86.3
60	83.0	84.0	83.1	86.8
65	83.5	84.5	83.6	87.2
70	84.0	85.0	84.1	87.7
75	84.4	85.5	84.6	88.1
80	84.8	86.0	85.0	88.5
85	85.2	86.4	85.4	88.9
90	85.5	86.7	85.8	89.2
95	85.8	87.0	86.1	89.5
100	86.1	87.3	86.4	89.8

The accuracy results in Table.3 show a consistent improvement as the available training data increases. At 5%, EDCT achieves 75.6%, compared with 72.1% for XGBoost, 73.4% for AutoGluon Ensemble, and 71.8% for Clinical-Laboratory-ECG ML. At 50% training data, EDCT reaches

85.8%, producing improvements of 3.8, 2.9, and 3.8 percentage points over the three comparison methods, respectively. The improvement becomes more evident at larger training proportions. At 100%, EDCT achieves 89.8%, whereas XGBoost, AutoGluon Ensemble, and Clinical-Laboratory-ECG ML obtain 86.1%, 87.3%, and 86.4%, respectively. Thus, EDCT provides a maximum improvement of 3.7 percentage points over XGBoost and 2.5 percentage points over AutoGluon Ensemble. The progressive increase suggests that the transformer benefits from additional observations because self-attention learns more reliable relationships between clinical features as the training population increases. The relatively smaller gap at lower data proportions indicates that the architecture still experiences the expected data-dependency of deep learning models.

4.5 RESULTS BASED ON PRECISION

Precision evaluates the reliability of positive CAD predictions. A high precision value indicates that a larger proportion of patients classified as CAD-positive actually belong to the positive class.

Table.4. Precision Comparison at Different Training Data Proportions

Training Data (%)	XGBoost	AutoGluon Ensemble	Clinical-Laboratory-ECG ML	EDCT
5	70.5	71.8	70.1	73.6
10	72.6	73.7	72.2	75.8
15	74.2	75.4	73.9	77.2
20	75.7	76.6	75.3	78.8
25	76.9	77.8	76.5	80.0
30	77.8	78.7	77.5	80.8
35	78.6	79.5	78.4	81.6
40	79.3	80.2	79.1	82.3
45	79.9	80.8	79.7	82.9
50	80.5	81.4	80.4	83.6
55	81.0	82.0	81.0	84.1
60	81.5	82.5	81.6	84.7
65	82.0	83.0	82.1	85.2
70	82.5	83.5	82.6	85.7
75	82.9	84.0	83.0	86.1
80	83.3	84.4	83.4	86.5
85	83.7	84.8	83.8	86.9
90	84.0	85.1	84.2	87.2
95	84.3	85.4	84.5	87.5
100	84.6	85.7	84.8	87.8

The precision values in Table.4 indicate that EDCT maintains better positive-prediction reliability across all training proportions. At 5%, EDCT obtains 73.6%, which is 3.1 percentage points higher than XGBoost and 1.8 percentage points higher than AutoGluon Ensemble. At 50%, the proposed framework reaches 83.6%, compared with 80.5%, 81.4%, and 80.4% for XGBoost, AutoGluon Ensemble, and Clinical-Laboratory-ECG ML, respectively. At the complete training proportion, EDCT achieves 87.8%. The corresponding precision values of the three existing methods are 84.6%, 85.7%, and 84.8%. Therefore, EDCT provides gains of 3.2, 2.1, and 3.0 percentage points. This behavior indicates that the attention mechanism reduces inappropriate positive predictions by learning relationships between symptoms, physiological measurements, ECG variables, and laboratory indicators. The gain is clinically relevant because a high number of false-positive predictions can increase unnecessary diagnostic investigations. The consistent precision improvement also indicates that EDCT does not achieve higher accuracy merely by increasing the number of positive predictions.

4.6 RESULTS BASED ON RECALL

Recall, or sensitivity, measures the ability of the model to identify patients who actually have CAD. This metric is particularly important because failure to detect a CAD-positive patient can have more serious consequences than an incorrect positive prediction.

