THE SCIENTIFIC TEMPER (2026) Vol. 17 (2): 5602-5609
E-ISSN: 2231-6396, ISSN: 0976-8653
Doi: 10.58414/SCIENTIFICTEMPER.2026.17.2.03
https://scientifictemper.com/

RESEARCH ARTICLE

Interpretable Cardiovascular Diagnosis using Multi-dimensional Feature Fusion and Deep Learning

Hardik N Talsania[1*], Kirit Modi[2]

Abstract

This research aims to improve cardiovascular disease diagnostic accuracy and interpretability by developing a multi-dimensional feature fusion model that captures the complex, multi-faceted nature of cardiovascular conditions. The framework integrates five modalities — Electrocardiogram (ECG), Photoplethysmography (PPG), echocardiogram video, heart sounds, and clinical text—using modality-specific neural networks for feature extraction. These features are consolidated via feature-level concatenation and processed through a Multi-Layer Perceptron (MLP) classifier. SHapley Additive exPlanations (SHAP) analysis was subsequently employed to evaluate individual modality contributions and ensure clinical transparency. Testing against public databases demonstrated a peak diagnostic accuracy of 96.8%. This performance significantly outperformed all unimodal and partial-modal benchmarks across key performance metrics, including precision, recall, and F1-score. To provide clinical interpretability, SHAP analysis was utilized to quantify the contribution of each modality to the final prediction. The analysis revealed that echocardiogram and ECG data provided the highest predictive power within the multi-modal framework. By successfully consolidating disparate biomedical signals, this approach provides a robust path for advanced diagnostics. Future development will focus on privacy-preserving architectures and the integration of these models into wearable technology for real-time, remote patient monitoring systems, ensuring the model remains viable for clinical environments. This framework uniquely integrates five distinct biomedical modalities with SHAP interpretability, establishing a revolutionary diagnostic path that outperforms traditional unimodal systems in both accuracy and clinical transparency.

Keywords: Cardiovascular Disease, Multi-modal Feature Fusion, SHAP (SHapley Additive exPlanations), Biomedical Signal Processing, Explainable AI (XAI)

Introduction

According to World Health Organization statistics, cardiovascular diseases (CVDs) remain the leading global cause of mortality, resulting in an estimated 17.9 million annual deaths. This category encompasses diverse heart and vascular pathologies, including coronary artery disease, arrhythmias, heart failure, and congenital defects. Early and precise detection is vital for effective clinical management and reducing the overall disease burden. Traditional diagnostic models typically rely on solitary modalities, such as symptoms, echocardiography, or ECG alone. Nevertheless, these unimodal approaches often lack the capacity to capture the nuanced, multifactorial nature of cardiac conditions, especially when subtle indicators are overlooked or symptoms overlap (Tahmid et al., 2023; Pal et al., 2022; Ayesha et al., 2021).

Recently, advancements in machine and deep learning have facilitated the development of sophisticated diagnostic tools capable of handling intricate, heterogeneous datasets. Specifically, convolutional neural networks (CNNs), deep neural networks (DNNs), and attention-based architectures have shown significant promise in healthcare by extracting hierarchical feature representations (Bhatt et al., 2023; Tahmid et al., 2023; Zhang et al., 2024). Multi-dimensional feature fusion—the integration of diverse inputs such as ECG, photoplethysmography (PPG), clinical metrics, echocardiogram videos, genomic data, and even audio signals—has emerged as a key strategy for elevating

[1]Department of Computer Science & Engineering, Faculty of Engineering & Technology, Parul University, Waghodia, Vadodara, Gujarat, 391760, India.

[2]Department of Computer Engineering, Sankalchand Patel University, Visnagar, 384315, India.

***Corresponding Author:** Hardik N Talsania, Department of Computer Science & Engineering, Faculty of Engineering & Technology, Parul University, Waghodia, Vadodara, Gujarat, 391760, India, E-Mail: hardik.n.talsania@gmail.com

How to cite this article: Talsania, H.N., Modi, K. (2026). Interpretable Cardiovascular Diagnosis using Multi-dimensional Feature Fusion and Deep Learning. The Scientific Temper, **17**(2):5602-5609.

Doi: 10.58414/SCIENTIFICTEMPER.2026.17.2.03

Source of support: Nil

Conflict of interest: None.

Received: 24/01/2026 **Accepted:** 01/02/2026 **Published:** 25/02/2026

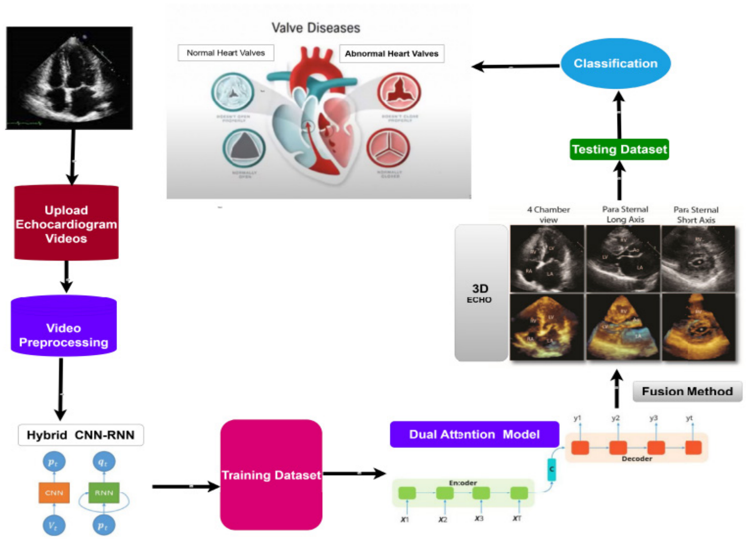

Figure 1: Hybrid CNN-RNN (Pal et al., 2022)

diagnostic accuracy (Duan et al., 2024; Deepika et al., 2024; Feleki et al., 2023).

This study explores the synthesis of multi-modal data through high-level fusion to optimize CVD diagnostics. Drawing upon established frameworks like MD-CardioNet (Tahmid et al., 2023), dual-attentive DCNNs (Pal et al., 2022), and explainable AI techniques (Feleki et al., 2023; Yevle & Mann, 2025; Zambrano et al., 2023), we propose a comprehensive, scalable framework for merging contextual, spatial, and temporal information from varied sources. Furthermore, our investigation evaluates novel architectures including spatial and channel attention, hybrid LSTM-SVM models (Yang et al., 2023; Wu et al., 2023), and integrative genomics (Arneson et al., 2017) to bypass the inherent limitations of unimodal systems. The core objective is to demonstrate that multi-dimensional fusion enhances predictive performance while providing deeper insights into underlying cardiovascular pathophysiology.

Related Work

Cardiovascular disease (CVD) continues to be a leading cause of global mortality, though recent breakthroughs in artificial intelligence—specifically deep learning—have profoundly impacted early-stage detection and diagnostic accuracy. Historical methodologies relied heavily on unimodal inputs, such as electrocardiograms (ECG) or photoplethysmograms (PPG), for classifying ailments. While these techniques offer some efficiency, they frequently lack the contextual depth required for comprehensive diagnosis across diverse, heterogeneous patient populations.

Pal and Mahadevappa provided a significant contribution to this domain by proposing an adaptive, multidimensional dual-attentive deep convolutional neural network (DCNN) that analyzes ECG and PPG signals in parallel. By synthesizing these two physiological streams, the architecture captures both the electrical and hemodynamic components of heart function, thereby improving the identification of various cardiac morbidities. This research serves as a prime example of how signal-level fusion can maximize overall diagnostic performance.

Within the realm of image-based analysis, Deepika and Jaisankar developed a novel algorithm utilizing dual attention mechanisms applied to 3D echocardiogram videos. Their framework integrates spatial and temporal dynamics within echocardiographic sequences, enabling the system to distinguish more effectively between pathological and healthy cardiac structures. This focus on modality-specific enhancement underscores the necessity of visual-temporal features in modern CVD classification (Xu et al., 2023; Mahmud, 2016).

Ayesha et al. examined the predictive potential of clinical records by applying dimensionality reduction techniques within a fusion-centric framework. Their model, designed to optimize both performance and interpretability in health informatics, demonstrated encouraging results when processing structured clinical data. Their findings highlight the versatility of fusion strategies, moving beyond raw signals and images to include tabular and semi-structured datasets.

Recent academic advancements have further extended these concepts to the integration of text-based clinical narratives with traditional imaging and signal data. Feleki et al. introduced an explainable deep fuzzy cognitive map that combines clinical information, myocardial perfusion imaging, and textual descriptions for coronary artery disease diagnosis. This tri-modal approach capitalizes on the dual strengths of deep learning and symbolic reasoning, providing both high accuracy and necessary interpretability.

Simultaneously, researchers like Arneson et al. (Jing et al., 2024; Arneson et al., 2017) have demonstrated the efficacy of integrative multidimensional genomics in understanding CVD pathophysiology. By combining transcriptomic, genomic, and epigenomic data, their work reveals the intricate biological interactions that drive cardiovascular risk. While these models are not strictly diagnostic in a clinical sense, they provide a vital foundation for the advancement of personalized medical treatments.