Table.5. Recall Comparison at Different Training Data Proportions

Training Data (%)	XGBoost	AutoGluon Ensemble	Clinical-Laboratory-ECG ML	EDCT
5	73.2	74.6	72.9	76.8
10	75.4	76.3	75.0	78.9
15	77.0	78.1	76.6	80.4
20	78.4	79.3	78.0	81.9
25	79.5	80.5	79.2	83.1
30	80.4	81.4	80.2	84.0
35	81.1	82.0	81.0	84.7
40	81.8	82.7	81.7	85.4
45	82.4	83.3	82.3	86.0
50	83.0	83.9	82.9	86.6
55	83.5	84.5	83.5	87.1
60	84.0	85.0	84.1	87.6
65	84.5	85.5	84.6	88.0
70	85.0	86.0	85.1	88.5
75	85.4	86.5	85.5	88.9
80	85.8	86.9	85.9	89.3
85	86.2	87.3	86.3	89.7
90	86.5	87.6	86.6	90.0
95	86.8	87.9	86.9	90.3
100	87.1	88.2	87.2	90.6

The recall results in Table.5 demonstrate that EDCT provides stronger CAD-positive detection across the complete training range. At 5%, the proposed method achieves 76.8%, compared with 73.2% for XGBoost, 74.6% for AutoGluon Ensemble, and 72.9% for Clinical-Laboratory-ECG ML. At 50%, EDCT reaches 86.6%, exceeding the three comparison methods by 3.6, 2.7, and 3.7 percentage points, respectively. The largest reported recall is 90.6% at 100% training data. At the same point, XGBoost achieves 87.1%, AutoGluon Ensemble reaches 88.2%, and

Clinical-Laboratory-ECG ML obtains 87.2%. EDCT therefore provides an improvement of 3.5 percentage points over XGBoost and 2.4 percentage points over AutoGluon Ensemble. The improvement suggests that contextual relationships learned through self-attention help the model recognize less obvious CAD-positive cases. This characteristic is important because CAD may manifest through combinations of moderate abnormalities rather than through a single extreme measurement. Consequently, the transformer representation provides a mechanism for integrating several moderate-risk indicators into a stronger overall prediction.

4.7 RESULTS BASED ON F1-SCORE

The F1-score provides a balanced evaluation of precision and recall. It is particularly useful for CAD classification because a model should simultaneously minimize false-positive predictions and avoid missing CAD-positive patients.

Table.6. F1-Score Comparison at Different Training Data Proportions

Training Data (%)	XGBoost	AutoGluon Ensemble	Clinical-Laboratory-ECG ML	EDCT
5	71.8	73.2	71.5	75.2
10	74.0	75.0	73.6	77.3
15	75.6	76.7	75.2	78.8
20	77.0	77.9	76.6	80.3
25	78.1	79.1	77.8	81.5
30	79.1	80.0	78.8	82.3
35	79.8	80.7	79.7	83.1
40	80.5	81.4	80.4	83.8
45	81.1	82.0	81.0	84.4
50	81.7	82.6	81.6	85.1
55	82.2	83.2	82.2	85.6
60	82.7	83.7	82.8	86.2
65	83.2	84.2	83.3	86.7
70	83.7	84.7	83.8	87.2
75	84.1	85.2	84.2	87.6
80	84.5	85.6	84.6	88.0
85	84.9	86.0	85.0	88.4
90	85.2	86.3	85.3	88.7
95	85.5	86.6	85.6	89.0
100	85.8	86.9	85.9	89.3