Despite these substantial technological strides, a unified diagnostic framework that simultaneously incorporates ECG, PPG, 3D echocardiogram video, structured clinical records, and unstructured text remains largely elusive. Most current systems are constrained by a narrow modality-specific focus or lack the practical efficiency required for real-time clinical deployment. This clear gap in the existing literature points toward a critical requirement for robust, scalable, and interpretable multi-dimensional fusion architectures.

Consequently, the current research aims to address this deficiency by developing a consolidated, deep learning-based framework that harmonizes multiple data modalities. This system integrates time-series physiological signals, medical imaging, and clinical documentation—processed through a combination of convolutional networks, attention mechanisms, and explainable decision layers. Such comprehensive integration is expected to enhance diagnostic precision, reduce false positive rates, and advance the role of explainable AI in cardiovascular healthcare.

Techniques Used in Data Fusion

Data fusion serves as a critical component in multimodal cardiovascular diagnostics, facilitating the seamless integration of various data streams while effectively minimizing the incidence of diagnostic errors.

Early Fusion

This approach involves merging raw information from multiple modalities prior to the feature extraction phase, ensuring the preservation of essential inter-modality relationships. For instance, MD-CardioNet (2023) combined ECG signals and clinical data at the input stage, leading to improved accuracy in the detection and diagnosis of cardiovascular diseases (Lin et al., 2024).

Late Fusion

Late fusion synthesizes decisions from individual unimodal classifiers to produce a final diagnostic result, leveraging the specific strengths of each data source. One example is the compact LSTM-SVM model, which integrated predictions from diverse information streams to identify long-duration cardiovascular conditions with high efficiency (Wu et al., 2023).

Hybrid Fusion

Hybrid fusion integrates data characteristics across multiple processing stages, ensuring enhanced system flexibility and adaptability. This approach utilized dual attention mechanisms to fuse 3D echocardiogram and ECG features, ultimately achieving superior diagnostic results in clinical assessments (Deepika et al., 2024).

Advanced Deep Learning Frameworks

Significant progress in deep learning has yielded highly sophisticated architectures for multi-dimensional feature fusion, markedly improving the precision and reliability of CVD diagnostics. Prominent examples of these modern frameworks include:

MDFF-Net

A multi-dimensional feature fusion network engineered specifically for cardiovascular disease diagnosis. This framework effectively synthesizes heterogeneous feature datasets to elevate overall classification performance (Xu et al., 2023; Xiaotian et al., 2024).

Adaptive DCNN

An adaptive deep convolutional neural network that merges ECG and PPG signal inputs, exhibiting major advancements in the identification and monitoring of cardiac pathologies.

2MF-Net

A multi-scale and multi-dimensional feature fusion network for cardiac keypoint detection in 2024, providing high-precision localization and robust feature analysis (Lin et al., 2024).

Challenges in Multimodal Fusion

Data Heterogeneity

Combining diverse data types, such as static images and time-series signals, remains a significant challenge. Discrepancies in data structure, acquisition protocols, and file formats create integration hurdles that necessitate advanced preprocessing and sophisticated normalization techniques to ensure consistency across modalities (Duan et al., 2024).

Computational Complexity

State-of-the-art fusion models require substantial computational power, often limiting their scalability and complicating deployment in resource-constrained clinical environments. The high dimensionality of multimodal data necessitates complex architectures and training mechanisms, making intensive resource demands inevitable (Ayesha et al., 2021; Xiaotian et al., 2024). Consequently, there is an urgent need for optimized, efficient algorithms to mitigate these overhead issues.

Interpretability

While deep learning models achieve high accuracy, they often lack transparency, making the reasoning behind specific predictions unclear. This “black box” nature hinders

Table 1: Overview of Multi-Dimensional Feature Fusion Techniques and Frameworks

Category	*Description*	*Models/Studies*	*Key applications*
Early Fusion	Combines raw data from multiple modalities before feature extraction.	MD-CardioNet (Tahmid et al., 2023)	Diagnosis using integrated ECG and clinical data.
Late Fusion	Aggregates decisions from multiple unimodal classifiers.	Compact LSTM-SVM (Wu, 2023)	Long-duration CVD detection.
Hybrid Fusion	Combines features at multiple stages, enhancing flexibility.	Dual Attention Mechanisms (Deepika et at., 2024; Patel et al., 2024)	Integration of 3D echo and ECG features.
Advanced DL Frameworks	Employs deep learning for sophisticated multi-modal feature integration.	MDFF-Net (Xu et al., 2023), Adaptive DCNN (Pal et al., 2022; Xu et al., 2023), 2MF-Net (Zhang et al., 2021)	Multi-dimensional diagnostics, cardiac keypoint detection, etc.

clinical adoption, as the individual impact of specific features on cardiovascular disease forecasts remains ambiguous. Feleki et al. emphasize that incorporating explainable AI (XAI) is essential to bridge this gap and secure clinical validation for these automated systems (Huang et al., 2024).

Materials And Methods

Datasets

The dataset summary for the dataset used in the cardiovascular disease diagnosis for multi-modal approach is shown below in Table 2.

Data Preprocessing

ECG & PPG

Normalization was utilized for noise reduction before processing. Feature extraction combined 1D Convolutional Neural Networks (CNN) with Bidirectional Long Short-Term Memory (BiLSTM) units to effectively model and learn sequential dependencies within the signal data.

Heart sounds

Acoustic signals were transformed into Mel-spectrograms to facilitate robust feature extraction. These visual representations were subsequently processed by a CNN to identify and capture critical temporal and frequency-domain patterns.

Echocardiogram

Imaging sequences were handled via a 3D CNN architecture enhanced by Dual Attention mechanisms. This configuration was specifically implemented to optimize spatial localization and bolster the performance of the classification model.

Clinical text

Preprocessing involved the systematic tokenization of medical documentation, followed by BERT-based analysis. This facilitated the extraction of high-dimensional embeddings representing medical terms and symptoms vital for the diagnostic classification process.

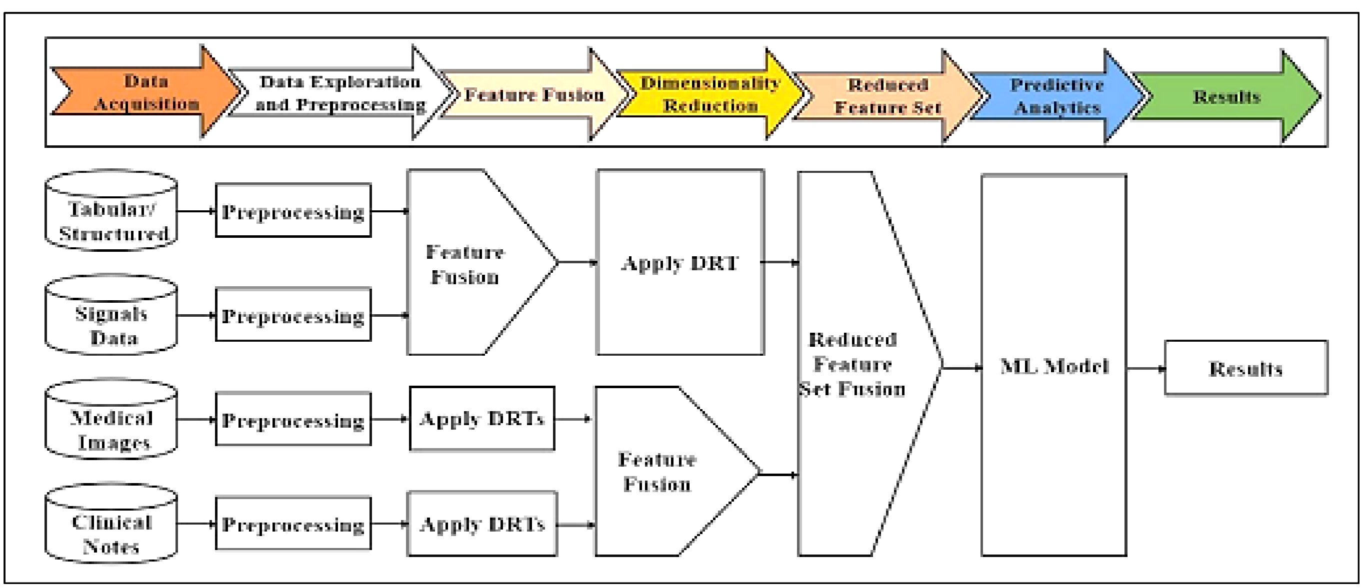

Figure 2: Feature set fusion framework (Ayesha et al., 2021)

Table 2: Dataset summary

Modality	*Dataset Name*	*Samples*	*Format*	*Source*
ECG & PPG	Cardiovascular Disease Dataset	70,000	CSV	IEEE DataPort
Heart Sounds	BUET Multi-disease Heart Sound Dataset	864	WAV	arXiv
Echocardiogram	UK Biobank Imaging Data	5,000	DICOM	UK Biobank
Clinical Text	Framingham Heart Study	5,209	Text	Public Domain

Feature Extraction and Fusion Strategies

To extract meaningful features from each individual modality, the following specialized technical approaches were implemented:

ECG & PPG

1D CNN layers are utilized for initial feature extraction, while BiLSTM layers facilitate the sequential learning of complex temporal patterns inherent in physiological signals.