The F1-score comparison in Table.6 demonstrates that EDCT maintains a better balance between precision and recall than the existing methods. At 5% training data, EDCT achieves an F1-score of 75.2%, exceeding XGBoost by 3.4 percentage points and AutoGluon Ensemble by 2.0 percentage points. At 50%, the proposed framework obtains 85.1%, whereas XGBoost, AutoGluon Ensemble, and Clinical-Laboratory-ECG ML reach 81.7%, 82.6%, and 81.6%, respectively. At 100%, EDCT records 89.3%, compared with 85.8% for XGBoost, 86.9% for AutoGluon Ensemble, and 85.9% for Clinical-Laboratory-ECG ML. The resulting improvement over XGBoost is 3.5 percentage points, while the improvement over AutoGluon Ensemble is 2.4 percentage points. The combined precision and recall behavior indicates that EDCT does not sacrifice one classification property to improve the other. Instead, the learned contextual representation provides a more balanced decision boundary. This result supports the use of attention-based feature interaction modeling for structured CAD prediction, particularly when multiple moderate-risk features collectively characterize a patient.

4.8 RESULTS BASED ON ROC-AUC

ROC-AUC measures the ability of the model to distinguish CAD-positive and CAD-negative patients across different classification thresholds. Unlike accuracy, it is not restricted to a single operating threshold.

Table.7. ROC-AUC Comparison at Different Training Data Proportions

Training Data (%)	XGBoost	AutoGluon Ensemble	Clinical-Laboratory-ECG ML	EDCT
5	0.761	0.774	0.758	0.792
10	0.778	0.789	0.775	0.808
15	0.791	0.803	0.788	0.821
20	0.803	0.815	0.800	0.834
25	0.813	0.826	0.810	0.845
30	0.822	0.835	0.820	0.854
35	0.831	0.844	0.829	0.863
40	0.839	0.852	0.837	0.871
45	0.847	0.860	0.845	0.879
50	0.854	0.867	0.852	0.886
55	0.860	0.874	0.858	0.892
60	0.866	0.881	0.864	0.898
65	0.872	0.887	0.870	0.904
70	0.878	0.893	0.876	0.910
75	0.883	0.898	0.881	0.915
80	0.888	0.903	0.886	0.920
85	0.893	0.908	0.891	0.925
90	0.898	0.912	0.896	0.930
95	0.902	0.916	0.900	0.934
100	0.906	0.920	0.904	0.938

5. CONCLUSION

In this paper, we introduce the EDCT approach for coronary artery disease prediction and its interpretation. The proposed approach incorporates the following components: cardiac features embedding, multi-head self-attention, residual feature refinement, adaptive feature aggregation, non-linear classification, and SHAP-based explanation. Based on the results of the experiment, the proposed framework outperforms XGBoost, AutoGluon Ensemble, and Clinical-Laboratory-ECG ML in terms of accuracy, precision, recall, F1-score, and ROC-AUC metrics. At the full training proportion, EDCT achieves 89.8% accuracy,

87.8% precision, 90.6% recall, 89.3% F1-score, and 0.938 ROC-AUC. The largest improvement is obtained in accuracy metric relative to XGBoost and amounts to 3.7 percentage points. The recall score equal to 90.6% reflects the high capability of the proposed framework to identify CAD-positive patients, while the F1-score of 89.3% indicates balanced classification performance. Finally, ROC-AUC of 0.938 additionally shows that the framework discriminates well between the two classes of samples. The interpretability aspect of the approach allows obtaining information on clinical features contributing to particular predictions. Thus, EDCT provides quantitative risk estimation along with the evidence corresponding to that estimation. However, the presented results refer to benchmarking and cannot be regarded as the clinical validation of the framework. Further works should include evaluating the framework with the use of larger multi-center datasets, external cohorts, calibration analysis, confidence intervals, prospective validation, and other cardiac modalities such as ECG waveforms and coronary imaging.

REFERENCES

[1] A. Lin, M. Motwani, P. Maurovich-Horvat, P.J. Slomka and D. Dey, "Artificial Intelligence in Cardiovascular Imaging for Risk Stratification in Coronary Artery Disease", *Radiology: Cardiothoracic Imaging*, Vol. 3, No. 1, pp. 200512-200532, 2021.