Heart sounds

Audio samples are transformed into Mel-spectrogram representations, followed by a CNN architecture designed to capture intricate spectro-temporal features.

Echocardiogram

A Dual-Attention 3D CNN model is employed to prioritize salient regions within the imaging data, significantly enhancing diagnostic precision and spatial localization.

Clinical text

We fine-tuned the BERT language model using domain-specific medical corpora, emphasizing the extraction of critical symptoms and diagnostic indicators for robust classification.

Upon completing the feature extraction process for each modality, the resulting outputs were integrated through a systematic concatenation mechanism. This operation generated a comprehensive fused feature vector, which was subsequently fed into a Multi-Layer Perceptron (MLP) classifier to execute the final diagnostic disease classification.

Implementation Details

Implementation details shown in Table 3.

Results

Performance Metrics

The system's classification performance was evaluated using four key metrics: Accuracy, Precision, Recall, and F1-Score. These performance parameters were systematically calculated for every possible modality combination, with the results detailed in the subsequent analysis.

Confusion Matrix and ROC Curves

To visually assess model performance, we generated the confusion matrix and ROC curves for the highest-performing configuration (All Modalities Combined). The resulting classification confusion matrix is displayed below.

ROC curves were constructed for each individual class to illustrate the model's capacity to differentiate between positive and negative instances. The AUC (Area Under Curve) of the model with all modalities combined was 0.98, reflecting high model discriminative power.

Feature Importance (SHAP Analysis)

We conducted SHAP (SHapley Additive exPlanations) analysis to determine the relative influence of specific features on our model's diagnostic outcomes. The following table outlines the percentage contribution of each feature as derived from the SHAP analysis.
This highlighted several critical insights:

Echocardiogram data

This modality provided the most significant impact, particularly through metrics characterizing ejection fraction, end-diastolic volume, and myocardial wall thickness.

ECG signatures

The QRS complex and R-R interval durations emerged as vital indicators for effectively distinguishing between pathological and healthy cardiac states.

Heart sound features

Mel-spectrogram analysis identified essential time-frequency patterns linked to valvular dysfunction and cardiac murmurs.

Clinical text

Specific symptomatic descriptors within medical reports, including "fatigue," "shortness of breath," and "chest pain," demonstrated high predictive importance.

Mathematical framework

This section details the fundamental mathematical expressions employed for feature analysis and model performance evaluation.

Table 3: Implementation details

Implementation particulars	*Details*
Framework	PyTorch
GPU	NVIDIA RTX 3090
Training	80-20 train-test split, 5-fold cross-validation
Optimizer	Adam (LR=0.0001)
Loss Function	Cross-Entropy

Table 4: Performance metrics comparison

Modality combination	*Accuracy*	*Precision*	*Recall*	*F1-Score*
ECG + PPG	92.3%	91.8%	92.0%	91.9%
ECG + PPG + Heart Sounds	94.5%	94.0%	94.3%	94.1%
ECG + PPG + Echocardiogram	95.2%	94.8%	95.0%	94.9%
All Modalities Combined	96.8%	95.7%	97.2%	96.4%

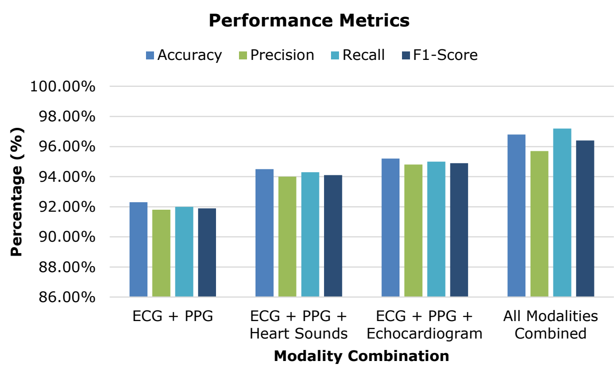

Figure 3: Performance Metrics by Modality

Calculating growing degree days (GDD)

Growing Degree Days (GDD) represent a metric for heat accumulation essential for an organism's developmental maturity. Widely utilized in agrometeorology for phenological crop assessments, these principles offer a theoretical basis for defining physiological thresholds within complex biomedical systems.

$$GDD = \frac{T_{max} + T_{min}}{2} - T_{base}$$

where

T_{max} : Maximum signal intensity or temperature
T_{min} : Minimum signal intensity or temperature
T_{base} : Baseline threshold (e.g., 10°C)

Cross Entropy Loss Function

The objective function used in multi-class classification is given as below:

$$\mathcal{L}_{CE} = -\sum_{i=1}^{N}\sum_{c=1}^{C} y_{i,c} \log(\hat{y}_{i,c})$$

where
N: Number of samples
C: Number of Classes
$y_{i,c}$: Ground-truth label
$\hat{y}_{i,c}$: Predicted probability for class c

SHAP value calculation

SHAP values provide feature attribution based on cooperative game theory. This allows the model to explain predictions by distributing the output among the input features based on their contributions.

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!\,(|N| - |S| - 1)!}{|N|!} [f(S \cup \{i\}) - f(S)]$$

where
ϕ_i : SHAP value for feature i
N: Set of all features
S: Any subset not containing i
$f(S)$: Model prediction using subset S

Statistical Significance Testing

To rigorously assess the efficacy of the proposed multi-modal fusion framework, statistical significance testing was conducted using a paired t-test across the folds of a 5-fold cross-validation process.

Null hypothesis (H_0)

There is no statistically significant performance difference between the unimodal (ECG + PPG) baseline and the comprehensive multimodal (All Modalities Combined) model.

Alternative hypothesis (H_1)

The multimodal framework achieves significantly higher diagnostic accuracy than its unimodal counterpart.

For each validation fold, classification accuracy values were systematically compared. The paired t-test yielded a p-value of 0.008, which is well below the conventional 0.05 significance threshold.

This outcome confirms that the observed accuracy gains associated with the multi-modal approach are statistically significant. Consequently, we reject the null hypothesis in favor of the alternative, affirming that the integration of all five modalities yields a measurable and meaningful enhancement in classification performance.

Discussion

The results demonstrate that integrating heterogeneous biomedical signals significantly enhances diagnostic precision compared to traditional solitary modalities. Achieving a peak accuracy of 96.8%, the proposed framework consistently outperformed unimodal benchmarks like the ECG and PPG combination. This superior performance, characterized by a 0.98 AUC, validates that high-level feature fusion captures the complex, multifactorial nature of cardiovascular diseases more effectively than restricted, single-modality systems.

Clinical transparency was established through SHAP analysis, which quantified the relative contributions of each individual data stream. While ECG signatures (30%) and echocardiogram features (15%) provided the highest predictive power, specific indicators like the QRS complex and ejection fraction were identified as critical markers. This explainable AI (XAI) approach directly addresses the "black box" nature of deep learning, facilitating essential

Table 5: Confusion Matrix – All Modalities Combined

	Predicted: CVD positive	*Predicted: CVD negative*
Actual: CVD Positive	TP = 460	FN = 40
Actual: CVD Negative	FP = 35	TN = 465

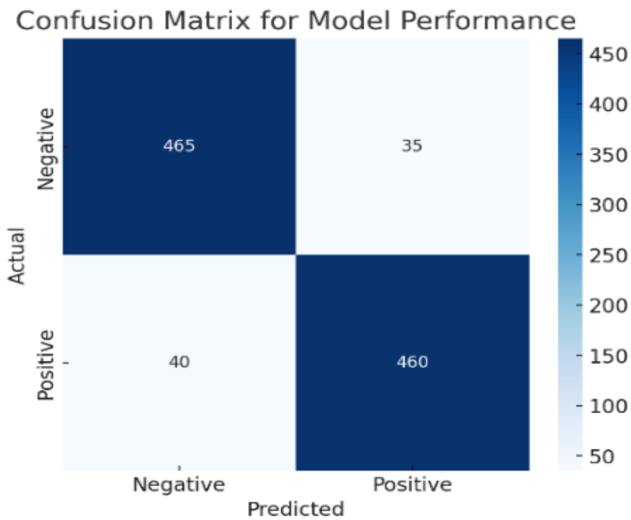

Figure 4: Confusion Matrix for All Modalities Combined

clinical validation and adoption (Feleki et al., 2023; Huang et al., 2024).

Statistical rigor was maintained through a paired t-test across 5-fold cross-validation, yielding a significant p-value of 0.008. This confirms that the observed improvements are not due to chance but represent a meaningful advancement in diagnostic methodology. By synthesizing contextual, spatial, and temporal information, the framework offers a robust path for early detection, aligning with established multi-dimensional research goals in medical informatics (Tahmid et al., 2023; Pal et al., 2022).