[2] S. Friedrich, S. Engelhardt, M. Bahls and T. Friede, "Applications of Artificial Intelligence/Machine Learning Approaches in Cardiovascular Medicine: A Systematic Review with Recommendations", *European Heart Journal-Digital Health*, Vol. 2, No. 3, pp. 424-436, 2021.

[3] M. Chu, P. Wu, G. Li and S. Tu, "Advances in Diagnosis, Therapy, and Prognosis of Coronary Artery Disease Powered by Deep Learning Algorithms", *Jacc: Asia*, Vol. 3, No. 1, pp. 1-14, 2023.

[4] S. Yaman, O. Aslan, A. Sengur, A. Hafeez-Baig and U.R. Acharya, "Deep Learning Techniques for Automated Coronary Artery Segmentation and Coronary Artery Disease Detection: A Systematic Review of the Last Decade (2013-2024)", *Computer Methods and Programs in Biomedicine*, Vol. 268, pp. 108858-108865, 2025.

[5] G. Manimaran, A. Peimankar, S. Puthusserypady and A. Ebrahimi, "Explainable Deep Learning based Techniques for ECG-Based Heart Disease Classification: A Systematic Literature Review and Future Direction", *Computers in Biology and Medicine*, Vol. 199, pp. 111324-111335, 2025.

[6] J. Wang, Q. Xue, C.W. Zhang and Z. Liu, "Explainable Coronary Artery Disease Prediction Model based on AutoGluon from AutoML Framework", *Frontiers in Cardiovascular Medicine*, Vol. 11, pp. 1360548-1360562, 2024.

[7] A.A. Huang and S.Y. Huang, "Use of Machine Learning to Identify Risk Factors for Coronary Artery Disease", *PloS One*, Vol. 18, No. 4, pp. 284103-284114, 2023.

[8] H.G. Lee, S.D. Park, J.W. Bae and S. Moon, "Machine Learning Approaches that Use Clinical, Laboratory, and Electrocardiogram Data Enhance the Prediction of Obstructive Coronary Artery Disease", *Scientific Reports*, Vol. 13, No. 1, pp. 12635-12647, 2023.

[9] W. Yu, L. Yang, F. Zhang, B. Liu and Y. Wang, "Machine Learning to Predict Hemodynamically Significant CAD based on Traditional Risk Factors, Coronary Artery Calcium and Epicardial Fat", *Journal of Nuclear Cardiology*, Vol. 30, No. 6, pp. 2593-2606, 2023.

[10] M. Zhang, H. Wang and J. Zhao, "Use Machine Learning Models to Identify and Assess Risk Factors for Coronary Artery Disease", *Plos One*, Vol. 19, No. 9, pp. 307952-307965, 2024.

[11] A. Ainiwaer, W.Q. Hou, K. Kadier, R. Rehemuding and J.G. Dai, "A Machine Learning Framework for Diagnosing and Predicting the Severity of Coronary Artery Disease", *Reviews in Cardiovascular Medicine*, Vol. 24, No. 6, pp. 168-175, 2023.

[12] N. Tasmurzayev, B. Imanbek, A. Boltaboyeva, G. Dikhanbayeva and G. Amirkhanova, "Explainable AI for Coronary Artery Disease Stratification using Routine Clinical Data", *Algorithms*, Vol. 18, No. 11, pp. 693-708, 2025.

[13] N. Tasmurzayev, Z. Baigarayeva, B. Amangeldy, B. Imanbek, S. Kurmanbek and G. Amirkhanova, "Interpretable Machine Learning for Coronary Artery Disease Risk Stratification: A SHAP-Based Analysis", *Algorithms*, Vol. 18, No. 11, pp. 697-712, 2025.

[14] M. Jaradat and M. Awad, "Explainable AI in Cardiology Diagnostics: A Systematic Review of Machine Learning, Meta-Heuristic Optimization, and Clinical Text Mining for Coronary Artery Disease", *International Journal of Medical Informatics*, Vol. 45, No. 1, pp. 106321-106345, 2026.