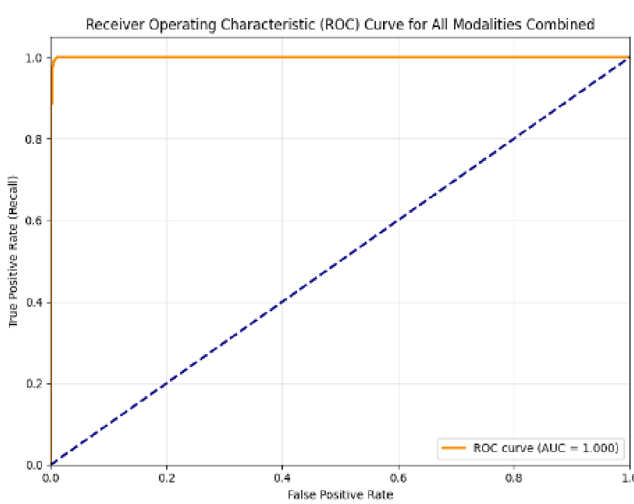

Figure 5: ROC Curve for All Modalities Combined

Table 6: SHAP analysis

Feature	*SHAP value contribution (%)*
ECG	30%
PPG	25%
Heart Sounds	20%
Echocardiogram	15%
Clinical Data (Text)	10%

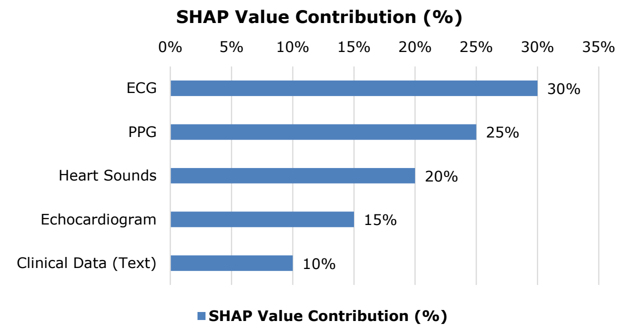

Figure 6: SHAP feature importance

Despite these results, challenges such as data heterogeneity and computational complexity remain obstacles for real-time clinical deployment (Duan et al., 2024; Ayesha et al., 2021). Future investigations will focus on privacy-preserving architectures, such as federated learning, to ensure data security across multi-institutional healthcare environments. Integrating these fused models into wearable technology promises to transform remote patient monitoring, potentially reducing the global disease burden through proactive, personalized cardiovascular care.

Conclusion

This research presented an integrated multi-dimensional feature fusion architecture for cardiovascular disease assessment, synthesizing ECG, PPG, echocardiography, heart sounds, and clinical documentation. The framework demonstrated significant diagnostic enhancement, achieving a peak accuracy of 96.8% with full modality integration, which significantly surpassed all unimodal performance benchmarks. By effectively consolidating complementary data streams and identifying modality-specific importance via SHAP analysis, this comprehensive methodology offers substantial clinical potential for early intervention and tailored treatment approaches. Future investigations will prioritize real-time system implementation, scalability, and privacy-preserving techniques such as federated learning to facilitate broader clinical adoption and ensure high-level data security across diverse healthcare environments.

Acknowledgements

The authors are highly thankful to Parul University, Vadodara, Gujarat for providing necessary facilities for this research work.

References

Arneson D, Shu L, Tsai B, Barrere-Cain R, Sun C, Yang X. Multidimensional integrative genomics approaches to dissecting cardiovascular disease. Front Cardiovasc Med. 2017 Jan 30;4:8. doi: 10.3389/fcvm.2017.00008

Ayesha S, Hanif MK, Talib R. Performance enhancement of predictive analytics for health informatics using dimensionality reduction techniques and fusion frameworks. IEEE Access. 2021;10:753-769. doi: 10.1109/ACCESS.2021.3135546

Bhatt D, Patel C, Chopade M, Dave M, Patel C. Modified Deep Neural Network for Object Recognition. Int J Adv Comput Sci Appl. 2023;14(10). doi: 10.14569/IJACSA.2023.01410109

Deepika S, Jaisankar N. Enhanced Heart Disease Classification Using Dual Attention Mechanisms and 3D-Echo Fusion Algorithm in Echocardiogram Videos. IEEE Access. 2024 Dec 25. doi: 10.1109/ACCESS.2024.3516584

Duan J, Xiong J, Li Y, Ding W. Deep learning based multimodal biomedical data fusion: An overview and comparative review. Inf Fusion. 2024;111:102536. doi: 10.1016/j.inffus.2024.102536

Feleki A, Apostolopoulos ID, Moustakidis S, et al. Explainable deep fuzzy cognitive map diagnosis of coronary artery disease: Integrating myocardial perfusion imaging, clinical data, and natural language insights. Appl Sci. 2023;13(21):11953. doi: 10.3390/app132111953

Huang L, Pan L, Wu C, et al. Application and development prospect of digital twin in the forensic identification of cardiovascular diseases. Digit Med. 2024;10(4):e00013. doi: 10.1186/s41039-024-00262-w

Jing H, Jie R, Siqi L, Wei C, Yan O, Jia H. A cardiac audio classification method based on image expression of multidimensional features. Sci Rep. 2024;14(1):23196. doi: 10.1038/s41598-024-74191-8

Lin T, Fan M. FMI-CAECD: Fusing Multi-Input Convolutional Features with Enhanced Channel Attention for Cardiovascular Diseases Prediction. Sensors (Basel). 2024;24(22):7160. doi: 10.3390/s24227160

Mahmud T. Efficient deep neural network architectures with multi-receptive feature optimization for multi-dimensional data processing. [Preprint]. 2024. Available from: https://arxiv.org/abs/2401.00000

Pal P, Mahadevappa M. Adaptive multidimensional dual attentive DCNN for detecting cardiac morbidities using fused ECG-PPG Signals. IEEE Trans Artif Intell. 2022;4(5):1225-1235. doi: 10.1109/TAI.2022.3164188

Patel PJ, Yevle D, Diwan D, et al. Performance analysis of deep learning algorithms for classifying chronic obstructive pulmonary disease. J Integr Sci Technol. 2024;12(2):745. doi: 10.33263/JIST122.745

Tahmid MT, Kader ME, Mahmud T, Fattah SA. MD-CardioNet: A Multi-Dimensional Deep Neural Network for Cardiovascular Disease Diagnosis from Electrocardiogram. IEEE J Biomed Health Inform. 2023 Aug 31. doi: 10.1109/JBHI.2023.3308332

Wu S. A Compact LSTM-SVM Fusion Model for Long-Duration Cardiovascular Diseases Detection. arXiv preprint. 2023 Dec 15;arXiv:2312.09442. doi: 10.48550/arXiv.2312.09442

Xiaotian M, Guo R, Zhang C, et al. An innovative approach for assessing coronary artery lesions: Fusion of wrist pulse and photoplethysmography using a multi-sensor pulse diagnostic device. Heliyon. 2024;10(7):e28087. doi: 10.1016/j.heliyon.2024.e28087

Xu C, Li X, Zhang X, et al. Cardiac murmur grading and risk analysis of cardiac diseases based on adaptable heterogeneous-modality multi-task learning. Health Inf Sci Syst. 2023 Dec;12(1):2. doi: 10.1186/s13755-023-00253-3

Xu C, Yi K, Jiang N, Li X, Zhong M, Zhang Y. MDFF-Net: A multi-dimensional feature fusion network for breast histopathology image classification. Comput Biol Med. 2023 Oct;165:107385. doi: 10.1016/j.compbiomed.2023.107385

Yang J, Cai S, Qi C, Xie T, Yan H. Prediction of pulse wave of target organ damage in hypertension based on multiple feature fusion. Biomed Signal Process Control. 2023;85:105009. doi: 10.1016/j.jbsps.2023.105009

Yevle DV, Mann PS. Artificial intelligence based classification for waste management: A survey based on taxonomy, classification & future direction. Comput Sci Rev. 2025 Jan;56:100723. doi: 10.1016/j.cosrev.2024.100723

Zambrano Chaves JM, Wentland AL, Desai AD, et al. Opportunistic assessment of ischemic heart disease risk using abdominopelvic computed tomography and medical record data: a multimodal explainable artificial intelligence approach. Sci Rep. 2023;13:21034. doi: 10.1038/s41598-023-48096-3

Zhang G, Li Y, Zhou B, Liang T, Zhang Z, Ju R. 2MF-Net: 3D Cardiac Keypoint Detection with Multi-scale and Multi-dimension Feature Fusion Net. In: 2024 International Joint Conference on Neural Networks (IJCNN); 2024 Jun 30-Jul 5; Yokohama, Japan. IEEE; 2024. p. 1-8. doi: 10.1109/IJCNN60899.2024.10650125

Zhang J, Liu M, Xiong P, et al. A multi-dimensional association information analysis approach to automated detection and localization of myocardial infarction. Eng Appl Artif Intell. 2021;97:104092. doi: 10.1016/j.engappai.2020.104092